

GESELLSCHAFT
FÜR INFORMATIK



Gesellschaft für Informatik e.V. (GI) (Hrsg.)

INFORMATIK 2021

Computer Science & Sustainability

27. September – 01. Oktober 2021
Berlin

Gesellschaft für Informatik e.V. (GI)

Lecture Notes in Informatics (LNI) - Proceedings

Series of the Gesellschaft für Informatik (GI)

Volume P-314

ISBN 978-3-88579-708-1

ISSN 1617-5468

Volume Editor

Gesellschaft für Informatik e.V. (GI)

Ahrstr.45, 53175 Bonn, Germany

gs@gi.de

Series Editorial Board

Andreas Oberweis, KIT Karlsruhe,

(Chairman, andreas.oberweis@kit.edu)

Torsten Brinda, Universität Duisburg-Essen, Germany

Dieter Fellner, Technische Universität Darmstadt, Germany

Ulrich Flegel, Infineon, Germany

Ulrich Frank, Universität Duisburg-Essen, Germany

Michael Goedicke, Universität Duisburg-Essen, Germany

Ralf Hofestädt, Universität Bielefeld, Germany

Wolfgang Karl, KIT Karlsruhe, Germany

Michael Koch, Universität der Bundeswehr München, Germany

Peter Sanders, Karlsruher Institut für Technologie (KIT), Germany

Andreas Thor, HFT Leipzig, Germany

Ingo Timm, Universität Trier, Germany

Karin Vosseberg, Hochschule Bremerhaven, Germany

Maria Wimmer, Universität Koblenz-Landau, Germany

Dissertations

Steffen Hölldobler, Technische Universität Dresden, Germany

Thematics

Agnes Koschmider, Universität Kiel, Germany

Seminars

Judith Michael, RWTH Aachen, Germany

© Gesellschaft für Informatik, Bonn 2021

printed by Köllen Druck+Verlag GmbH, Bonn



This book is licensed under a Creative Commons BY-SA 4.0 licence.

Vorwort

Die 51. Jahrestagung der Gesellschaft für Informatik bzw. aus Berlin, die [INFORMATIK 2021](#), steht unter dem Themenschwerpunkt „Computer Science & Sustainability“. Angesichts der pandemischen Situation wird die INFORMATIK 2021 wie schon im Vorjahr weitestgehend virtuell durchgeführt, mit Sendestudios im Weizenbaum-Institut und in der German International University in Berlin. Die Leitfragen der Konferenz sind: Welchen Beitrag kann die Informatik für eine nachhaltige Entwicklung leisten? Wie muss sich die Disziplin selbst aufstellen, um nachhaltiger zu agieren?

Themen wie Green IT, Ressourcenschutz, Einsatz intelligenter Technologien und Optimierung von Systemen beherrschen die Diskussion, wie nachhaltig die Informatik ist und wo sie Hilfestellungen für ein enachhaltigere Lebens-, Arbeits- und Wirtschaftsweise leisten kann. Dabei adressiert die INFORMATIK 2021 verschiedene Handlungsfelder in der Informatik an der Schnittstelle zur Nachhaltigkeit: ökologische, ökonomische, soziale und technologische. Handlungsleitend sind dabei die 17 Nachhaltigkeitsziele der Vereinten Nationen.

Die INFORMATIK 2021 widmen sich – wie die 50 Ausgabe zuvor – aktuellen Informatikthemen, die auf die Nachhaltigkeitsziele einzahlen und die sich im Workshopprogramm wiederfinden. Insgesamt mehr als 30 Workshops werden in diesem Jahr virtuell stattfinden. Die ausgewählten Beiträge werden in diesem Tagungsband dokumentiert. Die Workshops finden sich unter diesen Oberthemen:

- Nachhaltigkeit und Umweltinformatik
- Sicherheit und Zuverlässigkeit
- KI-Systeme
- Software- und Informationssysteme
- Digitale Kompetenz, Digitale Bildung
- Normung, Standardisierung und Testung

Im Rahmen der INFORMATIK 2021 findet auch die 35. Internationale Konferenz zu Umweltinformations- und Kommunikationstechnologie, die [enviroInfo 2021](#), statt, bei der Fragen der Nachhaltigkeit der Informatik und durch die Entwicklungen in der Disziplin seit je her Kern des wissenschaftlichen Austausches sind. Zentrale Frage der Konferenz ist: Umweltinformatik - Schreckgespenst oder Heilsbringer zur Erreichung der UN-Ziele für nachhaltige Entwicklung?

Dazu findet im Rahmen der INFORMATIK 2021 mit der [KI 2021](#) die 44. Jahrestagung des GI-Fachbereichs Künstliche Intelligenz statt, mit sieben Workshops sowie weiteren Special Meetings und Tutorials. Die KI2021 ist eine der wichtigsten europäischen KI-Konferenzen und bringt traditionell akademische und industrielle Forscher aus allen Bereichen der KI zusammen und bietet einen idealen Ort für den Austausch von Neuigkeiten und Forschungsergebnissen zu Theorie und Anwendungen.

Last but not least ist mit der SKILL 2021, auch die 11. Studierendenkonferenz Informatik für den wissenschaftlichen Nachwuchs, nun bereits seit mehreren Jahren fester Bestandteil

der Jahrestagung gehört. Die Beiträge der Teiltagungen werden in eigenen Tagungsbänden publiziert.

Die Workshops umrahmen das Hauptprogramm, das am Mittwoch (29.09.) den *Tag der Informatik* und Donnerstag (30.09.) den *Tag der digitalen Bildung* mit einer Vielzahl von Keynotes und Paneldiskussionen gesellschaftspolitische Themenstellungen erörtert. Grußworte halten u.a. Prof. Mike Hinchey, Präsident der International Federation of Information Processing (ifip) und Prof. Dr. Gabriele Kotsis, Präsidentin der Association of Computing Machinery (ACM), die Staatssekretäre Florian Pronold aus dem Bundesministerium für Umwelt, Naturschutz und nukleare Sicherheit sowie Prof. Dr. Wolf-Dieter Lukas aus dem Bundesministerium für Bildung und Forschung und der Präsidenten der Kultusministerkonferenz Britta Ernst.

Die einleitenden Joint Keynotes zum Tagungsmotto werden von Prof. Dr. Maja Göpel, Mit-Initiatorin von Scientists for Future und Prof. Dr. Kristian Kersting von der TU Darmstadt gehalten. Dazu erwarten uns Panels zu Themen rund um „Informatik und Nachhaltigkeit“ sowie zu den Themen digitale Transformation, der digitalen Bildung und digitalen Innovationen für die Gesellschaft.

Die INFORMATIK 2021 findet statt im Rahmen des 60. Jubiläumsjahres der ifip und wäre nicht denkbar ohne die Unterstützung unserer Partner – insbesondere des Weizenbaum-Instituts und des Fraunhofer Verbundes IUK-Technologie sowie von SAP, IBM, Google und Huawei. Ein ganz besonderer Dank gilt auch dem Team in den Geschäftsstellen der GI, das maßgeblich zum reibungslosen Gelingen der INFORMATIK 2021 beigetragen haben.

Berlin, im September 2021

Daniel Krupka,

Geschäftsführer der Gesellschaft für Informatik

Sponsoren

Wir danken den folgenden Unternehmen und Institutionen für die Unterstützung der Konferenz.

Google Cloud

SAP SE

Huawei Technologies Co., Ltd.

Google Deutschland GmbH

IBM Deutschland GmbH

De Gruyter/Oldenbourg Verlag

genua GmbH

Institut für Nachhaltigkeit

Springer-Verlag GmbH

adesso SE

Digitale Infrastrukturen (eine Initiative von eco-Verband der Internetwirtschaft)

Partner:



weizenbaum
institut



Tagungsleitung

Gesamtleitung:	Daniel Krupka, Gesellschaft für Informatik Cornelia Winter, Gesellschaft für Informatik
Leitung KI 2021:	Ralf Möller, Universität Lübeck
Leitung enviroInfo 2021:	Volker Wohlgemuth, htw Berlin
Leitung SKILL 2021	Michael Becker, Universität Leipzig Judith Michael, RWTH Aachen Thomas Riechert, HTWK Leipzig Ludger Porada, Gesellschaft für Informatik e.V.
Gesamtkoordination:	Alexander Scheibe, Gesellschaft für Informatik
Koordination Hauptprogramm:	Maike Klein, Gesellschaft für Informatik Julia Meisner, Gesellschaft für Informatik
Koordination	Ludger Porada, Gesellschaft für Informatik
Workshopprogramm:	Nikolas Becker, Gesellschaft für Informatik

Organisationsteam der Gesellschaft für Informatik

Maike Klein	Koordination Hauptprogramm & Programmbeirat / Moderation
Julia Meisner	Webseite & Social Media
Piet Fölster	Webseite & Social Media
Nikolas Becker	Workshops & Moderation
Ludger Porada	Workshops & Redaktion Proceedings
Karl-Heinz Künkel	Technische Umsetzung Tagungsband
Frithjof Nagel	Kommunikation / Moderation
Paula Böhme	Kommunikation / Moderation
Elisabeth Schaueremann	Technische Konzeption & Durchführung / Moderation
Rebeka Fülöp	Digital Innovation Challenge
Kristina Hatas	Moderation
Anna-Sarah Lieckfeld	

Programmbeirat

Hannes Federrath	Universität Hamburg (GI-Präsident)
Ulrike Lucke	Universität Potsdam (GI-Vizepräsidentin)
Volker Wohlgemuth	Hochschule für Technik und Wirtschaft (HTW) Berlin
Thomas Bendig	adesso SE
Nadine Bergner	Universität Dresden
Klaus Greve	Universität Bonn
Lorenz Hilty	Universität Zürich
Wieland Holfelder	Google
Alexander Rabe	eco - Verband der Internetwirtschaft
Sandra Schulz	Universität Hamburg
Stefan Ullrich	Weizenbaum Institut

Astrid Nieße
Ralf Möller
Lars Grunske

Universität Oldenburg
Universität Lübeck
HU Berlin

Workshopchairs

Andreas Abecker	Disy Informationssysteme GmbH
Nadine Bergner	TU Dresden
Carsten Brockmann	Deloitte
Claudia-Melanie Chituc	Leibniz-Institut für Bildungsforschung und Bildungsinformation
Jan de Meer	HTW Campus
Johannes Fährdrich	Hochschule für Polizei Baden-Württemberg
Peter Fischer-Stabel	Umwelt-Campus Birkenfeld
Matthias Goeken	Hochschule der Bundesbank
Maximilian Hüb	LRZ Garing
Wilfried Honekamp	Hochschule Stralsund
Ulrich John	Victoria University
Reinhard Kahle	Universität Tübingen
Natalja Kleiner	FZI-Forschungszentrum Informatik
Gerlinde Knetsch	Hochschule für Technik und Wirtschaft (HTW) Berlin
Ralf Kneuper	IU Internationale Hochschule - Fernstudium
Birgitta König-Riess	Universität Jena
Kerstin Lenk	Tampere University, Finnland
Ulf Leser	HUU Berlin
Daniel Loebenberger	Aisec.Fraunhofer, FG Krypto
Jorge Marx Gómez	University of Oldenburg
Arne Meier	Universität Hannover
Lydia Müller	Universität Leipzig
Stefan Naumann	Umwelt-Campus Birkenfeld
Jana-Rebecca Rehse	Universität Mannheim
Thorsten Schoormann	Universität Hildesheim
Michael Schulz	NORDAKADEMIE
Annika Selzer	SIT-Fraunhofer
Dehla Sokenou	FG TAV (wps-solutions)
Christoph Sorge	Universität des Saarlandes
Aleksandra Sowa	GI Fachgruppen PET, FoMSESS, SecMgt, EZQN (Springer)

Indra Spiecker gen.
Döhmann
Marion Steiner
Benjamin van Giffen
Eduard Wagner
Veronika Weiß
Axel Wiepke

Universität Frankfurt/M.
FG SECMGT
Uni St. Gallen (und GI-SIM)
TU Berlin
HS Magdeburg-Stendal
Universität Potsdam

Inhaltsverzeichnis

Workshop: Betriebliche Umweltinformationssysteme (BUIS)

Ali Akyol, Jorge Marx Gómez

Entwicklung eines neuartigen Umweltportals mit Bürger_innenforschung zur Beseitigung von Barrieren an Umweltinformationen 33

Doha Al-Nashash, Lukas Gerdes, Katharina Hennig, Theresa Hilger, Tugay Sonakalan, Max Spetzler, Christian Janßen, Cedrik Theesen, Jorge Marx Gómez

Ökologische und ökonomische Optimierung von Fahrzeugflotten mittels einer Fahrzeugvergleichssoftware 45

Marvin Banse, Florian Schmalriede, Oliver Theel, Andreas Winter

Environmental Wellbeing through Guerilla Sensing 57

Tanja Behrendt, Malena Ripken, Harald Warmelink, Thomas Klenke

Maritime Spatial Planning and Integrated Coastal Zone Management in Higher Education: Utilizing the MSP Challenge Serious Game 67

Dirk Bendlin, Moritz Werther Häckell, Marie-Antoinette Schwarzkopf, Alicja Julia Kucewicz, Jorge Marx Gómez

Informationsfluss in der Offshore Windenergie - Ein Referenzprozess für Betriebsführung und Instandhaltung 79

Viktor Dmitriyev, Christian Janßen, Andreas Solsbach

Enable students to acquire new hands-on software experience using the modular teaching assistance system 91

Malte Eilers, Simon Hemen, Sebastian Frers, Benjamin Stahr, Natalia Hölscher, Paula Bürmann, Christian Janßen, Jorge Marx-Gómez

An instant matching algorithm in the context of ride-hailing applications, using isochrones and social scoring 103

Volker Wohlgemuth, Abdullah Güccük <i>Konzeption und Entwicklung einer Weboberfläche zur Visualisierung umweltfreundlich erzeugter Energie in Industrienetzwerken</i>	115
Tobias Höper, Johannes Schering, Jorge Marx Gómez <i>Can smartphone generated cycling data contribute to the improvement of the bicycle infrastructure?</i>	127
Antonia Krebs, Björn Koch, Thomas Klenke <i>“Civic Energy Informer” - Facilitating assistance for actors in civic energy through a modular digital platform</i>	137
Alexandra Pehlken, Patrick Eschemann, Henriette Garmatter, Fabian Cyris, Astrid Nieße <i>Einsatz von Künstlicher Intelligenz in der Digitalisierung von Abfallverbrennungskraftwerken</i>	147
Barbara Rapp, Lea Freese, Lukas Cremers <i>Natural Computing zur Identifikation von Verwertungspfaden biogener Rohstoffe</i>	157
Barbara Rapp, Andreas Solsbach <i>BUIS in der Lehre: Theorie und Praxis Hand in Hand</i>	167
Bettina Steden, Jorge Marx Gómez <i>Kennzahlen zur nachhaltigen Gestaltung von Mobilität in Unternehmen durch Umstellung auf Elektromobilität</i>	175
Leonard Stepien, Frank Köster <i>Parametrizing Complex Co-Simulations to Support Decision Making in Mobility</i>	185
Heiko Thimm, Karsten Boye Rasmussen, Milan Tanik <i>Global Environmental Website Disclosure of Large Production Companies and their CDP Scores</i>	197
Marius Wybrands, Anne Dyck <i>Langfristige Wasserbedarfsprognosen in Theorie und Praxis</i>	207

Workshop: 2. Workshop Künstliche Intelligenz in der Umweltinformatik (KIUI-2021)

Andreas Abecker, Julian Bruns, Stefan Naumann <i>2. Workshop Künstliche Intelligenz in der Umweltinformatik (KIUI-2021)</i>	221
Nicole Ludwig <i>KIU-2021 Keynote: Uncertainty in Machine Learning for Environmental Research</i>	227
Marc-Alexander Lutz, Katharina Beckh, Jörg Kindermann, Juliane Schneider, Julia Walgern, Sebastian Pfaffel, Stefan Faulstich, Alisa Staack <i>Digitalisierungsworkflow zur Strukturierung und Standardisierung von Instandhaltungsinformationen von Windenergieanlagen</i>	229
Philipp Akharath, Jaqueline Altkrüger, Harkiran Sahota, Volker Herbort, Henrik te Heesen <i>Modellierung eines Photovoltaik-Fehlererkennungsansatzes unter Berücksichtigung von maschinellem Lernen</i>	251
Paul Tremper, Till Riedel, Matthias Budde <i>Spatial Interpolation of Air Quality Data with Multidimensional Gaussian Processes</i>	269
Achim Guldner, Sandro Kreten, Stefan Naumann <i>Exploration and systematic assessment of the resource efficiency of Machine Learning</i>	287
Philipp Hertweck, Hylke van der Schaaf, Desiree Hilbring, Jonas Weis, Tanja Liesch, Matthias Budde <i>Integration von KI-Algorithmen in Umweltinformationssysteme mittels SensorThings API</i>	301
Katharina Rudisch, Sebastian Jüngling, Ricardo Carrillo Mendoza, Ulrike K. Woggon, Ina Budde, Mario Malzacher, Karsten Pufahl <i>Paving the road to a circular textile economy with AI</i>	313

Julian Janßen, Paul Tremper, Till Riedel <i>Adaptives luftqualitätsgewichtetes Fahrradrouting mittels Land-use Regression auf Basis offener Daten</i>	321
Lukas Sinnwell, Julian Bruns, Matthias Budde, Andreas Abecker <i>A meta analysis of the status of AI in environmental computer science . . .</i>	333

9. Workshop Umweltinformatik zwischen Nachhaltigkeit und Wandel (UINW 2021)

Jan Tomaschek, Ulrike Kugler <i>Digitale Schlüsseltechnologien für eine Nachhaltige Digitalisierung . . .</i>	351
Felix Meyer, Joachim Rüter, Grit Behrens, Matthias Diehl <i>Intelligente PV-Modul Datenbank-Applikation für Predictive Maintenance von PV-Anlagen</i>	359
Malte Räuchle, Jochen Wittmann <i>Nachhaltigkeit in der Softwareentwicklung</i>	365
Klaus Schlender, Grit Behrens, Reiner Creutzburg <i>Maschinelles Lernen von CNN-Modellen zur Segmentierung von Störobjekten auf Gebäudefassaden auf Infrarot- und Farbbildern</i>	375
Matthias Nattke <i>Effizientere Flächennutzung in der Landwirtschaft mit vernetzten Drohnen</i>	389
Andreas Winter <i>Impulse für die Weiterentwicklung des Blauen Engels aus einer verhaltenswissenschaftlichen Perspektive</i>	401

Workshop: Computer Science for Biodiversity (CS4BIODiversity)

Clemens-Alexander Brust, Björn Barz, Joachim Denzler <i>Carpe Diem: A Lifelong Learning Tool for Automated Wildlife Surveillance</i>	417
--	-----

Jonas Höchst, Jannis Gottwald, Patrick Lampe, Julian Zobel, Thomas Nauss, Ralf Steinmetz, Bernd Freisleben <i>tRackIT OS: Open-source Software for Reliable VHF Wildlife Tracking . . .</i>	425
Dimitri Korsch, Paul Bodesheim, Joachim Denzler <i>Deep Learning Pipeline for Automated Visual Moth Monitoring: Insect Localization and Species Classification</i>	443
Julian Zobel, Paul Frommelt, Patrick Lieser, Jonas Höchst, Patrick Lampe, Bernd Freisleben, Ralf Steinmetz <i>Energy-efficient Mobile Sensor Data Offloading via WiFi using LoRa-based Connectivity Estimations</i>	461
Julia Böhlke, Dimitri Korsch, Paul Bodesheim, Joachim Denzler <i>Exploiting Web Images for Moth Species Classification</i>	481
Matthias Körschens, Paul Bodesheim, Christine Römermann, Solveig Franziska Bucher, Mirco Migliavacca, Josephine Ulrich, Joachim Denzler <i>Automatic Plant Cover Estimation with Convolutional Neural Networks . . .</i>	499
Dagmar Triebel, Ariane Grunz, Stefan Seifert, Anton Link, Gerhard Rambold <i>DiversityNaviKey, a Progressive Web Application for interactive diagnosis and identification</i>	517
Andreas Kohlbecker, Anton Güntsch, Norbert Kilian, Wolf-Henning Kusber, Katja Luther, Andreas Müller, Eckhard von Raab-Straube, Walter Berendsohn <i>A pragmatic approach to concept-based annotation of scientific names in biodiversity and environmental research data</i>	539
Daphne Auer, Paul Bodesheim, Christian Fiderer, Marco Heurich, Joachim Denzler <i>Minimizing the Annotation Effort for Detecting Wildlife in Camera Trap Images with Active Learning</i>	547
Dina Sharafeldeen, Mohamed Bakli, Alsayed Algergawy, Birgitta König-Ries <i>ISTMINER: Interactive Spatiotemporal Co-occurrence Pattern Extraction: A Biodiversity case study</i>	565

Samira Babalou, David Schellenberger Costa, Jens Kattge, Christine Römermann, Birgitta König-Ries <i>Towards a Semantic Toolbox for Reproducible Knowledge Graph Generation in the Biodiversity Domain - How to Make the Most out of Biodiversity Data</i>	581
--	-----

Workshop: Energie- und Ressourceneffizienz von Rechenzentren (E2R21)

Fernando Penaherrera, Alexandra Pehlken, Ralph Hintemann, Simon Hinterholzer, Björn Koch <i>Softwarearchitektur für die Bewertung des Ressourcenbedarfes in Rechenzentren durch ganzheitliche Ressourceneffizienz</i>	593
---	-----

Karsten Kutzer <i>Mehr Leistung, mehr Strom – und wie weiter?</i>	603
---	-----

Norbert Conrad, Michael Resch <i>Nachhaltigkeit und Abwärmenutzung an HPC-Zentren</i>	605
---	-----

Frederik Feike, Frank Dammel, Peter Stephan, Andreas Wolf, Christian Bischof <i>Abwärmenutzung eines warmwassergekühlten Hochleistungsrechners mit einer Wärmepumpe</i>	607
---	-----

Jan Zabloudil, Ernst Haunschmid, Herbert Störi <i>VSC-3 – a summary of the experiences with an HPC system immersed in oil</i>	609
---	-----

Workshop: Umweltinformationssysteme – Ein Beitrag zur digitalen nachhaltigen Transformation der Gesellschaft (UISDigi-Trans)

Stefan von Andrian-Werburg, Maria Theuring <i>Der UBA-Umweltatlas als Werkzeug des Umweltbundesamtes für die Öffentlichkeitsinformation</i>	613
---	-----

Christian Schneider, Klemens Mrogenda, Bettina Dibbern, Rene Höfer, Marlen Davis, Heike Sommer, Michael Bilo <i>Umweltinformationssysteme für den Naturschutz</i>	623
Janik Großmann, Werner Koch, Heino Rudolf <i>envVisio: Universelle Bereitstellung von Umweltdaten</i>	633
Markus Handschuh, Sönke Birk, Christoph Dreiser, Marc Förstler <i>Einschätzung des Lebensraumangebots für ausgestorbene Arten aus Fernerkundungs- und Forsteinrichtungsdaten am Beispiel des Haselhuhns (Tetrastes bonasia rupestris) im Nationalpark Schwarzwald</i>	647
Gregor Wengler, Karim Lindner, Theresa Gleich, Silvia Raschke <i>lebensmittelwarnung.de - Rückrufe und öffentliche Warnungen in Deutschland</i>	657
Raphaela Helbig, Sven von Höveling <i>Nutzerzentrierte Referenzmodellierung für Nachhaltigkeitskennzahlen</i>	667
H.-Josef Klimeczek <i>Die umweltgerechte Stadt</i>	681

Workshop: Workshop on Systems to Support Renaturation Projects (SyRePro 2021)

Yves Annanias, Marc Wahsner, Gerik Scheuermann, Daniel Wiegrefte <i>Report on the Correction of Erroneous Geometry Data in Land Reuse Projects</i>	695
Jörg Benndorf, Diego Alejandro Restrepo, Natalie Merkel, Andreas Knobloch, Martin Kressner, Adriana Guatame-Garcia, Harm Nolte, Stefan Möllerherm, Marinus Dalm <i>TRIM4Post-Mining: Transition Information Modelling for attractive Post-Mining Landscapes – A Conceptual Framework</i>	705
Kim Bürgl, Lea Reinhardt, Frank Binder, Lydia Müller, Andreas Niekler <i>Digitizing Drilling Logs - Challenges of typewritten forms</i>	709

Benjamin Kalloch, Toni Tontchev, Mario Hlawitschka <i>Object Detection and Classification in Digital Surface Models of the Lausitz Region in Germany</i>	719
Christoph Mengers, Christian Bender, Mario Hesse <i>Geodaten in der Öffentlichen Verwaltung</i>	723
Natalie Merkel, André John <i>Utilization and analysis of digital mining dump models</i>	739
Jonah Windolph, Robert Wehlitz, Theo Zschörnig, Bogdan Franczyk <i>Integrating External Data Sources into Internet of Things Architectures for Weather and Environmental Monitoring in Former Mining Areas</i>	743
 Workshop: Security, Datenschutz und Anonymisierung	
Ralf Kneuper, Annika Selzer, Marion Steiner <i>Security, Datenschutz und Anonymisierung</i>	755
K. Valerie Carl <i>Corporate Digital Responsibility: Evaluating Privacy and Data Security Activities on Company-level</i>	757
Annika Selzer, Ingo J. Timm <i>Chances and Limitations of Personal and Anonymized Data Processing</i>	773
Sarah Diel, Wendy Heereman, Dominik Spsychalski <i>The use of web tracking technologies and its compliance with the consent-requirement</i>	789
Raphael Antonius Frick, Martin Steinebach <i>Achieving Facial De-Identification by Taking Advantage of the Latent Space of Generative Adversarial Networks</i>	795
Marc Roßberger, Alperen Aksoy, Doğan Kesdoğan <i>How to apply Database Anonymity Notions to Mix Networks</i>	807

Erfan Koza

Eine empirische Kontentanalyse zur Ermittlung von praxisorientierten Optimierungsfeldern zur Resilienz- Erhöhung der IT-Systeme im Sinne der ganzheitlichen Betrachtung der Informationssicherheit 819

Workshop: International Workshop on Digital Forensics (WDF)

Sven Becker, Tommy Bergmann, Steffen Peldschus, and Dirk Labudde
Hypothesis-driven Case Analysis in 3D-Space as a Support for Forensic Casework 835

Marleen Mohaupt, Sieke Stoeter, Dirk Labudde
Automated Determination of Fingerprint Ridge Density and Fingerprint Size to Detect Sex Differences 847

Joshua Zöllner, Dmitry Petschke, Alexander Schinner, Kristin Weber, Manuel Mayer
A kernel driver modification to visualize and reconstruct data transfer between computer and USB mass storage devices 857

Maurice Meyer, Gunnar Auth, Alexander Schinner
A Method for Evaluating and Selecting Software Tools for Remote Forensics 867

Nicole Garbers, Michael Brodthage
Herausforderungen beim Einsatz Künstlicher Intelligenz zur Erkennung von im Zusammenhang mit Kinderpornografie stehenden Mediendateien . 879

Johannes Fähndrich, Frank Trollmann
AI Reading, or Automatic Semantic Decomposition into Knowledge Graphs and Symbolic reasoning through Marker Passing 891

Workshop: Post-Quanten Kryptographie in der Praxis (PQKP)

Wasilij Beskorovajnov, Roland Gröll

Requirements Engineering of the Cryptography within the Application Landscape of the German Banking Industry 905

Nouri Alnahawi, Alexander Wiesmaier, Tobias Grasmeyer, Julian Geißler, Alexander Zeier, Pia Bauspieß, Andreas Heinemann

On the State of Post-Quantum Cryptography Migration 907

Lennart Schulze, Jan-Rainer Lahmann

Evaluating Error Mitigation Strategies for Entangled Quantum States on Near-Term Quantum Computers 943

Workshop: Datenschutz im Diskurs (RuT2021)

Rüdiger Grimm, Gerrit Hornung, Christoph Sorge, Indra Spiecker genant Döhmann

Recht und Technik – Datenschutz im Diskurs 963

Stefan Grombacher, Tobias Straub

Alles akzeptieren oder Einstellung(en) ändern? – zum Stand der Praxis bei der Nutzung von Cookies 965

Sabrina Schomberg, Ernestine Dickhaut, Torben Jan Barev, Andreas Janson

Mithilfe von Privacy Nudging zu rechtsverträglichen Videokonferenztools 981

Mona Winau, Florian Klaus Kaiser, Marcus Wiens, Frank Schultmann, Indra Spiecker gen. Döhmann

Datenschutz durch Technikgestaltung und unternehmerische Strategie in der Digitalwirtschaft 999

Can Gümüs, Wolfgang Köhler, Christian Schultz, Christoph Rasche

Konzept eines Modells zur ganzheitlichen Datenschutzbetrachtung unter Anwendung von Data Mining 1019

Alexander Böhner, Dominik Herrmann <i>ExPD: Semi-automated Web Extraction of Personal Data</i>	1035
Jan-Philipp Stroscher, Marius Schnaubelt, Alejandro Sanchez Guinea, Stefan Fabian, Julius von Willich, Yasin Alhamwy, Maximilian Bauer, Oskar von Stryk, Max Mühlhäuser, Kurt Geihs, Uwe Klingauf, Gerrit Hornung <i>Technical & Data Protection Aspects of a Smart Digital Control Center for Smart Cities</i>	1047
Niklas Bunzel, Neele Zander, Robert Landwirth, Ann-Katrin Riedel <i>Adversarial Examples zum Selbstschutz? Der Fall biometrischer Gesichtserkennung im öffentlichen Raum</i>	1059
Martin Steinebach <i>Technische Herausforderungen bei der Umsetzung von Uploadfiltern</i>	1077

13. Workshop KI-basiertes Management und Optimierung komplexer Systeme (MOC 2021)

Maximilian Lowin, Cristina Mihale-Wilson <i>Towards Predictive Maintenance as a Service in the Smart Housing Industry</i>	1093
Sven Löffler, Ilja Becker, Franz Kroll, Petra Hofstedt <i>A Survey of Constraint Transformation Methods</i>	1107
Michaela Baumann <i>Improving a Rule-based Fraud Detection System with Classification Based on Association Rule Mining</i>	1121
Helge Nuhn <i>Organizing for temporality and supporting AI systems – a framework for applied AI and organization research</i>	1135
Niclas Mummert, Olga Lange, Paul Cioffica, Tobias Reutemann <i>Kontinuierliche Evaluierung eines KI-basierten interaktiven Systems</i>	1149

Workshop: Konzeptionelle Herausforderungen für die KI: Nachhaltigkeit

Klaus Mainzer

Künstliche Intelligenz und Nachhaltigkeit 1153

Workshop: Künstliche Intelligenz für kleine und mittlere Unternehmen (KI-KMU 2021)

**André Ebert, Anika Apel, Piotr Chodyko, Kyle Hiroyasu, Festina
Ismali, Hyein Koo, Julia Kronburger, Robert Pesch**

*Blueprint for a Production-Ready Information Retrieval System based on
Multi-Modal Embeddings* 1165

Nico Döring, Jonathan Kreiss, Thomas Schuster, Raphael Volz

Brezel-Cast: Verkaufsprognose von Backwaren 1177

Johannes Otterbach, Thomas Wollmann

*Chameleon: A Semi-AutoML framework targeting quick and scalable
development and deployment of production-ready ML systems for SMEs* . . 1185

**Deborah Walz, Andreas Wächter, Stefan Tomov, Konrad Heimbach,
Matthias Weigold**

Energieeffizientes Kaltstartverhalten spanender Werkzeugmaschinen 1193

**Maximilian Lowin, Domenic Kellner, Tobias Kohl, Cristina
Mihale-Wilson**

*From Physical to Virtual Leveraging Drone Imagery to Automate
Photovoltaic System Maintenance* 1201

Sophie Schäfter, Thorsten Zylowski

German Question Answering in CRM Systems 1213

Lukas-Walter Thié

A systematic literature review of machine learning canvases 1221

Thomas Schuster, Lukas Waidelich, Raphael Volz <i>Reifegradmodelle zur Bewertung Künstlicher Intelligenz in kleinen und mittleren Unternehmen</i>	1237
Sven Nitzsche, Brian Pachideh, Victor Pazmino, Norbert Link, Christoph Schauer, Lukas Theurer, Valentin Haas, Philipp Marquardt, Sergey Biniaminov, Jürgen Becker <i>Neuromorphic Vision mit Spiking Neural Networks zur Sturzerkennung im betreuten Wohnen</i>	1247
Domenic Kellner, Maximilian Lowin, Moritz von Zahn, Johannes Chen <i>Towards Designing a User-centric Decision Support System for Predictive Maintenance in SMEs</i>	1255

Workshop: (Agiles) Enterprise Architecture Management in Forschung und Praxis (EAM)

Eldar Sultanow, Carsten Brockmann, Christian Czarnecki <i>Unternehmensarchitekturmanagement im Zeitalter von Agilität und Künstlicher Intelligenz</i>	1263
Christian Koch, Benjamin Linnik, Frank Pelzel, Eldar Sultanow, Sebastian Welter, Sean Cox <i>A Reference Architecture for On-Premises Chatbots in Banks and Public Institutions</i>	1265
Johannes Ludsteck, Eldar Sultanow, Alina Chircu, Gebhard Herget, Matthias Seßler <i>An architecture for detecting infrastructure anomalies at Germany's Federal Employment Agency</i>	1283
Konstantin Ritschel, Adina Stenzel, Christian Czarnecki, Chin-Gi Hong <i>Realizing Robotic Process Automation Potentials An Architectural Perspective on a Real-life Implementation Case</i>	1303

Wilmjakob Johannes Herlyn <i>Application-Architecture of an ERP-System for the era of Industry 4.0</i> . . .	1313
Holger Cermann, Sebastian Kern, Florian Bertram, Martin Sander, Thomas Heimann, Rachel Proctor <i>Towards an agile and adaptive Enterprise</i>	1325
Andreas Becker <i>Enterprise Architecture für Netzwerkorganisationen</i>	1341
Philipp Klingmann, Albrecht Altenhoener, Carsten Brockmann, Selina Silbernagel <i>Organizational Performance Management – From Planning to Execution and Re-Alignment</i>	1351

Workshop: Datengetriebene Geschäftsmodelle für nachhaltige Wertschöpfungssysteme (NaWerSys)

Friedemann Kammler, Paul Christoph Gembarski, Thorsten Schoormann, Simon Hagen <i>Workshop: Datengetriebene Geschäftsmodelle für nachhaltige Wertschöpfungssysteme (NaWerSys)</i>	1361
Philipp Fukas, Oliver Thomas <i>Innovation by Information Technology Recombination: How Artificial Intelligence Progressive Web Apps Foster Sustainable Development</i>	1369
Michael Ganske, Katherina Meißner, Julia Rieck <i>Game-based Learning: Acquiring Supply Chain Management Skills Through Playing Computer Games</i>	1383
Christoph Heinbach, Oliver Thomas <i>Decarbonization in the Digital Platform Age – A Green Service Canvas for Small and Medium-sized Road Carriers</i>	1397

Hannes Jahn, Hans-Jürgen Pfisterer <i>Entwicklung eines Datenerfassungssystems zur Verwendung mobiler Speicher in lokalen Netzen, um erneuerbare Energien stärker zu nutzen . . .</i>	1411
Jonas Rebstadt, Henrik Kortum, Simon Hagen, Oliver Thomas <i>Towards a transparency-oriented and integrating Service Registry for the Smart Living Ecosystem</i>	1425

Workshop: IT-Governance und Strategisches Informationsmanagement (ITG-SIM)

Michael Freiberger, Sebastian Floerecke, Christoph Ertl, Alexander Herzfeldt <i>Auswahl einer IoT-Plattform für die Integration von operativen Anlagendaten an deutschen Flughäfen Vorgehen und Lessons-Learned . . .</i>	1441
Christoph Albers, Ralf Kneuper <i>Outsourcing im Kontext agiler Produktentwicklung – Ein Erfahrungsbericht</i>	1455
Christopher Julian Kern, Thomas Schäffer, Dirk Stelzer <i>Towards Augmenting Metadata Management by Machine Learning</i>	1467
Rodrigo Raab, David Treu, Sebastian Straßburg, Helmut Beckmann <i>State-of-the-Art der IT-Governance-Forschung</i>	1477
Sebastian Straßburg, Sebastian Kahlert, Helmut Beckmann <i>IT-Compliance - Themenschwerpunkte und Frameworks</i>	1487

Workshop: Stand und den Herausforderungen des Geschäftsprozessmanagements (ZuGPM)

Ralf Laue, Jana-Rebecca Rehse, Thorsten Schoormann <i>7. Workshop zum Stand und den Herausforderungen des Geschäftsprozessmanagements</i>	1499
---	------

Ralf Laue	
<i>100 Jahre graphische Geschäftsprozessmodellierung</i>	1501
Thomas Bauer	
<i>Effiziente Modellierung von Eskalationen und Stellvertretungen in Geschäftsprozessen</i>	1503
Lars Drewes, Volker Nissen	
<i>Der Einfluss bürokratischer Hürden auf die Prozessakzeptanz</i>	1517
Fabian Sauer, Manuel Wittig, Helmut Beckmann	
<i>Ein systematisches Literatur-Review des IT- Prozessmanagements</i>	1531
Franz Lehner, Sebastian Floerecke	
<i>Identifikation erfolgsrelevanter Faktoren von Cloud- basierten Geschäftsprozessmanagement-Lösungen</i>	1545
Clemens Schreiber	
<i>On the notion of diversity in the context of information aggregation in BPM</i>	1561
Marc Frerichs, Stephan Leible, Markus Nüttgens	
<i>Towards method- and tool-independent business process modeling</i>	1567
Felix Holz, Christoph Gibcke, Sven Erdmann, Michael Fellmann, Birger Lantow	
<i>Menschliche Faktoren in Workflow-Management-Systemen</i>	1573

Workshop: Digitale Kompetenz, Digital Literacy, Digital Skills

Dorina Gumm, Sonja Beer, Stephan Schumacher	
<i>Digitale Mündigkeit fördern in IT-nahen Studiengängen</i>	1591
John Ziegler, Alexander Brand, Julian Freyberg, Tim König, Wolf Schünemann, Marina Walther, Michael Gertz	
<i>EPINetz: Exploration of Political Information Networks</i>	1603
Ute Schmid, Anja Gärtig-Daug, Linda Müller, Alexander Werner	
<i>Grundkonzepte des Maschinellen Lernens für die Grundschule – Algorithmen, Biases, Generalisierungsfehler</i>	1611

Lukas Höper, Carsten Schulte

Datenbewusstsein im Kontext digitaler Kompetenzen für einen selbstbestimmten Umgang mit datengetriebenen digitalen Artefakten 1623

Workshop: Für Informatik begeistern - vom Kindesalter bis zum Abitur (Informatik-Begeisterung 2021)

Katja Bach, Nils Springob

BOB3 – Ein kleiner Roboter zum Einstieg in Informatik und Programmieren lernen 1635

Stefan Hildebrand, Carolin Neumann

Die BYTE Challenge – One bit at a time 1641

Erik Marx, Anne Hamann, David Baberowski, Christin Nenner, Antje Kazimiers, Thiemo Leonhardt, Nadine Bergner

EduInf - Education in Informatics 1645

Christine Günther, Christin Nenner

Das I in MINT von Anfang an 1649

Nadine Kirsch, Kerstin Reese, Verena Wolf

InfoLab Saar - Schülerlabor Informatik der Universität des Saarlande 1653

Mayr Franziska, Waldvogel Bettina, Schumann Nicole, Rüdüsüli Cornelia

INGA - Informatik Garten: Eine Initiative zur Förderung von Informatikkompetenzen im Kindergarten 1657

Annabell Brocker, Christina Schramm, Ulrik Schroeder

InfoSphere - Schülerlabor Informatik 1661

Korinna Bade, Stephanie Böhnke, Grit Marschik, Tobias Scheidat

Die intoMINT App als Einstieg in die Informatik 1665

Korinna Bade, Stephanie Böhnke, Andrea Hirschel

Make Up Your MINT: Multimediale Popup-Lernlabore 1669

Workshop: Hochschule 2031

- Markus von der Heyde, Armin Gerl, Rainer Groß, Frank Klapper, Andreas Knaden, Vera G. Meister, Simone Rehm**
Vision, Mission und Governance der Landesstrategien im Zeitalter der Hochschul-Digitalisierung 1675
- Christian Fleiner, Julian Grümmer, Goran Zvekan, Andreas Harth**
WireGraph - Entwicklung eines Lernspiels mit Prosumentenumgebung zur Förderung der fachlichen und digitalen Kompetenz von Wirtschaftsstudierenden 1689
- Christian Bunse, Lieven Nils Kennes, Jan-Christian Kuhr**
Distanzlabore in der ingenieurwissenschaftlichen Ausbildung 1701
- Harald Gilch, Imke Jungermann, Klaus Wannemacher**
Modellierung einer digitalen Hochschule post-Corona 1713

Workshop: Where is the Science in Data Science Projects? Online-Workshop über die Wissenschaftlichkeit von Vorgehensmodellen für Data-Science-Projekte (WISDAP)

- Jens Kaufmann, Stephan Kühnel, René Theuerkauf, Emal M. Alekosai, Sayed Hoseini, Uwe Neuhaus, Michael Schulz**
Where is the Science in Data Science Projects? 1729

Standardization Workshop on Industrial Automation and Control Systems (6thIACSWS'21)

- Jan de Meer, Karl Waedt, Axel Rennoch, Hans-Joachim Hof**
Presentation of the Programme of the 6th GI/ACM Standardization Workshop on Industrial Automation and Control Systems 1745

Josef Schindler, Erkin Kirdan, Karl Waedt <i>Secure OPC UA Server configuration for smart charging stations</i>	1753
Robert Altschaffel, Ivo Hempel, Oliver Keil, Josef Schindler, Martin Szemkus, Jana Dittmann, Matthias Lange, Karl Waedt, Yongjian Ding <i>Supporting Security in Industrial Automation and Control Systems using Domain-Specific Modelling</i>	1759
Asmaa Tellabi, Abdelbast Sabri, Christoph Ruland, Karl Waedt <i>Security aspects of FPGA and virtualization case studies</i>	1771
Christele Larissa Moussi Djeukoua, Timothee Guiraud, Edita Bajramovic, Josef Schindler, Karl Waedt <i>Secure Unidirectional Security Gateways for Industrie 4.0</i>	1781
Raman Barakat, Faruk Catal, Sascha Hackel, Axel Rennoch, Martin A. Schneider <i>Towards a certification scheme for IoT security evaluation</i>	1789
Dominik Rohrmus, Anja Simon, Alexander Willner, Markus Rentschler <i>Neutral Interoperability Testbeds</i>	1797

Workshop; Herausforderungen beim Testen moderner Softwaresysteme (TAV-2021)

Dehla Sokenou, Mario Friske, Baris Güldali, David Faragó, Mario Winter, Wilhelm Ruckdeschel <i>TAV-2021: Herausforderungen beim Testen moderner Softwaresysteme . . .</i>	1809
Matthias Hamburg, Stephan Weißleder <i>Der Glaubenskrieg über Testautomatisierung</i>	1813
David Georg Reichelt, Stefan Kühne, Wilhelm Hasselbring <i>Testselektion für Performanzregressionsbenchmarks in CI-Prozessen . . .</i>	1819
Jonas Klamroth, Max Scheerer, Oliver Denninger <i>Holistische Verifikation von Hybriden Quantenprogrammen durch Software Bounded Model Checking</i>	1825

Dierk Ehmke	
<i>Testdatenmanagement mit formalen Testdatenspezifikationen in einem großen Migrationsprojekt</i>	1831
Ralf Somplatzki	
<i>QA-Strategie oder Flickenteppich?</i>	1839
Mario Winter, Karin Vosseberg, Frank Simon	
<i>Softwaretest-Umfrage 2020 – (Wie) passt die Forschung zum Bedarf der Praxis?</i>	1845
Simon André Scherr, Frank Elberzhager, Adeline Silva Schäfer	
<i>Schnelllebige Qualitätssicherung im digitalen Ökosystem für mobile Dienste - vom ganz großen Schmerz, viralen Posts und dem Aufenthalt in der Bahn</i>	1853
Jens Bielefeldt, Kai-Uwe Basener, Siddique Reza Khan, Mozhdeh Massah, Hans-Werner Wiesbrock, Stefan Scharoba, Michael Hübner	
<i>Einsatz von Deep Learning für den Test Eingebetteter Systeme</i>	1859
Matthias Hamburg	
<i>Metamorphes Testen</i>	1865

Autorinnen- und Autorenverzeichnis

Workshop: Betriebliche Umweltinformationssysteme (BUIIS)

Entwicklung eines neuartigen Umweltportals mit Bürger_innenforschung zur Beseitigung von Barrieren an Umweltinformationen

Ali Akyol¹, Jorge Marx Gómez²

Abstract: Umwelt und Klimaschutz nehmen eine immer stärkere Rolle in der Gesellschaft ein. Es gibt viele Möglichkeiten, sich für den Umwelt- und Klimaschutz einzusetzen, allerdings wissen viele Bürger_innen nicht, wie sie dabei vorgehen sollen und besitzen zu wenig Informationen über ihre Umwelt, um eine Veränderung herbeirufen zu können. Durch die sich verschärfende Gesetzgebung zu Umweltinformationen sind Unternehmen und Kommunen aufgefordert, Umweltinformationen vorzuhalten. Die Bereitstellung der Umweltinformationen bringt Unternehmen und Kommunen in eine Situation der Ungewissheit, da diese nicht wissen, ob die bereitgestellten Informationen auch den Anforderungen der Bürger_innen gerecht werden. Über die Bürger_innenbedarfe an Umweltinformationen ist nur wenig bekannt. Zudem bestehen Hürden zum Zugang zu Umweltinformationen, da bestehende Umweltportale für Laien schwer verständlich sind. Dieser Beitrag bearbeitet die Identifikation der Bürger_innenbedarfe an Umweltinformationen.

Keywords: Umweltinformationen, Umweltportal, Bürgerpartizipation, Umweltsensoren

1 Einleitung

Die Bereitstellung von Umweltinformationen bekommt vor dem Hintergrund des zunehmenden Bewusstseins der Gesellschaft für Umwelt- und Klimaschutzthemen eine steigende Bedeutung [BM18]. Dies bestätigt sich auch am Beispiel der Schülerbewegung „Fridays For Future“, die 2018 von Greta Thunberg in Schweden ins Leben gerufen wurde. [KN19]

Laut der Europäischen Umweltagentur sterben in der EU mehr als 400.000 Menschen und weltweit mehr als 700.000 jährlich an den Folgen von Luftverschmutzung. In Deutschland fordert die Luftverschmutzung jährlich 70.000 Todesfälle. [EU20, Sv20, We21] Seit der Einführung einer verpflichtenden nichtfinanziellen Berichterstattung vom Deutschen Bundestag, sind Unternehmen und Kommunen aufgefordert, Umweltinformationen vorzuhalten, um den stetig steigenden Anforderungen von Politik und Öffentlichkeit gerecht zu werden [De17]. Allerdings haben erste Recherchen zu den Bedarfen von

¹ Universität Oldenburg, Department für Informatik, Abteilung VLBA, Ammerländer Heerstr. 114-118; 26129 Oldenburg, ali.akyol@uni-oldenburg.de

² Universität Oldenburg, Department für Informatik, Abteilung VLBA, Ammerländer Heerstr. 114-118; 26129 Oldenburg, jorge.marx.gomez@uni-oldenburg.de

Bürger_innen an Umweltinformationen ergeben, dass über die Bedarfe der Bürger_innen an Umweltinformationen nur wenig bekannt ist. Dies führt dazu, dass Unternehmen und Kommunen nicht wissen oder nur schwer abschätzen können, ob die bereitgestellten Informationen den Anforderungen der Bürger_innen gerecht werden. Ein weiteres Problem stellt der Prozess der Bereitstellung der Umweltinformationen dar. Unternehmen und Kommunen sind laut dem Umweltinformationsgesetz (UIG) dazu verpflichtet, Umweltinformationen bereitzustellen.

Die Bereitstellung von Informationen durch Anfragen zu Umweltinformationen werden bisher individuell durchgeführt, sodass die Unternehmen und Kommunen Personalressourcen dafür bereitstellen müssen. Hinzu kommen die Herausforderungen, dass Umweltinformationen für Laien nur schwer verständlich sind und weitere Beschreibungen benötigen, um die Informationen interpretieren zu können. Eine Hypothese dieser Arbeit besagt, dass Bürger_innen Barrieren zu Umweltinformationen ausgesetzt sind, welche im zweiten Kapitel näher erläutert werden. Im Rahmen dieser Forschung soll ein Umweltinformationsportal entwickelt werden, das vielseitig und von unterschiedlichen Stakeholdern verwendet werden kann. Die Key-User sind Bürger_innen, Unternehmen und Kommunen, die jeweils eine tragende Rolle in dem Umweltportal einnehmen.

2 Grundlegendes

Im folgenden Kapitel wird der Stand der Forschung zu Umweltinformationen und Umweltinformationsportalen ausgearbeitet.

2.1 Umweltinformationen

Als Folge der EU-Richtlinie 2003/EG wurde im Jahr 2005 das Umweltinformationsgesetz von der Bundesregierung verabschiedet und in nationales Recht umgewandelt, um den Zugang zu den Umweltdaten und -thematiken zu erweitern. [bm21] Unter dem Begriff Umweltinformationen werden Daten definiert, welche sich mit allen möglichen Informationen über Luft, Boden, Wasser oder weiteren natürlichen Lebensräumen befassen. Auch gehören die Indikatoren, wie Lärm, Strahlung oder Energie zu Umweltdaten. Zudem haben Umweltdaten in der Regel einen räumlichen Bezug. [Bu21]

Das Bundesministerium für Umwelt, Naturschutz und nukleare Sicherheit beschreibt, dass durch den Zugang der Öffentlichkeit zu umweltbezogenen Informationen das Umweltbewusstsein in der Bevölkerung und der Meinungs-austausch ermöglicht werden soll. [Bu21]

2.2 Umweltinformationsportal

Zu Beginn wird der Begriff Umweltinformationsportal, auch Umweltportal oder auch Umweltplattform definiert. In dieser Arbeit wird das Synonym Umweltportal verwendet. Unter einem Umweltportal wird eine Internetseite verstanden, die Informationen über die Umwelt zur Verfügung stellt. [PU20] Umweltportale stellen Informationen, wie Messreihen, Datenwerte, die strukturiert oder semi-strukturiert sind, zur Verfügung. [EZK10, Ab09]. Eines der Hauptziele von Umweltportalen ist, Umweltinformationen der Allgemeinheit zugänglich und nutzbar zu machen. Damit kann sichergestellt werden, dass Bürger_innen sich über die aktuelle Umweltlage einer Stadt oder Gemeinde informieren können. [KS10]

In Deutschland wurde die Bereitstellung von Umweltinformationen erstmals vom LfU (Landesamt für Umweltschutz Baden-Württemberg) im Jahr 1996 in digitaler Form zur Verfügung gestellt. [EST11]

2.3 Barrieren zu Umweltinformationen

Unter Barrieren werden Hindernisse verstanden, die einem Menschen den Zugang oder den Weg zu etwas versperren. Barrieren an Umweltinformationen können in unterschiedlichen Ausprägungen vorkommen. Dabei kann die Ursache der Barriere entweder an dem Menschen selbst liegen, sodass er sich selbst befähigen kann, diese Barriere zu überwinden, oder sie sind nicht selbst an den Barrieren zu Umweltinformationen verschuldet. Im Folgenden sind mögliche Barrieren aufgeführt:

1. Bürger_innen kennen das Umweltportal nicht, das Umweltinformationen bereitstellt.
 - a. In diesem Fall könnte die Person durch Eigeninitiative eine Recherche vornehmen und nach „Umweltportal“ im Internet suchen, um sich über die Umwelt zu informieren.
 - b. Eine weitere Möglichkeit besteht darin, dass die Person ihre Kommune oder Stadt anruft, die ihr den Link des regionalen Umweltportals nennen kann.
2. Bürger_innen verstehen die zur Verfügung gestellten Inhalte nicht.
 - a. Umweltinformationen sind nicht immer schnell verständlich, sodass die Bürger_innen darauf angewiesen sind, sich über die Werte weiter zu informieren, um zu verstehen, ob die angezeigten Werte gesundheitsschädlich sind oder nicht.
3. Bürger_innen müssen sich selbst über die Grenzwerte erkundigen.
 - a. Die Informationen müssen proaktiv von Bürger_innen auf dem

Umweltportal abgerufen werden. Grenzwertüberschreitungen, bekommen die Bürger_innen in den meisten Fällen nicht oder nur im Nachhinein mit.

4. Umweltinformationen werden im Alltag nicht integriert.
 - a. Die bestehenden Systeme ermöglichen es nicht, dass Bürger_innen stets über ihre natürliche Umwelt informiert sind. Der Zugang zu Umweltinformationen dauert zu lange und ist nicht einfach gestaltet, sodass wichtige Informationen die Bürger_innen nicht erreichen.
5. Umweltinformationen werden nicht gemessen.
 - a. Diese ist eine der größten und häufigsten Barrieren, die es zu überwinden gilt. Bürger_innen haben bei dieser Barriere keine Möglichkeit sich über ihre natürliche Umwelt zu erkundigen. Diese Barriere kann nur mit hohem Aufwand beseitigt werden, indem eigene Sensoren beschafft und aufgebaut werden.

3 Umweltinformationsportal

Es gibt drei Hauptakteure, die als Nutzer_innen für ein Umweltinformationsportal identifiziert wurden: Bürger_innen, Unternehmen und Kommunen. Für größere Unternehmen entstehen durch die zunehmend strengeren Gesetze und Vorschriften ein Aufwand, den sie zu bewältigen haben [De17]. Daher sind diese auch an einer Lösung interessiert, um Umweltinformationen vereinfacht an die zuständigen Behörden zu übermitteln. Eine Lösung soll im zu entwickelnden Umweltportal ebenfalls umgesetzt werden. Das Projekt bringt sowohl praktische als auch theoretische Vorteile für die Stakeholder und die Kommunikation untereinander wird zudem verbessert. Das Projekt setzt bei der Umsetzung des Umweltinformationsportals den Fokus auf die Bedürfnisse der Bürger_innen, die auch für Unternehmen und Kommunen eine hohe Relevanz haben [PU20].

3.1 Identifikation von Umweltinformationsbedarfen

Zur Identifikation der Bürger_innenbedarfen wurden sowohl quantitative als auch qualitative empirische Forschungsmethoden durchgeführt. Für die quantitative empirische Datenerhebung wurde eine Onlinestudie mit über 600 Teilnehmer_innen verwirklicht. Für die qualitative Datenerhebung wurden unterschiedliche Workshops mit Bürger_innen, Schüler_innen und Studierenden organisiert.

3.1.1 Onlinestudie

Seit den 1990er Jahren sind Online-Studien ein bewährtes Instrument zur Datenerhebung in der Markt- und Sozialforschung. Onlinestudien bringen die Vorteile, dass sie mit wenig Kosten umgesetzt werden können und die Auswertung der Ergebnisse schnell erfolgen kann. Die Ergebnisse der Teilnehmer_innen sind digital und in Echtzeit verfügbar. Zudem haben Onlineumfragen eine höhere Rücklaufquote als postalischen Umfragen, da sie von den Teilnehmer_innen ortsunabhängig ausgefüllt werden können. Neben den Vorteilen bringen Onlineumfragen auch Schwierigkeiten mit sich, da diese nur von Teilnehmern ausgefüllt werden können, die Zugang zum Internet haben. Nicht zu vergessen sind auch die Datenschutz- und Sicherheitsaspekte, die bei einer Onlineumfrage berücksichtigt werden müssen. [BC15]

Die durchgeführte Onlinestudie wurde so konzipiert, dass die Teilnehmer anonym teilnehmen konnten. Es wurden keine verpflichtenden Angaben erfragt, mit denen es möglich wäre, die Teilnehmer_innen anhand der Antwort zu identifizieren. Bei der Frage nach der Postleitzahl wurden nur die ersten vier Ziffern erfragt, welches den Radius des Wohnortes der Teilnehmer_innen erhöht.

Die Onlinestudie wurde durchgeführt, da nur wenig Wissen über die Nutzer_innen von Umweltportalen und ihren Bedarfen bekannt ist. Damit die Entwicklung sich an den Bedarfen der Bürger_innen orientieren kann. Mit den Ergebnissen der Studie können Anforderungen definiert werden, die als Funktionen im Portal implementiert werden können.

Für die durchgeführte Onlinestudie wurden zwei Kernfragen definiert.

1. *Welche Nutzergruppen besuchen Umweltportale?*
2. *Was sollten die Schwerpunkte eines Umweltportals sein?*

Die durchgeführte Onlineumfrage besteht aus 31 Fragen, die in fünf Fragegruppen aufgeteilt wurden. Die Umfrage wurde anonym durchgeführt, jedoch hatten die Teilnehmer_innen die Möglichkeit, ihre E-Mail zu hinterlegen, sodass sie bei Neuigkeiten zum Forschungsprojekt PUUK informiert werden können. [PU21a]

Tabelle 1. Fragegruppen der Onlineumfrage

Nr.	Fragegruppe
1.	Zu Ihrer Person
2.	Allgemeines Interesse zu Umweltdaten
3.	Für Sie relevante Umwelthinformationen
4.	Von Ihnen genutzte Umwelthinformationsportale
5.	Ein für Sie entwickeltes Umwelthinformationsportal sollte folgende Aspekte erfüllen

Die Umfrage wurde von 644 Teilnehmer_innen vollständig ausgefüllt (Eigene Quelle).

Durch die Onlinestudie konnten neue Erkenntnisse über die Umweltinformationsbedarfe der Bürger_innen gewonnen werden. Einige Annahmen wurden durch die Antworten der 644 Teilnehmer_innen bestätigt. So wurde z. B. bestätigt, dass die meisten Menschen an ihrer natürlichen Umwelt interessiert sind und Informationen darüber erhalten möchten. Der Großteil der Teilnehmer_innen, ca. 95 %, geben an, dass sie sich für Umwelt- und Klimaschutz interessieren. Außerdem geben 60 Personen an, dass sie in einer Umwelt- und Klimaschutzorganisation aktiv sind. Auch wurden neue Erkenntnisse darüber gewonnen, welche Begriffe eine Definition benötigen und welche nicht. Die Begriffe Stickstoffdioxid, PM2,5, PM10 und Schwefeldioxid sind die Begriffe, die am häufigsten eine Beschreibung oder Definition benötigen. Zudem gaben 428 Teilnehmer_innen an, dass sie Umweltportale aus Interesse zur Umwelt besuchen und 261 Teilnehmer_innen offenbarten, dass sie Umweltportale besuchen, um ihre Aktivitäten zu planen. Es konnten Erkenntnisse darüber gewonnen werden, welche Begriffe in einem Umweltportal definiert und beschrieben werden sollten. Außerdem gaben 537 der Teilnehmer_innen an, dass sie Umweltinformationen von Internetseiten beziehen, und 328 Personen nutzen soziale Netzwerke als Quelle für ihre Umweltinformationen. Zudem konnten Ergebnisse über die Häufigkeit der Besuche an einem Umweltportal gewonnen werden. Es gaben 66 Teilnehmer_innen an, dass sie Umweltportale täglich besuchen, und 187 Teilnehmer_innen gaben an, dass sie wöchentlich Umweltportale besuchen. Auch zu erwähnen ist, dass 222 Teilnehmer_innen angaben, Umweltportale gar nicht zu besuchen. Mehr als die Hälfte von den 222 Teilnehmer_innen sagen, dass sie kein bisheriges Wissen über Umweltinformationsportale haben (Barriere Nummer 1 in Kapitel 2.3). Als zweithäufigster Grund wurde angegeben, dass sie keine Zeit dafür haben. [PU21a]

3.1.2 Bürger_innenworkshop

An dem Bürger_innenworkshop haben ungefähr 40 Personen teilgenommen. Die Ziele des Workshops bestanden darin, die Ergebnisse der Umfrage zu intensivieren und anhand der Ergebnisse Anforderungen an das Portal mit den Teilnehmer_innen definieren zu können. Die Ergebnisse der Onlineumfrage haben ergeben, dass vier Themen besonders wichtig für die Bürger_innen sind. [PU21b]

1. Bürger_innenpartizipation am Umweltportal
2. Informationen über Maßnahmen zur Verbesserung und Erhalt der Umwelt
3. Mobilität und Umweltinformationen
4. Anforderungen an die Darstellung von Umweltinformationen

Diese Themen wurden gemeinsam mit den Teilnehmer_innen besprochen und diskutiert. Damit eine Diskussion möglich war, wurden die Teilnehmer_in vier Gruppen aufgeteilt und konnten in kleineren Gruppen miteinander diskutieren. Jede Gruppe wurde von Fachkundigen in dem Themenfeld moderiert. Am Ende wurden die Ergebnisse von den Moderator_innen im Plenum vorgestellt.

3.1.3 Schüler_innenworkshop

Es wurden zwei Schüler_innenworkshops jeweils im Rahmen einer 90-minütigen Unterrichtsstunde durchgeführt. Die Workshops fanden digital an unterschiedlichen Schulen und in unterschiedlichen Jahrgängen statt. Neben den Umweltinformationen zu Luftverschmutzung gaben die Schüler_innen an, auch eigene Profile in einem Portal erstellen zu wollen, damit sie die Informationen nicht immer neu suchen müssen. Die Bodenverunreinigung, wie bspw. der Nitratgehalt im Boden, war ebenfalls ein Thema, das angesprochen wurde. Messdaten in einer Karte, aber auch in tabellarischer Form wurden von den Schüler_innen als wichtig betrachtet. Eine einfache und verständliche Sprache war ebenfalls von hoher Relevanz, damit die Schüler_innen die Messerwerte und ihre Bedeutung verstehen. Besonders relevant fanden die Schüler_innen Informationen aus unmittelbarer Nähe angezeigt zu bekommen. Zudem machten die Schüler_innen Vorschläge, wie das Integrieren von Belohnungssystemen aussehen sollte, damit umweltverträgliches Verhalten belohnt werden kann. [PU21c]

3.1.4 Studierendenworkshop

Im Rahmen einer Gastvorlesung mit Studierenden im Masterstudiengang wurde ebenfalls ein Workshop durchgeführt. Ergänzend zu dem Schüler_innenworkshop wurde erwähnt, dass, Warnungen für Unwetter auf dem Umweltportal angezeigt werden sollen. Zur Darstellungsform gaben die Studierenden an, dass die Umweltinformationen möglichst übersichtlich sein sollen und Fließtext möglichst vermieden werden soll. Die Interpretation von Messwerten sollte mit Farbausoziationen dargestellt werden (rot = schlecht, grün = gut). Somit sind die Inhalte wesentlich einfacher zu verstehen. Benachrichtigungen in Echtzeit wurden ebenfalls als eine Anforderung angegeben, um eine Navigation anhand von Umweltdaten zu erhalten. Somit könnten beispielsweise gesundheitlich belastende Orte umfahren werden. Die Implementierung von Prognosen war ebenfalls eine Anforderung, die erwünscht war.

3.2 Bürger_innenforschung

Das BMBF beschreibt den Citizen Science als „eine win-win Situation für alle“, da beide, Forscher_innen und Bürger_innen, an der Zusammenarbeit profitieren. Forscher_innen kommen durch die hohe Anzahl der Personen in der Bevölkerung an ein Wissen, welches sie allein nicht oder mit nur sehr viel Aufwand erreichen würden, und Bürger_innen können sich an der aktuellen Forschung beteiligen. [BM21]

Die Europäische Union definiert Citizen Science als Bezug „[...] auf die Beteiligung der Öffentlichkeit an wissenschaftlichen Forschungsaktivitäten, wenn Bürgerinnen und Bürger aktiv mit ihren Werkzeugen und Ressourcen zur Wissenschaft beitragen“ [Bu17].

Die Citizen Science Ansätze, auch Bürger_innenforschung genannt, sind aus der heutigen

Forschung nicht mehr wegzudenken. Mit ihr ist die Forschung nicht mehr alleinstehend, sondern mit der Gesellschaft und dem Alltag eng verzweigt. Durch die bürger_innennahe Forschung stehen die Wissenschaftler_innen nicht mehr allein dar. Die Forscher_innen können die Bürger_innen in ihre Forschung aktiv mit einbinden. Die Zusammenarbeit der Forschenden und Bürger_innen kann sich dabei auf diverse Forschungsgebiete erstrecken. Eine ältere und einfache Art der bürger_innennahen Forschung ist z. B. das Zählen von bestimmten Tierarten, um den Artenbestand abschätzen zu können. [Co21, Le21]

3.3 Bürgeranforderungen an Umweltportal

Die gewonnen Ergebnisse aus den empirischen Forschungsmethoden werden im nächsten Schritt weiterverarbeitet, um Anforderungen an das zu entwickelnde Umweltportal zu definieren. Im Folgenden werden nur die Kernanforderungen dargestellt, sodass die Anforderungen keinen Anspruch auf Vollständigkeit für ein gesamten Umweltportal haben.

Tabelle 2. Kernanforderungen an Umweltportal

Nr. Anforderungsbeschreibung	Nummer der Barriere, die gelöst werden kann (siehe Kap. 2.3)
1. Intuitive Bedienbarkeit für alle Nutzer_innengruppen	2
2. Beschreibung zu den dargestellten Umweltinformationen	2
3. Darstellen von Grenzwerten	2
4. Zurverfügungstellung von Umweltinformationen in Echtzeit	4
5. Bereitstellung von Wetterdaten und Prognosen	4
6. Partizipation von Bürger_innen ermöglichen	5
7. Personalisierte Umweltinformationen anbieten	4
8. Ermöglichen von Benachrichtigungsfunktionen	3
9. Austausch zwischen den Bürger_innen ermöglichen	
10. Nutzer_innen klimafreundliches Verhalten zeigen	

Mittels der definierten Anforderungen können vier von fünf Barrieren gelöst werden. (Eigene Quelle)

Weitere Anforderungen, wie z.B. die Integration von Unternehmen und Kommunen, sind ebenfalls unter den Anforderungen zu finden.

3.4 Systemarchitektur

Aufbauend auf dem Dreischichtenmodell von Pagé, soll das zu entwickelnde System aus der Darstellungsschicht (Frontend), der Verarbeitungsschicht (Backend) und der Datenerhaltungsschicht (Datenbank) bestehen. [Pa96] Im Frontend sollen die Umweltinformationen den Nutzer_innen angezeigt werden. Das zu entwickelnde Umweltportal unterscheidet sich von herkömmlichen Umweltportalen, indem Umweltinformationen personalisiert dargestellt werden können. Die Besucher_innen haben die Möglichkeit, ein Profil anzulegen. Dabei sollen sie entscheiden können, welche Umweltinformationen sie interessieren. Die Daten werden in der Datenverarbeitungsschicht aufbereitet und in der Datenbank gespeichert, welche die Datenerhaltungsschicht ist.

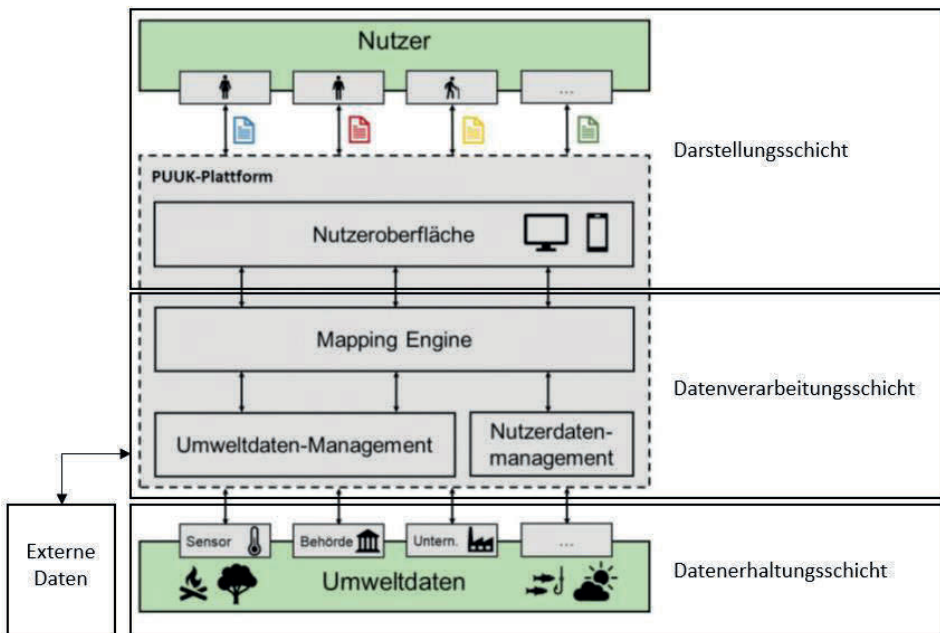


Abbildung 1 Systemarchitektur PUUK (eigene Darstellung)

In dem zu entwickelndem Umweltportal werden zusätzlich auch externe Daten mittels Schnittstellen integriert. Damit kann das zu entwickelnde Umweltportal Umweltinformationen darstellen, die es selbst nicht in der Datenbank gespeichert hat.

4 Zusammenfassung

Die Ergebnisse dieser Arbeit ermöglichen es, ein Umweltportal zu implementieren, welches auf die Bedürfnisse der Bürger_innen zugeschnitten ist. Durch die empirische Onlinestudie und die Workshops konnten Bürger_innenstimmen eingeholt werden. Diese Ergebnisse konnten als Grundlage verwendet werden, um Anforderungen für ein neues Umweltportal zu definieren, welches gemeinsam mit den Bürger_innen implementiert werden kann. Die Bürger_innen sollen das System nicht nur verwenden, sondern haben die Möglichkeit, Entscheidungen in der Funktionalität und Gestaltung des Umweltportals mitzutragen.

Die Identifikation von Barrieren zum Zugang zu Umweltinformationen ist eine Erkenntnis, die bisher in der Forschung nicht betrachtet wurde. An den Barrieren besteht noch weiteres Forschungspotenzial. Durch die Implementierung eines neuartigen Umweltportals können 4 von 5 Barrieren zu Umweltinformationen beseitigt werden. Durch den Zugang zu Umweltinformationen können die Bürger_innen für Umwelt- und Klimathemen weiter sensibilisiert werden. Das kann dazu führen, dass sie ihr Verhalten zur Umwelt neu reflektieren und langfristig verbessern. Ein weiterer Punkt ist, dass die Bürger_innen im aktuellen Stand nicht in der Lage sind, zu prüfen, ob die Luft in ihrer Umwelt für sie gut oder schlecht ist. Zudem können die Bürger_innen politische Entscheidungen für eine bessere Luftqualität in ihrer Umgebung fordern, wenn Grenzwerte überschritten werden. Die Barriere (Nr. 1) hat sich in der Onlinestudie bestätigt, in dem über die Hälfte der Teilnehmer_innen, die keine Umweltportale verwenden, als Grund angaben, dass sie davon nichts wissen.

Die gewonnenen Erkenntnisse aus dieser Arbeit werden für den Prototypen verwendet, um ein neuartiges Umweltportal zu entwickeln. Im nächsten Teil dieser Forschungsreihe wird tiefer auf das System und Partizipationsmöglichkeiten der Bürger_innen eingegangen. Dafür wurden Sensorboxen entwickelt, die mit dem Umweltportal kommunizieren können. Somit werden die Bürger_innen befähigt, nicht nur Umweltdaten zu erhalten, sondern auch selbst welche bereitzustellen.

Danksagung

Gefördert mit Mitteln der Deutschen Bundesstiftung Umwelt. Projekt Förderkennzeichen: 35152/01. Projekttitel: Portal für eine bedarfsgerechte Umweltinformationsbereitstellung von Unternehmen und Kommunen unter Bürgerbeteiligung

5 Literaturverzeichnis

- [Ab09] Abecker, A.: SUI - Ein Demonstrator zur semantischen Suche im Umweltportal Baden-Württemberg. Konzepte zur verbesserten

- Informationssuche in Umweltportalen auf Basis von Technologien des Semantic Web. <https://pudi.lubw.de/detailseite/-/publication/62352>, Stand: 09.12.2020.
- [BC15] Barz, H.; Cerci, M.: Online-Befragung: Frauen in Kunst und Kultur: Zwischen neuem Selbstbewusstsein und Quotenforderungen. Springer Fachmedien Wiesbaden, Wiesbaden, S. 95–106, 2015.
- [BM18] BMU: Umweltinformation. <https://www.bmu.de/themen/bildung-beteiligung/umweltinformation/>, Stand: 13.12.2020.
- [bm21] bmu.de: Richtlinie 2003/4/EG über den Zugang der Öffentlichkeit zu Umweltinformationen und zur Aufhebung der Richtlinie 90/313/EWG - BMU-Gesetze und Verordnungen. <https://www.bmu.de/gesetz/richtlinie-20034eg-ueber-den-zugang-der-oeffentlichkeit-zu-umweltinformationen-und-zur-aufhebung-der/>, Stand: 16.05.2021.
- [BM21] BMBF-Internetredaktion: Bürgerforschung - BMBF. <https://www.bmbf.de/de/buergerforschung-225.html>, Stand: 16.05.2021.
- [Bu17] Bunge, E.: Citizen Science in der Bibliotheksarbeit: Möglichkeiten und Chancen. b.i.t. online Verlag, 2017.
- [Bu21] Bundesministerium für Umwelt, Naturschutz und nukleare Sicherheit: Umweltinformation – Worum geht es? <https://www.bmu.de/themen/bildung-beteiligung/umweltinformation/umweltinformation-worum-geht-es/>, Stand: 16.05.2021.
- [Co21] Corey T. Callaghan Jodi J. L. Rowley, William K. Cornwell, Alistair G. B. Poore, Richard E. Major: Improving big citizen science data: Moving beyond haphazard sampling. <https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.3000357>, Stand: 17.05.2021.
- [De17]: Gesetz zur Stärkung der nichtfinanziellen Berichterstattung der Unternehmen in ihren Lage- und Konzernlageberichten (CSR-Richtlinie-Umsetzungsgesetz), 2017.
- [EST11] Ebel, R.; Sawade, A.; Tauber, M.: Semantische Suche im UIS Baden-Württemberg. Semantische Suche nach Umweltinformationen-Einführung, Stand: 10.12.2020.
- [EU20] Europäische Umweltagentur: Vorzeitige Todesfälle durch Luftverschmutzung 2020.
- [EZK10] Ebel, R.; Zetzmann, K.; Konstantinidis, S.: Umweltportale in Deutschland. Integriertes Umweltwissen für die Bürger. <https://docplayer.org/3448224-Umweltportale-in-deutschland.html>, Stand: 30.07.2020.

- [KN19] Koos, S.; Naumann, E.: Vom Klimastreik zur Klimapolitik: Die gesellschaftliche Unterstützung der „Fridays for Future“-Bewegung und ihrer Ziele: Forschungsbericht, 2019.
- [KS10] Krieger, R.; Schmitt, B.: Kommunale Umweltinformationssysteme als TYPO3-Erweiterung. <https://docplayer.org/9806287-Kommunale-umweltinformationssysteme-als-typo3-erweiterung.html>, Stand: 10.12.2020.
- [Le21] Lemmens Medien GmbH: Citizen Science Germany. https://www.citizen-science-germany.de/citizen_science_germany_dossier/mobile/index.html#p=2, Stand: 17.05.2021.
- [PU20] PUUK: Portal für eine bedarfsgerechte Umweltinformationsbereitstellung von Unternehmen und Kommunen unter Bürgerbeteiligung. <https://puuk-projekt.de/>, Stand: 30.07.2020.
- [PU21a] PUUK-Team: Ergebnisse der Umfrage. PUUK- Portal für eine bedarfsgerechte Umweltinformationsbereitstellung von Unternehmen und Kommunen unter Bürgerbeteiligung, 2021.
- [PU21b] PUUK-Team: Ergebnisse des Workshops #1 – Bürgerbedarfe. PUUK- Portal für eine bedarfsgerechte Umweltinformationsbereitstellung von Unternehmen und Kommunen unter Bürgerbeteiligung, 2021.
- [PU21c] PUUK-Team: Ergebnisse der Workshops #2 & #3 – Schülerbedarfe. PUUK- Portal für eine bedarfsgerechte Umweltinformationsbereitstellung von Unternehmen und Kommunen unter Bürgerbeteiligung, 2021.
- [Sv20] Sven: Mehr als 400.000 Tote durch Luftverschmutzung: EU-Kommission muss entschlossener handeln. Sven, 2020.
- [We21] Welle, D.: Coronavirus: Feinstaub verschlimmert COVID-19-Risiko | DW | 17.11.2020. <https://www.dw.com/de/coronavirus-feinstaub-verschlimmert-covid-19-risiko/a-55628089>, Stand: 15.05.2021.

Ökologische und ökonomische Optimierung von Fahrzeugflotten mittels einer Fahrzeugvergleichssoftware

Doha Al-Nashash, Lukas Gerdes, Katharina Hennig, Theresa Hilger, Tugay Sonakalan, Max Spetzler, Christian Janßen,¹ Cedrik Theesen,² Jorge Marx Gómez³

Abstract: Ein konstant wachsendes Verkehrsaufkommen, dessen klimabelastende Tendenzen und die daraus entstehenden Umwelt- und Ressourcenprobleme führen zu einem Umdenken innerhalb des Mobilitätssektors. Aufgrund dieses Umdenkens ist die Notwendigkeit einer nachhaltigen Vergleichssoftware ermittelt worden, welche unter der Forschungsfrage, welche Merkmale eine Vergleichssoftware aufweisen muss, um eine Fahrzeugflotte sowohl ökologisch als auch ökonomisch zu optimieren, untersucht wird. Diesbezüglich wird am Markt etablierte Fahrzeugvergleichssoftware analysiert, um den Unique Selling Point dieser Software zu identifizieren. Weiterhin wird eine darauf basierende Lösung, zur verbesserten Auswahl von Fahrzeugen hinsichtlich ökologischer und ökonomischer Aspekte, dargestellt sowie anhand fester Kriterien betrachtet und anschließend ein Ausblick über potenzielle, identifizierte Erweiterungsmöglichkeiten der sich in Entwicklung befindenden Software gegeben.

Keywords: Mobilität; Nachhaltigkeit; Vergleichssoftware; Fuhrparkmanagement

1 Einleitung

Mobilität unter dem Aspekt der Reduzierung des CO₂-Ausstoßes nachhaltiger zu gestalten, ist sowohl in der Politik als auch der Ökonomie zu einer zentralen Prämisse geworden. Zu dieser trägt neben dem hohen Anteil an CO₂-Emission von Fahrzeugen sowie dem Verbrauch endlicher, fossiler Energieträger auch ein wachsendes Umdenken in der Bevölkerung bei [MOW15].

Resultierend daraus hat die Bundesregierung eine Nachhaltigkeitsstrategie entwickelt, die auf der im Jahr 2015 beschlossenen *Agenda 2030* basiert. Zweck der *Agenda 2030* ist es, mittels ihrer 17 Nachhaltigkeitsziele die globale Entwicklung sozial, ökologisch und ökonomisch nachhaltiger zu gestalten, um Problematiken wie dem Klimawandel, der negativen Veränderung der Biodiversität sowie dem hohen ökonomisch verursachten Ressourcenverbrauch entgegenzuwirken [BM20]. Die darauf basierende Nachhaltigkeitsstrategie der

¹ Carl von Ossietzky Universität Oldenburg, Ammerländer Heerstraße 114-118, 26129 Oldenburg, Deutschland
christian.janssen@uni-oldenburg.de

² Carl von Ossietzky Universität Oldenburg, Ammerländer Heerstraße 114-118, 26129 Oldenburg, Deutschland
cedrik.theesen@uni-oldenburg.de

³ Carl von Ossietzky Universität Oldenburg, Ammerländer Heerstraße 114-118, 26129 Oldenburg, Deutschland
jorge.marx.gomez@uni-oldenburg.de

Bundesregierung befasst sich somit, neben der Bekämpfung von Hunger und Armut, auch hinsichtlich des Klimawandels und der Förderung der Energiewende [Pr21].

Mittels der Einführung von innovativen Antriebskonzepten gegenüber dem Otto- sowie Dieselmotorkraftstoffmotor soll eine klimafreundliche und somit nachhaltigere Mobilität gewährleistet werden. Um einen Umstieg auf nachhaltige Antriebsarten zu verstärken, fördert die Bundesregierung den Kauf von Elektro- und Hybridfahrzeugen zum einen mittels Kaufprämien, zum anderen durch steuerliche Opportunität sowie den Ausbau der Ladeinfrastruktur [BM21].

Vor allem der gewerbliche Sektor hat im Jahr 2020 insgesamt 62,8 Prozent aller neu zugelassenen Fahrzeuge erworben [Kr21]. Die Komplexität, aufgrund des stetig steigenden Angebots an Automobilen, erschwert hierbei besonders den Findungsprozess von gewerblich genutzten Fahrzeugen [Ho15].

Aus dieser Problematik heraus entstand die Notwendigkeit einer eigenständigen, webbasierten und plattformunabhängigen sowie echtzeitfähigen⁴ Vergleichssoftware. Eine solche Lösung wird im Kontext dieses Papers unter dem Namen *Faversy*⁵ implementiert.

Mittels dieser sollen Entscheidungen hinsichtlich einer klimafreundlicheren Zusammenstellung des Fuhrparks getroffen werden, um so dem Trend der stetig steigenden, verkehrsbedingten Treibhausgasemissionen entgegenzuwirken. Darüber hinaus soll die Auswahl nicht allein anhand von umwelttechnischen Aspekten getroffen werden, sondern im Gleichgewicht mit Merkmalen der Wirtschaftlichkeit und Ausstattung stehen, um so dem angedachten Einsatzumfeld des Fahrzeuges gerecht zu werden. *Faversy* unterstützt den Nutzer⁶ dabei, sowohl die ökologischen als auch die ökonomischen Vorteile diverser Fahrzeuge zu identifizieren und zu vergleichen, um letztendlich den Auswahlprozess effektiver, effizienter und einfacher zu gestalten.



Abb. 1: Logo Faversy

Gemäß der Fragestellung – *Welche Merkmale muss eine Vergleichssoftware aufweisen, um eine Fahrzeugflotte sowohl ökologisch als auch ökonomisch zu optimieren?* – wird in Kooperation mit dem Flächenversorger *Oldenburgisch-Ostfriesischer Wasserverband* (OOWV) untersucht, welchen Einfluss *Faversy* auf den Entscheidungsprozess nimmt. Das heißt, welche ökonomischen und ökologischen Faktoren sind ausschlaggebend, um eine solche Entscheidung zu treffen und welche Merkmale sowie Funktionen muss die Software aufweisen, um diese herauszustellen.

⁴ Neue Fahrzeuge und angepasste Daten können jederzeit ohne ein zu installierendes Update verfügbar gemacht werden.

⁵ Akronym, welches sich aus den Wörtern „**F**ahrzeug“, „**V**ergleich“ und „**S**ystem“ zusammensetzt.

⁶ Im Verlauf des Papers wird die männliche Sprachform bei personenbezogenen Pronomen und Substantiven verwendet. Dies dient lediglich der sprachlichen Vereinfachung und stellt keine Benachteiligung des weiblichen Geschlechts dar.

2 Methodisches Vorgehen

Anhand eines qualitativen Vorgehens wird im Folgenden der Nutzen eines Vergleichssystems und die daraus resultierenden Vorteile auf ökologischer, ökonomischer, rechtlicher und imagebezogener Ebene herausgearbeitet. Basierend auf den erarbeiteten Ergebnissen, wird die Software *Faversy* implementiert. Die hierfür genutzte Untersuchungsmethode ist die qualitative Inhaltsanalyse nach Mayring, welche, bezugnehmend auf die Forschungsfrage, Quellen systematisch erfasst [Ma15]. Die Quellen, welche die Vorteile einer Vergleichssoftware aufzeigen, werden so auf einen überschaubaren, wesentlichen Inhalt reduziert und in die in Tabelle 1 gelisteten Kategorien unterteilt, wodurch eine übersichtliche Struktur geschaffen wird [Ma15].

Ökonomisch	<ul style="list-style-type: none"> - Volle Kostenkontrolle - Visuelle Darstellung der Gesamtkosten [Br10] - Schnelle Identifizierung von Kostentreibern - Detaillierte Informationen zu Kostendifferenzen - Zeitersparnis durch sofort verfügbare Daten - Informierte Kaufentscheidung - Effizienz - Direkter Zuschnitt auf Präferenz mittels Filterparametern [Sc17] - Große Transparenz und Benutzerfreundlichkeit - Positive Auswirkung auf wirtschaftlichen Erfolg
Ökologisch	<ul style="list-style-type: none"> - Übersicht von alternativen Antrieben - Anregung auf Alternativen zurückzugreifen [Br10] - Direkter Vergleich des CO₂-Ausstoßes [Ha15] - Vergleich unterschiedlicher Antriebsarten [Fu19]
Rechtlich	<ul style="list-style-type: none"> - Umweltschonende Antriebstechnologien bis 2025 auf 40 Prozent steigern - Subventionen [Ha15] - Erfüllung aktueller und zu erwartender gesetzlicher Auflagen mittels Filterparameter [Mü10] - Sonderabschreibungen in Höhe von 50 Prozent für elektrische Nutzfahrzeuge [Fu19]
Imagebezogen	<ul style="list-style-type: none"> - Steigerung der Relevanz des Umweltgedankens - Senden von Signalen in Richtung nachhaltiger Zukunft [Ha15] - Nachhaltigkeit als potenzieller Wettbewerbsvorteil eines Unternehmens [Br10]

Tab. 1: Nutzen einer Fahrzeugvergleichssoftware

Die sich in der Tabelle 1 befindenden Vorteile, die eine Fahrzeugvergleichssoftware mit sich bringt, werden bei der Implementierung der Software *Faversy* berücksichtigt. Konkrete Funktionalitäten, um diese Vorteile in Form einer Software zu realisieren, werden im folgenden Abschnitt erhoben und ihre Umsetzung im Kapitel *Lösung* beschrieben.

3 State of the Art

Um die Unterschiede der am Markt konkurrierenden Anbieter gegenüber *Faversy* hervorzuheben und den Mehrwert *Faversys* gegenüber diesen aufzuzeigen, wird im Folgenden bereits am Markt etablierte Vergleichssoftware analysiert.

3.1 Analyse bestehender Vergleichssoftware

Zur Analyse der konkurrierenden Anbieter werden die folgenden Kriterien herausgearbeitet, die eine Vergleichssoftware beinhalten muss, um eine Unterstützung im Beschaffungsprozess darzustellen. Damit ein Überblick über die aktuelle Marktsituation erlangt werden kann, wird recherchiert, welche Anbieter eine Vergleichssoftware anbieten. Diese wird jeweils manuell auf die in der Tabelle 2 aufgelisteten Vergleichskriterien überprüft. Bezüglich der Analyse werden unterschiedliche Anwendungen betrachtet, welche webbasiert sind, das heißt ohne einen Installationsprozess über einen beliebigen Browser verwendet werden können. Darüber hinaus wird betrachtet, ob die Anwendungen sowohl über eine Filtermöglichkeit verfügen, um Fahrzeuge nach bestimmten Merkmalen zu selektieren, als auch, ob diese eine Vergleichsansicht für mehrere Fahrzeuge enthalten, damit diese gegenübergestellt werden können. Zusätzlich zum tabellarischen Vergleich wird berücksichtigt, ob ein grafischer Vergleich der Daten möglich ist. Dies bedeutet die Visualisierung von potenziellen Entscheidungsparametern, wie beispielsweise dem Verbrauch, anhand von adäquaten Diagrammen. Weiterführend sind die *Total Cost of Ownership (TCO)* ein wichtiges Vergleichskriterium, da sie die Anschaffungskosten sowie alle laufenden Betriebskosten umfassen. Daher gehört zu den Vergleichskriterien, welche konkurrierenden Anbieter eine *TCO*-Berechnung beinhalten und die Zusammensetzung dieser grafisch darstellen. Zur Bereitstellung von umfangreichen Informationen wird überprüft, welche Anbieter eine Detailansicht enthalten, um den Anwender mit detaillierten Informationen bei Kaufentscheidungen zu unterstützen. Zudem berücksichtigen die meisten Softwarelösungen nur Leasingoptionen und keine Kaufoptionen, obwohl diese für Unternehmen ebenso von Relevanz sein können. Darüber hinaus erhalten Unternehmen oftmals individuelle Rabatte, die einen Einfluss auf die Kosten nehmen und deswegen in einer Vergleichssoftware berücksichtigt werden sollten. Ein weiteres Kriterium bilden die Nutzerrollen. Das heißt, ob es eine Unterscheidung der Nutzer in Anlehnung an den Beschaffungsprozess gibt, wie zum Beispiel einen Flottenmanager oder Einkäufer, da somit der Beschaffungsprozess vereinfacht werden könnte.

Merkmalsausprägungen	ADAC-Autokatalog ⁷	Car Cost Expert ⁸	VimCar Fleet ⁹	Athlon TCO-Simulator ¹⁰	Autokostencheck.de ¹¹	Faversy
Webbasiert	✓	✓	✓	✓	✓	✓
Filtermöglichkeit	✗	✓	✗	✓	✓	✓
Vergleichsoption	✓	✓	✗	✓	✗	✓
Grafischer Vergleich	✗	✗	✗	✗	✗	✓
TCO-Berechnung	✓	✓	✓	✓	✓	✓
TCO-Diagramm	✗	✗	✗	✓	✗	✓
Detailansicht der Fahrzeugdaten	✓	✓	✗	✗	✗	✓
Berücksichtigung von Kaufoption	✓	✗	✗	✓	✗	✓
Individuelle Rabatte	✗	✗	✗	✗	✗	✓
Nutzerrollen	✗	✗	✗	✗	✗	✓

Tab. 2: Analyse der konkurrierenden Anbieter und *Faversy*

3.2 Schlussfolgerung

Aus der Analyse der bereits am Markt konkurrierenden Anbieter ergibt sich, dass jede Software Stärken und Schwächen aufweist, jedoch die Vergleichsansichten nur in tabellarischer Form realisiert werden und keine Möglichkeit besteht, die Daten auch grafisch gegenüberzustellen. Darüber hinaus können bei den bestehenden Softwarelösungen Fahrzeuge anhand von finanziellen sowie technischen Merkmalen verglichen werden, jedoch werden ökologische Aspekte weniger beachtet, obwohl diese, wie in Kapitel 1 beschrieben,

⁷ <https://www.adac.de/rund-ums-fahrzeug/autokatalog/marken-modelle/>

⁸ <https://schwacke.de/produkt/car-cost-expert/>

⁹ <https://vimcar.de/flottenmanagement>

¹⁰ <https://www.athlon.com/de/unternehmen/mobilitaetsloesungen/fuhrpark-management/tco-simulator>

¹¹ <https://www.autokostencheck.de/>

eine immer größere Bedeutung haben. Um die ökologischen Aspekte mehr in den Fokus zu stellen sollte eine Filter- oder Vergleichsmöglichkeit nach Antriebsart, Abgasnorm oder CO₂-Ausstoß bestehen, um umweltschonende Fahrzeuge anzuzeigen. Außerdem bietet keine Anwendung die Option unternehmensabhängige Rabatte bei der Berechnung des Anschaffungspreises einfließen zu lassen und diese Auswirkung auf die TCO zu berücksichtigen. Weiterführend bilden die in Tabelle 2 aufgeführten Merkmalsausprägungen einen Beitrag zur Anforderungserhebung für die Software *Faversy*, um sich von den konkurrierenden Mitbewerbern abzugrenzen. Auf die Umsetzung dieser Merkmalsausprägungen durch *Faversy* und den dadurch bestehenden Mehrwert wird in Kapitel 4 genauer eingegangen.

4 Lösung



Abb. 2: Nicht abschließender Ausschnitt möglicher Filterparameter

Wie im Kapitel *State of the Art* dargestellt, ist der Bedarf nach einer ökonomischen und ökologischen Fahrzeugvergleichssoftware erkennbar. Die derzeit bestehenden Lösungen implementieren die erhobenen Funktionen nur teilweise oder nicht zufriedenstellend. Sie stellen oft praktische und auch benötigte Funktionen bereit, wobei diese nicht als integrierte Produktkomponenten vorliegen und demnach für den Praxiseinsatz unzureichend sind. Auf Grundlage der Analyse von Mitbewerbern (vgl. 3.1) und der Erhebung von Anforderungen aus der Praxis (vgl. 3.2), in Kombination mit dem in Tabelle 1 identifizierten Potenzial einer umfassenden Fahrzeugvergleichssoftware, befindet sich die nachfolgend beschriebene Lösung *Faversy* in der Entwicklung.

Die Software zeichnet sich durch die ganzheitliche Abbildung des Beschaffungsprozesses, angefangen bei dem Suchen und Vergleichen von Fahrzeugen, über die Auswahl der passenden Lösung, bis hin zum Monitoring des Anschaffungsstatus, aus. Hierzu ist die Datenhaltung über eine relationale Datenbank realisiert, auf welche über SQL-Anfragen zugegriffen wird. Die Suche nach dem geeigneten Fahrzeug wird über die Filterfunktion der Software unterstützt. Mittels dieser ist es dem Nutzer möglich, die Datenbasis auf Grundlage der individuellen Bedürfnisse, beziehungsweise der speziellen Anforderungen des Einsatzumfeldes, zu filtern. Um den Aspekt der Nachhaltigkeit und gleichzeitig die angesprochenen Anforderungen und Bedürfnisse abzubilden, ist die Filterung sowohl anhand von Merkmalen der Ökologie als auch der Ökonomie und Ausstattung möglich (vgl. Abbildung 2). Da die Qualität der Daten von zentraler Bedeutung für die sinnvolle Nutzbarkeit der Software ist, müssen diese stets

um die Kommunikation sowohl im Kontext der Aufträge als auch im generellen Sinne zu ermöglichen. Durch diese anwendungsinterne Kommunikationsmöglichkeit wird der Austausch von zusätzlichen Informationen ohne eine weitere Plattform unterstützt, was zu einer Simplifizierung des Beschaffungsprozesses beitragen soll.

Im Rahmen der Effizienzsteigerung realisiert *Faversy* durch die Accountbindung außerdem individuelle und persistente Merk- und Vergleichslisten. Über verschiedene Sessions und Standorte hinweg kann der Nutzer Fahrzeuge hinzufügen und abrufen, um sie zu einem späteren Zeitpunkt nicht erneut suchen zu müssen. Die konkreten Rechte der verschiedenen Rollen sind in Abbildung 4 zu finden und auch durch das moderne, individuelle *User Interface* umgesetzt. Features, auf die der Nutzer mit seiner jeweiligen Rolle keinen Zugriff hat, werden ihm dementsprechend nicht dargestellt. Ziel ist es eine simple Bedienbarkeit zu gewährleisten, sodass der Nutzer sich intuitiv zurechtfindet und die Software eine direkte Unterstützung in der Abwicklung des Beschaffungsprozesses darstellt.

	Fahrermitbeweise	Einwähler	Fahrermitbeweise mit Adminrechten	Einwähler mit Adminrechten
Fahrzeugsuche	✓	✓	✓	✓
Fahrzeugvergleich	✓	✓	✓	✓
Fahrzeugdetailansicht	✓	✓	✓	✓
Auftragsübersicht	✓	✓	✓	✓
Auftrag erteilen	✓	✓	✓	✓
Auftrag löschen	✓	✓	✓	✓
Auftragsdetailansicht	✓	✓	✓	✓
Auftragsstatus bearbeiten	✓	✓	✓	✓
Nachrichtenfunktionalität	✓	✓	✓	✓
Nutzerverwaltung	✓	✓	✓	✓

Abb. 4: Rechte der verschiedenen Nutzerrollen

5 Evaluation

Durch die Evaluation der Software soll die Benutzerfreundlichkeit des *User Interface* überprüft werden. Zudem können durch die Überprüfung eines Usability-Experten mögliche Probleme für die Endanwender vorhergesagt sowie Verbesserungsmöglichkeiten identifiziert werden, sodass diese vor Veröffentlichung der Software behoben beziehungsweise eingearbeitet werden können. Da *Faversy* eine Software im Bereich der Mensch-Maschinen-Interaktion ist, wird für die Evaluation die Methodik *Cognitive Walkthrough* genutzt.

Zur Durchführung des *Cognitive Walkthrough* wird dem Usability-Experten die Benutzerschnittstelle mitsamt Logindaten sowie einer detaillierten Beschreibung der zu erfüllenden Aufgabe beziehungsweise einer Aufgabenfolge zur Verfügung gestellt [He03]. Dabei muss dieser selbst herausfinden, über welche Aktionen die Aufgaben erfüllt werden können. Während der Usability-Experte die Aufgabe ausführt, muss dieser bei jedem Schritt hinterfragen, ob die späteren Endanwender diesen auch als logischen nächsten Schritt ansehen würden. Dafür wird nach jedem ausgeführten Arbeitsschritt ein Protokoll angefertigt, in welchem dieser dokumentiert wird sowie, ob die Anwender die Relation zwischen den Aktionen und dem später gewünschten Endergebnis erkennen. Im Anschluss daran soll eine Mindestanzahl von zehn Mobilitätsexperten befragt werden, um Feedback bezüglich des Einflusses der Vergleichssoftware auf die Optimierung einer Fahrzeugflotte zu erhalten. Dazu wird ein leitfadengestütztes Interview geführt, welches themenbezogen und strukturiert aufgebaut ist. Der Leitfaden bestimmt den Verlauf, wodurch gewährleistet werden soll, dass der Fokus auf dem Thema liegt, aber dennoch eine offene Kommunikation mit genügend Spielraum

für andere themenbezogene Aspekte, die aus Sicht des Experten ebenso relevant sind, möglich ist [LT09]. Grundsätzlich sollen jedoch mindestens die folgenden Fragen durch die Mobilitätsexperten beantwortet werden:

- Sind die Filterparameter zur Fahrzeugsuche sinnvoll gewählt?
- Wie bewerten Sie die Darstellung und den Inhalt des grafischen Fahrzeugvergleichs?
- Wie bewerten Sie die Darstellung und den Inhalt des tabellarischen Fahrzeugvergleichs?
- Ist die Menge der erhobenen Fahrzeugdaten im Hinblick auf Vollständigkeit und Notwendigkeit sinnvoll gewählt?
- Sind die dargestellten Informationen in der Fahrzeugvorschau (diese beinhaltet die wichtigsten Fahrzeugparameter in kompakter Darstellungsform) sinnvoll gewählt?

Die Evaluation soll als iteratives Vorgehen durchgeführt werden. Nach Erlangen des Feedbacks werden die Optimierungsmöglichkeiten, welche von mehreren Mobilitätsexperten unabhängig voneinander genannt worden sind, umgesetzt, bevor eine weitere Iteration der Evaluation durchgeführt wird. Aufgrund dessen, dass sich *Faversy* noch in der Entwicklungsphase befindet, konnte bisher keine Evaluation durchgeführt werden. Zurzeit wird die Funktion des Fahrzeugvergleichs implementiert, worauf noch weitere folgen. Planmäßig soll die Implementierungsphase im August abgeschlossen sein und eine erste Version der Software zur Verfügung stehen. Der Evaluationszeitraum ist bis Mitte September vorgesehen, wobei in dem Zeitraum parallel das aus der ersten Iteration erhaltene Feedback umgesetzt werden soll.

6 Fazit

Der vorliegende Beitrag beschäftigt sich damit, welche Merkmale und Funktionen eine Fahrzeugvergleichssoftware aufweisen muss, um ökonomische Unternehmensziele mit einer ökologischen Fahrzeugflotte vereinbaren zu können. Basierend auf der untersuchten Literatur und den erhobenen Anforderungen aus der Praxis werden diverse Merkmale einer umfangreichen Fahrzeugvergleichssoftware identifiziert und in die, sich derzeit in der Entwicklung befindenden, Software *Faversy* integriert. Um den konkreten Einfluss der Software zu beurteilen, ist der derzeitige Entwicklungsstatus noch nicht weit genug fortgeschritten. Trotzdem kann auf Grundlage der durchgeführten Recherchen von einigen positiven Einflüssen des Produkts ausgegangen werden. So kann beim Vergleich von alternativen und herkömmlichen Fahrzeugen die Wahl auf ökologisch nachhaltigere Fahrzeuge fallen, da sich diese gegebenenfalls besser mit den ökologischen Zielen des Unternehmens vereinbaren lassen. Da *Faversy* eine webbasierte Software ist, kann diese ohne Installation verwendet und in verschiedene Unternehmensbereiche integriert werden. *Faversy* bietet Unternehmen, basierend auf den vorangegangenen Recherchen, eine Entscheidungshilfe beim Aufbau einer ökonomisch und ökologisch optimierten Fahrzeugflotte. Da eine nachhaltige Fahrzeugflotte immer mehr an Bedeutung gewinnt, kann zeitgleich damit die Akzeptanz

und Verbreitung einer Fahrzeugvergleichssoftware wie *Faversy* steigen, welche den Prozess der Entscheidungsfindung erleichtert.

Grundsätzlich soll mittels *Faversy* eine Vielfalt an alternativen Technologien in Fuhrparks deutscher Unternehmen gefördert werden. Dies wird vor allem mittels der Unterstützung bei der Entscheidungsfindung durch den benutzerfreundlichen Vergleich ermöglicht. Durch den Vergleich auf ökologischer Ebene und unter Berücksichtigung ökonomischer Aspekte kann ein positiver Erfolgsfaktor für Unternehmen geschaffen werden. Denn bei falscher Fahrzeugwahl kann der Fuhrpark zum Kostentreiber eines Unternehmens werden, welcher durch *Faversy* vermieden werden kann. [Sc17].

6.1 Ausblick

Aktuell befindet sich die Software im Entwicklungsprozess, sowohl die Softwaregrundstruktur als auch die dazugehörige Fahrzeugdatenbank sind implementiert. Weiterhin sind die verschiedenen Seitenansichten, wie beispielsweise der Fahrzeugvergleich oder die Fahrzeugsuche, umgesetzt. Mögliche Erweiterungen der zu entwickelnden Software *Faversy* sind zum einen die Mehrsprachigkeit, sodass die Software optional deutsch- oder englischsprachig eingestellt werden kann. Dies kann bei der Vermarktung von *Faversy* ein kritischer Erfolgsfaktor sein, da durch die Bilingualität auch internationale Märkte erschlossen werden können. Zum anderen ist eine mögliche Erweiterung das Anzeigen von Vertragswerkstätten durch eine interaktive Kartenansicht. In diesem Fall kann sich abhängig davon, ob sich eine Vertragswerkstatt in der Nähe des Nutzungsstandortes befindet, für oder gegen den Kauf eines Fahrzeugs einer jeweiligen Automarke entschieden werden. Ein weiteres Feature kann im Hinblick auf die benutzerspezifischen Angaben realisiert werden. So besteht die Möglichkeit für den Flottenmanager, Administrator und Einkäufer, die eigene Arbeitszeit innerhalb der Software *Faversy* anzugeben, sodass für andere Mitarbeiter deren Erreichbarkeit direkt innerhalb der Software ersichtlich ist.

6.2 Limitationen

Bei der Entwicklung von *Faversy* existieren einige Limitationen, die zu beachten sind. Eine wesentliche Limitation ist der Zugriff auf umfassende Fahrzeugdaten. Solche umfassenden und stets aktuellen Daten bilden die Basis für die Nutzung der Software. Diese müssen die neuesten Fahrzeugmodelle und technischen Daten sowie Fahrzeugkosten enthalten, damit ein sinnhafter Vergleich zwischen Fahrzeugen möglich ist. Außerdem muss der Datensatz kontinuierlich aktualisiert werden, sodass neue Baureihen sowie neue Fahrzeuge in der Software berücksichtigt werden. Dies ist vor allem für den Vergleich von neuen Antriebstechnologien, wie zum Beispiel bei Elektrofahrzeugen, ein wichtiger Aspekt. Demnach muss eine Datenbank geschaffen werden, in welche stets die neuesten Fahrzeugmodelle und die dazugehörigen technischen Daten automatisiert eingepflegt werden, welche wiederum über

die Software verfügbar sind. Hierfür müssten Daten aus verschiedenen Quellen extrahiert, aufeinander abgestimmt und dann in die *Faversy*-Datenbank importiert werden, was mit einem sehr hohen Aufwand verbunden ist. Für *Faversy* wurden verschiedene Datenquellen gesichtet und teilweise integriert, sodass die Funktionen implementiert und getestet werden können. Eine Integration von kostenpflichtigen Datensätzen, wie sie zum Beispiel vom *ADAC* oder vom *Kraftfahrt-Bundesamt* angeboten werden, ist zu prüfen. Eine weitere Limitation besteht in dem aktuellen Stadium der Entwicklung von *Faversy*. Da sich die Software noch in der frühen Entwicklungsphase befindet, ist der konkrete Einfluss auf die ökologische und ökonomische Optimierung von Fahrzeugflotten derzeit nicht ermittelbar, wodurch die Software zum aktuellen Zeitpunkt nicht vollständig evaluiert werden kann. Die Evaluation bezieht sich aktuell auf einzelne Funktionalitäten, welche im Sinne der agilen Softwareentwicklung fortlaufend angepasst werden. Die dargestellten potenziellen Vorteile der Anwendung stützen sich auf die Ergebnisse der Literaturrecherche und der Anforderungserhebung aus der Praxis, müssen allerdings noch am fertigen Produkt validiert werden.

Literaturverzeichnis

- [BM20] BMU, Bundesministerium für Umwelt, Naturschutz und nukleare Sicherheit: Die 2030-Agenda für Nachhaltige Entwicklung. [bmu.de/themen/europa-internationales-nachhaltigkeit-digitalisierung/nachhaltige-entwicklung/2030-agenda](https://www.bmu.de/themen/europa-internationales-nachhaltigkeit-digitalisierung/nachhaltige-entwicklung/2030-agenda), 2020. Zuletzt besucht 23. März 2021.
- [BM21] BMWi, Bundesministerium für Wirtschaft und Energie: Elektromobilität in Deutschland. <https://www.bmw.de/Redaktion/DE/Dossier/elektromobilitaet.html>, 2021. Zuletzt besucht 16. März 2021.
- [Br10] Brutschin, Christoph: Nachhaltiges Flottenmanagement im Unternehmen – Ein Wegweiser für Unternehmen. Bericht, Handelskammer beider Basel, 2010. S. 2 ff.
- [Fu19] FuhrparkConsulting: Elektromobilität im Flottenmanagement - Das Für und Wider. <https://fuhrparkconsulting.de/2020/08/12/elektromobilitaet-im-flottenmanagement>, 2019. Zuletzt besucht 23. März 2021.
- [Ha15] Hacker, Florian: Wirtschaftlichkeit von Elektromobilität in gewerblichen Anwendungen. Bericht, Bundesregierung Deutschland, 2015. S. 169.
- [He03] Hegner, Marcus: Methoden zur Evaluation von Software. IZ-Arbeitsbericht Nr. 29, Informationszentrum Sozialwissenschaften der Arbeitsgemeinschaft Sozialwissenschaftlicher Institute e.V. (ASI), Mai 2003. S. 23-26.
- [Ho15] Hommen, Mario: Welches Auto kaufen? Die Qual der Wahl! <https://www.pk.w.de/ratgeber/autokauf/welches-auto-kaufen#:~:text=Autofahrer%2C%20die%20auf%20der%20Suche,man%20am%20besten%20kaufen%20soll>, 2015. Zuletzt besucht 17. März 2021.
- [Kr21] Kraftfahrtbundesamt: Neuzulassungsbarometer im Dezember 2020. https://www.kba.de/DE/Statistik/Fahrzeuge/Neuzulassungen/MonatlicheNeuzulassungen/fz_n_MonatlicheNeuzulassungen_archiv/2020/202012_GIMonatlich/202012_

nzbarometer/202012_n_barometer.html?nn=2592390, 2021. Zuletzt besucht 17. März 2021.

- [LT09] Liebold, Renate; Trinczek, Rainer: Experteninterview. In: Handbuch Methoden der Organisationsforschung: Quantitative und Qualitative Methoden, S. 32–56. VS Verlag für Sozialwissenschaften, Wiesbaden, 2009.
- [Ma15] Mayring, Philipp: Qualitative Inhaltsanalyse. Beltz, 2015. S. 52 ff.
- [MOW15] Müller, D. K.; Ommen, N. O.; Woisetschläger, D. M.: Ein Segmentierungsansatz für die Adoption von Elektrofahrzeugen in Unternehmen. In: Entscheidungen beim Übergang in die Elektromobilität: Technische und betriebswirtschaftliche Aspekte. Springer Fachmedien Wiesbaden, Wiesbaden, S. 10–25, 2015.
- [Mü10] Münster, Marc: Flottenmanagement – wo sich Nachhaltigkeit konkret niederschlägt. https://www.mobilservice.ch/admin/data/files/news_section_file/file/2408/sanu-seminar-nachhaltiges-flottenmanagement.pdf?lm=1418801178, 2010. Zuletzt besucht 23. März 2021.
- [Pr21] Presse- und Informationsamt der Bundesregierung: Nachhaltigkeitsstrategie neu aufgelegt. <https://www.bundesregierung.de/breg-de/aktuelles/nachhaltigkeitsstrategie-2021-1873560#:~:text=Das%20Bundeskabinett%20hat%20die%20Deutsche,%C3%BCr%20zuk%C3%BCnftige%20Generationen%20lebenswert%20erhalten>, 2021. Zuletzt besucht 23. März 2021.
- [Sc17] Schäfer, Axel: Nachhaltigkeit in der Flotte. <https://flotte.de/magazine/flottenmanagement-magazin/2017/2/5/management/5856/nachhaltigkeit-in-der-flotte.html>, 2017. Zuletzt besucht 20. März 2021.

Environmental Wellbeing through Guerilla Sensing

Marvin Banse¹, Florian Schmalriede¹, Oliver Theel¹, Andreas Winter¹

Abstract: All over the world, individuals, companies, and institutions are exploiting the environment to gain an advantage for themselves. The damage to the environment affects people's health and environmental wellbeing. By Guerilla Sensing, we provide a platform to detect, spotlight, and monitor environmental pollution, that allows citizen to trigger a closed control loop and react to this exploitation by providing data reflecting the current environmental situation. It consists of an adaptive network of low-cost sensor nodes connected to an extensible platform, enabling complex analyses and detailed notifications. Guerilla Sensing can be used by citizens without any deep knowledge in measurements or computer science.

The paper introduces the approach and the components of this platform. Applied to environmental parameters such as radiation and particulate matter, Guerilla Sensing is exemplarily evaluated.

Keywords: Citizen Science; Guerilla Sensing; IoT; Sensor Nodes; EIS; Sensor Network

1 Motivation

Today's society is facing huge environmental and health problems due to accumulated pollution since the beginning of the industrial era [Br03]. People, companies, and governments seek wealth or profit, but they often compromise on environmental sustainability. State institutions formulate threshold values for pollutants and enforce their compliance to protect their citizen. But sometimes, these institutions, are either missing the technical equipment to detect threshold value violations or they are not willing to do so in order to avoid conflicts of interests. Citizens are often unsettled, frustrated, or suspicious because they do not trust in public measurements: measurements are carried out too infrequently, with unsuitable means or at the wrong locations.

Companies report their efforts to improve their environmental performance in sustainability reports [Is01; SN18]. Comprehensible and verifiable sustainability reports serve to optimize the companys' ecological reputation. For an increasing amount of customers this is an important buying criterion [BR00]. A verification of statements on environmental parameters by the public, the economy, the agriculture, etc. is often impossible for citizens and requires comprehensive knowledge of measurement methods and measurement technology. Thus, an easy to use and cost efficient means to measure, monitor, and analyze environmental data is dearly needed. Citizens without detailed technical knowledge must be enabled to flexibly adapt simple measurement devices to a variety of different environmental phenomena.

¹ <first name>.<last name>@uol.de - Carl von Ossietzky University, Department for Computer Science, Oldenburg, Germany

In summary, citizens and citizen scientists demand a low-cost and easy to use way to monitor and log their surrounding environmental situation, to spotlight and report environment influences, as well as to observe the effects as the consequences of their reports, or to increase transparency, or even for more complex research purposes.

This paper is organized as follows: After introducing existing works on citizen science-oriented acquisition and analyses of environmental data in Section 2 the approach of *Guerilla Sensing* is presented in Section 3. Section 4 exemplifies the realization of this approach, and Section 5 outlines a set of scenarios the *Guerilla Sensing* approach is applied to. Finally, Section 6 draws a conclusion.

2 Related Work

The direct surrounding environmental situation of citizens can be described by several parameters. These parameters include, e.g. the carbon dioxide concentration, radioactive pollution, and fine dust. The particular environmental parameters at a point in time can be quantitatively determined by measurements. Thus, individually measured values can be compared with each other or to reference values. Accordingly, the environmental situation, citizens are confronted with, can be classified by measuring and comparison. If there is a hypothesis of a harmed environmental condition, the characteristics of environmental parameters concerned can be measured and compared against specified reference values. Depending on this result of the comparison, the hypothesis is confirmed or refuted.

The environmental parameters fluctuate over time. Individual measurements may also exhibit various types of error. Often, several measurements taken over time are needed to come to a well-founded statement. Repeated manual measurements are time-consuming. With an autonomous solution for measuring environmental parameters, as in the citizen science projects *luftdaten.info* [Lu21] and *senseBox* [GI21b], the time required can be reduced and measurements can be performed consistently.

***luftdaten.info*:** With *luftdaten.info*, the *OK Lab Stuttgart* provides a platform for the collection and analysis of fine dust measurements. Interested citizens can participate in the collection of fine dust measurements by building and / or operating a sensor node. On the website of *luftdaten.info*, the construction of a sensor node is described by a construction manual. Additionally, a firmware for the sensor node is offered for download. Once put into operation, sensor nodes measure the fine dust load of the surrounding air and send the values measured to *luftdaten.info*. Currently, measured values are visualized on a map. On another map, showing the individual sensor nodes, logged-in users can register for notifications of fine dust limit violations. In addition to fine dust, *luftdaten.info* also measures temperature, relative humidity, pressure, and noise. The active, country-specific

offshoots of *luftdaten.info* called *influencair.be*², *airbg.info*³, and *luftdata.se*⁴ demonstrate the transferability and applicability of this citizen-centered idea.

The assembly instructions and the firmware for the sensor node make it possible for every citizen to participate in the data collection of *luftdaten.info*. Moreover, a sensor node of *luftdaten.info* is composed of freely-available and low-cost components. When citizens distrust the collected data, or wish to have the data collected at other times or at a different location, then they can perform measurements by themselves. With the help of the maps provided on the website, citizens can actively inform themselves about the current environmental situation or register for self-selected fine dust limit violations. This gives citizens another source of information besides the authorities to assess their own environmental situation. The *OK Lab Stuttgart* focuses on air data and / or noise data with *luftdaten.info*. Thus, *luftdaten.info*, in particular, covers environmental parameters influenced by traffic. In order to detect a variety of pollution types, a wider range of environmental parameters needs to be monitored. The *senseBox* approach from *GI@SchoolLab* has a flexible design and can be equipped with many different sensors so that a wide range of environmental parameters can be recorded.

senseBox: In 2010, the student and research laboratory *GI@SchoolLab* at the Westphalian Wilhelms University of Münster in Germany developed *senseBox* [GI21b]. With *senseBox*, the *GI@SchoolLab* primarily aims at supporting teachers in educating students in the collection and analysis of environmental data. In addition, *senseBox* is suitable for citizen science projects and environmental data collection. The core of a *senseBox* is a microcontroller called *senseBox MCU*. It can be complemented by a wide range of communication modules and sensors offered on the *senseBox* website. A *senseBox* can be equipped with scenario-specific hardware modules and can be freely programmed with the graphical programming language *Blockly*. In addition, a website also offers teaching materials and documentation. Different environmental parameters can be measured and published on the *openSenseMap* [GI21a] platform, where each registered *senseBox* is displayed on an map.

The flexible design of the *senseBox* and the free programming feature using *Blockly* allows many possible applications. Communication modules and sensors can be selected specifically for each scenario. This can potentially detect many types of pollution at different locations, based on different preconditions. Like *luftdaten.info*, *GI@SchoolLab* offers with *openSenseMap*, a map on which citizens can actively inform themselves about the current environmental situation. However, technical knowledge is required to program a *senseBox*. Furthermore, the hardware components of a *senseBox* are more expensive compared to other freely available ones on the market. In order to offer as many citizens as possible the possibility to record and classify their environmental situation, the required technical knowledge and the costs for the hardware components must be as low as possible.

² <https://influencair.be/>

³ <https://airbg.info/>

⁴ <https://luftdata.se/>

Especially, if there is a conflict of interest with authorities or polluters due to particular measurements, then these measurements cannot be performed undisturbed without further arrangements. Following the movement *Guerilla Gardening* [Mü11], in which participants act covertly to protest land scarcity, measurements can also be performed covertly. If the conflicting party does not know that a measurement is being made, then that party cannot disrupt or influence the measurement. But neither *luftdaten.info* nor *senseBox*, respectively *openSenseMap*, actively support covert measurements.

Even if measurements can be performed and published undisturbed, this does not lead to a stop of the observed pollution. Politics and authorities must be involved so that polluters can be fully identified and their undertakings be stopped. By the *Guerilla Sensing* approach, the ideas behind *luftdaten.info*, *senseBox* and *Guerilla Gardening* are combined, complemented, and extended to an overall concept to ensure the environmental wellbeing of citizens.

3 The Guerilla Sensing Approach

The *Guerilla Sensing* approach is based on a closed control loop (cf. Fig. 1) that aims at stopping environmental pollution. As indicated, a polluter has a harmful impact on the

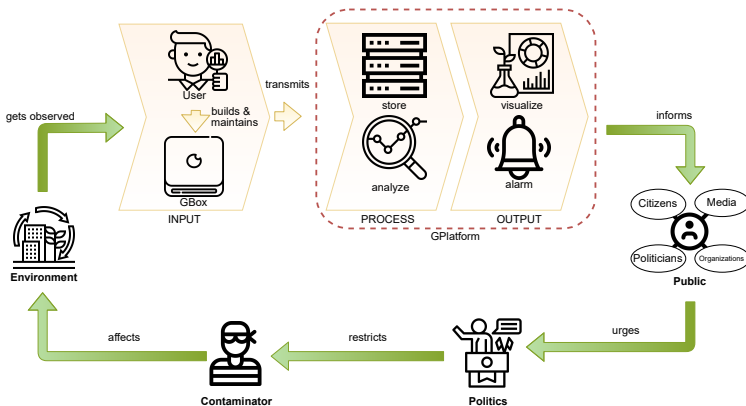


Fig. 1: The control loop behind *Guerilla Sensing*, taken from *Guerilla Sensing* [B121].

environment. Citizens who suspect environmental pollution would like to investigate and confirm or refute their suspicions. To do this, citizens can assemble and operate a sensor node, called *GBox*, scenario-specific from freely available and low-cost components to detect their environmental situation. The instructions needed to build a sensor node are provided on a website along with corresponding firmware. Thus, even citizens with no prior technical knowledge can build and operate an *GBox* on a small budget. If citizens are distrustful of the measurement data, they can perform their own measurements at any time.

Once in operation, a *GBox* autonomously monitors environmental parameters. The collected measurement data is sent from a *GBox* to the *GPlatform*. On the *GPlatform*, the data is

stored, processed, and published on a map, the latter possibly with an adjustable delay. By publishing the measurements, citizens can be informed about the environmental situation. With the adjustable delay of publication, measurements can be covertly taken over some period of time. Polluters have no chance to detect covered measurements via the *GPlatform*.

In case of over- or undershooting of reference values, previously registered citizens and institutions are notified immediately. Even without an active search, citizens and institutions can get informed about environmental pollution. As soon as the collected measurement data confirms a suspicion of environmental pollution, citizens and institutions can use the measurement data to motivate politics and authorities to take action. Politics and authorities are expected to follow up on the suspicion and, if pollution is present, to fully identify the polluters and stop their activities. Whether a polluting activity is stopped can in turn be verified using a *GBox*. Thus, the control loop of *Guerilla Sensing* is closed.

4 Guerilla Sensing Realization

An environmental information system matching the *Guerilla Sensing* approach has been developed by 11 students within a year.⁵ The system consists of three major components: 1) one or more *GBoxes*, 2) a *GPlatform* and 3) one or many *GClients* as shown in Fig. 2. Environmental parameters are observed by a *GBox*. A *GBox* transmits data to the *GPlatform*. A user interact with the *GPlatform* using a *GClient* via a webpage-based interface⁶.

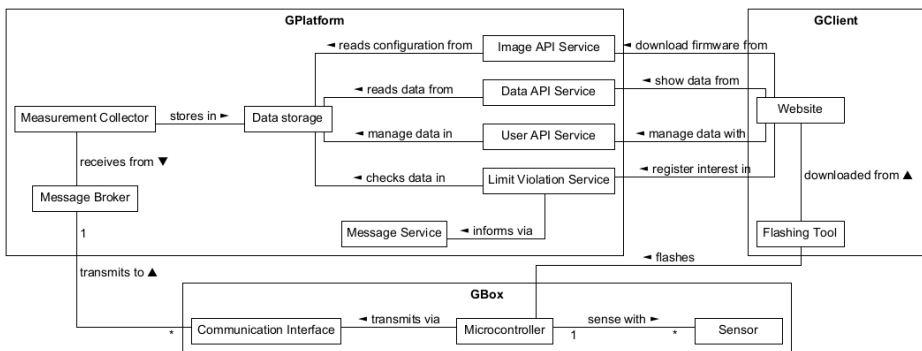


Fig. 2: *Guerilla Sensing* Architecture

A *GBox* (bottom part in Fig. 2) is composed of a microcontroller, mission-specific sensors, and a communication interface. Various microcontrollers compatible with the *Arduino* framework are supported such as the *ESP32*. Temperature, humidity, air pressure, soil

⁵ The authors would thank the project group *Guerilla Sensing* namely: Christian Blank, Florian Brandt, Gerald Fortmann, Martin Haase, Niklas Holtz, Dennis Kempf, Elmo K pker, Kevin Lang, Pascal S fken, and Philipp Thien for their work.

⁶ <https://www.guerilla-sensing.de/>

moisture, radioactivity, redox potential, pH-value, water conductivity, fine dust, and various gas sensors are supported. As communication interfaces, WiFi, LoRa, and GSM technologies are implemented. A user can build a *GBox* according to an environmental measuring scenario, he or she is interested in. The user gets various support by the platform website.

On the platform website, a user can select the desired components using a graphical configurator that exists in a simple and in an expert mode version. Based on the specified configuration, the *GPlatform* automatically creates a specific firmware and offers it for download. Using a *Flashing Tool*, easily obtainable by download, a user then can forwardly install the generated firmware on a *GBox*. Finally, a manual, also published by the website, helps with assembly of the already selected components in order to obtain a ready-to-use customized *GBox*. In Fig. 3, an assembled *GBox* is shown. The *GBox* consists of a *ESP32* microcontroller with a LoRa communication interface. The sensors are chosen such that the *GBox* can measure fine dust, temperature, and humidity.

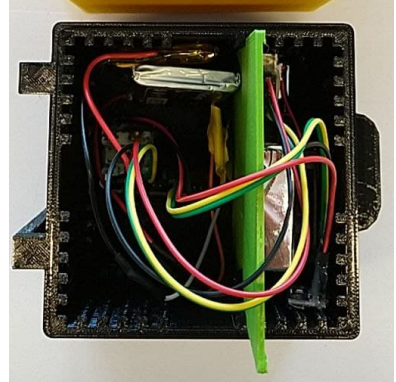


Fig. 3: A sample *GBox*.

Two batteries are connected to the *GBox* for autonomous operation. All components of the *GBox* are housed in a 3D-printed case.

A operating *GBox* transmits the measured data to the *Message Broker* of the *GPlatform* (top left part in Fig. 2). Communication with the *Message Broker* is done using the *MQTT* (Message Queuing Telemetry Transport) protocol. The messages received by the *Message Broker* include the unique identifier of the sending *GBox*. The *Measurement Collector* receives the messages and stores the received values together with identifier in a relational database (*Data storage*). Afterwards, the values can be accessed via the *Data API Service*.

With a *GClient* (top right part in Fig. 2), such as a personal computer or mobile phone, the user can access the website. The website uses the *Data API Service*, among others, to visualize the measured values on a map and to display trends over time. Also user-specific settings can be made on the website via the *User API Service*. Among other features, it is possible to delay the publication of measured values for a *GBox* and to conceal its true location for some amount of time. These features allow the covert observation of environmental parameters if doing so openly would hamper the measuring process. As a consequence, there is no chance to detect hidden measurements via the platform itself. Finally, people can register themselves on the website for information or events, he or she is interested in, e.g., limit violations. The *Limit Violation Service* checks cyclically whether any exceedance has been measured. If there is a limit violation, then the *Limit Violation Service* notifies all people who have registered for the respective limit violation via the *Message Service* using an e-mail. The limits to which a user can subscribe, like the recommendation of the European Union for fine-dust concentration [EU08], are stored in the *Datastorage*.

5 Guerilla Sensing Application

The *Guerilla Sensing* approach has already been applied to some case studies, in order to exemplarily evaluate its functionality (i.e. prove its concept) in the context of a large city. The scenarios described here are supposed to show the suitability of important design goals of *Guerilla Sensing*, the functionality in practice of a *GBox*, the *GPlatform*, and the communication techniques connecting them. The used sensors are all off-the-shelf and low-cost. Unfortunately their measurements are not too scientifically accurate. But if the observed environmental parameter changes significantly, the sensors are capable of sensing this deviation. Furthermore, the *GBox* offers the necessary interfaces to connect high-quality sensors, in case an application actually demands it.

Scenario A: Measuring the radiation of a castor transport: Nuclear power plants produce radioactive waste. This waste is often transported over railways to a safe disposal facility. Some people, living or working close to the transportation route, do fear radiation overexposure. To protect people and environment from radiation overexposure, radioactive waste is incarcerated in massive highly radiation-protecting hulls (called castors). In this case study, we want to verify whether castors actually do their dedicated jobs. Furthermore, the exact date of a castor transportation is typically classified, thus, the setup was spontaneously done within a couple hours, evaluating the simple and fast adaptation of *Guerilla Sensing*.

In the scope of this scenario, *GBox* is equipped with a radiation sensor of *RadiationD-v1.1(caJOE)* type [Ac19]. The sensor measures β -radiation in range of 100 mR/h to 1800 mR/h and γ -radiation in range of 20 mR/h to 120 mR/h. It is placed in a cardboard box, together with an ESP32 microcontroller. As communication technique, WiFi-technology is used. In sum, the costs for the *GBox* for this specific scenario added up to 75€ [BI21, p. 254]. This includes the microcontroller, communication module, sensors, housing, wiring, and battery. The box is placed on a leveled wall in 1.2 m height behind a 4 mm thick glass window. The railway track is straight line-of-sight in a distance of 46 m.

On November 3, 2020, at 20:35, the castor train has passed the *GBox* in the described distance. Figure 4 shows the radiation measurements of that day. Obviously and luckily, the *GBox* did not measure any significant changes in radiation in moment of the trains passage (arrow in Fig. 4). Thus, the castor seemed to work as supposed, shielding the radioactive waste as that no significant amount of radiation was released. We, thus, are confident that no harm neither for people nor the environment stemmed from the castor in the observed track section. *Guerilla Sensing* worked flawlessly despite the spontaneous character of the scenario.

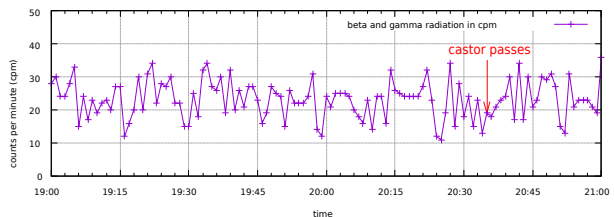


Fig. 4: Timeplot over β - and γ -radiation. At 20:35 a castor passes.

Scenario B: Measuring and logging fine dust particles near two public schools: Schools in large cities are often close to major roads. Parents worry about the health of their children and question their children’s school to being a safe place. A schools’ administration often does not know how to cope with those fears or is afraid of too many fine dust particles forcing the school to be temporarily closed.

Two *GBoxes* equipped with an SDS011 [No15] fine dust sensor each, were installed in *guerilla*-like operations. This was necessary due to the fact that a school in a potentially highly fine dust polluted area refrained from any cooperation and explained their behavior by bizarre and administrative hindrances. The *GBoxes* operated exposed to the forces of nature, on battery supply and in the danger of being recognized. Different communication techniques had to be used due to different characteristics at the measuring location. The sensor measured fine dust PM10 (particles smaller than $10\ \mu\text{m}$) and PM2.5 (particles smaller than $2.5\ \mu\text{m}$) separately in a range from $0\ \mu\text{g m}^{-3}$ to $999.9\ \mu\text{g m}^{-3}$. The *GBoxes* were hidden outdoors in about 2 m height near the school premises. In sum, the costs for each *GBox* for this specific scenario added up to 85€ [B121, p. 254]. This includes the microcontroller, communication module, sensors, housing, wiring, and battery.

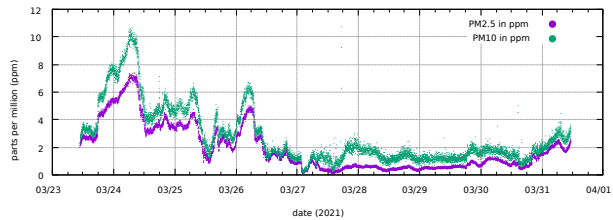


Fig. 5: PM2.5 and PM10 concentration near a school over eight days.

The *GBoxes* operated exposed to the forces of nature, on battery supply and in the danger of being recognized. Different communication techniques had to be used due to different characteristics at the measuring location. The sensor measured fine dust PM10 (particles smaller than $10\ \mu\text{m}$) and PM2.5 (particles smaller than $2.5\ \mu\text{m}$) separately in a range from $0\ \mu\text{g m}^{-3}$ to $999.9\ \mu\text{g m}^{-3}$. The *GBoxes* were hidden outdoors in about 2 m height near the school premises. In sum, the costs for each *GBox* for this specific scenario added up to 85€ [B121, p. 254]. This includes the microcontroller, communication module, sensors, housing, wiring, and battery.

In case of the first measurement series, shown in Figure 5, a *GBox* was connected to the *GPlatform* via *GSM*. The fine dust concentration was measured every minute over a duration of eight days. This way, the *GBox* worked about 24 h before the battery was drained and had to be replaced. Thereby, the PM2.5 fine-dust concentration was at its maximum in the morning of Wednesday, the 24th of March, of $7\ \mu\text{g m}^{-3}$, which was not even close to the limit value of $25\ \mu\text{g m}^{-3}$. The PM10 fine dust concentration was $10.5\ \mu\text{g m}^{-3}$ at that time. The European Union recommends an annual average concentration in atmosphere below $25\ \mu\text{g m}^{-3}$ for PM2.5 and $40\ \mu\text{g m}^{-3}$ for PM10 [EU08].

In case of the second measurement series, a *GBox* was connected to the *GPlatform* via *LoRaWAN* through a *LoRaWAN* gateway [Ha18]. The fine dust concentration was measured every 50 s for a duration of two days. This way, the *GBox* worked for about 36 h before the battery was drained and had to be replaced. Again, the measurements never came even close to the recommended limit values.

In conclusion, the *GBoxes* measured fine dust concentration far below the limit values, recommended by the European Union. At the time of measurement, the observed schools were uncritical in regard of fine dust. The *GBoxes* stayed intact and undiscovered over the long measuring interval and were capable of measuring fine dust even in low concentrations. Both

communication techniques worked successfully, proving the flexibility of *Guerilla Sensing* in this regard. As lower values were observed potentially due to reduced traffic in pandemic times, measurements should be repeated, when society is „back-to-normal.“ Nevertheless, the scenario shows how the *Guerilla Sensing* approach can be used such that all parties involved benefit: the parents by now having their peace of mind and school administration by being assured that they truly offer a secure learning location for their students.

6 Conclusion

The *Guerilla Sensing* approach allows to measure, log, analyze, and report a variety of environmental parameters. By using differing application scenarios, we could successfully prove the idea and realization of *Guerilla Sensing*. It is very flexible according to communication techniques, as it supports WiFi, GSM, and LoRa technology. The *GBox* is small enough to be effectively hidden and can work autonomously without any wired energy or wired communication supply. By that, unrecognized *guerilla*-like operations are possible, as we have demonstrated by Scenario B. Together with a solid housing, this enables the *GBox* to run for indoor as well as outdoor or segregated measurement scenarios.

Due to the fact that the *GPlatform* automatically generates the individual firmware of a *GBox*, no programming skills are required to run a sensor. The components of a *GBox* are off-the-shelf hardware that can be ordered at common online stores for a small amount of money (typically between 40€ and 130€, depending on the set of required sensors). The *GPlatform* assists in the assembly of the *GBox* by detailed step-by-step-guides. A simple assembly process consists of connecting wires, which truly requires only basic knowledge.

The *GPlatform* sends a e-mail to interested entities, if a legal limit is exceeded. It shows all measured data to all the platform users, if a *GBox* operator wishes to do so. *GPlatform* users can inform themselves and interested entities about the environmental situation in their region when *GBoxes* are present.

In conclusion, the *Guerilla Sensing* approach enables citizen to resist against environmental exploitation by providing automated functions to observe, discover, log, visualize, and report to politics & authorities in charge, demanding or urging them to act. In this respect, we regard *Guerilla Sensing* as a valuable asset in difficult times.

References

- [Ac19] Actpohomoc: M4011 Geiger Tube Specification, Github, 2019, URL: <https://github.com/SensorsIot/Geiger-Counter-RadiationD-v1.1-CAJ0E->.
- [AG20] AG, S.: Datasheet Sensirion SCD30 Sensor Module, 1.0, May 2020, URL: https://www.sensirion.com/fileadmin/user_upload/customers/sensirion/Dokumente/9.5_CO2/Sensirion_CO2_Sensors_SCD30_Datasheet.pdf.

- [BI21] Blank, C.; Brandt, F.; Fortmann, G.; Haase, M.; Holt, P.; Holtz, N.; Kempf, D.; K pker, E.; Lang, K.; S fken, P.; Thien, P.: Dokumentation zur Projektgruppe Guerilla Sensing Oldenburg and Beyond, 2021.
- [BR00] Bansal, P.; Roth, K.: Why Companies Go Green: A Model of Ecological Responsiveness. *Academy of Management Journal* 43/4, pp. 717–736, Aug. 2000, URL: <https://doi.org/10.5465/1556363>.
- [Br03] Briggs, D.: Environmental pollution and the global burden of disease. *British Medical Bulletin* 68/1, pp. 1–24, Dec. 2003, ISSN: 0007-1420, eprint: <https://academic.oup.com/bmb/article-pdf/68/1/1/25151995/ldg019.pdf>, URL: <https://doi.org/10.1093/bmb/ldg019>.
- [DZL21] Deng, S.; Zou, B.; Lau, J.: The Adverse Associations of Classrooms’ Indoor Air Quality and Thermal Comfort Conditions on Students’ Illness Related Absenteeism between Heating and Non-Heating Seasons—A Pilot Study. *International Journal of Environmental Research and Public Health* 18/4, p. 1500, Feb. 2021, URL: <https://doi.org/10.3390/ijerph18041500>.
- [EU08] EU: DIRECTIVE 2008/50/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 21 May 2008 on ambient air quality and cleaner air for Europe./, June 2008, URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX%3A32008L0050&qid=1419250736800>.
- [GI21a] GI@SchoolLab: opensensemap.org, 2021, URL: <https://opensensemap.org>.
- [GI21b] GI@SchoolLab: senseBox, 2021, URL: <https://sensebox.de/de/about-us.html>.
- [Ha18] Haxhibeqiri, J.; Poorter, E. D.; Moerman, I.; Hoebeke, J.: A Survey of LoRaWAN for IoT: From Technology to Application. *Sensors* 18/11, p. 3995, Nov. 2018.
- [Is01] Isenmann, R.; Lenz, C.; M ller-Merbach, H.; Reitz, C.: Nachhaltigkeitsberichterstattung: Internet als Medium zur Integration der Gesch fts- und Umweltberichterstattung von Unternehmen. *Wirtschaftsinformatik Proceedings* 2001 58/, 2001.
- [Lu21] Lutz, J. A.: luftdaten.info, 2021, URL: <https://luftdaten.info>.
- [MQ21] MQTT.org: MQTT, 2021, URL: <https://mqtt.org>.
- [M 11] M ller, C.: Guerilla Gardening und andere Strategien der Aneignung des st dtischen Raums. In: *Eigensinnige Geographien*. VS Verlag f r Sozialwissenschaften, pp. 281–288, 2011.
- [No15] Nova Fitness Co., L.: Laser PM2.5 Sensor specification, Oct. 2015, URL: <http://inovafitness.com/en/a/chanpinzhongxin/95.html>.
- [SN18] Stewart, R.; Niero, M.: Circular economy in corporate sustainability strategies: A review of corporate sustainability reports in the fast-moving consumer goods sector. *Business Strategy and the Environment* 27/7, pp. 1005–1022, Mar. 2018, URL: <https://doi.org/10.1002/bse.2048>.

Maritime Spatial Planning and Integrated Coastal Zone Management in Higher Education: Utilizing the MSP Challenge Serious Game

Tanja Behrendt¹, Malena Ripken¹, Harald Warmelink², Thomas Klenke¹

Abstract: Digitalization is gaining increasing attention in Higher Education (HE). The integration of digital tools into instructional settings is particularly challenging. However, it offers many opportunities to improve the learning process of students, especially in interdisciplinary teaching scenarios such as teaching sustainable usage of space and resources i.e. for the coastal zones and the marine areas. Worldwide, Marine Spatial Planning (MSP) and Integrated Coastal Zone Management (ICZM) are much needed approaches to manage and organize the increasing use of the sea and coastal areas. Both are complex fields that are attracting more and more attention in interdisciplinary HE. Correspondingly designed, the module ‘Planning and Management of Coastal Zones and Sea Basins’ at the University of Oldenburg, Germany, provides a case for integrating digital tools into HE. In 2020, the digital serious game ‘MSP Challenge’ was used in an online learning format. This interactive and collaborative tool supports informed decision making based on real and simulated data, comparable to business (decision) processes based on environmental information systems (EIS). Therefore, the MSP Challenge game fosters not only the understanding of the complex topic but additionally methodological skills which can be transferred to the usage EIS. While playing, students become able to (1) evaluate and simulate impacts of uses on coastal and marine environments, (2) describe the main interactions in ecosystems, (3) conceptualize information for sectoral or integrated MSP and (4) reflect on the role and use of data. In the presented case master students studying “Water and Coastal Management” participated in the module. Moreover, the digital serious game and the interdisciplinary topics of MSP and ICZM provides additional opportunities to explore subtopics (e.g. IT-related) from other disciplinary perspectives.

Keywords: Higher Education, Innovation, Serious Gaming, Maritime Spatial Planning, Integrated Coastal Zone Management.

1 Introduction

Digitalization is gaining increasing attention in Higher Education (HE). Different aspects are considered in digitalization strategies of HE institutions, mainly to foster digital skills of future professionals. The integration of digital tools into instructional settings is particularly challenging but it implies many opportunities to improve the learning process of students especially in interdisciplinary teaching settings such as teaching sustainable usage of space and resources, i.e. of coastal zones and marine areas.

¹ COAST- Centre for Environment and Sustainability Research, University of Oldenburg, Ammerlaender Heerstr. 114 – 118, 26129 Oldenburg, Germany, t.behrendt@uol.de

² Academy for Games and Media, Breda University of Applied Sciences, Mgr. Hopmansstraat 2, 4817 JS Breda, the Netherlands, Warmelink.H@buas.nl

European seas, and especially the North Sea and Baltic Sea belong to the busiest seas for maritime industries worldwide [Ha08]. Their shared resources represent a crucial asset but also a mutual territorial challenge to European countries. Also, coastal zones show high socio-economic activity with high vulnerability (especially to climate change) and a continuously changing environment. Marine Spatial Planning (MSP) and Integrated Coastal Zone Management (ICZM) are much needed approaches to manage and organize the use of the sea and coastal areas to prevent uncoordinated planning and management. Both coastal zones and marine areas are complex systems with an important ecology, as well as numerous human uses and activities. Various challenges arise from local to national to international levels. MSP and ICZM are complex fields of societal action. Accordingly, they constitute interdisciplinary fields of study with significantly overlapping important topics with respect to theoretical concepts, disciplinary approaches and problem-driven interdisciplinary integration in HE. Students in their future careers will encounter complex issues (such as sustainable development of marine and coastal environments) and must be able to make informed decisions.

Even though it is also an already legally binding mechanism within the EU, MSP is still relatively new and has no classical textbook for knowledge transfer. Educational innovations and trials are therefore of enormous importance to communicate current research and challenges of the sectors involved in both MSP and ICZM.

Serious/simulation games (SG) can be used to facilitate student engagement, provide opportunities to deepen mutual understanding, and explore and integrate new ideas and solutions, also for water, sea and ocean issues [Me20]. SG are collaborative tools that make use of the technologies of games and the principles of play to achieve objectives that are valued not for their intrinsic value (i.e., merely for the sake of entertainment) but for the extrinsic value achieved by the consequences, such as the fact that engagement and feedback in play is a good condition for learning and change [Ma16].

In this paper the goal is to address the challenges related to MSP and ICZM education through interdisciplinary teaching with an innovative SG approach. In other words, the goal is to integrate a SG within a MSP and ICZM module efficiently and effectively, i.e., making use of the different SG features to encourage a MSP and ICZM learning process. The SG will be evaluated concerning its relevance to the intended learning outcomes of a MSP/ICZM-HE module.

2 Methods and Educational Setting

SGs offers novel opportunities for problem-based interdisciplinary education for mutual understanding, exchange of ideas, development of new approaches and understanding the complexity of different activities. SGs are of particular value in teaching because the direct interaction of students and teachers with the subject matter is carried out through trial and error, practically and interactively. As such playing a SG is a form of collaborative discovery learning.

2.1 MSP Challenge

Since 2011, under the name ‚MSP Challenge‘ different SGs in support of MSP have been developed and applied, i.e., for education, social learning, stakeholder engagement and decision-support [Ab19], [Ke18], [Ma14]. MSP Challenge is an interactive and collaborative tool for informed decision making to experience the complexity of possible actions of marine planning and coastal management based on real data. It combines role-play, game technology, geodata and simulation models to offer a planning-oriented learning tool for MSP and a communicative environment that makes players think, talk and interact.

The MSP Challenge Simulation Platform was developed to further support the development of digital SGs concerning particular sea basins [Sa20]. The simulation platform allows players to creatively operate it for scenario development, and/or for multi-player SG sessions. This means that the platform supports multiple purposes such as planning scenario exploration, co-design, validation or policy-oriented learning [Ab19]. The simulation platform helps users or players to reflect on their individual and combined decisions using real geodata and simulation models for energy production, shipping and environment.

MSP Challenge communicates through Open Geospatial Consortium standards to obtain geodata relevant to the edition at hand. For the North Sea edition this means using WMS hosted by a GeoServer installation to obtain both vector and raster data from several different sources, notably EMODnet Human Activities, Copernicus Marine Service and bespoke sources (full list on https://community.mspschallenge.info/wiki/Data_sources). Subsequently, MSP Challenge uses much of this data as well as data created by the users/players themselves for energy, shipping and marine ecosystem simulations. In turn these simulations create new and altered raster layers and key performance indicators (KPI). The ecosystem simulation communicates with a Ecopath-with-Ecosim model specific to the sea basin at hand [St20]. The shipping and [Gr19] energy [Hu18] simulations are bespoke applications.

In the North Sea Edition up to 8 countries can be played with 2-6 players per country. The players are confronted with the general goal of developing the EEZ of the country they represent considering e.g. „Blue Growth“ and „Sustainability“. The exact interpretation and implementation of this goal is subsequently of course up to the players to discover as they play. The interaction and effects between human activities, geosystem and ecosystem are simulated (incl. Feedback loops) and visualized using heat maps and indicators.

2.2 MSP / ICZM Higher Education Module

The module ‘Planning and Management of Coastal Zones and Sea Basins’ is an established master course at the Carl von Ossietzky University of Oldenburg, Germany. It consists of two lectures “Marine Spatial Planning” and “Integrated Coastal Zone Management”. Within the module, a heterogeneous group with students from all over the

world with an interdisciplinary background - learns together. Adding to the variety of previous knowledge and experiences, the current studies of the students vary as well: the module is elective within certain master programs focusing on natural science, economics or management. The module covers the complex interrelations between the natural marine and coastal spaces on the one hand, and human activities on the other.

The main goal of module is to enable students to develop a theoretical, practical and critical understanding of the notion of MSP and ICZM and the wider scientific and governance framework in which it is being implemented.

2.3 Methodological-didactical concept

The two lectures are generally separated lecture series (separated ICZM and MSP issues) with a significant overlap in the intended skills. As an innovative approach of the module, the two lectures have been linked via the MSP Challenge SG. The learning process and experiences of the gaming events were reported and reflected by the students on the associated Wordpress blog and in their final report (final assignment).

To ensure a significant learning impact of the SG, the following considerations contributed to the final time structure of the module. First, before the SG started, an overview of the topics needed to be given in which important terms would be clarified. Second, students needed an introduction to the SG involving a synchronous tutorial with asynchronous self-learning task. Third, students needed to be able to share their ideas, strategies and conclusions for the SG with the whole group. Based on their reflection, new or adjusted strategies could then be developed and applied in the next SG session.

The SG experience was one of the cornerstones of the module. Therefore, in the middle of the lecture series two SG days have been scheduled. The SG was designed as a role-play: the players represented different stakeholders of five European Countries along the North Sea. They received the task to develop their country along personal priorities keeping the EU policy concerning Blue Growth, Renewable Energy and Environmental Protection in mind. As the whole module was realized in an online setting, different virtual meeting rooms (main room, group rooms, additional virtual meeting rooms) have been prepared for communication, discussion and debriefing. The main room was used to meet with the whole group at the beginning and in the middle of the day for guidance and clarifying open questions and at the end of the day for a joint evaluation of the resulting plans and simulations.

3 Results

To achieve the goals of the module and to develop an adequate instructional design, intended competences and learning outcomes for the students were defined and the impact

of the different learning activities (lectures, gaming, wordpress, final report) on the learning outcomes has been evaluated.

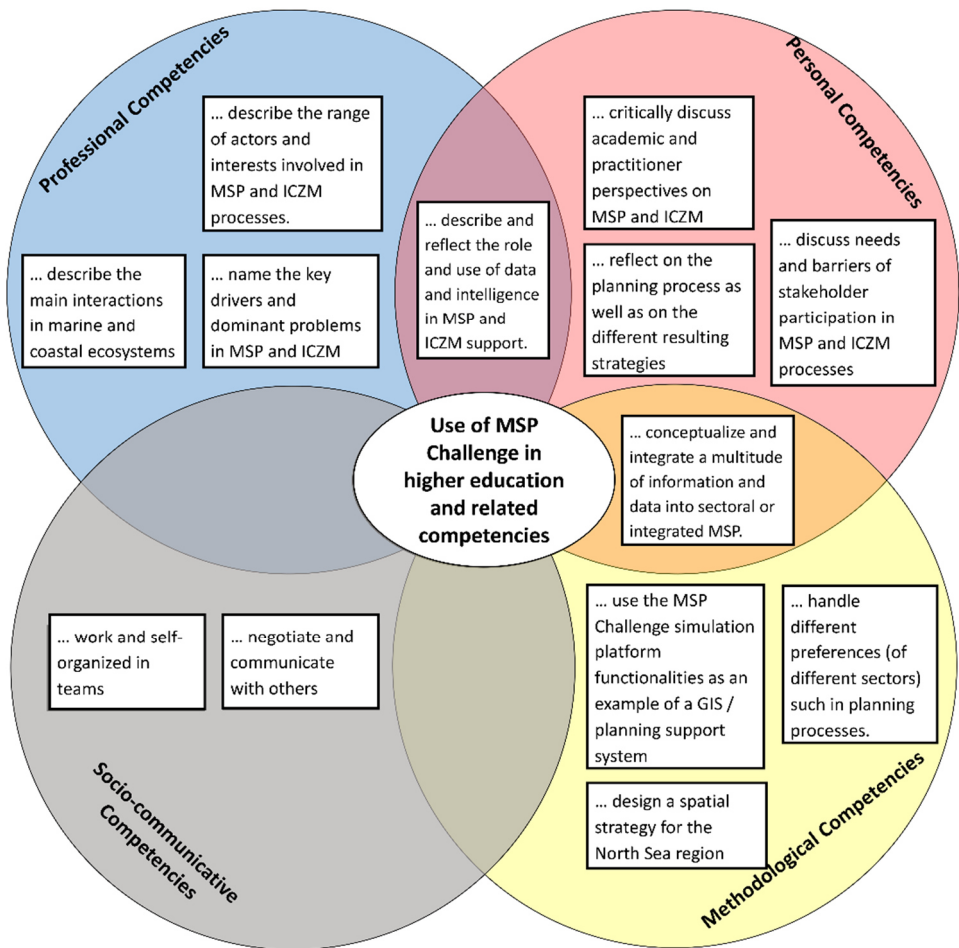


Fig. 1: Learning outcomes of the module "Planning and Management of Coastal Zones and Sea Basins" addressed by the MSP Challenge game depending on different competence areas: professional competencies (blue), personal (red), methodological (yellow) and socio-communicative competencies (grey) and the interfaces between personal and professional competencies (violet) and personal and methodological competencies (orange).

Among 21 different competences and skills that the students shall gain during the module, thirteen learning outcomes have been identified to be supported by the MSP Challenge SG (with varying intensity). Fig. 1 shows these learning outcomes. To receive a detailed impression of the impact, the competences are distinguished into the competence areas professional, personal, socio-communicative and methodological competencies. The use

of the MSP Challenge SG in the module addresses different learning outcomes within all four competence areas. A specifically high impact is anticipated for the development of personal, social-communicative and methodological competencies. While using the SG, students learn to evaluate and simulate impacts of uses on coastal and marine environments. This requires to (a) describe the main interactions in ecosystems, (b) conceptualize information for sectoral or integrated MSP and (c) reflect on the role and use of data (simulation, models, AI).

Concerning the improvement of the learning process for an efficient integration of the SG into the HE module, the impact of the SG approach is evaluated in more detail: the different learning outcomes are attributed to the different features included in the MSP Challenge SG software and the SG play process. The main features are summarized in Tab.1 and 2 and their attribution to the intended learning outcomes are given in Tab. 3.

The features delivered by the MSP Challenge game are distinguished between „mechanical features“ and “dynamical features”. Mechanical features are structurally included in the SG software. These features are the basis for the application of the SG in an educating setting. The players work with a map of the North Sea including important information as different layers. They are organized (logged in) as groups related to their countries, able to draw plans for their EEZ, able to communicate with other countries, and receive automatic feedback (i.e. restrictions or requirements). Based on the created plans, the simulations calculate their impact on the different data layers. The simulations can be started, accelerated or stopped by the teacher depending on the didactical need. The simulation results are visualized in terms of pressures and indicators in maps and diagrams. The time view control allows to go back and forth in time to evaluate the evolution of a specific dataset.

Dynamical features are those introduced by the SG play process due to interactions between different mechanical features or different players, or through intervention, support or tasks by the teacher. An important aspect in the SG play process is the role play. Each player has a different responsibility for his/her country with a goal for the considered sector (based also on the knowledge gained during the lectures). Every player is able to create a plan (e.g. to declare an environmental protection area) giving it also an implementation date facing the lack of time and space. A multitude of data and maps can be evaluated for making decisions. This includes learning about the management of data with certain hurdles as lack of data, different scales in time and space, incorrect, incomplete and unavailable data which is also discussed within the whole group. By communicating with other group members and with other groups (i.e. included approval system) a consensus has to be reached.

The role of the teacher and the forum discussions (course concept) to intensify the impact of the dynamical features is crucial. In break-out moments and debriefing sessions throughout the whole gaming process the students are guided. Within these joint sessions results are shown and evaluated, interactions and impacts are highlighted (including the limitations of simulations concerning the considered mechanism) and reasoning of the

different individual decisions are discussed. A good time management to balance the need of time for planning and the need for guided reflection is crucial at this point. Improving the communication between the players could save time and enhance the outcome, for example by playing the game face-to-face or giving more guidance on communication paths.

Mechanics			
Simulations	Visualization	Tools	Organization
ecology, shipping, energy models	real data/ simulated data	planning tool for planning process	team organization
running simulation (incl. speeding up/pause)	maps with different layers	time based planning and simulation	approval system
simulation results	Time view control	drawing feature (points, lines, shapes/polygons)	communication option in approval system
limitations (e.g. resolution, considered mechanisms)	Key Performance Indicator (KPI)	knowledge base/wiki	restrictions and feedback mechanism
	pressures		

Tab. 1: Mechanical features introduced by the MSP Challenge game.

Dynamics		
Role Play	SG play Process	Course Concept
players representing role	predefined objectives / goals	support of players by teacher
proactively discussing	management of data	apply knowledge gained in lectures
teachers as legislative international actor	mapping	"break-out"-moments;
barriers as lack of time and space	evaluation of available information	debriefing
group of 2-3 students	strategy development	evaluation process
	Personal priorities	

Tab. 2: Dynamical features introduced by the MSP Challenge game.

Learning Outcome	Mechanical Features	Dynamical Features
...describe the range of actors and interests involved in MSP and ICZM processes.	different layers (management, activities, ecology), knowledge base; restriction feedback mechanism	role play (player representing role); defined objectives/goals
...describe the main interactions in marine and coastal ecosystems	i.e. marine ecosystems: simulation, maps, KPI, time view control (evolution during time)	role play (proactiveley discussing),
...name the key drivers and legal context in Europe	knowledge base / wiki; layers; approval system;	predefined objectives (country specific)/goals (EU policy); role play
...to describe and reflect on the role and use of data and intelligence in MSP and ICZM support.	MSP Challenge based on real data, simulations (incl. limitations, e.g. resolution, considered mechanisms)	management of data (different scales - time & space, incorrect, incomplete, contested, out-of-date and unavailable data)
...reflect on examples of land-sea interactions	pressures; ports; shipping and energy simulation / recreation	
...use the MSP Challenge simulation platform functionalities, as an example of a GIS / planning support system.	drawing feature (points, lines, shapes/polygons), time-based planning and simulation, real data and simulations, lack of relevant data	
...handle different preferences (i.e. sectors) in planning processes.	different data layers - non static layers that can be changed /simulations;	role play, support by teacher (i.e. synergies), gaming procedure: objectives that need to be reached;
...design a spatial strategy for the North Sea region	layers (existing situation), planning tool, approval system	SG task incl. defining objectives/ steps; evaluation of available information, mapping
...conceptualize and integrate a multitude of information and data into sectoral or integrated MSP	large amount of real-data and simulated data available in MSP Challenge, knowledge base	depending on personal/national strategies; apply knowledge gained in lectures; evaluate available data

...discuss needs and barriers of stakeholder participation in MSP and ICZM processes.	lack of time and space / speeding simulation up/pause; communication option in approval system	role play incl. barriers as lack of time and space
...reflect on the planning process as well as on the resulting strategies.	running simulation / results (immediately) visualized in different maps (time lapse); KPIs	"break-out"-moments; debriefing; evaluation process
...work and self-organize in teams, e.g., making role and task division and coordinate tasks	team organization / log in as team; possibility to take over others plan in the team	role play / group work: group of 2-3 students responsible for an EEZ of one European country
...negotiate/communicate with others (countries and stakeholders).	approvals by other countries; communication tab	role play (players representing different stakeholders); strategy development; personal priorities

Tab. 3: Attribution of MSP Challenge features to the intended learning outcomes of the higher education module. The colors indicate the different competence areas (see Fig.1).

4 Reflection

In terms of digitalization of Higher Education an example of a SG approach was presented. The applicability of the MSP Challenge simulation platform for MSP/ICZM education was evaluated in terms of its impact on the intended learning outcomes. Our initial evaluation shows a high potential for the SG approach in this setting. Overall, the module successfully utilized the ‘MSP Challenge’ SG to highlight the complex interrelation of coastal and maritime spaces for the benefit of interactive and interdisciplinary higher education. A limited impact of the game is seen for the reflection of land-sea-interaction. Further effects might be simulated (energy demand/supply interactions; agriculture pollution and other land uses) to increase the knowledge gain.

In the presented case master students studying “Water and Coastal Management” participated in the module. However, the classes are open to students from various backgrounds and disciplines, as the digital SG and the interdisciplinary topics of MSP and ICZM provides additional opportunities to explore subtopics from other disciplinary perspectives. With its inherent simulation and information tools it fosters not only the understanding of the complex topic but additionally methodological skills such as the collection and evaluation of relevant environmental data. Furthermore, multiple IT relevant components of the digital SG comprise opportunities to further develop, modify or adjust it (in terms of software development) in appropriate educational settings.

Further implementation of the SG approach needs a continuous evaluation to measure the learning progress supported by the game and to introduce further optimization. The associated evaluation of the progress in intended learning outcomes (rated by the students) remains to be analyzed.


Acknowledgements: This research is part of the project “Gamification of Higher Education” “Innovative Lehr- und Lernkonzepte: Innovation Plus” of the Ministry of Science and Culture of Lower Saxony, Germany. The authors would especially like to thank Igor Mayer, Xander Keijser, Lodewijk Abspoel, as well as the MSP Challenge Team at BUAs, the Netherlands and all supporters of the study and participating students.

References

- [Ab19] Abspoel, L.; Mayer, I.S.; Keijser, X; Warmelink, H., Fairgrieve, R., Ripken, M.; Abramic, A.; Kannen, A.; Cormier, R.; Kidd, S. Communicating Maritime Spatial Planning: The MSP Challenge Approach. *Marine Policy* 2019 <https://doi.org/10.1016/j.marpol.2019.02.057>
- [Gr19] de Groot, P., Boode, W., Pereira Santos, C., Warmelink, H., & Mayer, I. (2019). A shipping simulation through pathfinding: SEL within the MSP Challenge simulation platform. Paper presented at GAME-ON'2019, Breda, Netherlands. Available online: <https://pure.buas.nl/en/publications/a-shipping-simulation-through-pathfinding-sel-within-the-msp-chal>
- [Ha08] Halpern, B.; Walbridge, S.; Selkoe, K.; Kappel, C.; Micheli, F.; D’Agrosa, C.; Bruno, J.; Casey, K.; Ebert, C.; Fox, H.; et al. A global map of human impact on marine ecosystems. *Science* 2008, 319, 948–953.
- [Hu18] Hutchinson, K., Warmelink, H.J.G., Boode, W., Pereira Santos, C.A., & Mayer, I.S. (2018). An offshore energy simulation through flow networks: CEL within the MSP Challenge 2050 simulation game platform. In D. Claeys, & V. Limere (Eds.), *ESM 2018 proceedings* (pp. 157-163). Eurosis. Available online: <https://pure.buas.nl/en/publications/an-offshore-energy-simulation-through-flow-networks-cel-within-th>
- [Ke18] Keijser, X.; Ripken, M.; Mayer, I.; Warmelink, H.; Abspoel, L.; Fairgrieve, R.; Paris, C. Stakeholder Engagement in Maritime Spatial Planning: The Efficacy of a Serious Game Approach. *Water* 2018, 10(6), 724. <https://doi.org/10.3390/w10060724>
- [Ma14] Mayer, I.S.; Zhou, Q.; Abspoel, L.; Keijser, X. Gaming the Future of the Ocean: The Marine Spatial Planning Challenge 2050. In *SGDA 2014: Serious Games Development and Applications*; LNCS 8778; Springer: Cham, Switzerland, 2014; pp. 150–162.
- [Ma16] Mayer, I.S. *Playful Organisations & Learning Systems*; NHTV Breda University of Applied Science: Breda/The Hague, The Netherlands, 2016; ISBN 978-90-825477-0-2. Available online: <https://pure.buas.nl/en/publications/playful-organisations-amp-learning-systems>
- [Me20] Medema, W.; Chew, C.; Adamowski, J.; Mayer, I.; Wals, A. E. J. (2020). Understanding Game-based Approaches for Improving Sustainable Water Governance : The Potential of Serious Games to Solve Water Problems. (Special issue ed.) *MDPI* 2020. <https://doi.org/10.3390/books978-3-03928-763-5>

- [Pe20] Pereira Santos, C.; Warmelink, H.; Boode, W.; de Groot, P.; Hutchinson, K.; Goncalves, M.; Steenbeek, J.; Mayer, I. A digital game-based simulation platform for integrated marine spatial planning: design challenges and technical innovations. *The Journal of Ocean Technology* 2020, 15(2), 78-100. https://www.thejot.net/article-preview/?show_article_preview=1163
- [St20] Steenbeek, J., Romagnoni, G., Bentley, JW., Serpetti, N., Heymans, JJ., Goncalves, M., Pereira Santos, C., Warmelink, H., Mayer, I., Keijsers, X., Fairgrieve, R., & Abspoel, L. (2020). Combining ecosystem modeling with serious gaming in support of transboundary maritime spatial planning. *Ecology and Society*, 25(2), 1-24. [21]. <https://doi.org/10.5751/ES-11580-250221>

Informationsfluss in der Offshore Windenergie - Ein Referenzprozess für Betriebsführung und Instandhaltung


Dirk Bendlin ¹, Moritz Werther Häckell², Marie-Antoinette Schwarzkopf³, Alicja Julia Kucewicz⁴, Jorge Marx Gómez⁵

Abstract: Diese Arbeit untersucht, als Vorarbeit zum Forschungsprojekt WiSA Big Data (FKZ: 03EE3016E) das Geschäftsumfeld von Betriebs- und Instandhaltungsprojekten in der Offshore Windenergie. Speziell soll dabei eine generische und möglichst allgemeingültige Prozesslandkarte, angelehnt an das Institut für Asset Management (IAM) erstellt werden. Im Weiteren soll ein Kernprozess für die Betriebsführung und Instandhaltung von Offshore Windparks in der Business Process Model and Notation (BPMN) abgebildet werden. Die Hauptbeteiligten der Betriebsführung und Instandhaltung, werden exemplarisch in tabellarischer Form dargestellt. Dies wird auf Basis einer Literaturrecherche entwickelt. Mit diesem generischen Modell kann das Anforderungsprofil des im Forschungsprojekt geplanten Software Demonstrators domänenspezifisch gestaltet werden.

Keywords: Betriebsführung und Instandhaltung, Offshore Windenergie, Referenzprozess, BPMN, IAM, Rollen und Kernaktivitäten

1 Motivation

Die Betriebsführung und Instandhaltung eines Offshore Windparks schlägt mit 25-40% Anteil an den Stromgestehungskosten zu Buche [PFR17] und umfasst den zeitlich größten Bereich des Projektlebenszyklusses. Dabei geht es darum, den Windpark über 20 bis 25 Jahre Betriebszeit, technisch kosteneffizient instand zu halten und zu betreiben. Aufgrund des langen Zeitraums dieser Projektphase ergibt sich die Möglichkeit, Daten zu sammeln

¹ Carl von Ossietzky University of Oldenburg, Business Informatics (Very Large Business Applications), Ammerländer Heerstr. 114-118, Oldenburg, 26129, dirk.bendlin@uni-oldenburg.de, 
<https://orcid.org/0000-0003-0278-2248>

² Ramboll Deutschland GmbH, Asset Management, Jürgen-Töpfer-Straße 48, Hamburg, 22763, moritz.haekell@ramboll.com

³ Ramboll Deutschland GmbH, Asset Management, Jürgen-Töpfer-Straße 48, Hamburg, 22763, marie.schwarzkopf@ramboll.com

⁴ Ramboll Deutschland GmbH, Asset Management, Jürgen-Töpfer-Straße 48, Hamburg, 22763, alicja.kucewicz@ramboll.com

⁵ Carl von Ossietzky University of Oldenburg, Business Informatics (Very Large Business Applications), Ammerländer Heerstr. 114-118, Oldenburg, 26129, jorge.marx.gomez@uni-oldenburg.de

und Prozesse, Abläufe, technische Komponenten und das Zusammenspiel mit Vertragspartnern durch Datenanalysen zu optimieren. Innerhalb der Betriebsführung und Instandhaltungsphase stehen verschiedene Sensordaten aus den Windturbinen, Fundamenten oder der Umspannstation (Offshore Substation oder OSS) zur Verfügung, gleichzeitig werden Wartungs- und Inspektionsprotokolle erstellt.

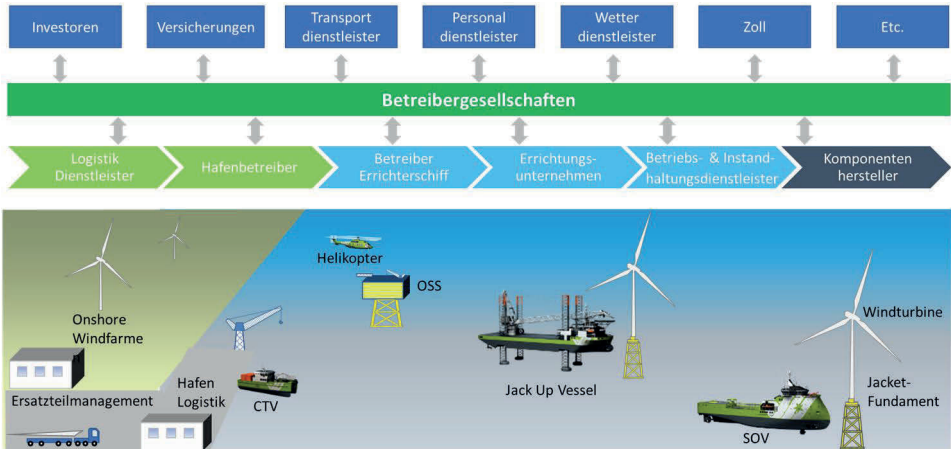


Abb. 1: Schnittstellen innerhalb der Betriebsführung und Instandhaltung von Offshore Windparks [WHHF16, GL 13] ©Offshore Logistik Bilder GRS Global Renewable Shipbrokers GmbH⁶

Eine Herausforderung liegt in der komplexen Infrastruktur und Logistik, die für die Betriebsführung und Instandhaltung der Offshore-Windparks notwendig ist [RØSH17]. Abb. 1 gibt einen Überblick über mögliche Projektbeteiligte [WHHF16]. Verstärkt werden diese Herausforderungen durch Restriktionen aufgrund von Umwelteinflüssen, wie signifikante Wellenhöhe, Windgeschwindigkeiten und den Abstand des Windparks zur Küste, die sich aufgrund von Schiffs- Betriebsgrenzen auf die Erreichbarkeit des Windparks auswirken [ELK14].

2 Methoden

Wilde und Hess [WH06], teilen die Wirtschaftsinformatik in zwei Hauptparadigmen zur Ergänzung des Wissensstandes in der Forschung auf: Design Science zur Erweiterung menschlicher und organisatorischer Grenzen durch (IT-) Artefakte und Behavioral Science als wissenschaftlicher Ansatz zur Erklärung menschlicher und organisatorischer Verhaltensweisen. Diese Arbeit ist der Design Science nach Hevner zugeordnet, da im Ergebnis des Forschungsprojektes WiSA Big Data ein Software Demonstrator zur Entscheidungsunterstützung der Instandhaltungs- und Betriebsführung von Offshore Windparks als Artefakt erstellt werden soll. Nach Hevner [HC10, HMPR04] unterteilen

⁶ OSS (Offshore Substation), CTV (Crew Transfer Vessel), SOV (Service Operating Vessel)

sich die Geschäftsanforderungen in Personen, Organisationen und Technologien. Die Identifizierung des Projektumfeldes ist kritisch für eine erfolgreiche Implementierung relevanter Artefakte, als Teilforschungsergebnisse zur Bereicherung des wissenschaftlichen Erkenntnisstandes [HC10]. Aus diesem Grund wurden Veröffentlichungen recherchiert, die sich mit diesen Aspekten des Geschäftsumfeldes der Betriebsführung- und Instandhaltungsprojekten innerhalb der Offshore Windenergie beschäftigen. Diese Arbeit legt einen Fokus darauf ein Prozessmodell für das Geschäftsumfeld abzuleiten. Speziell für Geschäftsprozesse hat sich die BPMN etabliert. Laut Venera [CVen12] sind BPMN sowie die Unified Modelling Language (UML) Aktivitätsdiagramme gleichermaßen geeignet, um Geschäftsprozesse in ihrem Umfang darzustellen. Ziel des Referenzprozesses ist es das Zusammenspiel der Hauptbeteiligten innerhalb der Wartung der Offshore Assets darzustellen. Die Bestrebung liegt hierbei darauf kein detailliertes Ablaufdiagramm für Softwareentwicklung, sondern ein tieferes Verständnis der generischen Arbeitsabläufe unter den Hauptprozessbeteiligten zu erarbeiten. Diese Arbeit nutzt BPMN analog zu Joschko [JWP13], mit einem vergleichbaren Anforderungsprofil.

Das Referenzprozessmodell soll ausreichend generisch gehalten werden, um den geplanten Softwaredemonstrator im Forschungsprojekt WiSA Big Data innerhalb des Geschäftsumfeldes verorten zu können. Auf der anderen Seite beinhaltet es noch ausreichend Komplexität, um repräsentativ für die betriebliche Realität zu sein.

3 Untersuchung des Geschäftsumfeldes der Betriebsführung und Instandhaltung von Offshore Windparks

Die Hauptbestandteile des Offshore Windparks, auf die im späteren Teil dieser Arbeit eingegangen wird, werden weitläufig als Assets bezeichnet [JAnd08, AWK07, IDin14]. Assets stellen sachliche Vermögenswerte dar. Die Verwaltung dieser Assets wird als Asset Management bezeichnet. Als Grundlage dieser Untersuchung wurde eine Prozesslandkarte des Instituts für Asset Management verwendet [IAM12]. Als Basis für eine Schärfung dieser Prozesslandkarte wurden vor allem folgende Vorarbeiten genutzt: Wegener beschreibt zusammen mit verschiedenen Marktteilnehmern detailliert die Schnittstellen innerhalb eines Offshore Windpark Projektes [WHHF16]. Ergänzend dazu wurden vom Garrad Hassan [GL 13] typischen Schnittstellen schematisch aufgezeigt. Diese Vorgaben der IAM in Verbindung mit den theoretischen Grundlagen, wurden in enger Abstimmung mit den Industriepartnern des Forschungsprojektes WiSA Big Data auf ein notwendiges Mindestmaß gekürzt. Zusammen mit den Arbeiten von Greiner [GAJ15] im Forschungsprojekt GOWOG lässt sich so Abb. 2 ableiten.

Die ursprüngliche Prozesslandkarte der IAM stellt oben die äußeren Einflüsse auf die Organisation dar. Das Modell umfasst dabei sechs Kernbereiche. Der erste Bereich „Strategie und Planung“ umfasst die organisationalen Ziele und ihre Erreichung im „Asset Management“. Vereinfacht wird er in unserem Modell als „Management/

Assetmanagement“-Sicht zusammengefasst. Der Bereich Asset Management Entscheidungsfindung (Decision Making) befasst sich mit der Einbindung und dem Einfluss von Entscheidungsprozessen auf den Asset- Lifecycle. So können sich beispielsweise Entscheidungen der Transport- und Installationsphase auf die Betriebsführung und Instandhaltungsphase auswirken. In dieser Phase konzentriert sich das Forschungsprojekt WiSA Big Data vor allem auf klassische Entscheidungsprozesse und aktuelle sowie zukünftig mögliche datengetriebene Entscheidungsprozesse. Der Bereich Lebenszyklus- Delivery stellt das dauerhafte und konstante Management der relevanten Risiken und Informationen im Asset Management sicher. Für WiSA Big Data steht hier vor allem das Datenmanagement und der Datenfluss als werthaltige Ressource im Mittelpunkt. Der nächste Punkt, die „Asset Informationen“, umfassen alle Rohdaten, die innerhalb der Betriebsführungs- und Instandhaltungsphase von Offshore Windparks anfallen. Von unten fließen diese Asset Informationen in den Prozess mit ein, beispielsweise in Form von Sensordaten. Der Bereich „Organisation und Menschen“ stellt die Organisation und relevante Unterstützungsfunktionen wie Warenwirtschaft dar. Die letzte Gruppe „Risiko und Überprüfung“ schließt schließlich den Kreis durch ein kontinuierliches Überwachen der Risiken und eine damit gekoppelte Anpassung der strategischen Ziele.

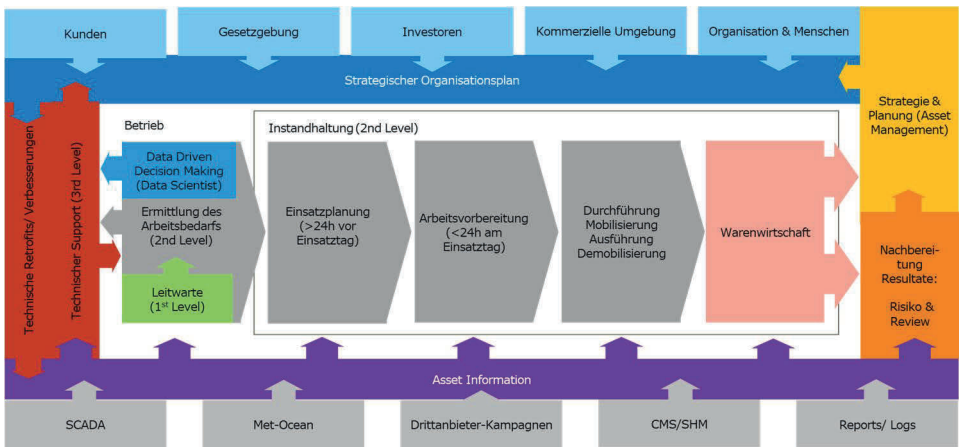


Abb. 2: Prozesslandkarte Betriebsführung und Instandhaltung eigene Darstellung Vergleiche [WHHF16, GAJ15, IAM12, GL 13]

Im Folgenden werden die Bestandteile eines Offshore Windparks anhand des Schemas in Abb. 2 nach Menschen, Organisation und Technik gegliedert für das Beispiel dargestellt.

3.1 Technologie - Strukturen und Komponenten eines Offshore Windparks

Asset Informationen beinhalten Daten verschiedener Gewerke, die in diesem Abschnitt näher vorgestellt werden sollen, eine Übersicht bietet Abb. 3.

Die Energieerzeugung der Offshore Windparks findet durch ihr Herzstück, die Windenergieanlagen (WEA) statt. Sie sind durch feste Gründungsstrukturen mit dem Baugrund verbunden. Die Gründungsstrukturen oder Fundamente, wie die weitverbreiteten Monopile- oder Jacketfundamente sorgen dafür, dass die Windenergieanlagen innerhalb ihrer 20- 25 Jahre Lebensdauer, den Umgebungsbedingungen (z.B. Extremwellen und Windgeschwindigkeiten, Baugrundveränderungen usw.), widerstehen können [MSei14, ÁBA21, Rc20]. Neben den Gründungsstrukturen gehören die Innerparkverkabelung, die die WEA mit der Offshore Umspannstation verbindet zum Balance of Plant (BoP).

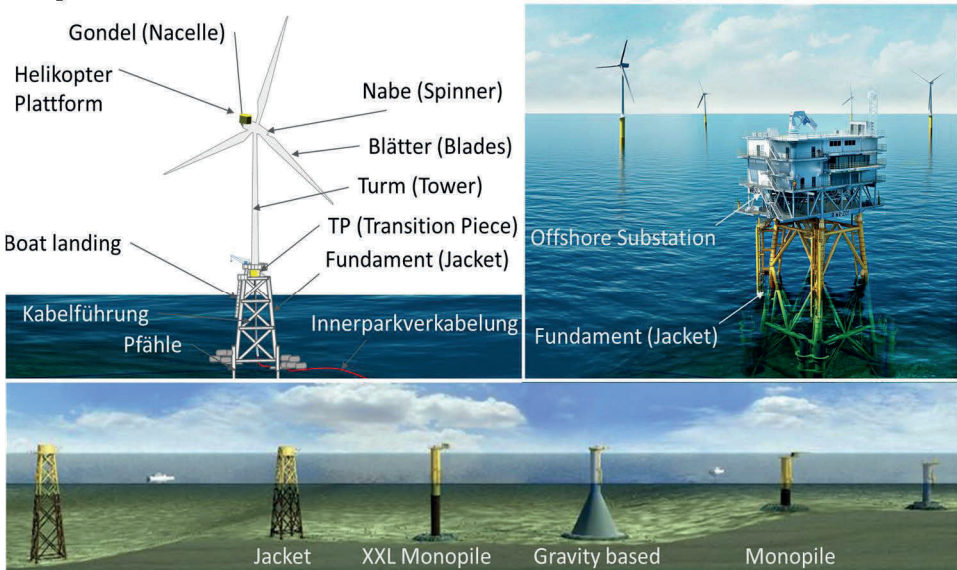


Abb. 3: Hauptkomponenten eines Offshore Windparks ©Offshore Logistik Bilder GRS Global Renewable Shipbrokers GmbH, © Fundament und OSS Bilder Ramboll

3.2 Organisation eines Offshore Windparks

Im Kontext von Offshore-Windparks können wesentliche Teile der täglichen Instandhaltung- und Betriebsführung in drei Haupt- Serviceinstanzen strukturiert werden [WHHF16]. Diese Instanzen werden hier vereinfacht als First-, Second- und Third-Level-Support bezeichnet, entsprechen den Strukturen, die in IT- oder Call-Centerstrukturen zu finden sind [AM17]. Mit steigendem Level nimmt die Anzahl der Aufgaben ab und die Komplexität zu. Service-Tickets können innerhalb dieser Struktur zur Lösung in den jeweils nächsten Service Level eskaliert werden [AM17].

Leitwarte und Einsatzzentrale (First Level Support): Um den Betrieb und die Instandhaltung von Offshore Windparks sicherzustellen, muss der Windpark

kontinuierlich überwacht werden. Leitwarten und Einsatzzentralen stellen die Zustandsüberwachung vom Festland oder bemannten Offshore Umspannstationen sicher. Die Leitwarte empfängt alle Störungen und Fehler aus dem Windpark, typischerweise kann die Mehrzahl oder 80% der Störungen durch Fernwartung bzw. -entstörung gelöst werden [WHHF16]. Wenn dies nicht möglich ist, eskaliert der First Level die Aufgabe an den Second Level Support. Im Gegenzug erhält der First-Level-Support vom Second-Level-Support alle notwendigen Informationen über geplante Arbeiten.

Service Leitung (Second Level Support): Der Second-Level-Support stellt die Serviceaufträge zusammen und kontrolliert deren fachgerechte Abwicklung [WHHF16]. Somit ist er die Instanz, die Service Tickets erstellt, On- und Offshore- Aktivitäten organisiert, die Verfahrensweisung und Risikobeurteilung für die Tätigkeiten definiert (Risk Assessment and Method Statement – RAMS) und die Arbeiten vorbereitet wie z. B. durch die Verwaltung und Bestellung von Ersatzteilen, Technikern und Transportmitteln (z. B. Onshore, per Schiff oder Hubschrauber). Diese Supportebene ist eng mit dem First und Third Level Support verbunden, um bei komplexeren technischen Problemen, erforderliche Maßnahmen zu definieren sowie sich mit unterstützenden Funktionen wie Ersatzteilmanagement, Einkauf oder Logistikdienstleistern abzustimmen. Dieses Level löst ca. 15% der Störungen [WHHF16].

Technischer Support (Third Level Support 3): Der Third-Level-Support setzt sich aus technischen Spezialisten zusammen. Die Spezialisten können nach Komponenten aufgeteilt sein oder nach technischen Kompetenzbereichen wie etwa Großkomponenten, Hydraulik, Nieder- und Hochspannungselektrik. Speziell bei langfristigeren Optimierungen und der Planung eines Großkomponententausches oder für technische Verbesserungen oder Nachrüstungen (Retrofits) sind die Ingenieure des technischen Supports gefragt. Gegebenenfalls sind auch externe Dienstleister beteiligt. Die dritte Unterstützungsebene ist auch für die Erstellung und Bewertung langfristiger Inspektions- und Servicestrategien zuständig. Die Aufgabe dieses Bereiches besteht in der technischen Unterstützung des Second-Level-Supports, der Erstellung von Root Cause Analysen (RCA) und der Aufdeckung von systematischen Fehlerquellen [WHHF16]. Es handelt sich hier um hochgradig spezialisierte Aufgaben, die im Schnitt ca. 5% der Störungen ausmachen [WHHF16].

Asset Management: Beim Asset Management geht es darum aus Assets Werte zu schaffen und die Ziele einer Organisation zu erreichen. Asset Management als Disziplin umfasst somit das Geschäfts- und Finanzmanagement und das Management der technischen Betriebsführung und Instandhaltung. Wissen und Werkzeuge im Asset Management ermöglichen Organisationen nachzuweisen, dass sie ihre Assets optimal verwalten. Diese Information hat einen großen Wert für viele Interessengruppen, seien es Eigentümer, Kunden, Aufsichtsbehörden, Vertragspartner oder die breite Öffentlichkeit [IAM12]. Das Asset Management interagiert mit allen drei Service Leveln hinsichtlich der Erreichung der Unternehmensziele, sowie um strategische Entscheidungen in den Betrieb zu integrieren und korrigierend bzw. optimierend zu steuern.

3.3 Menschen - Rollen und Hauptprozessbeteiligte

Für diese Arbeit bleibt der Fokus auf der Betreiberorganisation, die im Kern mit den drei Service Leveln beschrieben ist. Besnard und andere [BFT13, DM14, SM18, SM19, ELK14, DLD15] behandeln dabei vor allem den als Support Level 2 bezeichneten Bereich der Arbeitsvorbereitung und Offshore Entstörung. Neben Kosten und Umweltbedingungen spielt in diesen Modellen vor allem die Logistik eine große Rolle.

Weitere spezifischere Rollen werden von Mette [MVH17] definiert: Techniker, Qualitäts- und Instandhaltung, Sanitäter, Gesundheit und Sicherheit (Health and Safety), Management Offshore Einsätze, Management Offshore-Dienstleister und Arbeitsmediziner. In einer späteren Arbeit ergänzt Mette noch um Onshore Management [MVH+18]. Sahnoun [SBM15] ergänzt um Ingenieure, die in seinem Modell allerdings höher qualifizierte Techniker abbilden. Davon abgeleitet zeigt Tab. 1, eine kondensierte exemplarische Liste von Hauptakteuren sowie deren In- und Outputs, die in Rücksprache mit den Industriepartnern des Forschungsprojektes WiSA Big Data erstellt wurde.

Hauptrollen	Ziele	Inputs	Outputs
Management: Finance (e.g. Geschäftsleitung)	Kontrolle der Zielerreichung der Geschäftsstrategie	Gesetzliche Anforderungen	Geschäftsstrategie High Level KPI, Anforderungen e.g. an wiederkehrende Prüfungen.
Asset Management (e.g. Windfarm Asset Manager)	Zielerfüllung der Geschäftsstrategie	Festlegung relevanter Key Performance Indikatoren (KPI)	Generische, Asset spezifische KPI
Data Scientists (e.g. Data Analyst)	Verbesserung und Auswertung der Datenlage	Rohdaten, Reports	Reports, Dashboards und Entscheidungshilfen
Level 1 Support – Leitwarte (e.g. Techniker)	Fernentstörung der Assets, Kontrolle der Tätigkeiten im Park.	Fehlermeldungen der Assets	Fehlereskalation zum Level 2 Support
Level 2 Support (e.g. Service Site Manager)	Vor-Ort Entstörung der Assets	Vom 1st Level Eskalierte Fehlermeldungen	Ticket- und Reporterstellung 72h Grobplanung, 24h Feinplanung, 12h Schichtplan. Eskalation von Tickets zum Level 3 Support
Level 3 Support (e.g. Technischer Fachexperte)	Langfristige Zuverlässigkeit der Assets.	Technische Anfragen Level 2,	Technische Anforderungen an Level 1 und 2 Support., Optimierung von Einsatzgrenzen und Techn. Verbesserungsmaßnahme

Tab. 1 Hauptrollen innerhalb der Betriebsführung- und Instandhaltung von Offshore Windparks [WFFF16, GAJ15, JAnd08, IAM12]

Der Kernprozess wurde aus verschiedenen Quellen erarbeitet. Neben den Wartungsstrategien [AWKA06, Bun07, CMMM15], den Umweltbedingungen [DCM13, DMR13, DLTJ15] und der Logistik [RV10, The14, EC 12, DDL15, RØ15] waren die Arbeitsvorbereitung und Einsatzplanung [ELK14, BFT13, WHHF16, GAJ15] die wichtigsten Inputparameter. Das Zusammenspiel der verschiedenen Rollen wurde in einem beispielhaften Prozess in Abb. 4 zusammengefasst.

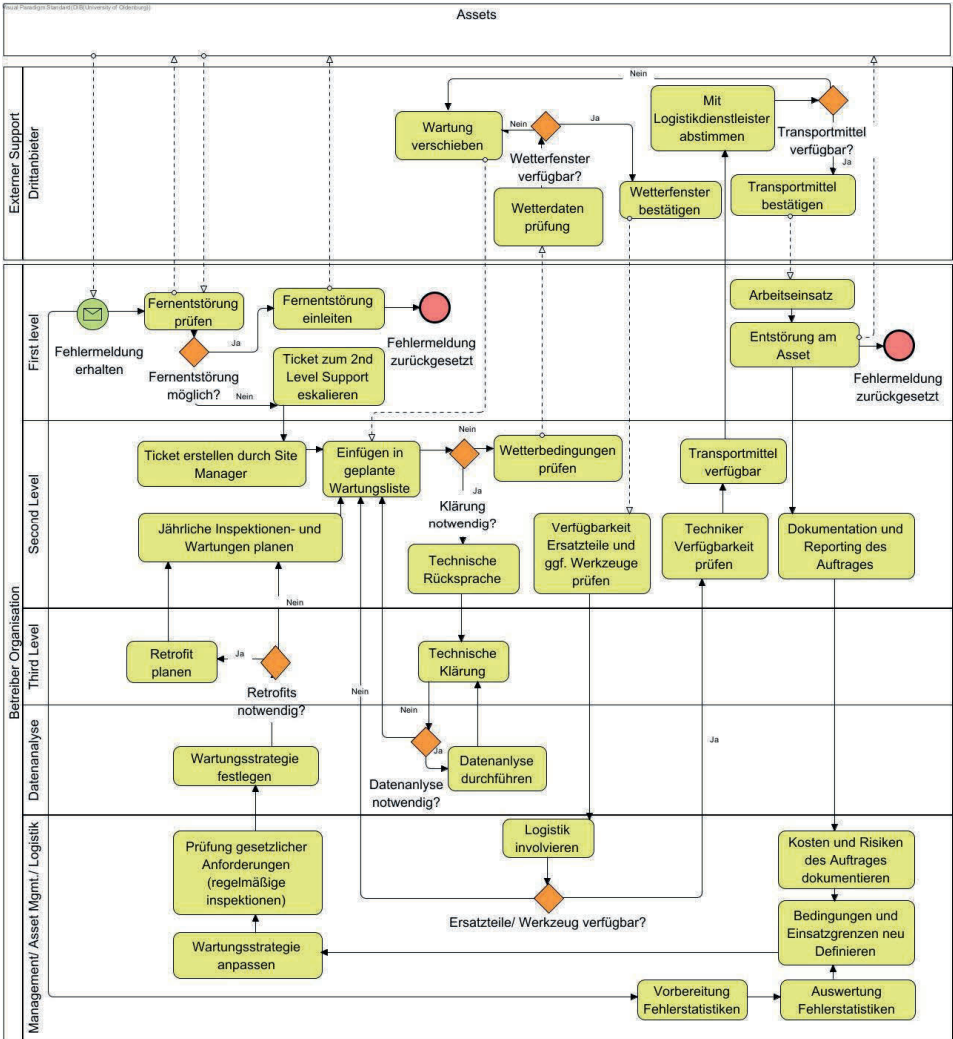


Abb. 3 Referenzprozess für die Betriebsführung und Instandhaltung von Offshore Windparks

Diese Arbeit stellt nur einen Auszug der möglichen Betreiberorganisationen dar.

Aufgrund des Fokus auf der Einfachheit und Allgemeingültigkeit dieser Ergebnisse wurde auf einen höheren Detaillierungsgrad verzichtet. Diese fehlende Tiefenschärfe macht es notwendig die Ergebnisse für individualisierte Untersuchungen anzupassen.

4 Schlussfolgerung

Durch diese Arbeit wird die wissenschaftliche Erkenntnisbasis um ein generisches Prozessmodell für das Geschäftsumfeld im Bereich Betriebsführung und Instandhaltung von Offshore Windpark Projekten ergänzt. Im Gegenzug zu anderen Forschungsarbeiten mit einem Fokus auf Risiken und Projektunsicherheiten kann die Beschreibung der verwendeten Technik, der Kernprozess und die Beschreibung der Hauptbeteiligten breitgefächert Verwendung finden. Die Ergebnisse bieten aufgrund ihrer generellen Gültigkeit eine gute Grundlage, um verschiedenste Forschungsfragen im Kontext des Geschäftsumfeldes zu bewerten. Wie hoch der Nutzen dieser Modelle für Entscheidungsfindung ist, konnte daher nicht bewertet werden. Um dieses ausgiebiger zu bewerten, müssen zukünftige Forschungen die erarbeiteten Ergebnisse nutzen, um domänenspezifische Anforderungsprofile zu erstellen. Im Forschungsprojekt WiSA Big Data werden diese Ergebnisse als Grundlage für das Anforderungsprofile eines Software Demonstrators genutzt. Wie von Hevner beschrieben, sollte neben dem Geschäftsumfeld auch der aktuelle Wissensstand der Forschung in Bezug auf Methodiken untersucht werden. Durch die Verbindung dieser beiden Paradigmen lassen sich die relevantesten Forschungsfragen am besten verorten. [HC10]. Diese Fragestellungen wurden innerhalb dieser Arbeit nicht adressiert.

Das in dieser Arbeit untersuchten Geschäftsumfeld ist die Vorarbeit zu einem domänenspezifischen Softwareanforderungsprofil. Dadurch kann die bestehende Wissensbasis um neue domänengerechte Modelle und Methoden erweitert werden.

5 Danksagung

Das Projekt WiSA Big Data wird vom Bundes Ministerium für Wirtschaftsförderung (BMWi) unter dem Förderkennzeichen (FKZ): 03EE3016E gefördert. Verschiedene Forschungs- und Industriepartner gaben uns in diesem Projekt Informationen zu Ihren Prozessen- und Abläufen, wir bedanken uns für Ihren Support. Bei Global Renewable Shipbrokers bedanken wir uns für die Abbildungen der Schiffslogistik und bei Ramboll für die 3D Abbildung von Fundamenten und der Umspannstation.

6 Literaturverzeichnis

[ÁBA21] Álamo, G. M.; Bordón, J. D.; Aznárez, J. J.: On the application of the

- beam model for linear dynamic analysis of pile and suction caisson foundations for offshore wind turbines. *Computers and Geotechnics* 134, S. 104107, 2021.
- [AM17] Abts, D.; Mülder, W.: *Grundkurs Wirtschaftsinformatik. Eine kompakte und praxisorientierte Einführung*. Springer Fachmedien Wiesbaden, Wiesbaden, 2017.
- [AWK07] Andrawus, J.; Watson, J.; Kishk, M.: Wind Turbine Maintenance Optimisation. *Principles of quantitative maintenance optimisation. Wind Engineering* 2/31, S. 101–110, 2007.
- [AWKA06] Andrawus, J. et al.: The Selection of a Suitable Maintenance Strategy for Wind Turbines. *Wind Engineering* 6/30, S. 471–486, 2006.
- [BFT13] Besnard, F.; Fischer, K.; Tjernberg, L. B.: A Model for the Optimization of the Maintenance Support Organization for Offshore Wind Farms. *IEEE Transactions on Sustainable Energy* 2/4, S. 443–450, 2013.
- [Bun07] Principles for condition-based maintenance of wind turbines. Adopted by the Expert Consulting Committee of the German Wind Energy Association (BWE e.V.) on 21st September 2007, 2007.
- [CMMM15] Carroll, J. et al.: Availability Improvements from Condition Monitoring Systems and Performance Based Maintenance Contracts. In (The European Wind Energy Association (EWEA) Hrsg.): *EWEA Offshore 2015. Conference*, 2015.
- [CVen12] Venera Geambasu, C.: BPMN vs. UML Activity Diagram for business process modelling. *Journal of Accounting and Management Information Systems* 11, S. 637–651, 2012.
- [DCM13] Dinwoodie, I.; Catterson, V. M.; McMillan, D.: Wave height forecasting to improve off-shore access and maintenance scheduling. In (Institute of Electrical and Electronics Engineers (IEEE) Hrsg.): *IEEE Power and Energy Society general meeting (PES)*, 2013. 21 - 25 July 2013, Vancouver, BC, Canada. IEEE, Piscataway, NJ, S. 1–5, 2013.
- [DDL15] Dalgic, Y. et al.: Optimum CTV fleet selection for offshore wind farm O&M activities. In (Nowakowski, T. et al. Hrsg.): *Safety and reliability. Methodology and applications, proceedings of the European Safety and Reliability Conference, ESREL 2014, Wrocław, Poland, 14-18 September 2014*. CRC Press/Balkema, Leiden, The Netherlands, S. 1177–1185, 2015.
- [DLD15] Dalgic, Y. et al. Hrsg.: *The influence of multiple working shifts for offshore wind farm o&m activities - Strathow-OM Tool*, 2015.
- [DLTJ15] Dalgic, Y. et al.: Investigation of optimum jack-up vessel chartering strategy for offshore wind farm O&M activities. *Ocean Engineering* 95,

- S. 106–115, 2015.
- [DM14] Dinwoodie, I.; McMillan, D.: Operational strategies for offshore wind turbines to mitigate failure rate uncertainty on operational costs and revenue. *IET Renewable Power Generation* 4/8, S. 359–366, 2014.
- [DMR13] Dinwoodie, I. et al.: Development of a Combined Operational and Strategic Decision Support Model for Offshore Wind. *Energy Procedia* 35, S. 157–166, 2013.
- [EC 12] EC Harris: Offshore Wind Cost Reduction Pathways. Supply Chain Work Stream, 2012.
- [ELK14] Endrerud, O.-E.V.; Liyanage, J.P.; Keseric, N. Hrsg.: Marine logistics decision support for operation and maintenance of offshore wind parks with a multi method simulation model, 2014.
- [GAJ15] Greiner, S. et al.: GOWOG. German Offshore Wind Operation Guide. Leitfaden für die technische Betriebsführung, 2015.
- [GL 13] GL Garrad Hassan: A Guide to UK Offshore Wind Operations and Maintenance, 2013.
- [HC10] Hevner, A.; Chatterjee, S.: Design Research in Information Systems. Springer US, Boston, MA, 2010.
- [HMPR04] Hevner, A. R. et al.: Design Science in Information Systems Research. *MIS Quarterly* Vol. 28 No. 1, S. 75–105, 2004.
- [IAM12] Asset Management – an anatomy, www.theIAM.org/AMA, 2012.
- [IDin14] Dinwoodie, I.: Modelling the operation and maintenance of offshore wind farms. Doctoral thesis, 2014.
- [JAnd08] Andrawus, J.: Maintenance optimisation for wind turbines. Doctoral thesis, 2008.
- [JWP13] Joschko, P. et al.: Modelling and Simulation of Offshore Windfarm - Mapping and Analysis of relevant O&M Process. In (Page, B. et al. Hrsg.): *EnviroInfo 2013 - Environmental Informatics and Renewable Energies. 27th International Conference on Informatics for Environmental Protection; proceedings of the 27th Conference on Environmental Informatics - Informatics for Environmental Protection, Sustainable Development and Risk Management; September 2-4, 2013, University of Hamburg, Germany. Shaker, Aachen, S. 1–12, 2013.*
- [MSei14] Seidel, M.: Substructures for offshore wind turbines. Current trends and developments. In (Institut für Stahlbau, H. et al. Hrsg.): *Festschrift Peter Schaumann. Gewidmet Univ.-Prof. Dr.-Ing. Peter Schaumann zur Vollendung seines 60. Lebensjahres. Technische Informationsbibliothek u.*

Universitätsbibliothek, 2014.

- [MVH⁺18] Mette, J. et al.: Healthy offshore workforce? A qualitative study on offshore wind employees' occupational strain, health, and coping. *BMC public health* 1/18, S. 172, 2018.
- [MVH17] Mette, J. et al.: "It's still a great adventure" - exploring offshore employees' working conditions in a qualitative study. *Journal of occupational medicine and toxicology (London, England)* 12, S. 35, 2017.
- [PFR17] Pfaffel, S.; Faulstich, S.; Rohrig, K. Prof. Dr.: Performance and Reliability of Wind Turbines. A Review. *Energies* 11/10, S. 1904, 2017.
- [Rc20] Raktate, T.; choudhary, R.: Design of Monopile Foundation for Offshore Wind Turbine. *E3S Web of Conferences* 170, S. 1024, 2020.
- [RØ15] Raknes, N. T.; Ødeskaug, K.: Optimal Scheduling of Maintenance Tasks and Routing of a Joint Vessel Fleet for Multiple Offshore Wind Farms. Master Thesis, Trondheim, 2015.
- [RØSH17] Raknes, N. et al.: Scheduling of Maintenance Tasks and Routing of a Joint Vessel Fleet for Multiple Offshore Wind Farms. *Journal of Marine Science and Engineering* 1/5, S. 11, 2017.
- [RV10] Rodriguez, M. A.; Vecchietti, A.: Inventory and delivery optimization under seasonal demand in the supply chain. *Computers & Chemical Engineering* 10/34, S. 1705–1718, 2010.
- [SBM15] Sahnoun, M. et al.: Modelling and simulation of operation and maintenance strategy for offshore wind farms based on multi-agent system. *Journal of Intelligent Manufacturing*, S. 1–17, 2015.
- [SM18] Seyr, H.; Muskulus, M.: Using a Langevin model for the simulation of environmental conditions in an offshore wind farm. *Journal of Physics: Conference Series* 1104, S. 12023, 2018.
- [SM19] Seyr, H.; Muskulus, M.: Decision Support Models for Operations and Maintenance for Offshore Wind Farms: A Review. *Applied Sciences* 2/9, S. 278-302, 2019.
- [The14] Sharing lessons learned and good practice in offshore transmission. Summary report, 2014.
- [WH06] Wilde, T.; Hess, T.: Methodenspektrum der Wirtschaftsinformatik.doc. Arbeitsbericht 02/2006, 2006.
- [WHHF16] Wegener, J. et al. Hrsg.: Praxishandbuch Schnittstellenmanagement Offshore Wind. EEHH, Hamburg, Hamburg, 2016.

Enable students to acquire new hands-on software experience using the modular teaching assistance system

Viktor Dmitriyev,¹ Christian Janßen² and Andreas Solsbach³

Abstract: As long as learning management systems (LMS) are offering a predefined set of functionalities and could be not flexible enough for changes and improvements. Especially, it could be difficult to enable usage of a new software tool, which is required only by a special teaching curriculum and when change requests are relevant only for a particular course of study. Thus, incorporating innovative technologies into existing learning platforms could be a very challenging task. However, this problem can be addressed through the latest IT technologies in combination with innovative teaching methodologies. In this work, students are supported in acquiring better hands-on experience using real software in real scenarios. It is done in a way that such scenarios and tools are directly embedded into a teaching curriculum. The introduced modular teaching assistance system helps to simplify installation, configure, and data wrangling phases and enables students to focus on primary tasks, which is solving scenario problems, rather than spend valuable time and deal with software as such. On one hand, the system takes care of various problems with installation and configuration for students, on the other hand, the system is modular and enables adaptation of an innovative teaching experience by being flexible and extensible by design.

Keywords: Teaching; CEMIS; Teaching Platform; Software

1 Introduction

Digitization and Industry 4.0 are both buzzwords that have been used for almost a whole decade [KW21]. In addition, the digital transformation processes continue to force companies in looking for new ways to reconfigure and redesign their business strategies [Sc17]. In this context, companies often ask themselves how they have to implement the technological change in order to remain successful with their products and services without jeopardizing their economic success. With a view to further future topics, the implementation of sustainability aspects in the business strategy are essential. Reducing CO₂ consumption not only takes place on a product level (e.g. *What materials should be use for production or how it should be manufactured, packed and shipped?*). With regard to the business strategy and operational processes, sustainability is a company wide reconfiguration task.

¹ University of Oldenburg, Department of Computing Science, Business Information Systems / VLBA, Oldenburg, Germany, viktor.dmitriyev@uni-oldenburg.de, <https://orcid.org/0000-0001-5661-4587>

² University of Oldenburg, Department of Computing Science, Business Information Systems / VLBA, Oldenburg, Germany, christian.janssen@uni-oldenburg.de, <https://orcid.org/0000-0003-0338-2852>

³ University of Oldenburg, Department of Computing Science, Business Information Systems / VLBA, Oldenburg, Germany, andreas.solsbach@uni-oldenburg.de, <https://orcid.org/0000-0002-8799-0478>

At this point research and education comes into foreground and have an important impact for the mediation of a sustainable mindset for companies in times of the ongoing digital transformation. At the University of Oldenburg, different sustainability related courses are offered and addressed to students with different academic backgrounds (e.g. Business Informatics or Sustainability Economics and Management). *Teaching sustainability in universities is challenged by the rapidly changing requirements and standards in the field of sustainability [...]* [Ra17]. The Covid-19 pandemic showed, that digitization in companies and at the educational institutions is not yet completed and far away from being implemented in a way, which will fulfill requirements. A number of steps must be performed to adapt teaching processes for new circumstances, which includes a teaching of practical skills as well as teaching an ability to combine theory and practice to work with real software tools and real scenarios [Ga20]. The requirements to switch to working at home with the necessary processes and equipment (e.g. workstation, video conference systems, working place, etc.) is a challenge for all involved participants. The changeover from presence teaching and face-to-face meetings to digital online teaching requires more advanced and innovative concepts, methodologies and tools in order to support self-study of the students and facilitate the teaching of practical knowledge [KB21]. For this reason, the entry barriers to use any required software tool must be kept as low as possible for the students.

In this paper, we look at the challenge for students who must be enabled to obtain not only theoretical knowledge but also practical knowledge. As an example, sustainability course materials will be used. For this reason, we developed a software tool to support students in the context of self-learning. Within the course *Sustainability Informatics* at the University of Oldenburg research oriented teaching methods will be used to empower students hands-on experience research intentions. One of the teaching core tasks inside the course is to set up a sustainability report. This will be also the use case of this research contribution. For that reason, sustainability data needs to be analyzed and Key Performance Indicators (KPIs) have to be set up. Usually, multiple software tools and systems will be used for this purpose and must be installed. This could be a time-consuming and frustrating experience for students in case something does not work (e.g. the installation process fails). Also, students could have the different technical knowledge and need to be supported in different ways. Our provided solution takes up this problem.

The modular teaching assistance system using for teaching hands-on experience offers a centralized infrastructure.

The rest of the paper is structured as follows: Section 2 describes existing online teaching platforms and the research-oriented teaching approach in the course *Sustainability Informatics*. Section 3 characterize challenges in teaching hands-on experience. The following section 4 shows our provided solution and the architecture of the software tool. Section 5 discusses the use of the tool based on a constructed use case. Finally, section 6 concludes the paper.

2 Online Teaching Platforms and Research-oriented Teaching Methodologies

2.1 Online Teaching Platforms

There are a whole set of software systems, which are used in teaching and for teaching. For instance, such systems could offer students virtualized laboratories to run experiments [DDI17] or support generic online-teaching processes. E-Learning platforms such as Moodle^{4a}, Stud.IP^{4b}, Canvas^{4c}, Open EdX^{4d}, Oppia^{4e}, ELMS Learning Network^{4f} and many others are primarily used to deliver content to the student and establish teacher-to-student communication. Such platforms are also known as a learning management system. This kind of platforms are sophisticated and provide a huge set of features and functionalities, which are designed to address a teaching process as a whole. Following Elabnody theories, several features of LMS support students (or further users) in their learning paths: tracking and reporting, assessment and testing tools, communication and collaboration or mobile learning [E116]. Despite being very sophisticated such platforms are centrally managed and offer multiple predefined set of features for online teaching experience.

As long as these platforms offer a predefined set of functionalities, they are not flexible enough for changes and improvements. Especially, it could be difficult, when change requests are only relevant for a particular teaching curriculum. Thus, it could be a challenging task to incorporate the teaching of innovative technologies into existing learning platforms. This problem can be addressed using technology or teaching methodologies, see the section 2.2. For instance, for teaching in-memory database and their applications in real-world driven applications various approaches such as bachelor or master theses, seminars, projects groups were performed [Dm16]. There are also teaching curricula, which require the usage of specialized software tools as such. For instance, using virtual laboratories to simulate experiments in natural sciences could be a helpful resource to facilitate knowledge acquisition by students [DDI17, A114]. In the case of LMS, despite existing functionalities that could be extended via diverse plugins and modules, even this couldn't be a proper solution to the problem, as long as such changes are applied to the whole teaching system at once and influence all participants, which is not always desirable and sometimes even not possible.

The introduced in the paper modular teaching assistance systems help students to acquire hands-on experience and flattens the learning curve by taking care of installation, configuration, and data wrangling phases. The system offers to each student a personalized workspace⁵ with an extensible number of software modules to be used (e.g. database, tools to create dashboards, etc.). The solution also enables to adapt teaching curriculum for a

⁴ (a) <https://moodle.org/> (b) <https://studip.de/> (c) <https://github.com/instructure/canvas-lms>
(d) <https://open.edx.org/> (e) <https://github.com/oppia/oppia> (f) <https://github.com/elmsln/elmsln>

⁵ This particular workspace consists of multiple modules and could be seen as a personalized sandbox for individual experiments

new software tool, because the modularity of the teaching assistance systems is part of its architecture.

2.2 Research-oriented Teaching Methodologies

The purpose of higher education at universities is to support and sensitize the students thinking and labor in academic behaviors. Educational institutions such as universities have the responsibilities to guarantee the highest standards of scholarships and in particular, the conceptual development of innovative, modern research methods. This may help students to bring their research methods and skills such as scientific research, writing, spelling and teamwork as a process in mind by using new approaches [Ra17]. This reflects also Humboldt's idea of a modern and future oriented university. At the University of Oldenburg different research oriented teaching programs such as FLIF+ ⁶ tries to increase the standards of student self-oriented education by connecting research and learning concepts. Following this point of view, research oriented teaching and learning can be seen as an enabler to support students and also teachers in gaining scientific knowledge and also teaching hands-on experience. In times of a global pandemic, flexible learning methodologies as described in [ASAEKB21] and the strengthening of students self-responsibility could be a key factor.

Within the course *Sustainability Informatics* a research oriented and flexible teaching approach as described in [WHP20] will be used. On the basis of the overarching task of preparing a sustainability report, the students formulate independently own research questions, conduct and evaluate the findings and finally present the findings in a manner they prefer. Our developed software tool supports the students during their work and causes that the students are able to concentrate on the essentials - the analysing, interpretation and presentation of findings.

3 Challenges in Teaching Hands-On Experience

As it was mentioned earlier, to support students in getting their own hands-on experience the real tools and real scenarios (or simplified versions of such) should be integrated into a teaching curriculum. The primary student learning objective in such cases is not only to get familiar with a particular software tool, but also apply acquired theoretically knowledge. In most of the tutorials for getting such type of hands-on experience the very first step would be to install and configure the required software tool and get yourself familiar with it. Of course, an ideal case would be to use pre-installed software, which utilizes hardware located in a laboratory room. However, this particular option is not always feasible and may even not be applicable, for instance in the case of an online course. Thus, the very first step

⁶ FLIF+ in German stays for *Forschungsbasiertes Lernen im Fokus Plus*. URL - <https://uol.de/forschen-at-studium>

should be simplified as much as possible, to help students to overcome initial problems on a learning curve much easier. In addition, students can concentrate on their essential steps of the research oriented learning procedures. A process of a software installation could be simplified and even sometimes automated (e.g. using software tools such as Vagrant⁷, Docker⁸ or any other available virtualization approach). These virtualization tools are also useful for teaching a number of subject in computer engineering curricula (e.g. computer networking, operating systems, etc.) [Ci21, Hu19]. However, as long this particular step on the way to get a hands-on-experiments represents a clear “blocker” it is crucial to accomplish this step as fast as possible without much straggle on the side of a student. The problem with any installation gets even worse, when a particular teaching curriculum requires multiple software tools and systems to be used. As long as the installation should be conducted by the student, it could lead to a frustrating experience. Thus, instead of spend his/her time on a task, students could struggle with the installation process of a new software tool. There are a number of ways how to overcome or at least minimize aforementioned problems – (a) provide very detailed installation and configuration instruction; (b) automate installation and configuration; (c) use cloud-based or comparable 3rd party services; (d) user own dedicated infrastructure. Fig. 1 provides a relative comparison between given 4 variants with respect to user friendliness and control over provided tools for teaching hands-on-experience.

When variant (a) is applied, it should be mentioned that maintaining all these instructions, manuals, and setup tutorials from year to year could be a tedious task. That leads to additional workloads for teaching staff. This overhead happens due to software changes and updates. Students usually do not have a homogeneous hardware and software setup of their hardware, which is supposed to be used in this particular variant. It also must be noticed, that adding new software tools to the existing curriculum may lead to unforeseeable complications on the side of the student’s hardware.

Variant (b) could help to overcome issues of the variant (a), but still got own pre-requisites. There are two main pre-requisites – hardware setup that fits and basic knowledge of bash/command-line scripting. The last is crucial, because it is usually not possible to test and automate all possible variants, and instructions for script adjustments are provided.

Variant (c) could solve all issues of the (a) and (b), but major drawbacks of this particular variant are the following – request official contractors to offer 3rd party solutions and in case of community or academy licensing, it could be that some parts of the tools and systems used within the study will be deprecated/removed by the vendor without any notifications. Basically, with variant (c) one has less control over the infrastructure in comparison to the (a) and (b). The variant (d) could all aforementioned issues but requires that “technical complexity” of the installation and configuration will be hidden from the students and they will be offered only access to required tools upon request. Besides this variant solves issues of the variants (a), (b), and (c), it also allows instructors to maintain technical complexity for the students. For instance, in the case when a database instance should be used with

⁷ <https://www.vagrantup.com/>

⁸ <https://www.docker.com/>

User acceptance	High	cloud-hosted services for teaching hands-on-experience (c)	self-hosted services for teaching hands-on-experience (d)
	Low	installation using step-by-step tutorials and manuals (a)	automated installation on users' hardware using scripts (b)
		Low	High
		Control over provided software	

Fig. 1: Provides a comparison between 4 given variants ((a), (b), (c) and (d)) with respect to the user acceptance and control over provided software for teaching a hands-on-experience

particular data inside it, this data could be pre-imported, which will allow students to directly start working with data and overcome typical strangles with a data import process.

This contribution implement the variant (d) for the course *Sustainability Informatics*. It demonstrates how it is implemented and how the solution can be extended with new tools. The work also demonstrates, how such solution helps to maintain complexity of the hands-on-expertise exercises of a particular curriculum.

4 Architecture and Implementation

The *mta-system* (**modular teaching assistance system**) consists of 4 main components – mta backend, reverse proxy, local Docker Images registry and personalized workspaces. The mta backend is responsible for communication with user, creation of services in form of predefined Docker containers. The crucial part of the mta backend is called *integration utilities*. They are responsible for creation of individualized services and configuration of the reverse proxy. The reserve proxy is responsible for managing incoming HTTP/HTTPS requests for reach offered services and proxy these request towards requested service. The Docker Images Local Registry is basically a collection of pre-created Docker images, which

are used as base to create user specific Docker containers upon request. The workspace of a user consists of a number of encapsulated services individually run for each user. Usually, one Docker container is created for each user and services. However, it also could be that one Docker container is shared by multiple users. In order to save resources on the server, running containers will be stopped at some point and users will have to “start” them manually in order to resume their work on the task. This process is also handled by the backend part of the mta-system. The architecture of the system showed on Fig. 2. The latest version of the source code of the mta-system can be found in the public git-repository⁹. The mta-system is modular by design. It means that a new service in form of a Docker container can be easily added and hosted by the system.

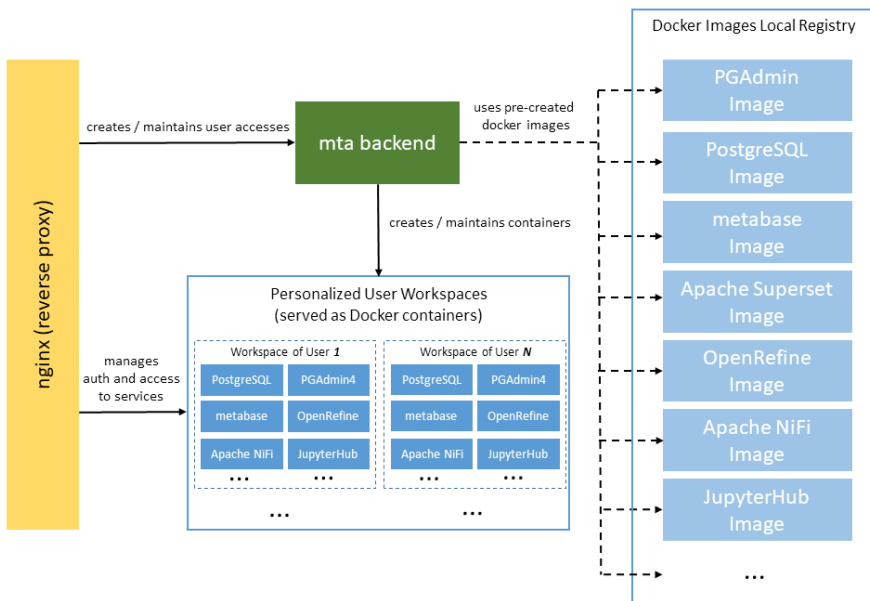


Fig. 2: Demonstrate the architecture of the hands-on experience teaching assistance system for students. It uses *nginx* as a reverse proxy to manage HTTPS inquiries to various services (e.g. *metabase*). The mta-system uses special Docker images to create user-specific services based on them, it also manages user access and running containers⁹

However, the drawback of such an approach is that only online-based tools and services could be used to get hands-on experience. This leads to the point that in some cases tools with a limited number of features could be used, which could be acceptable, as long as such tools are normally dedicated to demonstrate and teach core principles, rather than provide a detailed introduction of a particular technology.

The system works as follows: students register themselves in the modular teaching assistance

⁹ The mta-system implementation can be found here - <https://github.com/vdmtriye/mta-system>

system. After registration was successfully completed, access to various services should be provided. These services are offered and accessed via the mta-system. The typically user view on the system is shown on Fig. 3.



Fig. 3: Demonstrate how a particular service (e.g. OpenRefine) could be accessed in the user interface. Typically, the user interface of the system provides credentials, access links and further required instructions

However, not all services are accessible and visible to the students at once. It is not always necessary, because not all services are used at the same time. Thus, services could be accessed on demand. Demand for a particular service could be coupled with a teaching curriculum. For instance, it could be that a database with a dataset should be immediately accessed by students to start working on a use case, but a visualization tool could be developed later on when the first parts of the use case would be completed.

5 Use Case and Discussion

One of the given tasks in the course *Sustainability Informatics* is to analyze material flows. The mta-system used the freely availed implementation of the web-based Sankey diagram called “SankeyMATIC (BETA)”. Fig 4 shows an example of a diagram with a flow of a household budge. The “Sankey diagram” solution is deploy with a Docker container and encapsulated into a service on the mta-system to be delivered upon request to each student individually.

Besides using web-based Sankey diagram, other use cases may require an import of a dataset into a database for further analysis using SQL. For this purpose, the PGAdmin4¹⁰ tool was used by students to import required data. Beside that, the mta-system provided a specialized data import feature out of the box, to reduce the import overheads. Such feature could be useful, if an import of a data should not be a part of a teaching assignment. The mta-system offered a DBMS (e.g. PostgreSQL) to store required datasets to each student as well. As long as databases do not usually offer visualization capabilities, further visualization software (e.g. Sankey Diagram, business intelligence visualization tools like Apache Superset) were offed to students via the mta-system. This allows students to design and visualize desired

¹⁰ pgAdmin 4 is a management tool for PostgreSQL database- <https://www.pgadmin.org/>

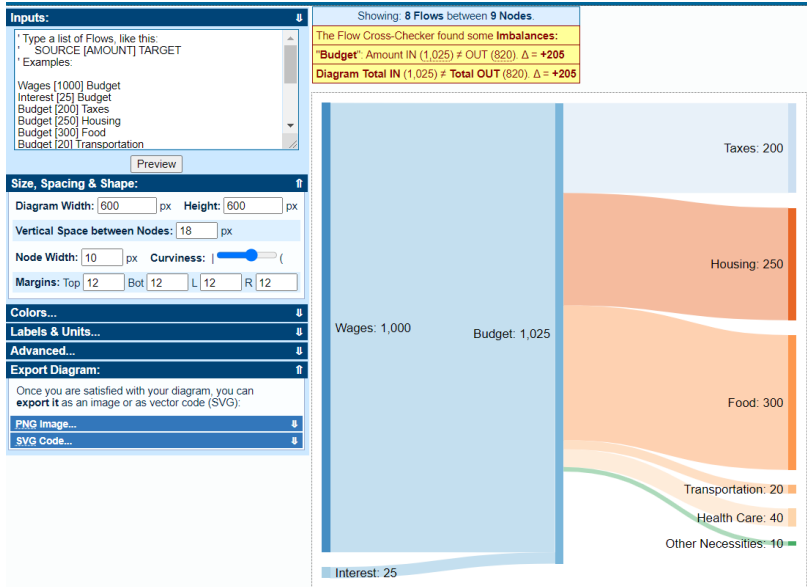


Fig. 4: Shows the web-based implementation of the “sankey diagram”. The diagram shows an example of a household budget flow. The “sankey diagram” solution is deployed using the special Docker image and served via the mta-system as a service

reports and dashboards, using any given data. For advanced classes, which may require programming experience, hosted within JupyterHub an interpreter of a Python programming language and a predefined dataset (e.g. machine-readable sustainability report) was offered. It let students to run simple analysis and create simple prediction model (e.g. estimate predefined key performance indicators (KPIs) for a particular sustainability report and dashboard).

It also has to be mentioned, that the mta-system offers an interface to maintain individualized workspaces. It allows students to start/stop offered within the system services, which leads to better utilization of the server resources (e.g. services that are not used by a student will be stopped after a while, but could be always be resumed).

In general, students that participated in offered to them assignments and used the mta-system for this purpose performed slightly better during the final examination in comparison to other students, who didn't use this option. It was measured by comparing the final scores of the students who used the mta-system with other of the students. However, it has to be mentioned that it happened due to the deeper understanding of the theoretical foundation and deeper involvement in teaching process. Which was possible due to practical assignments. Usage of the system was optional because instructions to care out installation locally (on own PCs) for a required software tool were provided as well. Nevertheless, the usage of

the mta-system allowed to simplify an overall hands-on software experience process for students and helped teaching faculty members to gain more control over this process.

6 Conclusion and Outlook

As it was mentioned in the work most of the learning management systems offer a predefined set of functionalities. Extending functionality of such systems only for a particular teaching curriculum could sometimes be not feasible. For instance, because required changes could only be relevant for a particular study course. In this work, the mta-system was introduced, which supports students in acquiring better hands-on experience using real software in real scenarios. The mta-system was directly integrated into a teaching curriculum, as an additional software tool and not as a replacement for LMS. The introduced modular teaching assistance system helps to simplify installation, configure, and data wrangling for students. It also simplifies software tools support process, which is usually carried out by teaching faculty member. On one hand, the system overtakes problems with installation and configuration, on the other hand, the system was designed to be modular and to enable adaptation of an innovative teaching experience. The mta-system is modular by design, thus a new tool required by a teaching curriculum could be potentially integrated and used. For example, another tool that can be potentially integrated is `stplanr` [LE18]. `stplanr` is a package for sustainable transport planning with R and its geo-visualization could be potentially enabled for further development from students on the mta-system. The system will be further developed and used in other teaching curricula. Also, further software tools will be adapted to be used in container-driven individualized workspaces.

Bibliography

- [Al14] Ali, Numan; Ullah, Sehat; Rabbi, Ihsan; Alam, Aftab: The Effect of Multimodal Virtual Chemistry Laboratory on Students' Learning Improvement. In (De Paolis, Lucio Tommaso; Mongelli, Antonio, eds): *Augmented and Virtual Reality*. Springer International Publishing, Cham, pp. 65–76, 2014.
- [ASAEKB21] Ait Si Ahmad, Hana; El Kharki, Khadija; Berrada, Khalid: Agility of the Post COVID-19 Strategic Plan on Distance Learning at Cadi Ayyad University. An Opportunity Towards a Total Digital Transformation of the University. In (Agrati, Laura Sara; Burgos, Daniel; Ducange, Pietro; Limone, Pierpaolo; Perla, Loredana; Picerno, Pietro; Raviolo, Paolo; Stracke, Christian M., eds): *Bridges and Mediation in Higher Distance Education*. Springer International Publishing, Cham, pp. 199–213, 2021.
- [Ci21] Cisternino, Antonio; Ducange, Pietro; Tonello, Nicola; Vallati, Carlo: Leveraging Cloud Infrastructures for Teaching Advanced Computer Engineering Classes. In (Agrati, L.; Burgos, D.; Ducange, P.; Limone, P.; Perla, L.; Picerno, L.; Raviolo, P.; Stracke, Ch., eds): *Bridges and Mediation in Higher Distance Education*. Springer International Publishing, Cham, pp. 256–270, 2021.

- [DDI17] Daineko, Yevgeniya; Dmitriyev, Viktor; Ipalakova, Madina: Using virtual laboratories in teaching natural sciences: An example of physics courses in university. *Computer Applications in Engineering Education*, 25(1):39–47, 2017.
- [Dm16] Dmitriyev, Viktor; Wagner vom Berg, Benjamin; Stamer, Daniel; Sandau, Alexander; Giesen, Nils; Siewert, Jens; Marx Gómez, Jorge: Teaching In-Memory Computing and Big Data Related Skills at University. In: *Multikonferenz Wirtschaftsinformatik (MKWI) 2016 (Brand II)*. volume 2, pp. 729 – 740, 2016. SBN 978-3-86360-132-4 (Druckausgabe, Gesamtwerk).
- [El16] Elabnody, M.: A Survey Of Top 10 Open Source Learning Management Systems. *International Journal of Scientific & Technology Research*, 5:7–11, 2016.
- [Ga20] Gamage, Kelum AA; Wijesuriya, Dilani I; Ekanayake, Sakunthala Y; Rennie, Allan EW; Lambert, Chris G; Gunawardhana, Nanda: Online Delivery of Teaching and Laboratory Practices: Continuity of University Programmes during COVID-19 Pandemic. *Education Sciences*, 10(10):291, 2020.
- [Hu19] Huang, Albert: Teaching, learning, and assessment with virtualization technology. *Journal of Educational Technology Systems*, 47(4):523–538, 2019.
- [KB21] Kóvári, Edit; Bak, Gerda: University Students’ Online Social Presence and Digital Competencies in the COVID-19 Virus Situation. In (Agrati, Laura Sara; Burgos, Daniel; Ducange, Pietro; Limone, Pierpaolo; Perla, Loredana; Picerno, Pietro; Raviolo, Paolo; Stracke, Christian M., eds): *Bridges and Mediation in Higher Distance Education*. Springer International Publishing, Cham, pp. 158–169, 2021.
- [KW21] Kagermann, Henning; Wahlster, Wolfgang: Digitalisierung der Produktion: Zehn Jahre Industrie 4.0. In: *Frankfurter Allgemeine Zeitung*. 29-03-2021.
- [LE18] Lovelace, Robin; Ellison, Richard: stplanr: A Package for Transport Planning. *The R Journal*, 10(2):7–23, 2018.
- [Ra17] Rapp, Barbara; Solsbach, Andreas; Gizli, Volkan; Janßen, Christian: Teaching ICT-based Sustainability Reporting by Using a Research-oriented Teaching Concept. In (Daniel, Nukpezah; Mensah, Adelina; Ofori, Benjamin; Rapp, Barbara; Marx Gómez, Jorge, eds): *Dialogue on Sustainability and Environmental Management*. Shaker Verlag, Düren, pp. 1–8, 2017.
- [Sc17] Schallmo, Daniel; Rusnjak, Andreas; Anzengruber, Johanna; Werani, Thomas; Jünger, Michael, eds. *Digitale Transformation von Geschäftsmodellen: Grundlagen, Instrumente und Best Practices*. Springer Fachmedien Wiesbaden, Wiesbaden, 2017.
- [WHP20] Wulf, Carmen; Haberstroh, Susanne; Petersen, Maren, eds. *Forschendes Lernen: Theorie, Empirie, Praxis*. Springer Fachmedien Wiesbaden, Wiesbaden, 2020.

An instant matching algorithm in the context of ride-hailing applications, using isochrones and social scoring

Malte Eilers,¹ Simon Hempen,² Sebastian Frers,³ Benjamin Stahr,⁴ Natalia Hölscher,⁵
Paula Bürmann,⁶ Christian Janßen,⁷ Prof. Dr.-Ing. habil. Jorge Marx-Gómez⁸

Abstract: Ride-sharing nowadays is a term most people have heard about. It is an efficient way to reduce both climate detrimental emissions and the overall traffic within large cities. A key aspect of ride-sharing is to match drivers and passengers according to a set of attributes. By matching people based on static and dynamic settings such as social preferences, this paper shows a fast and efficient approach for solving this problem. As a key difference to other solutions the core of the developed instant matching algorithm is built up on isochrones. These are used to match trips on specific geographical constraints, like the maximum detour and time until a passenger's pickup. Furthermore, social metrics allow optimization based on individual preferences in the users trip selection. The foregoing steps result in a comparative score to allow an optimal set of trips as an output.

Keywords: ride-sharing; ride-hailing; ride-pooling; dial-a-ride-problem; instant matching algorithm; isochrones; person matching; social scoring; sustainability

1 Introduction

Through the rise of mobility, metropolitan areas face traffic congestion problems. Especially in highly populated cities, this can lead to increased air pollution [CLC17]. Seen globally, traffic-related emissions increased by 68 percent from 1990 to 2015, accounting for 24 percent of the total amount of released greenhouse gases in the year 2015 [Sc18]. Though a multidisciplinary approach is needed to tackle this problem, one way to reduce the number of cars on the road and subsequently the overall pollution is ride-sharing [Ja17]. Since 2010, ride-sharing services have been expanding the possibilities of passenger transportation more and more [Ç17]. Established ride-sharing solutions like BlaBlaCar focus on arranging shared long-distance trips for individuals [CC]. Uber on the other hand is more like a cab company by matching individuals with licensed drivers [BP14]. This excludes the everyday short-distance travel by individuals within major cities from the potentially

¹ University of Oldenburg, 26129 Oldenburg, Germany malte.eilers1@uol.de

² University of Oldenburg, 26129 Oldenburg, Germany simon.hempen@uol.de

³ University of Oldenburg, 26129 Oldenburg, Germany sebastian.frers@uol.de

⁴ University of Oldenburg, 26129 Oldenburg, Germany benjamin.stahr@uol.de

⁵ University of Oldenburg, 26129 Oldenburg, Germany natalia.hoelscher@uol.de

⁶ University of Oldenburg, 26129 Oldenburg, Germany paula.buermann@uol.de

⁷ University of Oldenburg, 26129 Oldenburg, Germany christian.janssen@uol.de

⁸ University of Oldenburg, 26129 Oldenburg, Germany jorge.marx.gomez@uol.de

pollution-reducing benefits of ride-sharing. Challenges in the realization of a system for short-distance travel range from matching participant's preferences to providing a reliable system that optimizes matches in real-time [Ag10]. At this point the algorithm comes into the foreground. The aim of this paper is to present an algorithm for matching participants dynamically in real-time. By determining the accessibility using isochrones, an efficient matching between driver and passenger in short-distance travel range can be achieved. In addition to spatial factors, the matching is also influenced by social preferences. To create an understanding of the topic, current matching algorithms for dynamic ride-sharing are presented in Chapter 2. Chapter 3 explains how the matching algorithm works. This is followed by an evaluation using a simulation to validate various scenarios in Chapter 4 and finally a conclusion in Chapter 5.

2 State of the art

The matching process in ride-sharing can be subordinated to the Dial-a-Ride Problem (DARP). A matching process for short-distance destinations is dynamic and thus increases the difficulty of solving the problem. The solutions for the dynamic ride-sharing problem are various. In literature many solutions for matching algorithms use heuristics [Wo04, Ja86, AGK20]. This helps to reduce complexity as DARP is considered NP-hard [CAF16]. One approach includes a time window the matching algorithm. Time windows contain the maximum time of a customer's trip and the time of pickup, in order to ensure the waiting time may only be within a certain time window. The algorithm minimizes waiting time as far as possible [Ja86]. To match the users, Schreieck et al. developed an algorithm which checks whether trip offers are close to the departure and destination of trip requests. A trip request is matched with a trip offer if they are close to each other. This approach has low calculation time, but does not consider obstacles like parks and rivers into the matching [Sc16]. Some solutions to the Dial-a-Ride-Problem also include social aspects such as gender, employment status or sociality. Aydin et al. assign different weights to the social aspects, depending on the interests of the customer. When a passenger sends a request, the Needleman-Wunsch algorithm is used to determine which offered routes are similar to the passenger's route. For remaining drivers an average of the social aspects is calculated. The users are matched with the highest corresponding average of the social score. The algorithm matches participants with similar routes and social aspects [AGK20]. Other objectives when solving the ride-matching problem include minimizing total system travel cost [SX13], maximizing number of matches [SX13, NJ16] and maximizing distance savings [NJ16].

The algorithm described in Chapter 3 is carried out according to the methodology of Design Science Research with the model of Österle and Becker. This four-step model - Analysis, Design, Evaluation, Diffusion - was followed to create a better performing information system, in this case the matching algorithm for ride-hailing [Ö11]. In contrast to other research this paper extends the matching process by using isochrones around the pick-up and drop-off points to pay more attention to the environmental conditions and further

explores the application of using social metrics in the objective function by updating driver preferences on the decisions made.

3 Functionality of the matching algorithm

The main task of the matching algorithm is to select the best trips for a passenger from all available trips previously calculated within the system using a routing engine. Firstly, after the passenger requested a ride, the whole set of trips gets preselected. This is done by fixed filter parameters, which allow to sort out all the trips that violate these general constraints. To reduce the set of trips even more, isochrone matching is used next to match the geographical conditions. Once the pre-filtering of the trips is completed, secondly a score is calculated to compare the resulting trips. In addition to the detour for the driver and the waiting time for the passenger, the score also includes social preferences. The matching algorithm follows a holistic concept including several steps which are pictured in figure 1.

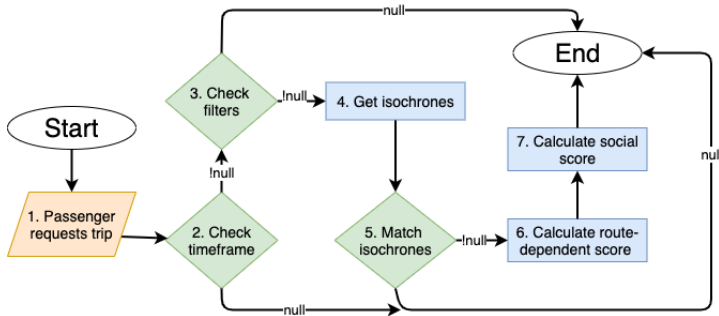


Fig. 1: Flow diagram of the matching algorithm process (Source: own representation)

3.1 Preselection of the trips

In the first step of the matching algorithm in figure 1 the passenger requests a trip. Afterwards fitting trips are filtered based on the fixed parameters shown in step two and three. The time of departure (step two), extended by six hours⁹, is used to create a valid timeframe and remove trips too far in the future. Then in step three the filter options (contact or friend, user rating and number of seats) are applied to exclude unwanted trips.

3.2 Isochrone Matching

A further reduction of possible trips is achieved in the fourth and fifth step from figure 1, based on isochrones. Isochrones are polygons indicating an area reachable from the central

⁹ The six hour time frame can be adjusted as needed and will be used within this implementation to maintain a reasonable number of possible trips.

point in a certain time or distance [Gal1]. The isochrones are calculated with the help of a routing engine in step four from figure 1 and result in a high reduction of trips for the matching process. Thus, the calculation of isochrones represents a decisive step in the matching algorithm. Trips involving a large detour for the driver are no longer taken into account in matching in step five. By spatially limiting the pick-up and drop-off points of the passenger, the driver is prevented from driving past the destination of the planned route in order to deliver the passenger at the drop-off point.

3.2.1 Route reduction through isochrones

For the consideration of the driver’s detour, a representation for a time interval is chosen. It means that as soon as a node of the route is contained within this polygon, the driver can reach the center of the polygon with a maximum detour of twice the value of the isochrone, turn around and continue the planned route at the same nodal point. However, this is an idealized value, since in reality the driver would still have to search for a turning point. This estimation can only determine whether the points can be reached at all, with a maximum detour. An exact determination of the specific detour of the driver is not possible, because the point’s within the polygon can not be determined or a possibly optimized route with a shorter detour can occur due to a changed routing. Figure 2 shows the possible changes compared to the initial route.

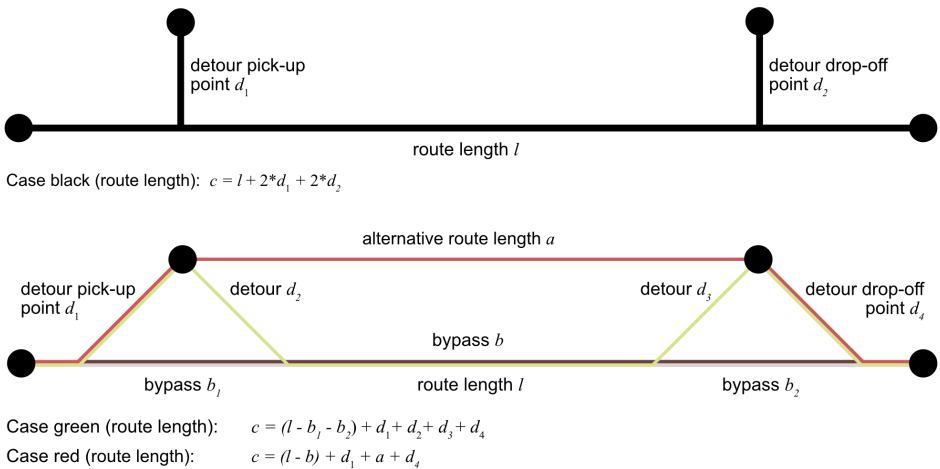


Fig. 2: Options for changing the route (Source: own representation)

For the red and green case, the length of sections $b1$ and $b2$, as well as a and b , are not presented for the determination of the absolute route length. However, these changes are always shorter than the route variant shown in the black case, whose partial routes are all

estimated to the maximum or, in the case of route length l , have already been calculated. The route selection is based on the principle of whether the system can guarantee the driver that the detour is always smaller than the driver prefers. In addition, in this step the maximum detour can be compared to a specified maximum system value or a dynamic value based on the route length.

3.2.2 Estimation of the maximum detour

To estimate the detour more precisely, several isochrones with an increasing range are calculated around the start and end points of the passenger. Four isochrones with a range of 1.25, 2.5, 3.75 and 5 minutes are used. The individual polygons reflect the doubled detour of 2.5, 5, 7.5 and 10 minutes. For each route, the smallest polygon around the start and end point is now determined by finding a nodal point of the route. By adding the intervals of the two determined isochrones, the maximum deviation from the driver's optimal route can be estimated. An example calculation of this deviation can be seen in Figure 3.

The first route, Route I, intersects the 3.75 minute isochrone formed around the passenger's

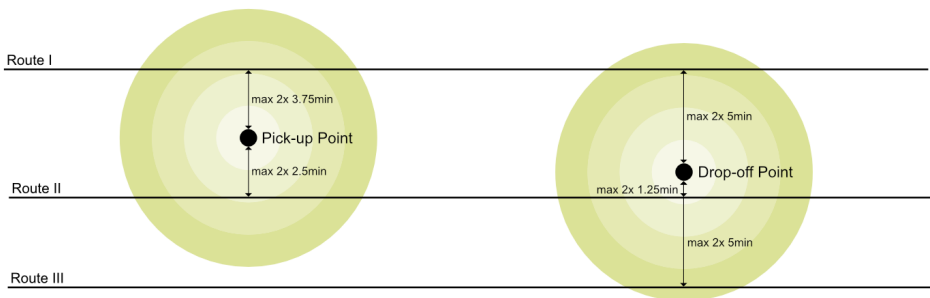


Fig. 3: Estimation of the detour (Source: own representation)

starting point. At the end point, this route intersects the isochrone describing a maximum detour of 10 minutes. As a result, the maximum detour of the trip is:

$$7.5min + 10min = 17.5min > 15min$$

If the driver of Route I specifies, for example, a maximum detour of 15 minutes, based on the calculation of the isochrones, it cannot be guaranteed that this driver requirement will be met. The route is removed from the further matching procedure. For Route II, a maximum detour of 7.5 minutes can be calculated by the isochrones:

$$5min + 2.5min = 7.5min < 15min$$

In this case, the estimation can ensure that the identical requirement of the driver is maintained during the matching procedure. The route is therefore still considered during the further matching procedure. For the case of Route III, if a route does not intersect any of

the isochrones at either start or end point, the route will also not be considered in the further steps due to the lack of maximum estimation. Another problem may arise if the direction of the driver's movement is opposite to the direction of the passenger. For example, the driver may have a small detour from certain route points to both points of the passenger, but the driver's route passes the passenger's drop-off point first. In this case, the maximum estimated detour must be increased due to the turnaround during the trip and the trip may need to be removed from the matching process. To solve this problem, only the isochrones around the pick-up point of the driver are initially considered. If the route does not intersect any of these isochrones, the process can be terminated at this point since no estimation is possible. If the route intersects an isochrone of the pick-up point, the first route point seen from the route's starting point is cached. For the determination of the drop-off point isochrone, the route is limited to the subsection between this stored route point and the end point of the route. Despite this consideration, the method has weaknesses in determining points that are close to each other and therefore can be reached via the identical route point. The solution to this problem is challenging without a concrete reachability determination of those points. However, this would increase the complexity and computation time to the extent that these cases remain in the matching process.

3.2.3 Spatial limits of pick-up and drop-off points

In addition to the calculation, the location of the pick-up and drop-off point of the trip is considered in relation to the location of the start and end points of the rides. This extra step is intended to prevent the driver from driving past his actual destination to drop off a passenger. Furthermore, this ensures that the driver always drives in the intended direction and does not have to drive in the opposite direction first. The method used is based on pure vector calculations due to their high-performance implementation. Figure 4 shows the steps of the method. To obtain the area in which both the pick-up and the drop-off point can be located, the air line between the start and end point of the trip is calculated first.

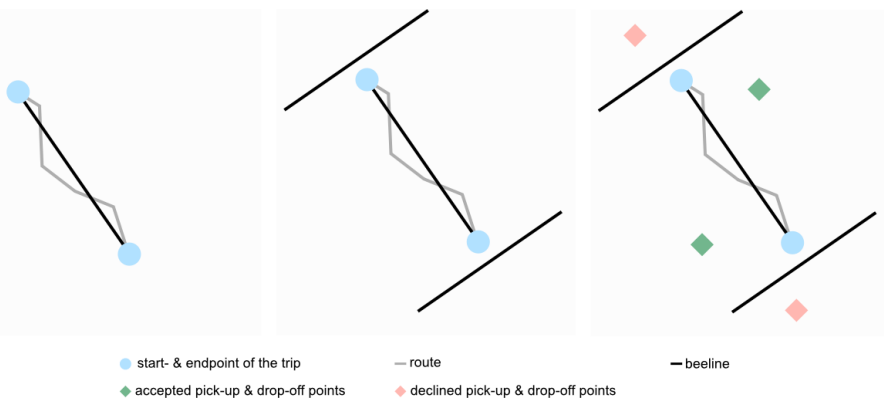


Fig. 4: Spatial limits of pick-up and drop-off points (Source: own representation)

The air line is rotated 90 degrees at both points to create a square area between the start and end point. Due to the simpler calculations compared to more complex shapes such as convex contours, a square shape is used. To include points in the close vicinity of the start and end point, the line is shifted away from the start and end point by 20 percent of the air lines length respectively. With a five kilometer air line between the routes start and end point, the line shift would integrate points with a one kilometer air line into the search area. In the context of short-distance trips, this range can be considered reasonable. For trips that initially move away from the end point of the trip, the route point with the greatest distance to the end point is selected. If after these steps either the passenger's pick-up or drop-off point is located outside these defined boundaries, the trip is excluded from the matching process. Otherwise the trip is considered in the final step of the process.

3.3 Calculation of a route score

After applying the matching in Paragraph 3, the best routes are selected in the sixth and seventh step from figure 1 based on a set of metrics. The metrics can be divided into route-dependent (step six) and individualized social metrics (step seven). Route-dependent metrics include absolute detour time, the proportional increase in route length compared to the original route and time to passenger's pickup. Since the route reduction using isochrones only estimates a maximum time detour and not the mentioned route-dependent metrics, a routing engine is used in an additional step to accurately recalculate the routes including the passenger's pick-up and drop-off point. Subsequently, the metrics are determined based on the difference between the original route and this newly recalculated route. Although the routing engine has a disadvantage in the computing time of the new routes compared to other methods such as approximations, the accuracy of the passenger's pick-up time in particular is a key argument for the integration into this algorithm. In order to compare the different metric types, the various metrics get mapped on a range of 0 to 1, where 1 is the best result. Optimizing these metrics benefits both the driver and passenger. Route-dependent metrics take up 67.5 percent of the score, where each metric has a 22.5 percent share of it. These values have been chosen for the first version of this algorithm, since the route-dependent metrics have a higher relevance for a well functioning match. Social metrics add a drift of suggested routes toward passenger's preferences regarding their driver to enhance the user experience. These include smoking, type of relationship, gender, user rating and willingness to either talk or hear music during the trip. Social metrics take up 32.5 percent of the score, whereby each metric has a 5.4 percent share of it. Except for user rating and type of relationship, these metrics get dynamically optimized based on the passenger's previously selected trips using weights. Every category of a metric has a weight $\omega \in [0, 0.05416]$ attached to it, all weights of one metric sum up to 0.05416. For first-time passengers, the weights of a metric are evenly distributed, which means exemplary for three categories, that at the beginning ω is equal to 0.01805. For adjusting the weights to the preferences of a passenger, the softmax function is used to scale the current weights of a metric to the interval $[0, 1]$, but these maintain their ratios to each other [Nw18]. Then, for each selection

that matches the currently selected category, a small value $\varepsilon \in \mathcal{R}$ is added to ω or subtracted for any mismatch. To compensate that a passenger's choice is sometimes the consequence of choices offered, ε is multiplied by $(1/f)$ where f is the frequency of the category's occurrence in the options offered. If the relative frequency of a category in options offered was 1, this compensation is not performed, because the passenger would have no choice. To adjust the resulting weights to the score's calculation, the softmax function is used again. Then the weight of the selected category is multiplied by the proportion of its metric, which is 5.4 percent. In this way, a score is calculated for each route, taking into account route-dependent and social metrics. Based on this, the best matches are returned to potential passengers.

4 Evaluation

An evaluation of the algorithm is carried out, according to the methodology of Design Science as mentioned in paragraph 2. It should be ensured that the instant matching algorithm generates appropriate matches for both the driver and the passenger and that its outputs provide optimized solutions to the aspects defined in paragraph 3. To achieve this, a test followed by an evaluation is mandatory [He04]. First, the general solution space of matching will be outlined. The evaluation is based on the score criteria set out in paragraph 3 and is carried out using different test scenarios. Those test scenarios are intended to check whether the desired behavior of the algorithm occurs and thus verify its functionalities. Due to the short duration of the project and the COVID-19 pandemic, there is no opportunity of in-person rides as well as to evaluate the influence of social metrics. Furthermore, there are no suitable data sets to allow independent evaluation, so randomly generated trips are used to validate the algorithm. The algorithm has been developed in a first stable version and is based on a database of randomly generated routes. A total of 5292 routes have been created, which constitute the basis for the evaluation. By using various filters, the pool of all available trips is significantly reduced, which means not all routes are used for the matching process.

Table 1 and Table 2 show which routes are put out after a request from the passenger. For this purpose, the desired start and end position of the passenger, as well as the preferred start time, are transferred and then the calculated scores are evaluated. The first test scenario was requested on 04/14/2021 at 06:34:47. Results of the request are shown in table 1. Table 2 shows the output for a request on 04/02/2021 at 16:00:00. The delay of the driver results

Route ID	Score	Detour time (h:m:s)	Detour distance (%)	Time until pickup (h:m:s)
2464	0.79587	00:05:35	35	00:19:12
2363	0.58698	00:08:23	38	03:40:48
2594	0.13314	00:10:58	46	04:48:04

Tab. 1: First test scenario

from the detour time and the distance the driver has to travel to pick up the passenger. Both detour time and distance have a 22.5 percent influence on the matching score. In table 2 it is

Route ID	Score	Detour time (h:m:s)	Detour distance (%)	Time until pickup (h:m:s)
432	0.77405	00:08:44	25	01:07:12
444	0.59168	00:09:32	23	00:37:12
286	0.46281	00:09:15	35	00:35:38
366	0.25393	00:09:41	37	02:30:00
510	0.16085	00:09:34	35	06:00:36

Tab. 2: Second test scenario

shown that the routes with IDs 286, 366 as well as 510 get a score below 0.5. On the other hand, the score for the routes with IDs 432 and 444 is above 0.5. The detour distance is smaller for the drivers in relation to the other three routes and is therefore rated better in the score. Especially the comparison between routes 286 and 444 clearly shows the influence of the detour distance on the score, as the other metrics show similar values. In table 1, the influence of detour time is illustrated, as the percentage detour distance of routes 2464 and 2363 differs by only three percent. Since the detour time for route 2363 is significantly higher, the score is proportionally lower. By comparing the two routes with IDs 2464 and 2363 in the first test scenario shown in table 1, it demonstrates the calculation of scores is influenced by the waiting time of the passenger. Despite similar values in terms of detour, route 2363 receives a lower score due to its long waiting time for the passenger. This is equally clear in the comparison of routes 286 and 366 from table 2. Both the interests of the driver and the passenger are considered in the calculation of the score. Since a large number of available drivers is essential for a dynamic ride-sharing system, the interests of the driver are weighted with a total of 45 percent. The time to pick up is included in the calculation with 22.5 percent. As a result, trips whose values are lower in terms of distance and duration of detour are displayed preferentially in the matching. If the comparison between two possible routes is similar in terms of the detour, as is the case for routes 286 and 366 in the second test scenario, it will become clear that the route with low waiting time receives a significantly better score. Considering the performance of the matching algorithm, it should not take longer than ten seconds to keep the user's attention [Ni94]. The test scenarios in table 1 and 2 take about one second on average, which is below the threshold. The reduction out of the 5292 offered routes by the pre-filtering processes has to be emphasized here. In the first test scenario 17 possible routes are considered and nine of them get filtered out because of their detour. The second test scenario takes 12 possible routes but only needs to recalculate five routes. Only by reducing the routes through the matching process, this acceptable result can be achieved for the user.

5 Conclusion

The paper represents a matching algorithm in the context of ride-hailing to join a passenger with a driver using isochrones and social scoring. The algorithm's internal evaluation points out the success of isochrone matching as an efficient way to represent both, the drivers and

passengers interests while still getting suitable results in terms of the drivers detour. It is a lightweight implementation mostly limited by external sources such as the used routing engine, which handles the isochrone calculation as well as the routing itself done within the matching process. Since this evaluation is based on a locally hosted engine, processing times might increase when running the operations on an external server. It is the only bottleneck in terms of speed though, having the greatest impact on the process.

To optimize the performance of this algorithm in a next version, tweaking the static weights might result in a better output since those are currently equally distributed. Furthermore, the evaluation must be conducted in a public study to obtain meaningful results for both matching itself and the influence of social factors.

Bibliography

- [Ag10] Agatz, N. A. H.; Erera, Alan; Savelsbergh, Martin; Wang, X.: Sustainable Passenger Transportation: Dynamic Ride-Sharing. ERIM Report Series Reference No. ERS-2010-010-LIS, 01 2010.
- [AGK20] Aydin, Omer Faruk; Gokasar, Ilgin; Kalan, Onur: Matching algorithm for improving ride-sharing by incorporating route splits and social factors. PLOS ONE, 15(3):1–23, 03 2020.
- [BP14] Bonazzi, Riccardo; Poli, Michaël: Beyond Uber. Business model considerations for alternatives to traditional taxis. Proceedings of the 11th Conference of the Italian Chapter of the Association of Information Systems (ITAIS) : Digital Innovation and Inclusive Knowledge in Times of Change 2014, 11 2014.
- [CAF16] Campbell, I.; Ali, M. M.; Fienberg, M. L.: Solving the dial-a-ride problem using agent-based simulation. South African Journal of Industrial Engineering, 27:143–157, 11 2016.
- [CC] Consumers, Health, Agriculture; (Chafea), Food Executive Agency: Exploratory study of consumer issues in online peer-to-peer platform markets: Task 4 – Case study: BlaBlaCar. Technical report, Consumers, Health, Agriculture and Food Executive Agency (Chafea).
- [CLC17] Chang, Yu; Lee, Yong; Choi, Sungsup: Is there more traffic congestion in larger cities? -Scaling analysis of the 101 largest U.S. urban centers. Transport Policy, 59:54–63, 10 2017.
- [Ga11] Gamper, Johann; Böhlen, Michael; Cometti, Willi; Innerebner, Markus: Defining Isochrones in Multimodal Spatial Networks. Proceedings of the 20th ACM International Conference on Information and Knowledge Management, p. 2381–2384, 2011.
- [He04] Hevner, Alan R.; March, Salvatore T.; Park, Jinsoo; Ram, Sudha: Design Science in Information Systems Research. MIS Quarterly, 28(1):75–105, 2004.
- [Ja86] Jaw, Jang-Jei; Odoni, Amedeo R.; Psaraftis, Harilaos N.; Wilson, Nigel H. M.: A heuristic algorithm for the multi-vehicle advance request dial-a-ride problem with time windows. Transportation Research Part B: Methodological, 20(3):243–257, 1986.

-
- [Ja17] Jalali, Roozbeh; Koochi-Fayegh, Seama; El-Khatib, Khalil; Hoornweg, Dan; Li, Heng: Investigating the Potential of Ridesharing to Reduce Vehicle Emissions. *Urban Planning*, 2(2):26, 6 2017.
- [Ni94] Nielsen, Jakob: *Usability Engineering*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 11 1994.
- [NJ16] Nourinejad, Mehdi; J.Roorda, Matthew: Agent based model for dynamic ridesharing. *Transportation Research Part C: Emerging Technologies*, 64:117–132, 03 2016.
- [Nw18] Nwankpa, Chigozie; Ijomah, W.; Gachagan, A.; Marshall, Stephen: Activation Functions: Comparison of trends in Practice and Research for Deep Learning. *CoRR*, abs/1811.03378, 2018.
- [Sc16] Schrieck, Maximilian; Safetli, Hazem; Siddiqui, Sajjad Ali; Pflügler, Christoph; Wiesche, Manuel; Krcmar, Helmut: A Matching Algorithm for Dynamic Ridesharing. *Transportation Research Procedia*, 19:272–285, 2016. *Transforming Urban Mobility*. mobil.TUM 2016. International Scientific Conference on Mobility and Transport. Conference Proceedings.
- [Sc18] Scheelhaase, Janina; Grimme, Wolfgang; O’Sullivan, Marlene; Naegler, Tobias; Klötzke, Matthias; Kugler, Ulrike; Scheier, Benedikt; Standfuß, Thomas: Climate Protection in the Transport Sector — Current Examples. *Wirtschaftsdienst*, 98(9):655–663, 2018.
- [SX13] Santos, Douglas O.; Xavier, Eduardo C.: Dynamic Taxi and Ridesharing: A Framework and Heuristics for the Optimization Problem. *International Joint Conference on Artificial Intelligence Procedia*, 13:2885–2891, 08 2013.
- [Wo04] Wolfer Calvo, Roberto; de Luigi, Fabio; Haastруп, Palle; Maniezzo, Vittorio: A distributed geographic information system for the daily car pooling problem. *Computers & Operations Research*, 31(13):2263–2278, 2004.
- [Ç17] Çetin, Tamer: *The Rise of Ride Sharing in Urban Transport: Threat or Opportunity?* Intech, Hamburg, 1 edition, 01 2017.
- [Ö11] Österle, Hubert; Becker, Jörg; Frank, Ulrich; Hess, Thomas; Karagiannis, Dimitris; Krcmar, Helmut; Loos, Peter; Mertens, Peter; Oberweis, Andreas; Sinz, Elmar J.: Memorandum on design-oriented information systems research. *European Journal of Information Systems*, 20(1):7–10, 2011.

Konzeption und Entwicklung einer Weboberfläche zur Visualisierung umweltfreundlich erzeugter Energie in Industrienetzwerken

Volker Wohlgemuth¹ und Abdullah Güccük²

Abstract: Die Industrie ist eines der wichtigsten Treibhausgasemittenten in Deutschland. Entsprechend sind Maßnahmen für die Senkung von Emissionen von großer Bedeutung. Die Mitwirkung der Industriegebiete ist hierbei gefordert. Überzeugende und motivierende Lösungsansätze sind nötig, um Unternehmen in Industriegebieten für den Umweltschutz und für Nachhaltigkeitsthemen zu sensibilisieren. Die entwickelte Weboberfläche für das Unternehmensnetzwerk Motzener Straße in Berlin visualisiert die umweltfreundliche Produktion von Energie im Industriegebiet mit dem Ziel, andere Betriebe zum Mitmachen zu motivieren und das Bild der Industrie in der Stadt zu verbessern. Das Industriegebiet soll als Vorbild für weltweit andere Industriegebiete dienen, die Produktion von umweltfreundlicher Energie in Industriegebieten fördern sowie das gemeinsame Nutzen von dieser Energie antreiben.

Keywords: Industriesymbiose, Industrienetzwerke, Eco-Visualisation

1 Einleitung

Mit fast 84% Anteil ist der Energiesektor die größte Emissionsquelle in Deutschland [Em21]. Darunter gehört die Industrie mit 15% Anteil und ist somit einer der wichtigsten Treibhausgasemittenten [Em21]. Umgerechnet ergibt sich ein Wert von ca. 108 Mio. t CO₂-Äquivalente, die die Industrie im Bereich Energie sparen kann. Um dieses Sparpotenzial vollständig ausschöpfen zu können, sind Industrienetzwerke wichtige Startpunkte, da sie aus einem Verbund von Unternehmen bestehen und einen gemeinsamen Standort teilen. Der Abfall eines Unternehmens kann für ein anderes Unternehmen eine wertvolle Ressource sein. So kann durch die gemeinsame Lage eine schnelle Koordination und ein schneller Austausch stattfinden. Das verschafft den jeweiligen Unternehmen einen wirtschaftlichen und ökologischen Vorteil. Ein weiterer Vorteil kann auch bei der Strom- und Wärmeerzeugung gewonnen werden. Mehrere Unternehmen eines Industrienetzwerkes könnten gemeinsam umweltfreundliche(n) Wärme/Strom erzeugen und diesen verbrauchen [Is14].

¹ Hochschule für Technik und Wirtschaft Berlin, Fachbereich Ingenieurwissenschaften – Technik und Leben, Studiengang Betriebliche Umweltinformatik Wilhelminenhofstraße 75A, 12459 Berlin, Volker.Wohlgemuth@HTW-Berlin.de

² Hochschule für Technik und Wirtschaft Berlin, Fachbereich Ingenieurwissenschaften – Technik und Leben, Studiengang Betriebliche Umweltinformatik, Wilhelminenhofstraße 75A, 12459 Berlin

1.1 Problemstellung

Das Unternehmensnetzwerk Motzener Straße in Berlin ist ein Gewerbegebiet, in dem versucht wird, eine koordinierte, umweltfreundliche und wirtschaftliche Zusammenarbeit zwischen Unternehmen zu leisten. Nahezu 200 Betriebe sind in diesem Gebiet ansässig [Un20]. Eine koordinierte Zusammenarbeit bringt ein beträchtliches Potenzial zum Schonen von natürlichen Ressourcen mit sich. Allerdings ist eine Zusammenarbeit von Unternehmen nicht immer einfach. Die Aufteilung und Begleichung der entstehenden Kosten, der Wille bzw. die Motivation zur umweltfreundlichen Herstellung von Produkten und Dienstleistungen und die Komplexität der Koordinierung von einem derartig großen Industriestandort sind einige Herausforderungen, die gemeistert werden müssen [Sc94]. Insbesondere stellt die Sensibilisierung der Unternehmen für eine ökologische und nachhaltige Produktion eine reale und große Herausforderung dar. Die Gewinnmaximierung und die damit verbundene Kostenminimierung sind Prinzipien von wirtschaftlichen Betrieben, die ökologische Prinzipien und Ziele weniger wichtig erscheinen lassen. Genau an dieser Stelle müssen Maßnahmen ergriffen werden, die die Unternehmen antreiben und für eine ökologische und nachhaltige Produktion leiten.

1.2 Zielsetzung

Im Gewerbegebiet Motzener Straße wird bereits auf umweltfreundlichem Wege Wärme und Strom produziert, jedoch nicht in Koordination. Diese produzierte Energie soll erfasst, visualisiert und veröffentlicht werden. Hierfür ist die Ausstattung der jeweiligen Anlagen zur Erzeugung von Energie mit geeigneter Technik vorgesehen. Die in Echtzeit ausgelesenen Daten sollen über eine noch zu installierende Leitung an einen Server weitergeleitet und in einer Datenbank gespeichert werden. In einer Webseite soll das Ergebnis visualisiert und der Öffentlichkeit zugänglich gemacht werden. Dabei ist es wichtig, dass die Daten den jeweiligen Unternehmen zugeordnet werden, um entsprechende Grafiken erstellen zu können. Diese Informationen sowie die folgenden Ziele stammen aus einem persönlichen Interview mit dem Unternehmensnetzwerk (U. Misgeld, persönliche Kommunikation, Berlin, 25.09.2020).

Auf Basis dieser Visualisierung soll der Grundbaustein für weitere umweltfreundliche Aktionen im Gewerbegebiet geschaffen werden. Sie soll dazu beitragen, dass die Öffentlichkeit ein positiveres Bild von Industriegebieten in Städten bekommt. Eine Studie aus dem Jahre 2018 zur Messung der Akzeptanz von Industrie zeigt, dass die Akzeptanz der Industrie in der Öffentlichkeit steigen sollte [BZ18a]. V.a. ist die Akzeptanz der Industrie in den Augen den jüngeren Menschen deutlich geringer als bei älteren Menschen [BZ18b]. Die zu erstellende Visualisierung soll zeigen, dass die Industrie auch „grün“ produzieren kann. Folglich soll die Anerkennung und Akzeptanz der Industrie in der Öffentlichkeit gesteigert werden. Jüngere Menschen sollen für industrielle Aktivitäten und Projekte begeistert werden. Hinzu kommt, dass die Energieeffizienz von solch einem Unternehmens-

verbund gesteigert wird. Zusätzlich soll dieses Projekt als Beispiel für andere Industriernetzwerke dienen und somit den Einsatz von erneuerbaren Energien in anderen Gewerbegebieten fördern. Unternehmen sollen für eine stärkere Zusammenarbeit im Netzwerk motiviert werden.

Dementsprechend muss die Visualisierung Jahresvergleiche ermöglichen, Betrachter überzeugen und Sparpotentiale darstellen. Eine Gegenüberstellung des Jahresbedarfs und die Potentiale für weitere CO₂-Einsparungen sind Beispiele für solch eine Visualisierung. Zielgerichtete und aussagekräftige Grafiken müssen erstellt werden. Letztendlich soll ein Konzept zur Visualisierung von umweltfreundlich erzeugter Energie in Industriernetzwerken erarbeitet und ein erster Prototyp entwickelt werden.

2 Grundlagen

2.1 Das Industriegebiet

Das Industriegebiet Motzener Straße befindet sich in Berlin-Marielfelde und zählt mit mehr als 100 ha Gewerbe- und Industriefläche zu den größten Industriegebieten in Tempelhof-Schöneberg (Berlin) [Wi20]. Mit einer hohen Anzahl an ansässigen Betrieben und ca. 5000 Beschäftigten hat das Gewerbegebiet Motzener Straße die höchste Industriequote unter den drei bezirklichen Gewerbegebieten. Knapp 80% der Beschäftigten und fast 50% der ansässigen Unternehmen werden dem produzierenden Gewerbe zugeordnet [Wi20].

2.2 Das Unternehmensnetzwerk Motzener Straße

Das Unternehmensnetzwerk Motzener Str. besteht seit 15 Jahren und ist ein Verein mit mehr als 60 Mitgliedern [Wi20]. Die Initiative wurde mit der Intention gegründet, eine Kooperation zwischen den Unternehmen zu organisieren und zu koordinieren, um somit Wettbewerbsvorteile zu verschaffen und die Entwicklung des Standortes zu gewährleisten [Un20].

Durch regelmäßig stattfindende Treffen wird den ansässigen Unternehmen ermöglicht, sich gegenseitig kennenzulernen, Erfahrungen und Beziehungen auszutauschen sowie neue Möglichkeiten zur Zusammenarbeit zu bilden. Zudem werden seitens des Netzwerkes Aktionen koordiniert, die u.a. den gemeinsamen Einkauf von Energie und Wärme, den gemeinsamen Bezug von Dienstleistungen sowie weitere Zusammenarbeiten umfassen, die den teilnehmenden Unternehmen eine bessere Marktposition verschaffen sollen. Ein Gebietskinderkarten für Mitarbeiter des Industriegebiets, die Kooperation mit naheliegenden Schulen für die Vergabe von Ausbildungsplätzen sowie die Zusammenarbeit mit Hochschulen für die Durchführung von zukunftsorientierten Projekten sind einige weitere Tätigkeiten, die die Synergie im Industriegebiet fördern sollen [Wi20].

2.3 Industriesymbiose

Das Industriegebiet und Unternehmensnetzwerk können als eine sogenannte *Industriesymbiose* bezeichnet werden. Unter einer industriellen Symbiose wird die zwischen Betrieben und Unternehmen stattfindende Zusammenarbeit zur Koordinierung von Energie- und Stoffströmen verstanden [VMI12]. Die industrielle Symbiose verbindet Unternehmen und Betriebe aus unterschiedlichen Branchen in einem kollektiven Ansatz zum Wettbewerbsvorteil [Ch06]. Ausgetauscht werden in solch einem Netzwerk Energie, Materialien, Wasser oder Nebenprodukte aus betrieblichen Prozessen. Abfälle oder Nebenprodukte, die in einem Betrieb als Abfall gelten, könnten in solch einem Netzwerk für andere Betriebe wichtige Ressource darstellen.

2.4 Datenvisualisierung

Visualisierung wurde schon vor der Schriftsprache als Mittel zur Kommunikation verwendet [WGK15]. Das menschliche Auge ist das wichtigste Instrument bei der Analyse von Daten [Ru15a]. Daher ist die Visualisierung von Daten von zentraler Bedeutung, wenn es darum geht, Daten zu analysieren. Definiert wird das Visualisieren in der Literatur als „[...]etwas bildhaft darstellen[...]“ [Se11]. Visualisierung ist die Kommunikation von Informationen mittels grafischer Darstellungen [WGK15a]. Daten werden in eine visuelle Form umgewandelt und bspw. als ein Diagramm abgebildet. Somit ist es für den Menschen einfacher aus den Daten Informationen zu erkennen. Das menschliche Gehirn kann ein Bild bzw. eine Visualisierung deutlich schneller verarbeiten als ein Text, wohingegen die Analyse eines Textes von der Lesegeschwindigkeit abhängig ist [WGK15a].

2.5 Auswahl der Visualisierungstechnik

Es gibt zahlreiche Visualisierungstechniken, wie z.B. konventionelle sowie moderne Methoden. Als geeignet wurden die konventionellen Diagramme eingestuft. Mit Hilfe von Säulen-, Balken-, Flächen-, Linien-, Streu- und Kreisdiagrammen können der Verbrauch bzw. die Produktion von Energie einfach für den Anwender verdeutlicht werden. Zudem ermöglichen diese Diagramme diverse Vergleiche. Hinzu kommt, dass diese miteinander kombiniert werden können, um vielfältige Darstellungen zu konstruieren. Außerdem sind diese Diagramme weit verbreitet und auch für Laien in diesem Gebiet einfach zu verstehen. Diese und viele weitere Argumente sind wahrscheinlich auch der Grund dafür, dass viele Energie-Dashboards, diese Diagramme zur Visualisierung verwenden.

Moderne Methoden, wie das Venn-, Sunburst- oder Chord-Diagramm werden nicht verwendet, da diese nicht sinnvoll mit der Datenbasis dieser Arbeit eingesetzt werden können. Venndiagramme werden verwendet, um Mengen und deren Beziehungen in der Mathematik zu beschreiben [Gl03]. Sunburstdiagramme werden oft für die Verbildlichung von Hierarchien eingesetzt [Ka18], während Chorddiagramme geeignete Mittel darstellen, um Abhängigkeiten zwischen Objekten zu visualisieren [Ti06]. Die grafische Darstellung von

Verbrauchs- bzw. Produktionsdaten von Energie beinhaltet keine Mengenbeziehungen, Hierarchien oder Abhängigkeiten, als dass diese sinnvoll eingesetzt werden können.

Heatmaps können angewendet werden, um besonders umweltfreundliche bzw. -unfreundliche Bereiche auf einer Karte des Industriegebiets zu markieren. Diese Technik wird im Konzept eingesetzt. Die genaue Verwendung wird im nächsten Kapitel erläutert. Mittels Sankeydiagramme kann der Energiefluss von der Produktion bis hin zu dem Verbrauch bei den Betrieben dargestellt werden. Sankeydiagramme repräsentieren Visualisierungswerkzeuge, wenn es darum geht, Material- und Energieflüsse zu visualisieren, um Einsparpotentiale zu entdecken [Ev17].

Schlussendlich werden moderne Diagramme aus den zuvor beschriebenen Gründen nicht verwendet. Konventionelle Diagramme finden in der Entwicklung wegen ihrer Bekanntheit und Einfachheit Verwendung.

3 Konzept

3.1 Anforderungsanalyse

Für die Bestimmung der Anforderungen wurden zunächst die Stakeholder bestimmt. Dazu gehören die Betriebe, das Unternehmensnetzwerk, die Fördermittelgeber und die Unternehmen, die den Prototyp weiter entwickeln sollen. In Gesprächen wurden die Anforderungen bestimmt und in funktionale und nicht-funktionale Anforderungen unterteilt. Tabelle 1 und 2 zeigen diese Anforderungen:

Funktionale Anforderungen	
1	Die Oberfläche umfasst insgesamt 10 Unterseiten: Dashboard, Energie, Heatmap, Sankeydiagramm, Ranking/Top Ökobetriebe, Emissionen, Standort, Über uns, Hilfe und Mitmachen lohnt sich.
2	Ein Header mit dem Logo, Titel und Menüpunkten wird erstellt.
3	Die Startseite der Oberfläche umfasst das Dashboard mit den wichtigsten Informationen und Grafiken zum Grünen Kraftwerk.
4	Das Dashboard wird in drei Teile unterteilt: Die Navigation, die Visualisierungen und sonstige Informationen.
5	Die Navigation wird auf der linken Seite platziert. Hier wird das Navigationsmenü, das gerade aktiv ist, farblich hervorgehoben, damit der Benutzer weiß, in welchem Menüpunkt dieser sich gerade befindet.
6	Den Hauptteil der Startseite bekommt das Dashboard. Hier sollen die wichtigsten vier Kennzahlen präsentiert. Weiterhin werden der aktuelle Strombedarf, die umweltfreundliche Erzeugung und der kumulierte Wert für die

Funktionale Anforderungen

	künftige Produktion abgebildet. Wichtige Darstellungen zu Emissionen, teilnehmenden Betrieben und zum Standort werden hier auch abgebildet.
7	Die Sonstigen Informationen werden auf der rechten Seite der Oberfläche platziert. Hier werden Informationen zu den Top Ökobetrieben, zum Mitmachen, aktuelle Themen und Kontaktdaten bereitgestellt.
8	Unterseite Energie: Unter diesem Menüpunkt werden grafische Darstellungen über die Produktion der einzelnen Anlagen und der gesamten Produktion dargestellt.
9	Unterseite Energie: Hier werden Filtermöglichkeiten angeboten, die es den Benutzern erlauben, Echtzeitdaten und Daten aus der Vergangenheit zu begutachten. Diverse Vergleiche werden hier aufgestellt.
10	Unterseite Energie: Mit Hilfe von Tabs (Reiter) kann der Benutzer zwischen einzelnen Anlagen wechseln. Jede im Industriegebiet benutzte Anlage erhält einen eigenen Bereich mit diversen Grafiken über die umweltfreundliche Produktion.
11	Unterseite Heatmap: Unter diesem Reiter wird ein Heatmap bereitgestellt. Dieser wird die grünen Bereiche (umweltfreundliche Produktion und Verwendung) zwischen den roten Bereichen (Verwendung von konventioneller Energie) farblich unterscheiden.
12	Unterseite Sankey: Hier wird ein Sankeydiagramm bereitgestellt, welches den Energiefluss visualisieren soll. Dabei wird zwischen produziertem und gekauftem umweltfreundlichem Strom sowie gekauftem konventionellen Strom unterschieden. Der Weg dieser Energien soll bis hin zu den Betrieben dargestellt werden.
13	Unterseite Ranking: Eine Tabelle mit den 10 umweltfreundlichsten Betrieben im Industriegebiet wird vorgestellt.
14	Unterseite Emissionen: Diverse Grafiken über Emissionen werden hier repräsentiert. Dabei geht es primär darum, die gesparten Emissionen und das Potential zu zeigen.
15	Unterseite Standort: An dieser Stelle soll eine Karte des Industriegebietes vorhanden sein. Dieser soll die installierten Anlagen auf der Karte darstellen. Zudem sollen die im Industriegebiet durchgeführten umweltfreundlichen Aktionen abgebildet werden.
16	Unterseite Mitmachen: Kontaktinformationen, Kontaktformular und Fragen und Antworten zum Mitmachen für andere Betriebe werden hier bereitgestellt.
17	Unterseite Mitmachen: Für Privatpersonen werden hier Möglichkeiten zum Teilen der Oberfläche auf sozialen Medien angeboten. Zudem können Benutzer hier mit Spenden die ehrenamtlichen Tätigkeiten des Netzwerks unterstützen.

Tab. 1: Funktionale Anforderungen

Nicht-funktionale Anforderungen	
1	Die Oberfläche wird einfach verständlich, klar strukturiert bzw. übersichtlich, farbenfroh und benutzerfreundlich aufgebaut.
2	Wenig Text, große Überschriften, einheitliche Farben und Legenden sollen die Benutzerfreundlichkeit, die Attraktivität und somit die Benutzbarkeit steigern.
3	Die Gemeinsamkeiten, Vorteile und Erfahrungen, die aus der Recherche nach ähnlichen Projekten und Dashboards identifiziert wurden, werden verwendet, um eine einfache und schnell erlernbare Oberfläche zu gestalten.
4	Das Laden der Oberfläche und der Grafiken soll so wenig wie möglich Zeit in Anspruch nehmen.
5	Die Oberfläche soll so wenig wie möglich Ressourcen verbrauchen.
6	Die Oberfläche wird so entwickelt, dass diese durch fremde Entwickler einfach zu verstehen ist.
7	Die Anwendung wird keine kostenpflichtigen Bibliotheken, Programmiersprachen, Frameworks oder andere technische Entwicklungsmittel enthalten. Die Oberfläche muss möglichst plattformunabhängig gestaltet werden.
8	Die Webseite soll per SSL-Verschlüsselung gesichert werden.

Tab. 2: Nicht-funktionale Anforderungen

Die zu entwickelnde Weboberfläche soll nach der durchgeführten Anforderungsanalyse drei grundsätzliche Ziele erfüllen (U. Misgeld, persönliche Kommunikation, Berlin, 25.09.2020):

1. Die Öffentlichkeit soll ein positiveres Bild von der Industrie in der Stadt erhalten,
2. nichtteilnehmende Betriebe im Gewerbegebiet sollen mobilisiert werden und
3. das Industriegebiet soll als Vorbild für andere Gewerbegebiete dienen.

Um diesen Zielen gerecht zu werden, muss die Anwendung einerseits Privatpersonen und andererseits Personen aus Betrieben mit Leitungsposition ansprechen. Die Weboberfläche muss für beide Zielgruppen überzeugend, motivierend und benutzerfreundlich gestaltet werden.

Zusätzlich wurde für die Konzeption nach Industriegebieten gesucht, die ein ähnliches Projekt aufweisen können. Hierbei wurde ein Gewerbegebiet [Gk20] identifiziert, das aber nur eine einzige Grafik beinhaltet und somit nicht weiter relevant ist. Auch ist die Suche nach ähnlichen wissenschaftlichen Arbeiten und Projekten erfolglos geblieben. Stattdessen wurden zahlreiche Projekte identifiziert, die Verbrauchs- und Produktionsdaten von Energiedaten visualisieren. Als Beispiel kann das Dashboard des irischen Netzbetreibers EirGrid [Ei21], des Fraunhofer Institutes [FI20], des französischen Elektrizitätsübertragungsnetzbetreibers RTE [RT20] oder das Dashboard der Agora Energiewende

[Ag20] genannt werden. Charakteristisch für diese Dashboards sind große Überschriften, farbenfrohe und verschiedene Grafiken, wenig Text und viele Filtermöglichkeiten. Dazu kommt, dass diese Dashboards mit minimalen Klicks verschiedene Grafiken liefern. Zeitverlust durch langes Laden von Grafiken oder durch Wechsel zwischen mehreren Seiten sind nicht vorhanden. Grafiken werden oft im Hintergrund geladen, so dass die Dashboards eine benutzerfreundliche Nutzung darstellen. Es wurde versucht, diese charakteristischen Merkmale bei der Konzeption zu beachten und eine einfach verständliche, überzeugende und sensibilisierende Weboberfläche zu gestalten.

3.2 Datenbasis

In dem Industriegebiet sollen die erzeugten Strommengen und Wärmemengen erfasst werden. Dabei sollen die in dem Gebiet genutzten Stromerzeugungsanlagen wie Blockheizkraftwerk- (BHKW), Photovoltaik- und/oder Brennstoffzellen-Anlagen und Wärmeerzeugungsanlagen wie Wärmepumpen-, Pelletkessel- und/oder Geo- und Solarthermie mit geeigneter Technik ausgerüstet werden. Diese Technik, eine Zählerinrichtung (Industriestromzähler o.ä.) mit Schnittstelle zur Datenauslesung, soll die Daten an eine Datenbank weiterleiten. Für die Verbindung der Zählerinrichtung und der Datenbank wird ein Netzwerkanschluss oder SIM-Karte verwendet. Die Daten sollen in einer noch zu modellierenden und aufzubauenden Datenbank gespeichert werden. Standortdaten sollen ebenso übermittelt werden, um die Daten später den Anlagen/Betrieben zuordnen zu können

Aktuell ist noch keine Technik installiert, die die Daten aus den Betrieben zu einer Datenbank senden könnte. Daher gibt es keine echten Daten, die für das Projekt verwendet werden können. Aus diesem Grund wird in diesem Projekt mit Dummydaten gearbeitet.

Der Fokus liegt bei dem Konzept zur Visualisierung. Die Produktion von Energie durch die umweltfreundlichen Anlagen (im Gesamten und Einzelnen), der Energiebedarf, der Energiemix, die Anzahl der teilnehmenden Betriebe, die Emissionen, Anzahl und Typen der einzelnen installierten umweltfreundlichen Anlagen und viele weitere Beispieldaten werden für das Konzept erstellt und in der Entwicklung verwendet. Das Konzept und das entwickelte Prototyp sollen als ein erster Vorschlag für die Umsetzung einer Visualisierung dienen. Die im Konzept dargestellten und visualisierten Daten werden vom Unternehmensnetzwerk entweder übernommen, angepasst oder abgelehnt. Entsprechend sind der Visualisierung und der Bestimmung der Daten zur Visualisierung keine Grenzen gesetzt. Dies bedeutet auch, dass die Daten die erhoben werden sollen, zum größten Teil nach diesem Konzept bestimmt werden.

3.3 Entwurf der Oberfläche

Die Oberfläche wurde als ein sogenanntes Dashboard konzipiert. Auf der linken Seite ist das Navigationsmenü. den restlichen Platz belegen die einzelnen Grafiken:



Abb. 2: Das Dashboard: die wichtigsten Informationen

Im Dashboard werden dem Benutzer die wichtigsten Informationen vorgestellt. Dazu gehören u.a. die gesparten Emissionen, die aktuelle umweltfreundliche Energieproduktion im Gewerbegebiet, die prozentuale Anzahl der teilnehmenden Betriebe und die aktuellen Emissionen mittels einer Tachoanzeige. Auf der rechten Seite des Dashboards werden die Top-Ökobetriebe, aktuelle Ankündigungen und die Kontaktinformationen des Gewerbegebiets angezeigt. Ein grüner Button, das zum Mitmachen animieren soll, ist ebenso vorhanden. Möchte der Benutzer mehr Informationen zu der jeweiligen Grafik erhalten, so kann dieser mit dem „Mehr Dazu“-Button zu den Details wechseln.

Abbildung 3 zeigt einen Auszug aus der Oberfläche, welche die detaillierten grafischen Informationen zur umweltfreundlichen Energieproduktion im Gewerbegebiet anzeigt. Hier werden deutlich mehr Grafiken dargestellt. Ebenso sind Filtermöglichkeiten vorhanden. Außerdem kann sich hier der Benutzer die Produktion der einzelnen Anlagen, aber auch die der gesamten Produktion anzeigen lassen.

Eine Heatmap, die die „grünen“ Bereiche (umweltfreundliche Produktion) und die „roten“ Bereiche (konventioneller Strom) des Industriegebiets darstellt, ist ebenso vorhanden. Mittels eines Sankeydiagramms ist es möglich, den Energiefluss des gekauften und produzierten Stroms in die Betriebe zu verfolgen. Eine Rankingtabelle, die die zehn umweltfreundlichsten Betriebe darstellt, wurde auch bereitgestellt. Damit wird das Ziel verfolgt, Betriebe mit vielen umweltfreundlichen Aktionen zu loben und folglich ein besseres Image zu verschaffen.



Abb. 3: Detaillierte visuelle Informationen zur umweltfreundlichen Energieproduktion

Dies soll andere Betriebe motivieren, mehr umweltfreundliche Aktionen zu tätigen und an der grünen Produktion teilzunehmen. Zudem sollen mittels unterschiedlichen Grafiken Informationen zu den Emissionen präsentiert werden. Vergleiche, Livedaten und/oder ein Tacho mit Warnungsfunktion bei zu hohen Emissionen sollen hier untergebracht werden.

4 Implementierung

Die Oberfläche soll in naher Zukunft in die Webseite des Unternehmensnetzwerks integriert werden. Daher ist es von Vorteil, wenn die Oberfläche so wenig wie möglich Abhängigkeiten aufweist. Frameworks werden daher nicht verwendet. Da keine Datenbank existiert, aus denen die Daten abgefragt werden können, ist auch kein Datenbankmanagementsystem vorhanden.

Die Oberfläche wurde mit reiner HTML, CSS und JavaScript entwickelt. Das weitverbreitete Frontend-Framework Bootstrap wurde verwendet, um auf das einfache Grid-System zuzugreifen. Für die einzelnen Grafiken findet Google Visualization API (Google Charts) Anwendung. Zudem wurde Simple PHP involviert, um in allen Bereichen der Webseiten vorkommenden Code (Header, Navigation etc.) einfach einzubinden. Letztendlich ist die Oberfläche einfach und schnell in jede technische Infrastruktur zu integrieren.

5 Fazit und Ausblick

Dieses Paper handelt von einem Konzept, welches die Visualisierung der umweltfreundlichen Produktion von Energie in einem Industriegebiet vorstellt. Sie soll andere Betriebe und Industriegebiete dazu motivieren, ihre Produktionen mittels umweltfreundlicher Energie durchzuführen. Außerdem soll damit ein nachhaltiges und effizientes Bild der Industrie in der Stadt hervorgehoben werden.

Die Oberfläche wurde klar strukturiert und verständlich gestaltet, so dass die Bedienung einfach ist und den gesetzten Zielen gerecht wird. Wichtige Informationen wurden besonders kenntlich gemacht und die Option zur Detailansicht wurde gegeben. Zudem wurden diverse Filtermöglichkeiten geboten. Die Gemeinsamkeiten und Vorteile zahlreicher Energiedashboards wurden bei der Entwicklung mitbeachtet. Mit dieser Arbeit wird ein Weg eingeleitet, der es den Industriegebieten ermöglicht, umweltfreundliche Energieproduktion der Öffentlichkeit zu repräsentieren.

In der Zukunft könnte die Oberfläche in mehrere Sprachen übersetzt werden, so dass sie weltweit begutachtet werden kann. Außerdem könnte das Dashboard zu einer Pinnwand umgebaut werden, so dass jeder Benutzer über die gestalterische Konfiguration selbst entscheiden kann. Ein Benutzerkonto für jedes Unternehmen hätte den Vorteil, dass diese die Grafiken nur mit ihren Daten befüllen lassen können. So würde jedes Unternehmen selbst sehen, wo es noch Sparpotential hat. Das aktuelle Dashboard ist nur für die Stromproduktion ausgelegt. Gleichfalls wäre eine Funktion vorstellbar, die zwischen Strom- und Wärmeproduktion unterscheidet und entsprechend der Auswahl die jeweiligen Grafiken erzeugt.

Danksagung

Das Unternehmensnetzwerk wird für die Umsetzung der in der Zielsetzung beschriebenen Projektes durch die Senatsverwaltung für Wirtschaft, Energie und Betriebe in Berlin im Rahmen der Umsetzung des Masterplan Industriestadt Berlin 2018-2021 gefördert.

6 Literaturverzeichnis

- [Ag21] Agorameter – Agora Energiewende, www.agora-energiewende.de/service/agorameter/chart/power_generation/30.04.2021/03.05.2021/, 03.05.2021
- [BZ18a] Bogenstahl, C; Zinke G.: Der Industrieakzeptanzindikator – Messung der Akzeptanz von Industrie in Deutschland, vgl. S.2, iit – Institut für Innovation und Technik in der VDI/VDE Innovation + Technik GmbH, 2018
- [BZ18b] Bogenstahl, C; Zinke G.: Der Industrieakzeptanzindikator – Messung der Akzeptanz von Industrie in Deutschland, vgl. S.8, iit – Institut für Innovation und Technik in der VDI/VDE Innovation + Technik GmbH, 2018

- [Ch06] Chertow, M.: *Industrial Symbiosis: Literature and taxonomy*. Annual Review of Energy and The Environment 25/00, S. 313-337, 2000.
- [Ei21] Smart Grid Dashboard – Eirgrid Group, www.smartgriddashboard.com, 03.05.2021
- [Em21] Emissionsquellen - Umweltbundesamt, www.umweltbundesamt.de/themen/klima-energie/treibhausgas-emissionen/emissionsquellen, 28.04.2021
- [Ev17] Evergreen, S.D.H.: *Effective Data Visualization: The Right Chart for the Right Data*, vgl. S.409f., SAGE Publications, 2017.
- [FI21] Energy Charts – Frauenhofer - Institut, energy-charts.info, 03.05.2021
- [Gk21] Grünes Kraftwerk – Frankfurter Osten nachhaltiges Gewerbegebiet, www.frankfurter-osten.de/kraftwerk/, 28.04.2021
- [Gl03] Glassner, A.: *Ven and Now*. IEEE Computer Society Press 23/03, S. 82-95, 2003.
- [Is14] Isenmann, Ralf.: *Industriesymbiosen: Kooperationen auf Wegwerfbasis. Ökologisches Wirtschaften - Fachzeitschrift* 29/14, S. 28-29, 2014.
- [Ka18] Kapadiya, A.A.: *Real-time User Query Analytics Using Sunburst Charts With D3.js*. SSRN Electronic Journal, 2018.
- [RT21] France's electric consumption data – Réseau de Transport d'Electricité, www.rte-france.com/en/eco2mix, 03.05.2021
- [Ru15a] Runkler, T.A.: *Data Mining: Modelle und Algorithmen intelligenter Datenanalyse*, vgl. S.37, Springer Vieweg, 2015.
- [Sc94] Schwarz, E.J.: *Unternehmensnetzwerke im Recycling-Bereich*, vgl. S.133 ff., Deutscher Universitätsverlag, 1994.
- [Se11] Seifert, J.W.: *Visualisieren Präsentieren Moderieren: Der Klassiker*, S.11, Gabal Verlag, 2011.
- [Ti06] Tintarev, N.; Rostami S.; Smyth B.: *Hierarchical Edge Bundles: Visualization of Adjacency Relations in Hierarchical Data*. IEEE Transactions on visualization and computer graphics, 12/06, S. 741-748, 2006.
- [Un21] Über uns – UnternehmensNetzwerk Motzener Straße, www.smartgriddashboard.com/, 03.05.2021
- [VM112] von Hauff, M.; Müller-Christ, G; Isenmann, R: *Industrial Ecology Management: Nachhaltige Entwicklung durch Unternehmensverbände*. Vgl. S. 59f, Springer Gabler, Wiesbaden, 2012.
- [Wi20] Wirtschaftsstandort Motzener Straße – UnternehmensNetzwerk Motzener Straße, www.motzener-strasse.de/wp-content/uploads/2019/07/Imagebroschuere-Netzwerk-Motzener-Str.-2019-web-3.pdf, 30.04.2021
- [WGK15a] Ward, M.; Grinstein, G; Keim, D.: *Interactive Data Visualization: Foundations, Techniques, and Applications*, vgl. S.1, Second Edition, CRC Press, 2015.

Can smartphone generated cycling data contribute to the improvement of the bicycle infrastructure?

Potentials of a gamification app to support decision processes in a city

Tobias Höper¹, Johannes Schering² and Jorge Marx Gómez³

Abstract: Smartphone generated data could make an important contribution to learn more about cycling behaviors to support the decision making process. As part of the *Oldenburg Bicycle Challenge* a huge amount of app based data was gathered in the German Oldenburg region. In combination with weather and counting data new knowledge on bicycle use can be generated. Parameters as distances and durations of the trips, hourly values, the distribution of the trips to recreational and utilitarian purposes, brakings and waiting times are analyzed and discussed as part of this work. GPS data of cyclists is not only gathered by smartphone apps but also by bike sensors that are attached to the bicycle as many projects has shown (e.g. *ECOSense*). As part of the conclusion, the advantages and disadvantages of both approaches regarding acquisition of participants, effort for data collection, data quality, potentials for cycling data analysis etc. will be discussed.

Keywords: Bicycle Application, gamification, bicycle data, smart cycling

1 Introduction and Motivation

Cycling gains in importance in the mobility behavior of people in Germany and Europe. Between 2002 and 2017, the amount of inner German bicycle traffic volume has increased for more than 13 percent. The bicycle traffic performance has increased up to 37 percent at the same time [Mi19]. The Corona pandemic leads to a further boom and a growing interest in cycling activities. A sales increase in the German bicycle industry of about 60.9 percent from 2019 to 2020 can be perceived [ZIV21]. In 2020, about one million more bicycles were sold compared to the year before what is an increase of about 35 percent. People has spent nearly 1.000€ per bicycle what is more than doubled compared to 2010 (460€) [Th21]. At the same time the German government shows further ambitions in the promotion of cycling. New objectives were set as part of the recently published *National Cycling Plan 3.0* (Nationaler Radverkehrsplan NRVP 3.0). Germany is aiming at to

¹ University of Oldenburg, Business Informatics (VLBA), Ammerländer Heerstraße, 26129 Oldenburg, tobias.hoep@uni-oldenburg.de

² University of Oldenburg, Business Informatics (VLBA), Ammerländer Heerstraße, 26129 Oldenburg, johannes.schering@uni-oldenburg.de

³ University of Oldenburg, Business Informatics (VLBA), Ammerländer Heerstraße, 26129 Oldenburg, jorge.marx.gomez@uni-oldenburg.de

become a cycling country. Other objectives are that cycling kilometers are doubling until 2030, to motivate more people to decide to use the bicycle because of the growing attractiveness and a safe infrastructure with a decreasing number of killed cyclists in traffic accidents [Bm21]. To make cycling more attractive and to supply the growing numbers of cyclists an appropriate bicycle infrastructure, an efficient and target oriented planning process is necessary. Cycling data that is gathered by smartphones could make an important contribution to support decision making processes in this context [Sc14].

2 Research Questions and Procedure

An actual running project that is based on smartphone generated tracking data is the *Oldenburg Bicycle Challenge*⁴. Based on a gamification approach the participants within a radius of 25 kilometers around Oldenburg/Germany have the chance to win prizes and to compete with other cyclists (e.g. colleagues as part of an inner company challenge) for recording their bicycle routes by an app. A huge part of the data set will be further analyzed and presented as part of this research. The app data was further enriched with weather data and data of the stationary bicycle counters of the city of Oldenburg with a following interpretation. In a next step, the analysis of the app data will be evaluated and compared to the data collection by mobile bicycle sensors and the relating data analysis part as the research project *ECOSense*⁵ did. This research will answer the following two research questions:

- Which new knowledge can be gathered regarding bicycle use by the analysis of the GPS data of a bicycle app in combination with weather and stationary counting data?
- Which potentials and challenges show the gathering of cycling data by an app compared to a mobile sensor that is attached to the bicycle?

To answer the research questions an explorative data analysis of the historical bicycle data from 2020 and 2021 was conducted based on the process model *Standard Process for Data Mining* (CRISP-DM). The preprocessing includes the removal of measuring errors, inaccurate GPS measurements with bad Delution of Precision (DOP)⁶ values and bicycle trips with too high average speed levels (above 45 km/h). In a next step for each single trip general attributes as total distances or durations were calculated. The number and coordinates of brakings and waiting times were determined whenever possible. In addition, it was identified whether a route had the same start and end point and which share the routes had within the city boundaries of Oldenburg. A methodology was

⁴ <https://www.ol-fahrradchallenge.com/>

⁵ <https://ecosense.mein-dienstrad.de/>

⁶ The DOP indicates how well a given satellite constellation is suited for measuring position at a given location.

developed to distinguish between leisure and everyday routes in the recorded routes. Thus, a route needed to fulfill the following characteristics to be classified as a leisure trip:

- Trip is outside the rush hour (Monday - Thursday: 04:00 - 09:00 as well as 15:00 - 17:00, Friday: 04:00 - 09:00 as well as 12:00 - 17:00) and is at least 5 km long
- Trip is on weekend (Saturday with min. 5 km length or Sunday regardless of the length)
- Trip is on a holiday
- Trip is a round trip, the length of which exceeds the average trip length of 8.7 km

A round trip is identified when a route has the same start and end point. The distance should be above the average trip length represented by the data, since even short errand trips, e.g. to a supermarket, can have the same start and end point. The bicycle trips were matched to the appropriate municipal counting stations mentioned above by a tolerance range of 15 meters distance to the counter. In a last step a final data set with all describing attributes was generated and aggregated with the weather data. The bicycle trips were visualized on heatmaps regarding route choice, speed levels and brakings. Fig. 1 shows a heatmap of points in the Oldenburg municipal area with decreased speed levels. Further visualizations such as pie or bar charts were created for the purpose of investigating distributions and correlations.



Fig. 1. Points in the Oldenburg city area with decreased speed levels.

3 Findings

3.1 Data Analysis

The results of the data analysis show that the distances of the bicycle trips of the bicycle app users are relatively long (8.7 km) compared to mobile sensor systems as *ECOSense* (around 3- 3.5 km per trip, depending on the weather situation). The duration of the app users is about 32 minutes (*ECOSense*: around 15:16 – 18:09 min). Fig. 2 shows the amount of trips per hour of the start of the trip. Many people using the app in the morning hours around 7 am (9.05%) and in the afternoon hours around 3 pm (9.14%) and 4 pm (9.58%). The cycling distances at weekends tend to be longer (12.2 km) compared to working days (8 km). At official holidays the distances are even increasing compared to the weekends (16.4 km).

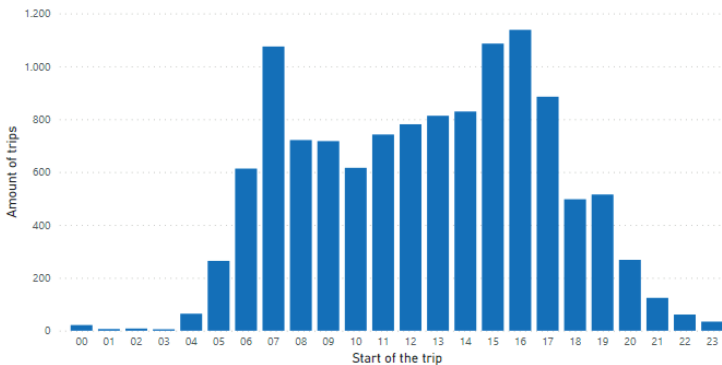


Fig. 2: Amount of bicycle trips per hour of the start of the trip.

Bicycle trips were divided into leisure and everyday trips using the developed methodology of their route characteristics. Thus, it was possible to determine that 40 percent of the trips could be leisure trips, accounting for 62.5 percent of the kilometers traveled. Leisure trips were significantly longer on average at 15 km than everyday trips at 5.3 km but had only insignificantly higher average speeds. In the period of the first three of the six analyzed challenges for which data on age and gender were available, around 30 percent of the participants are between 26 and 35, around 20 percent are between 46 and 55 years old. The average age of over 43 years was significantly higher than the age average recorded in similar studies with smartphone-generated bicycle data [FBL18][Ch 11]. Male participants are slightly overrepresented (56.1 percent) and cycled 63.26 percent of all kilometers of the bike challenges. Bad weather conditions may have an influence on the distances and durations. Bicycle trips under rainy conditions tend to be shorter (7.9 km) compared to moderate weather conditions (9 km). Fig. 3 shows the average trip lengths under different precipitation intensities, categorized according to the precipitation intensity ranges for rain of the German Weather Service (DWD) [Dw21]. The heavier the

precipitation, the shorter the recorded cycling distances. Thus, the distances are shortened by 25 percent with heavy precipitation of 10-50 mm. The average trip length reduces to 6.75 km and the trip duration 24:52 min. It seems to be that wind and temperature have no remarkable influence on bicycle use of the app users.

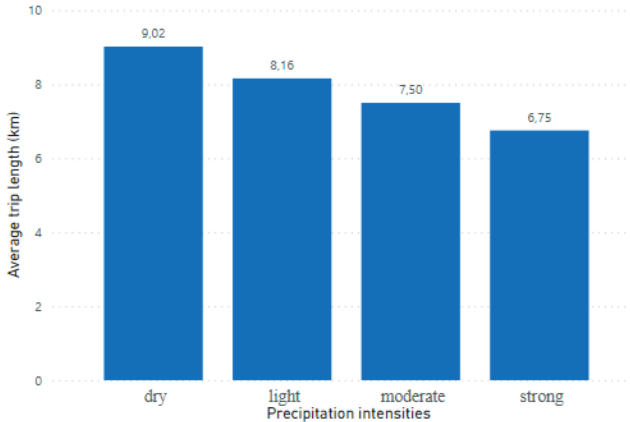


Fig 3. Average trip lengths under different precipitation intensities

On average, the app users braked 3-4 times per kilometer on their ride. At night between 00:00 and 06:00, there were fewer braking events than during the rest of the day. The waiting times as identified by the applied methodology assumed a share in a bicycle trip of about 10 percent. During the day between 07:00 and 18:00, the waiting times for cyclists tend to be longer. When looking at the traffic volume of the bicycle app trips assigned to the bicycle counters, compared to the actual traffic volume measured at the counters, it became apparent that the relative distribution of the number of trips over the day had similarities. But the distribution often differed more strongly and the peaks in traffic volume were at different times. Thus, the bicycle app data cannot accurately represent the prevailing traffic volume. Only 20 percent of all routes recorded by the app could be assigned to a counting station at all. This shows that stationary measuring stations alone are not sufficient to measure the traffic volume.

3.2 Comparison of smartphone-generated data vs. sensor-collected data

The recording of smartphone-generated data usually requires significantly less effort than data recording with a bike sensor. There are numerous tracking apps available in the app stores (e.g. Strava⁷, Bike Citizens⁸, Naviki⁹) but the conditions for collecting and transmitting the data must be clarified in advance with the app provider and potential

⁷ <https://www.strava.com/?hl=de>

⁸ <https://www.bikecitizens.net/de/>

⁹ <https://www.naviki.org/de/naviki/>

participants. For *ECOSense*'s sensors the entire sensor system, including storage processes and encryption, had to be developed. This also enabled sources of error. Only about 13% of all trips could be evaluated in the first round of data collection [Ec21]. Nevertheless, the final product provides reliable GPS data. In addition, only small usage costs were incurred for the use of the app and the provision of support, while *ECOSense* initially required €92.899,10 in funding to develop the sensors [Ec21]. Redeveloping a bicycle tracking app can be also expensive. It should be noted that *ECOSense* was a feasibility study with a first prototype and the costs can be significantly reduced in case of mass production. To participate in the *ECOSense* project, participants first had to register and were then selected based on their age, gender, and zip code to form a group of participants representative of the city of Oldenburg [Ec21]. The sensors had to be installed and removed from each bicycle at the start and end of the data recording. For smartphone generated data, only a registration was required before data could be generated immediately after installation of the app, resulting in a significantly lower inhibition to participate. However, a selection of participants in the *ECOSense* project offers the advantage of forming a supposedly representative group of participants before participation. When recording data via a smartphone app, there is a risk that the results are less representative due to an overrepresented group of participants, as already pointed out in related studies [FBL18]. In addition, these apps are often only used by a small number of users [FL17]. Even selected participant groups do not necessarily guarantee more representative data. While tracking via smartphone app in the case of the *Oldenburg Bike Challenge* must be started or stopped manually before and after each ride, which can sometimes be forgotten by participants, the sensors permanently record distances travelled by bike. However, by obtaining the appropriate permissions, automated recording of routes would also be possible, see Google Maps. A clear disadvantage of smartphone-generated data was the data quality (partially low measurement frequencies which depend on the type of smartphone and may vary). For a detailed analysis, a measurement frequency of 1 Hz would be desirable. In the sensor approach, the measurement frequency was 24 Hz (acceleration, magnetometer, gyroscope) and 1 Hz (all other sensors) and could be adapted to changing requirements depending on the type of the respective sensors. In addition, only GPS data could be smartphone-generated, whereas the *ECOSense* project also measured position and acceleration, as well as air pressure, temperature and other environmental parameters. In addition, the prototype could be extended by further sensors.

Overall, both approaches have advantages as well as disadvantages. To simply recording bicycle routes, smartphone-generated data represents the simpler approach with considerably less effort, while the approach of sensor-collected data offers opportunities in the gathering of additional information, for example about the state of bike paths.

4 Conclusion

Within the scope of this work, comprehensive findings regarding the cycling behavior of the app users could be obtained. The data does not represent the everyday cycling situation on the cycle paths since the gamification approach of the Challenge motivates many recreational cyclists with longer journeys to record their routes. Nevertheless, this data should not be disregarded for future traffic planning, as conspicuities in speed decreases or braking hotspots can indicate lacks in the bicycle infrastructure. Overall, it can be concluded that the analysis of cycling data is not a solution for the automated detection of weak points in the infrastructure but can be used as support in the decision process. The results of the data analysis provide orientation for traffic planners who always need to consider the situation at corresponding locations in detail and then initiate concrete actions. Nonetheless, further incentives need to be created to motivate more people to record their - both utilitarian and recreational – bicycle trips.

The Oldenburg Bicycle Challenge is part of the European research project *Bicycles and Intelligent Transport Systems* (BITS)¹⁰. The project is funded since 2019 by the Interreg VB program in the North Sea Region. The main objective is to build up a European open cycling data initiative which includes the testing and implementation of diverse ITS solutions. With the *CyclingDataHub*¹¹ a new open data portal for cycling data will be implemented by the Province of Antwerp. As part of the project, the VLBA department developed the *Bicycle Data*¹² website that makes European cycling data and relevant Key Performance Indicators (KPIs) of the gamification app and other ITS solutions available. The project enables the collaboration between world leading countries in cycling promotion (NL, DK, BE) and neighboring countries that want to learn and to advance in the field of smart cycling applications (DE, UK). BITS is also aiming at a reduction of CO₂ emissions of nine percent and the increase of bicycle use in the North Sea Region of ten percent.

Acknowledgements

The Oldenburg Bicycle Challenge is part of the project BITS (Bicycles and Intelligent Transport Systems) which is supported by the Interreg North Sea Programme (Priority 4, Promoting green transport and mobility) of the European Regional Development Fund of the European Union.

Disclaimer:

This paper reflects only the author's view and the Interreg North Sea Region is not responsible for any use that may be made of the information it contains.

¹⁰ <https://northsearegion.eu/bits/>

¹¹ <https://cycle-data-hub-provincieantwerpen.hub.arcgis.com/>

¹² www.bicycle-data.de



European Regional Development Fund



EUROPEAN UNION



5 Literature

- [Bm21] BMVI, Bundesministerium für Verkehr und digitale Infrastruktur: Nationaler Radverkehrsplan 3.0. Fahrradland Deutschland 2030. 2021.
- [Ch11] Charlton, B., Hood, J., Sall, E., & Schwartz, M.; Bicycle Route Choice Data Collection using GPS-Enabled Smartphones. In Transportation Research Board 90th Annual Meeting. 2011.
- [Dw21] DWD, Deutscher Wetterdienst, <https://rcccm.dwd.de/DE/service/lexikon/Functions/glossar.html?lv2=101812&lv3=101906>, accessed: 21/07/2021.
- [Ec21] ECOSense: <https://ecosense.mein-dienstrad.de/>. accessed: 21/07/2021.
- [FBL18] Francke, A. Becker, T.; Lissner, S.: Big Data im Radverkehr. Ergebnisbericht „Mit Smartphones generierte Verhaltensdaten im Radverkehr“, TU Dresden, Professur für Verkehrsökologie und Professur für Verkehrspsychologie, Dresden, 2018.
- [FL17] Francke, A., Lissner, S.: Big data in bicycle traffic. Technical Paper BMVI. 2017.
- [Mi19] MiD, Mobilität in Deutschland: Analysen zum Radverkehr und Fußverkehr. Study of infas, DLR, IVT and infas 360 on behalf of BMVI. (FE-Nr. 70.905/15). Bonn, Berlin. 2019.
- [Sc14] Schelewsky, M., Jonuschat, H., Bock, B., Stephan, K.: Smartphones unterstützen die Mobilitätsforschung. Springer Fachmedien. Wiesbaden. 2014.

- [Th21] Thureau, J., Deutsche Welle DW. World Bicycle Day: Germany sees COVID bike boom. <https://www.dw.com/en/world-bicycle-day-germany-sees-covid-bike-boom/a-57762676>, accessed: 25/07/2021.
- [ZIV21] ZIV, Zweirad-Industrie-Verband, https://www.zivzweirad.de/fileadmin/redakteure/Downloads/Marktdaten/PM_2021_10.03._Fahrrad-_und_EBike_Markt_2020.pdf, accessed: 03/05/2021.

“Civic Energy Informer” - Facilitating assistance for actors in civic energy through a modular digital platform

Antonia Krebs¹, Björn Koch¹ and Thomas Klenke¹

Abstract: Although civil society is a key driver for the energy transition, civic energy initiatives remain a niche competitor within the energy market. Facilitating the active involvement of citizens and communities, and subsequently the uptake of civic energy initiatives, requires a profound source of available information on the underlying processes, benefits, and possible risks such initiatives encompass. Data collected on practical experience from over 20 diverse civic energy pilots in seven European states over the past five years have highlighted the complexity of information demand the civic energy process entails. Due to this diversity and the distinctiveness of each pilot, this paper reports on the development of the “Civic Energy Informer”, a modular digital platform that covers these diverse information demands in order to offer specific advice for various stakeholders. This presents a foundation for promoting the facilitation of civic energy initiatives by providing well-founded insights and tools for potential adopters.

1 Introduction

While it has been frequently established that civil society is a key driver for the energy transition, civic energy initiatives remain a niche competitor within the European energy market [BS18]. Nonetheless, civic energy and its related, synonymously used concepts of ‘community energy’ or ‘local energy’ have become a noteworthy aspect of the energy landscape, representing a shift towards a decentralised and democratic renewables-based energy system [VS15, Mc19, Va16]. This shift is epitomised by the European Commission’s ‘Clean Energy for all Europeans Package’ and the subsequent recast of the Renewable Energy Directive and the Electricity Directive [Cl21]. These present an opportunity to further secure the role of civic energy by including definitions of community energy, alongside the entailed rights and the need to address existing barriers for a level playing field within the energy market [Di21a, Di21b]. While the practical application of these directives’ national transpositions is debatable [FM21], it could present an opportunity for further enshrining civic energy initiatives in regional energy strategies.

Although the opportunities that civic energy offers to the realisation of a successful inclusive and just energy transition [AB14, Ro00], defining and thus actively facilitating

¹ A. Krebs · B. Koch · T. Klenke
Carl von Ossietzky Universität Oldenburg, COAST
Ammerländer Heerstr. 114-118, 26129 Oldenburg, Germany
E-Mail: antonia.krebs@uol.de; bjoern.koch@uol.de; thomas.klenke@uol.de

civic energy poses a difficult task due to the inherently complex and heterogenous nature of civic energy initiatives [Ba16, DS18]. This complexity is highlighted by the European Economic and Social Committee, who outlines civic energy as “decentralized renewable energy generation owned (at least 50%) or operated by citizens, local authorities, charities, NGOs, farmers, cooperatives, or SMEs, creating a stream of local value that can stay within the region” [Eu15]. To enable the uptake of civic energy in the European energy landscape therefore requires a system that takes into consideration the broad variety of involved stakeholders, business models, energy systems, as well as the benefits such initiatives can deliver to the involved local communities. Facilitating the active involvement of citizens and communities, and subsequently the uptake of civic energy initiatives, thus necessitates a profound source of available information on the underlying processes, benefits, as well as possible risks such initiatives encompass.

On this basis, McGovern et al. [Mc19] investigated the need for a digital transformation that fosters civic energy uptake by offering a guided process management approach that supports not only certain aspects regarding the market roll out of established civic energy initiatives but also offers practice-based advice for every stage of the ‘civic energy cycle’ (CEC) [MK18]. Making information and communication about civic energy initiatives readily available for a heterogenous group of citizens and governments through digital means could therefore allow for enhanced participation and exchange that can foster acceptance and subsequently accelerate an inclusive and just energy transition [Mc19, SGS18].

Building on this potential for fostering civic energy initiatives through digitalisation on the one hand, and the identified demand for a variety of information and advice that derived from the reported practical experience of numerous distinct civic energy pilot projects on the other hand, this paper reports on the development of a modular digital platform that aims to facilitate assistance for civic energy initiatives. Based on this premise, a digital “Civic Energy Informer” platform is presented that offers initial advice for the diverse stakeholders involved in civic energy and caters to the specific information demands that each phase of the CEC entails. The objective is to create a foundation for promoting the facilitation of civic energy initiatives by providing a substantiated digital information source as an incentive for advancing the active engagement of existing and potential future adopters.

2 Motivation

Framing civic energy as a continuous improvement process, which captures aspects of civic energy initiatives that go beyond techno-economic processes, as proposed by McGovern and Klenke [MK18], allows to streamline heterogenous, multi-stakeholder initiatives and thus advances the facilitation of civic energy uptake. This continuous improvement process is reflected in the CEC, which categorises key assignments involved in the development of civic energy initiatives into four main phases with manageable

corresponding process stages [MK18]. Additionally, civic energy is characterised by its heterogeneous forms of organisational-, legal-, and business models, as well as by its variety of stakeholders involved, the intended community participation and benefits, its possible ‘activities’ (e.g. energy generation, supply, distribution, and/or consumption, etc.), and not to mention the different forms of energy technologies (e.g. solar, wind, heat networks, or e-mobility, etc.) [CU20, WD08]. Hence, civic energy initiatives can vary strongly between these characteristics and the underlying motivation of engaging in civic energy [AB14, Ra14], e.g. economic incentives, environmental and climate protection, or social cohesion. Consequently, each civic energy initiative in itself faces a myriad of different opportunities and barriers [Wa08, Ko16, Ro08], making the development of a universal support and advice structure challenging - if not impossible.

Due to this variety and complexity of civic energy, using a “toolbox” approach could cover essential information needs of multi-stakeholder partnerships for relevant processes in the CEC [Mc19]. Yet, the available tools predominately accommodate the technical requirements of energy production and delivery – to date, a digital toolbox framework that caters to all phases of the CEC is not readily available. As McGovern et al. [Mc19] established, the question when developing a digital toolbox for civic energy “[...] should focus on ‘what’ to support for ‘whom’ rather than ‘how’ or ‘by which means’” (p.686). In order to address the multifaceted dynamic of civic energy initiatives, a modular digital platform is deemed as the most appropriate approach to address the diverse information demands such initiatives entail. This approach allows to align a toolbox framework with the heterogeneity of civic energy stakeholders and the myriad of characteristics named above, which require distinct information to address specific barriers and opportunities. Furthermore, the creation of a modular digital platform addressing the entire CEC would further allow relevant intersection of information for several different factors of civic energy initiatives, e.g. the ‘who’/‘for whom’, and ‘what’.

Building on this premise, this paper reports on the data collected in the Interreg North Sea Region Programme ‘COBEN’ project [In21] on practical experience from over 20 diverse civic energy pilots in seven European states (i.e. the Netherlands, U.K.-Scotland, Denmark, Germany, Belgium, Norway, and Sweden) over the past five years. Existing data from surveys on business models, risk management, procurement guidelines and legal frameworks have highlighted the complexity of information demand that the civic energy process entails. This complexity and the distinctiveness of each civic energy initiative – covering a range of different pilots, e.g. solar cooperatives, district heating networks, e-mobility services, and wind energy landscapes – forms the basis of the modular digital platform presented in this paper. Based on the feedback on the experiences in the pilot projects, the identified information demands were then structured into relevant modules to align with the CEC.

3 Digital Platform - Criteria and Design

As mentioned above a Civic Energy Informer should be built as a toolbox framework where only selected tools will be available. While a perfect and universal Civic Energy Informer would identify and address all digitalisation needs of any possible civic energy initiative, a real-world Civic Energy Informer would have to focus on a variety of specific initiatives and the tools available. Its main task is to collect all needed tools based on this focus as well as to implement the underlying logic and information flow. This way an Informer can guide the user through the CEC and its decision by collecting the data needed for the tools at a specific point, feeding these tools with the data they need, and presenting the results back to the user. These results should be added to the set of collected data (user data), so the Civic Energy Informer can reduce the amount of data to be put in by the user.

Furthermore, this toolbox can be implemented as a “smart” toolbox (a tool selection decider) which narrows down the tools to be displayed to the user based on previous data entered by the user. This can be used to guide the user through the decision processes to be taken down the line when exploring the CEC and help the user focus on the relevant information only by hiding non-relevant tools for the user’s civic energy initiative. Figure 1 shows the basic platform design of a Civic Energy Informer.

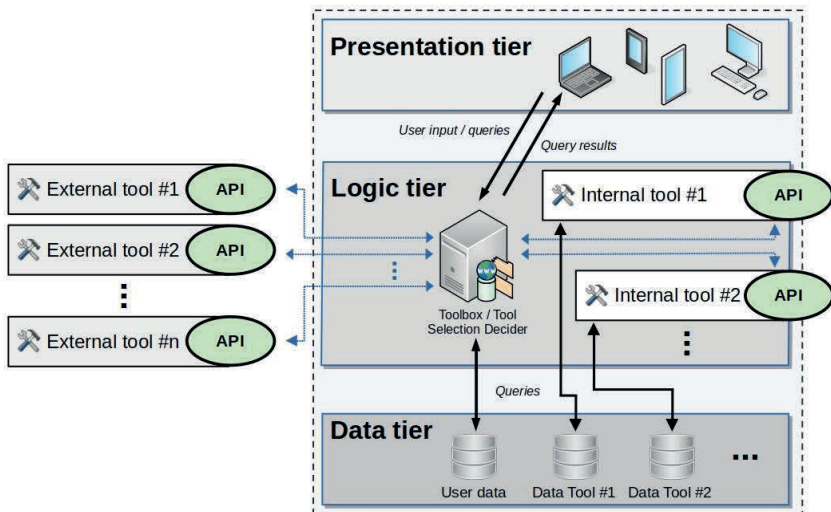


Fig. 1: Civic Energy Informer platform design adapted from the three-tier model by Pehlken, Koch and Kalverkamp, 2019 [PKK19]

The tools might be included into a Civic Energy Informer to build a monolithic and closed piece of software. But in order for the toolbox framework to develop its full potential, a

more open and distributed approach seems more promising. Outsourcing these tools as web services accessible through APIs has many advantages, like providing highly optimised tools for special requirements and conditions of an initiative as well as a larger pool of tools to choose from. In addition, it becomes easier to create your own Informer according to your needs and, above all, enables a uniform representation of the data across the entire presentation layer. Of course, the APIs of the web services need to be well documented, both technically and in terms of the focuses within the CEC. The latter in particular will be important in deciding which tools should be selected for the toolbox and presented to the user. Since the system is based on an open architecture, which prefers and focuses on the use of API calls, an integration of data from Corporate Environmental Management Information Systems into a Civic Energy Informer - and vice versa - can easily be done.

4 The “Civic Energy Informer” Platform

Within the COBEN project a prototype of a Civic Energy Informer is being developed as a proof-of-concept. The data and information collected within the COBEN project form the basis for the data set to be included into the platform and will be accessed via API calls from web services. The platform itself will be implemented as an interactive website and to avoid possible problems with the data protection regulations no user data will be stored outside the user session or on any external server.

To ensure the meeting of relevant information demand, the Civic Energy Informer is structured into relevant tools, or “modules”, to match with the CEC. The Civic Energy Informer aligns its derived modules with the four main phases of the CEC – Initiation, Planning, Roll-Out, Reflection and Adoption – as illustrated in Figure 2.

The initial input is asked to determine specific output of relevant Output section “modules”. These will be based on three segments of input data:

- Who is using the tool and inquiring information – which also implies the structure of the civic energy initiative;
- What kind of energy source is in the centre of the civic energy initiative;
- Which of the four main ‘phases’ of the civic energy cycle (CEC) is the initiative in/ would the user like to receive information about.

What modules to include in which stage is derived from the shared practical experiences of the COBEN pilots. For instance, based on the user input data, modules (i.e. coloured hexagons) for the enquired phase (e.g. Roll-out) the Civic Energy Informer will present relevant modules that aim to support the implementation of civic energy initiatives, such as risk management strategies, advice about tendering processes, and the delivery of community benefits.

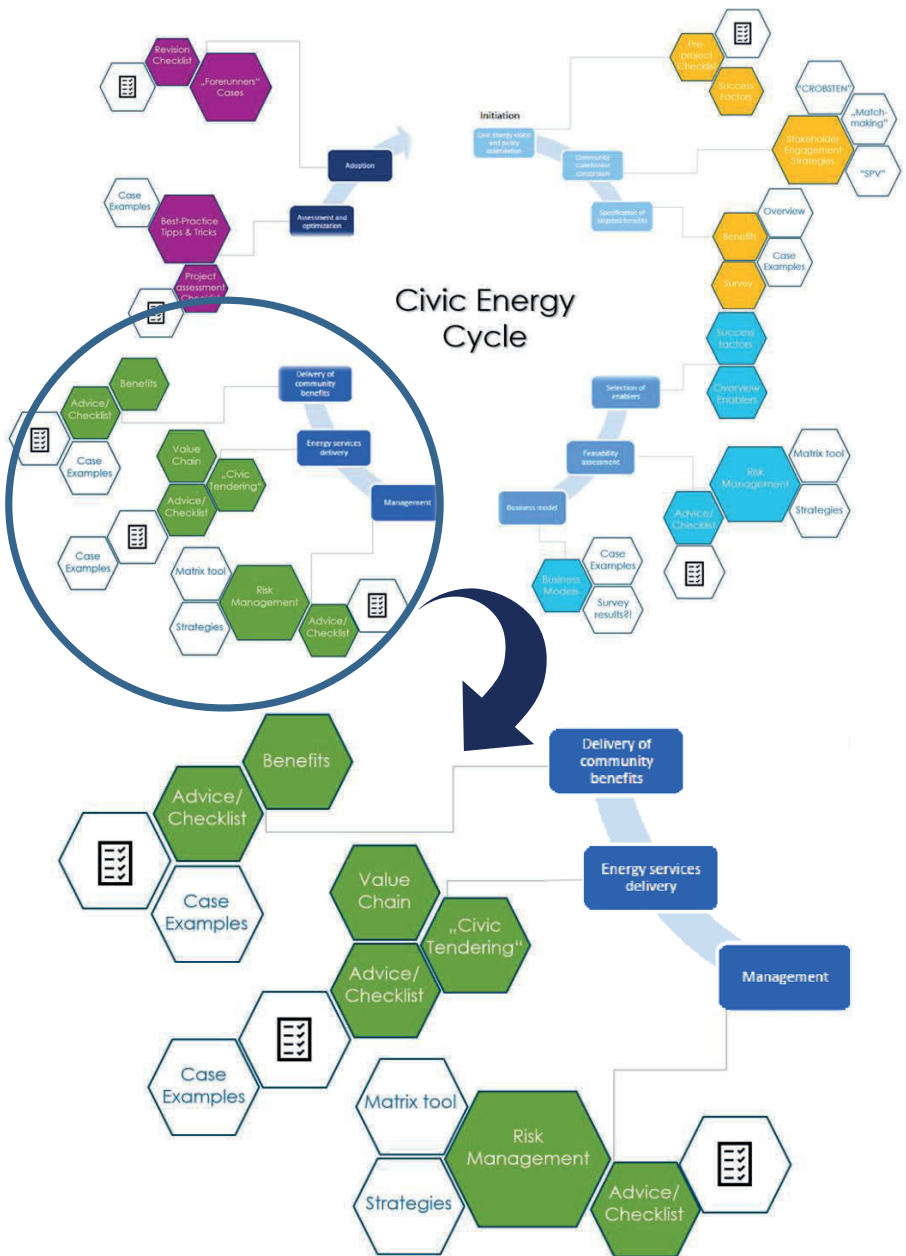


Fig. 2: Civic Energy Informer modules aligned with civic energy cycle adopted from McGovern and Klenke, 2018 [14]

On an operational level this means the information offered within the Civic Energy Informer platform is developed by active stakeholders through systematic reflection on which modules are suitable in each phase and their corresponding process stages. Hence, future operational adopters can derive information and tools, e.g. either for their own civic energy developments or for consulting purposes etc., from proven practice experiences of various operational actors, such as public authorities and businesses, which were actively involved in the COBEN project. In practice this means that, e.g. a municipality can receive information about the risks that might occur with developing a district heating network and can then access tools to support their risk management and other relevant issues that other municipalities have already identified and tackled. However, the Civic Energy Informer is not solely aimed for local authorities, but can also support citizens by offering information on, e.g. what to consider when planning on developing a cooperative solar park, or advice already active cooperatives in different civic energy process management challenges. Besides public authorities, citizens, and cooperatives, the role of the business and industry sectors is also represented in the platform, as these are a key actors in civic energy initiatives, e.g. as heat energy source for district heating networks. Hence, the Civic Energy Informer is covering most central stakeholders that are or could potentially be involved in the promotion of renewable, decentralised civic energy solutions.

5 Discussion and Conclusion

The Civic Energy Informer platform derives from the shared experiences of diverse pilot projects over an extended period of time, identifying common as well as case specific prospects and barriers throughout each phase and process stage of the CEC. In this, the platform that is presented in this paper is unique and fills in existing knowledge gaps as well as the practical demand for collective information, as McGovern et al. [Mc19] identified. By providing well-founded, experience-based insight and practical tools to cover the information demands that different actors and future adopters might encounter throughout the civic energy process, this Civic Energy Informer platform presents a foundation for promoting the facilitation of civic energy initiatives. The modular approach described above can thus offer an achievable way of delivering relevant advice and support tools to new and existing civic energy initiatives, while not neglecting the large variety of information demands and heterogeneity of the actors involved within a civic energy process.

However, since the high variety of possible information requests this platform might be facing due to the input variables, it will not be able to cater to every need the user might request. Since the scope of pilot project whose practical experience informed the development of the Civic Energy Informer was limited to the European North Sea Region and within this encompassed a diverse, yet not complete, set of civic energy pilots. Thus, one might argue that the platform cannot present appropriate advice for pilots that are effectively deviating from this scope. Furthermore, one could question the effectiveness of such information platform - that only encompasses a relatively general information

demand at the current stage - to significantly contribute power shift within the European energy structures [DS18]. Nonetheless, considering this Civic Energy Informer as an expandable knowledge base, additional information on the growing scope of civic energy pilots could be integrated and thus, progressively improve the quantity as well as quality of available information for future adopters. Hence, the Civic Energy Informer in itself could be regarded as a continuous improvement process similar to the CEC it is based on [MK18].

Another aspect to note is the omission of considering the policy factors in the platform. Policies are necessary enablers for civic energy initiatives and can – if present – enhance the capacity and facilitate the uptake of civic energy [BS18]. However, the fact that the legislative landscape in which civic energy initiatives are situated is frequently changing presents a problem for a tool such as the outlined Civic Energy Informer. Addressing the individual legislative support structures and/or barriers would be useful, yet problematic to integrate as those might vary in different EU member states as well as on the different regional and local levels [Wa08]. A possible solution could be an EU-wide implementation of legislative support structures that enable the facilitation of civic energy initiatives. While this might have been the intention with the recast of the Renewable Energy Directive and the Electricity Directive, the national transpositions of these directives differ notably between the member states and thus a fundamental, pan-European enabling structure is still needed. Additionally, if the policy context is to be considered in this digital platform, it would require the availability of long-term resources to keep the needed policy data up-to-date.

Nonetheless, the Civic Energy Informer modular platform tackles the existing knowledge demand and offers initial advice and guidance throughout the entire CEC, rather than just addressing specific isolated aspects of the civic energy process, as it has been frequently the case [Mc19]. Thus, the presented Civic Energy Informer offers a source of incentive for fostering civic energy through advancing the active engagement of interested stakeholders by providing a complex yet practical information platform that addresses the innate heterogeneity of civic energy.

Acknowledgements The case studies this paper is based on are part of the COBEN project, partly funded by the EU Interreg North Sea Region Programme, whose financial support we kindly acknowledge and would like to thank the project partners for their invaluable input and collaboration.

References


- [AB14] Arentsen, M., and Bellekom, S.: Power to the people: local energy initiatives as seedbeds of innovation?. *Energy, sustainability and society* 4/1, pp. 1-12, 2014.

- [Ba16] Bauwens, T.: Explaining the diversity of motivations behind community renewable energy. *Energy Policy* 93, pp. 235-249, 2016.
- [BS18] Burke, M. J., and Stephens, J. C.: Political power and renewable energy futures: A critical review. *Energy Research & Social Science* 35, pp. 78-93, 2018.
- [CI21] Clean energy for all Europeans package, https://ec.europa.eu/energy/topics/energy-strategy/clean-energy-all-europeans_en , accessed: 18.06.2021.
- [CU20] Caramizaru, A., and Uihlein, A.: Energy communities: an overview of energy and social innovation. EUR 30083 EN, Publications Office of the European Union, Luxembourg, 2020.
- [Di21a] Directive (EU) 2018/2001 of the European Parliament and of the Council of 11 December 2018 on the promotion of the use of energy from renewable sources, <http://data.europa.eu/eli/dir/2018/2001/oj> , accessed: 16.06.2021.
- [Di21b] Directive (EU) 2019/944 of the European Parliament and of the Council of 5 June 2019 on common rules for the internal market for electricity and amending Directive 2012/27/EU, <http://data.europa.eu/eli/dir/2019/944/oj> , accessed: 16.06.2021
- [DS18] Delina, L. L., and Sovacool, B. K.: Of temporality and plurality: An epistemic and governance agenda for accelerating just transitions for energy access and sustainable development. *Current opinion in environmental sustainability*, 34, pp. 1-6, 2018.
- [Eu15] European Economic and Social Committee: Changing the future of energy: Civil society as a main player in renewable energy generation. 2015. <https://www.eesc.europa.eu/eu-renewable-energy-directive>
- [FM21] Fleming, R. C., and Mauger, R.: Green and Just? An Update on the ‘European Green Deal’. *Journal for European Environmental & Planning Law*, 18/1-2, pp. 164-180, 2021.
- [In21] Interreg North Sea Region Programme COBEN, <https://northsearegion.eu/coben/> , accessed: 18.06.2021.
- [Ko16] Koirala, B. P., et al.: Energetic communities for community energy: A review of key issues and trends shaping integrated community energy systems. *Renewable and Sustainable Energy Reviews*, 56, pp. 722-744, 2016.
- [Mc19] McGovern, G., et al.: A Digitalization Profile for Civic Energy Processes. In (J. Marx Gómez et al. Hrsg.): *Smart Cities/Smart Regions–Technische, wirtschaftliche und gesellschaftliche Innovationen*. Springer Vieweg, pp. 683-691, 2019.
- [MK18] McGovern, G., and Klenke, T.: A process approach to mainstreaming civic energy. *Energies*, 11, 2914, 2018.
- [PKK19] Pehlken, A., Koch, B., and Kalverkamp, M.: Assessment of Reusability of used car part components with support of decision tool RAUPE. In (Pehlken, A., et al. Hrsg.): *Cascade Use in Technologies - Internationale Konferenz zur Kaskadennutzung und Kreislaufwirtschaft - Oldenburg 2018*, Springer Vieweg, pp. 75-82, 2019.
- [Ra14] Radke, J.: A closer look inside collaborative action: civic engagement and participation in community energy initiatives. *People, Place & Policy Online*, 8/3, pp. 235-248, 2014.

- [Ro00] Roseland, M.: Sustainable community development: integrating environmental, economic, and social objectives. *Progress in planning* 54/2, pp. 73-132, 2000.
- [Ro08] Rogers, J. C., et al.: Public perceptions of opportunities for community-based renewable energy projects. *Energy policy*, 36/11, pp. 4217-4226, 2008.
- [SGS18] Schwanholz, J., Graham, T., and Stoll, P. T.: *Managing Democracy in the Digital Age*. Springer, 2018.
- [Va16] Van Der Schoor, T., H. et al.: Challenging Obduracy: How Local Communities Transform the Energy System. *Energy Research and Social Science* 13/2016, pp. 94–105, 2016.
- [VS15] Van Der Schoor, T., and Scholtens, B.: Power to the people: Local community initiatives and the transition to sustainable energy. *Renewable and sustainable energy reviews* 43, pp. 666-675, 2015.
- [Wa08] Walker, G.: What Are the Barriers and Incentives for Community-Owned Means of Energy Production and Use?. *Energy Policy*, 36/12, pp. 4401–4405, 2008.
- [WD08] Walker, G., and Devine-Wright, P.: Community Renewable Energy: What Should It Mean?. *Energy Policy*, 36/2, pp. 497–500, 2008.

Einsatz von Künstlicher Intelligenz in der Digitalisierung von Abfallverbrennungskraftwerken

Ein Praxisbericht

Alexandra Pehlken ¹, Patrick Eschemann², Henriette Garmatter³, Fabian Cyris⁴,
Astrid Nieße⁵

Abstract: Die thermische Verwertung von Abfällen in Abfallverbrennungskraftwerke gestaltet sich als komplexe Aufgabe, da das vorrangige Ziel die Abfallbehandlung vor dem Ziel der Energiegewinnung steht. Insbesondere die Heterogenität des aus verschiedenen Abfällen bestehenden Brennstoffes stellt vielseitige Anforderungen an den Betrieb. Zur Optimierung der Betriebsführung von Dampferzeugern in Abfallverbrennungskraftwerken auf Basis der Maximierung des Durchsatzes, der Optimierung der Strom- und Wärmeproduktion sowie der Maximierung der Standzeit der Kraftwerkskomponenten unter Berücksichtigung der Emissionsgrenzwerte werden die im Kraftwerk erhobenen Messdaten eingesetzt sowie Bilddaten erhoben.


Diese werden aufbereitet und vortrainierten neuronalen Netzen zugeführt, mit dem Ziel verschiedene Materialien zu klassifizieren, um Informationen über die Abfallzusammensetzung und dem Heizwert zu erhalten.

Keywords: KI, Bilderkennung, Kraftwerk, Abfall, Verbrennung

1 Einleitung

Die nachhaltige Nutzung von Rohstoffen ist eine der zentralen Herausforderungen unserer Gesellschaft. Insbesondere in der Energieversorgung und in der Produktion von Verbrauchsgütern ist eine Effizienzsteigerung bei der Verwendung von Rohstoffen notwendig, um für zukünftige Generationen eine Versorgung mit endlichen Rohstoffen sicherzustellen.

In der EU-Abfallrahmenrichtlinie [Ri08] wurden 2008 die Rahmenbedingungen für die zukünftige Abfallwirtschaft festgelegt und 2012 im Kreislaufwirtschaftsgesetz [Ge12] in

¹ OFFIS – Institut für Informatik, Escherweg 2, 26121 Oldenburg, pehlken@offis.de, 
<https://orcid.org/0000-0003-1798-8679>

² OFFIS – Institut für Informatik, Escherweg 2, 26121 Oldenburg, patrick.eschemann@offis.de

³ Institut für Kraftwerkstechnik und Wärmeübertragung, Leibniz Universität Hannover, An der Universität 1, 30823 Garbsen, garmatter@ikw.uni-hannover.de

⁴ EEW Energy from Waste GmbH, Schöniger Straße 2-3, 38350 Helmstedt, Fabian.Cyris@eew-energyfromwaste.com

⁵ OFFIS – Institut für Informatik, Escherweg 2, 26121 Oldenburg, astrid.niesse@offis.de

Deutschland umgesetzt. Ziel der Richtlinie und des Gesetzes ist, die Kreislaufwirtschaft zu fördern und die Rohstoffverwertung umweltschonend zu gestalten. Aufgrund des seit 2005 geltenden Deponierungsverbotes in Deutschland für unbehandelte Siedlungsabfälle wird der größte Anteil aller anfallenden Siedlungsabfälle thermisch in einer Abfallverbrennungsanlage behandelt [Ve01]. Als übergeordnetes Ziel werden bei der thermischen Verwertung in den Abfällen enthaltene Schadstoffe reduziert und neutralisiert sowie das Volumen der Abfälle verringert. Allerdings lässt sich in einem Abfallverbrennungskraftwerk auch der positive Nebeneffekt nutzen, die in den Abfällen gebundene chemische Energie zurück zu gewinnen und als Elektrizität und Wärme der Bevölkerung und der Industrie zur Verfügung zu stellen. Neben der Volumenreduktion, die von erheblichem ökonomischem Interesse ist, stellt somit die energetische Verwertung einen weiteren gewichtigen Grund für die thermische Verwertung – auch unter der Maßgabe einer umfassenden Energieeffizienzbetrachtung – dar. Abfallverbrennungsanlagen tragen somit einen großen Beitrag zur den 17 Nachhaltigkeitszielen der Vereinten Nationen bei, da sie zum Beispiel durch Abfallbehandlung erreicht lebenswerte Städte zu schaffen. Außerdem schafft die Kreislaufwirtschaft, dessen Teil die Abfallverbrennungsanlagen sind, Einkommen und schützt die Gesundheit. Allerdings kann hier die Effizienz noch gesteigert werden.

Die Umsetzung einer effizienten thermischen Verwertung in einem Abfallverbrennungskraftwerk gestaltet sich als komplexe Aufgabe. Insbesondere die Heterogenität des aus verschiedenen Abfällen bestehenden Brennstoffes stellt vielseitige Anforderungen an den Betrieb. Die Universität Hannover und das OFFIS – Institut für Informatik Oldenburg erarbeiten gemeinsam eine Methode, um unter Einhaltung der Emissionsgrenzwerte den Durchsatz zu steigern und somit ineffiziente Anlagen zu verdrängen. Zusätzlich wird durch eine geeignete Betriebsführung die Maximierung von Wirkungsgrad und Standzeit der Bauteile angestrebt um sowohl die Nachhaltigkeit als auch die Wirtschaftlichkeit der thermischen Abfallverwertung zu steigern. Die Maximierung des Wirkungsgrades ermöglicht die optimale Nutzung der im Abfall chemisch gebundenen Energie. Die Bauteilstandzeit wird maßgeblich von Korrosion durch Bestandteile der Abfälle bestimmt. Eine Verlängerung der Bauteilstandzeit durch eine geeignete Betriebsführung reduziert ressourcenintensive Instandhaltungsmaßnahmen und steigert die Anlagenverfügbarkeit.

Zur Optimierung der Betriebsführung werden derzeit kaum Methoden aus dem Bereich der künstlichen Intelligenz angewandt. Der vorliegende Beitrag zeigt erste Ergebnisse des Forschungsprojektes Abfallverbrennungskessel 4.0⁶, welches in Zusammenarbeit mit einer realen Abfallverbrennungsanlage, der EEW Energy from Waste (EEW), durchgeführt wird. Die im laufenden Betrieb erhobenen Sensordaten werden mit Hilfe von Data-Mining und Machine-Learning Verfahren in Kombination von Bilderkennung

⁶ Gefördert durch das Bundesministerium für Wirtschaft und Energie aufgrund eines Beschlusses des Deutschen Bundestages, FKZ 03EE5038

ausgewertet und zur Vorhersage des Heizwertes und der Brenneigenschaften des Abfalls eingesetzt.

2 Stand der Technik in Abfallverbrennungskraftwerken

In Abfallverbrennungskraftwerken werden Abfälle thermisch verwertet. Das primäre Ziel der thermischen Verwertung ist die Reduzierung und Neutralisierung der enthaltenen Schadstoffe sowie die Volumenreduzierung. Zusätzlich wird in den Abfallverbrennungskraftwerken die chemisch gebundene Energie zurückgewonnen und in Elektrizität und Wärme umgewandelt [Kr17].

Die im Abfallverbrennungskraftwerk als Brennstoff eingesetzten Abfälle verfügen über sehr heterogene Brenneigenschaften [Ch11]. Die Abfalleigenschaften sind zum Zeitpunkt der Aufgabe des Abfalls in die Verbrennung unbekannt und unterliegen unvorhersehbaren Schwankungen. Die Betriebsführung eines Abfallverbrennungskraftwerks muss daher laufend auf die momentanen vorliegenden Abfalleigenschaften angepasst werden. Aufgrund des heterogenen Brennstoffes werden in Abfallverbrennungsanlagen robuste Feuerungssysteme und spezielle Abgasreinigungstechniken eingesetzt. Abbildung 1 zeigt den Aufbau einer typischen Abfallverbrennungsanlage mit Rostfeuerung in Horizontal-Bauweise. Der angelieferte Abfall wird in einem Abfallbunker zwischengelagert, wo er mit Greifern zur Homogenisierung durchmischt wird. Mit Hilfe der Greifer gelangt der Abfall über einen Aufgabetrichter auf einen Rost, wo er unter Zufuhr von Luft verbrennt. Die dabei entstehenden Rauchgase durchströmen einen Dampferzeuger und gelangen von dort in die Abgasreinigung, wo sie so gereinigt werden, dass sie die Anlage unter Einhaltung der gesetzlich festgelegten Emissionsgrenzwerte die Anlage verlassen. Die Abgaswerte werden online von der Aufsichtsbehörde überwacht. Im Dampferzeuger geben die Rauchgase Wärme an ein Wasserdampf-Gemisch ab, das in einem Wasserdampfkreislauf zur Erzeugung von Strom genutzt wird. Gleichzeitig wird im Wasserdampfkreislauf auch Energie zur Bereitstellung von Fernwärme oder Prozessdampf für Industriekunden ausgekoppelt.

Aufgrund des heterogenen Brennstoffes müssen Dampferzeuger von Abfallverbrennungskraftwerken spezielle Anforderungen erfüllen. Zum einen resultieren diese Anforderungen aus den im Rauchgas enthaltenen Schadstoffen. Einige Schadstoffe können im Dampferzeuger durch hohe Temperaturen bei gleichzeitig langsamen Rauchgastemperaturen effizient zerstört werden. Dieser Prozess ist ein zentraler Bestandteil der Abgasreinigung in einem Abfallverbrennungskraftwerk und muss somit im Dampferzeugerbetrieb jederzeit gewährleistet sein. Zum anderen führt die heterogene Zusammensetzung des Abfalls zu Schwankungen von Temperaturen, Drücken, Rauchgasmengen und der produzierten Dampfmenen. Um auf diese Schwankungen möglichst gut zu reagieren werden zahlreiche Messdaten erhoben. Diese werden teilweise für eine automatisierte Feuerleistungsregelung verwendet. Viele Messdaten können jedoch aufgrund ihrer komplexen Abhängigkeiten in der Feuerleistungsregelung nicht

berücksichtigt werden. Diese Messdaten werden verwendet, um kurzfristig manuelle Eingriffe in die Feuerleistungsregelung vorzunehmen und langfristige Analysen zur Zustandsbewertung des Dampferzeugers vorzunehmen.

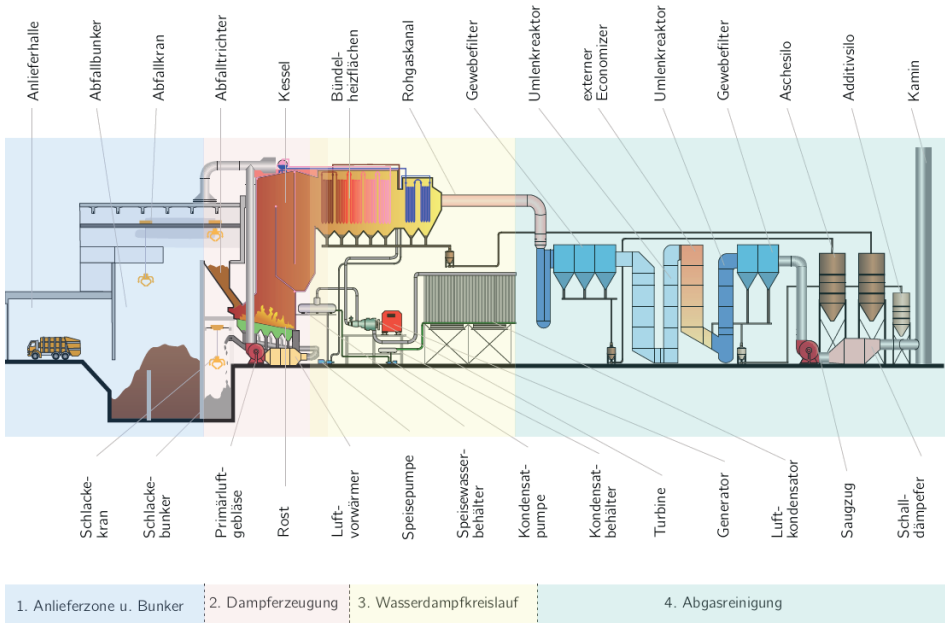


Abb. 1: Aufbau eines Abfallverbrennungskraftwerks mit Rostfeuerung in Horizontal-Bauweise [Cy21]

3 Anwendung von KI im Kontext der Bilderkennung von Abfallstoffen)

Aufgrund der hohen Heterogenität des Abfalls schwankt der zu erwartende Heizwert zwischen 8–20 MJ/kg⁷. Generell kann eine Abfallverbrennungsanlage den Verbrennungsprozess mit einem niedrigen Heizwert dauerhaft aufrechterhalten. Problematischer als zu hohe oder zu niedrige Heizwerte sind unerwartete, größere Schwankungen, die Gegenmaßnahmen nach sich ziehen. Hierbei werden durch Eingriffe in die Feuerleistungsregelung z.B. durch Änderungen der zugeführten Verbrennungsluft

⁷ Gemäß Auswertung vorliegender Sensordaten

oder durch Einbringung von Zusatzbrennstoffen (Heizöl oder Gas) die Effekte des schwankenden Heizwerts ausgeglichen.

Die Bilderkennung kann hier genutzt werden, um automatisiert und frühzeitig eine Abschätzung zum voraussichtlichen Heizwert des Abfalls zu gewinnen und die Verbrennung des Abfalls möglichst vorausschauend und effizient zu gestalten. Durch den Zeitverzug zwischen Abfallaufgabe und der eigentlichen Verbrennung vergehen durchschnittlich 87 Minuten, in denen der aufgegebenene Abfall hinsichtlich des zu erwartenden Heizwerts analysiert und der Kraftwerksbetrieb entsprechend vorausschauend angepasst werden kann.

Im nachfolgenden wird der Einsatz von Transfer-Learning und das Konzept zum Training eines eigenen, auf die Problemstellung zugeschnittenen neuronalen Netzes vorgestellt.

3.1 Transfer-Learning

Für den Ansatz des Transfer-Learnings wurde das im Jahr 2014 von Simonyan und Zisserman [SiZi14] vorgestellte vgg16-Netz verwendet. Dieses wurde mit 14 Millionen Bildern vortrainiert, um 1000 unterschiedliche Objektklassen zu erkennen [Im21]. Ein Abgleich der Liste erkennbarer Objekte mit typischerweise in Abfall vorkommenden Gegenständen zeigt u. a. folgende Übereinstimmungen auf:

- Bierflasche, Bierglas, Flasche, Autoreifen, Tetrapak, Plastiktüte und andere

Relevant ist u. a. die Erkennung von Plastikbeuteln. Bei ersten Untersuchungen zeigte sich, dass das vortrainierte Netz diese zuverlässig erkennen kann. Die Akkuratheit hängt von der benutzereingestellten Genauigkeitsschranke und der Filtergröße, mit der das Bild durchsucht wird ab. Bei einer hohen Genauigkeitsschranke, treten kaum Fehlerkennungen auf (hohe Präzision). In Vorversuchen mit einer Genauigkeitsschranke >80 zeigt sich folgende Konfusion-Matrix:

True Positives (TP): 10	False Positives (FP): 0
False Negatives (FN): 3	True Negatives (TN): 0

$$\text{Präzision} = \frac{TP}{TP+FP} = 1$$

$$\text{Trefferquote} = \frac{TP}{TP+FN} \approx \sim 0.76$$

$$\text{F1-Score} = 2 * \frac{\text{Präzision} * \text{Trefferquote}}{\text{Präzision} + \text{Trefferquote}} \approx 0.86$$

Die Vorversuche wurden auf Abfallbilder mit guter Auflösung durchgeführt, vgl. Abbildung 2.

Durch weiteres Training auf die Farbe des Beutels und des somit zu erwartenden durchschnittlichen Inhalts wären Heizwertschätzungen theoretisch denkbar. Die Erkennung ist jedoch abhängig von der Güte des Bildes und der Abfalldifferenzierbarkeit.



Abb. 2: Erkennung von Plastikbeuteln [De13]



Abb. 3: Originaldarstellung und Erkennung von Abfall im Trichterbereich [EE21]

Auf Abbildung 2 können je nach eingestellter Genauigkeit mit einer Sicherheit von >80 % Müllbeutel erkannt werden, wohingegen das vgg16-Netz auf Abbildung 3 keine sinnvollen Erkennungen ausgibt. Mit teilweisen hohen Wahrscheinlichkeiten erkannte das Netz Seekühe, Krawattenknoten oder Mäntel (F1-Score < 0.1).

Die durchgeführten Experimente weisen somit einerseits auf die Notwendigkeit besserer Videoaufnahmen im Kraftwerk, andererseits auf die Notwendigkeit eines speziell auf die Problemstellung trainierten neuronalen Netzes hin.

3.2 Konzept zum Training eines neuronalen Netzes

Grundsätzlich wird das neuronale Netz zur Regressionsanalyse verwendet, da dieses auf der letzten Ebene mit einem Neuron einen diskreten, kontinuierlichen Wert vorhersagen soll. Um ein neuronales Netz auf diese Aufgabenstellung zu trainieren, wird ein Trainingsdatensatz benötigt. Dieser muss Abfallbilder und die Zielgröße (Regressand) enthalten.

Als Regressanden eignen sich folgende zentrale Untersuchungsgrößen zum Verbrennungsvorgang:

- Emissionswerte (Quecksilber Hg, Schwefeldioxid SO₂, Kohlenstoffmonoxid CO, Kohlenstoffdioxid CO₂, Stickoxide NO_x, Ammoniak NH₃)
- Frischdampfmassenstrom in kg/s
- Heizwert in MJ/kg

Die Emissionswerte und der Frischdampfmassenstrom werden dauerhaft durch Sensoren im Dampferzeuger gemessen. Der Heizwert kann nicht direkt durch Sensoren erhoben werden, sondern wird unter Verwendung erhobener Messdaten bilanziert [Be12]. Obgleich der Heizwert für die vorausschauende Betriebsführung die interessanteste Größe darstellt, ist er aufgrund der Bilanzierung die Untersuchungsgröße mit der größten Unsicherheit. Diese Unsicherheit entsteht, da zahlreiche Messdaten in die Bilanzierung einfließen, welche jeweils mit einer Messungenauigkeit behaftet sind. Zusätzlich treten im Kraftwerksbetrieb immer wieder Defekte bei Sensoren auf. Entsprechend können fehlerhafte Messwerte in die Heizwertberechnung einfließen und diese beeinflussen.

Nach Stand der Technik werden Convolutional Neural Networks (CNN – zu deutsch Faltende neuronales Netz) zur Bilderkennung verwendet. Diese weisen vergleichbare Konnektivitätsmuster von Neuronenbahnen des menschlichen Gehirns auf und sind daher geeignet, um Aspekte bzw. Objekte in einem Bild zu identifizieren [Hou21]. Durch verschiedene Filter, die über ein Bild gelegt werden, werden vorwiegend Strukturen bspw. Kanten erkannt. Fraglich ist, ob das Auftreten von mehr kantigen oder runden Konturen, einen Zusammenhang zum Heizwert bzw. den übrigen Untersuchungsgrößen hat.

3.2.1 Automatisierte Erstellung der Bilddaten

Die Bilddaten stellen den Input für das neuronale Netz dar. Eine Abfallmenge zwischen 3-7 t wird regelmäßig durch einen Greifer in einen Trichter aufgegeben. Von dort wird der Abfall in den Verbrennungsraum gefördert. Die Abfallaufgabe wird durch eine Videokamera ganztägig aufgezeichnet.

Zur Verarbeitung des Videostreams wurde ein Pythonskript geschrieben, welches automatisiert die Abwurfsequenzen in Video- und Einzelbildformat extrahiert. Mithilfe der für Bildverarbeitung frei nutzbaren Programmibliothek OpenCV [Bra00] wird die Extraktion relevanten Bildsequenzen aus dem Videostream der Überwachungskamera nach folgendem Schema automatisiert:

1. Bewegungserkennung durch Differenzbildung zweier aufeinanderfolgender Frames.
2. Umwandlung des differenzierten Bildes in Graustufen und Weichzeichnung mit Gaußschem Weichzeichner.
3. Einstellung eines Schwellenwertes für die bereinigte Bilddifferenz. Jeder Bildpunkt wird mit dem eingestellten Schwellenwert verglichen. Ist der Pixelwert kleiner als der Schwellenwert, wird er der Farbe Schwarz zugeordnet, ist er größer, wird er der Farbe Weiß zugeordnet.
4. Dilatation bzw. Glättung von Linien/Kanten, die die Schwellenwerterkennung negativ beeinflussen können.
5. Ermittlung und Zählung aller erkannten Konturen im Bild.
6. Überschreitet die Anzahl der sich bewegenden Konturen einen experimentell bestimmten Schwellenwert, wird die Extraktion ausgelöst.
7. Bei Beginn der Extraktion, wird per Optical Character Recognition (OCR) [Ho21] die Zeitsignatur aus dem ersten Frame des Abwurfs ausgelesen. Für den Abwurf wird ein entsprechend benannter Ordner erstellt, in dem die Extraktionsdaten abgelegt werden.
8. Sinkt die Anzahl der sich bewegenden Konturen unter die untere Schwelle, wird die Extraktion beendet.

Fehlerkennungen werden durch eine Datennachbehandlung ebenfalls automatisiert entfernt. Erkannt werden Fehlerkennungen gemäß folgender zwei Bedingungen:

- Zwei Abwürfe sind zeitlich zu nah nacheinander
- Anzahl der einzelnen Frames eines erkannten Abwurfes weicht im Vergleich zu den übrigen Abwürfen deutlich ab

Bei Tests an einem 8 Stunden Videostream wurden 32 verwertbare Abwürfe extrahiert. Dies stimmt mit den Daten aus der Leitzentrale des Kraftwerks zu tatsächlich erfolgten

Abwürfen überein. Die Verarbeitung des Videostreams läuft auf der hiesigen Hardware mit 3-5-facher Geschwindigkeit.

Bezogen auf den Labeln-Prozess von Trainingsdaten sind zur Erstellung eines 10.000 Bilder umfassenden Datensatz die Verarbeitung von 2500 Stunden Rohvideomaterials aus dem Kraftwerk notwendig. Gegebenenfalls kann der bestehende Trainingsdatensatz durch die sogenannte Image Data Augmentation erweitert werden. Bei dieser Technik werden die vorhandenen Daten durch Pixelverschiebung, Rotation und Zooming bearbeitet, um zusätzliche Trainingsdaten aus den bereits vorhandenen Daten zu generieren. In einer Untersuchung zu dem Thema von Shorten und Khoshgoftaar [SK19] zeigt sich, dass dieser Ansatz helfen kann, die Stabilität eines neuronalen Netzes zu verbessern und das Overfitting zu reduzieren.

4 Diskussion und Ausblick

Aufgrund der Heterogenität des Inputs in die Abfallverbrennungskraftwerke ist es schwierig das Verbrennungsverhalten vorherzusagen. Die Thematik ist extrem komplex und es gibt sehr wenig parallele Forschungsaktivitäten. Auch die Übertragbarkeit von Erkenntnissen aus gängigen Kraftwerken ist aufgrund der Unterschiede in den Brennstoffen nicht möglich. Erschwerend kommt hinzu, dass das vorrangige Ziel der Abfallverbrennungskraftwerke die Abfallbehandlung ist und Emissionsgrenzwerte eingehalten werden müssen. Zukünftig werden Emissionsgrenzwerte weiter verschärft werden, um die Abfallverbrennung noch umweltschonender zu gestalten. Gleichzeitig ist es ein gesellschaftliches Ziel die Effizienz der Rückgewinnung der im Abfall gebundene Energie zu erhöhen, um die Energieversorgung nachhaltiger zu gestalten. Die Kreislaufwirtschaft ist ein unverzichtbarer Bestandteil unserer Gesellschaft und trägt wesentlich zur Erreichung unserer Nachhaltigkeitsziele bei. Abfall muss umweltschonend behandelt werden, damit keine gefährlichen Reststoffe unsere Umwelt schädigen.

Vor diesem Hintergrund ist es notwendig, neue Methoden für die Analyse des Verbrennungsverhaltens des Brennstoffs Abfall zu entwickeln. Methoden aus dem Bereich der Künstlichen Intelligenz sind bisher aufgrund der Heterogenität des Inputmaterials sehr wenig eingesetzt. Wie im Beispiel des Transfer Learnings gezeigt, bieten Bildanalysen von Abfällen großes Potential, um vor der Verbrennung einen Überblick über die zu erwartenden Verbrennungseigenschaften zu generieren. Gleichzeitig zeigt das Beispiel des Transfer Learnings, dass die Entwicklung spezielle auf Abfälle zugeschnittener Methoden erforderlich ist um hier einen signifikanten Beitrag zur Vorhersage leisten zu können. Durch die hier vorgestellte Methodik zur Generierung von Trainingsdaten für ein speziell auf die Bilderkennung von Abfällen zugeschnittenes Neuronales Netz wird ein großer Beitrag zur Entwicklung neuer KI-gestützter Methoden für die Abfallverbrennung geleistet. Die Methodik wird laufend erweitert und trainiert, so dass weitere Ergebnisse in naher Zukunft zu erwarten sind.

5 Literaturverzeichnis

- [Be12] Beckmann, M.; Scholz, R.; Pohl, M. (2012): Bilanzierung und energetische Bewertung von Verfahren zur Abfallbehandlung. In: Energie aus Abfall 9, S. 147–192.
- [Bra00] Bradski G. The OpenCV Library. Dr Dobbs's Journal of Software Tools 2000
- [Ch11] Christensen, Thomas H. (Hg.) (2011): Solid Waste Technology & Management. Volume 1 and 2. University of Denmark. Chichester, United Kingdom: Wiley.
- [Cy21] Cyris, F.: Prozessgütemonitoring der Wärmekraftumwandlung von Abfallverbrennungskraftwerken. Dissertation. Leibniz Universität Hannover 2021
- [De13] Decker, F. Altkleidersammlung, 2013, CC-BY-SA-3.0, Lizenz: CC-BY-SA-3.0, <https://de.wikipedia.org/wiki/Datei:Altkleidersammlung.JPG>
- [EE21] EEW Energy from Waste GmbH, eigene Bilder mit freundlicher Genehmigung, 26.4.2021
- [Ge12] Gesetz zur Förderung der Kreislaufwirtschaft und Sicherung der umweltverträglichen Bewirtschaftung von Abfällen (Kreislaufwirtschaftsgesetz - KrWG): KrWG, 2012.[Ho21] Hofstaetter, Pytesseract – A Python wrapper for Google Tesseract, <https://github.com/madmaze/pytesseract>, Stand 27.04.2021
- [Hou21] HOU, Yafei, et al. A Study of Throughput Prediction using Convolutional Neural Network over Factory Environment. In: 2021 23rd International Conference on Advanced Communication Technology (ICACT). IEEE, 2021. S. 429-434.
- [Im21] ImageNet, <http://www.image-net.org/>, Stand 26.04.2021
- [Kr17] Kranert, Martin (2017): Einführung in die Kreislaufwirtschaft. Wiesbaden: Springer Fachmedien Wiesbaden.
- [Ri08] Richtlinie 2008/98/EG des Europäischen Parlaments und des Rats über Abfälle und zur Aufhebung bestimmter Richtlinien, 2008.
- [SK19] Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. Journal of Big Data, 6(1), 1-48.
- [SZ14] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
- [Vo01] Verordnung über die umweltverträgliche Ablagerung von Siedlungsabfällen (Abfallablagerungsverordnung): AbfAbIV, 2001.

Natural Computing zur Identifikation von Verwertungspfaden biogener Rohstoffe

Barbara Rapp¹, Lea Freese² und Lukas Cremers³

Abstract: In diesem Paper stehen der Aspekt der Kaskadennutzung, am Beispiel der Ölpflanze Raps, sowie relevante Vertreter aus dem Bereich des Natural Computing im Vordergrund. Es werden verschiedene Nutzungspfade des biogenen Rohstoffs aufgezeigt, wobei verschiedene Möglichkeiten der Endnutzung abgedeckt werden. Die für die Optimierung relevanten Eigenschaften werden abgeleitet und die Anforderungen an die Zielfunktion benannt. Anschließend werden Evolutionsstrategien, Partikelschwarmoptimierung, das künstliche Bienenvolk und das Simulated Annealing als typische Vertreter des Natural Computing auf ihren möglichen Einsatz im Rahmen einer multikriteriellen Optimierung untersucht.

Keywords: Multi-Agenten System, Kaskadennutzung, biogene Rohstoffe, Natural Computing, multikriterielle Optimierung

1 Einleitung

Im Gegensatz zu den Energiequellen Wind und Sonne ist die Energie in biogenen Rohstoffen gebunden und steht somit bedarfsbezogen sowie zeitlich disponibel und permanent zur Verfügung. Die Frage, die sich bei der Nutzung jedoch stellt, ist: Soll die produzierte Biomasse als Energieträger, Lebens- oder Futtermittel verwendet werden? Die Antwort ist einfach: alles und das sogar gleichzeitig. Möglich wird dies indem die Nutzung derart orchestriert wird, dass das vorhandene Potential vollständig und optimal ausgeschöpft wird und dabei zeitgleich Nutzungskonkurrenzen vermieden werden. Dies kann beispielsweise über eine Kaskadennutzung erreicht werden.

In der Realität müssen dafür viele verschiedene Akteure mit unterschiedlichen Interessen (Anbieter und Nachfrager von Ressourcen, Dienstleistungen (z.B. Transporte), Prozess(schritt)en) zusammengebracht, zeitliche und räumliche Restriktionen der betrachteten Region sowie verschiedene (primäre und sekundäre) Ziele beachtet werden. Eine optimale Vernetzung kann also nur automatisiert, genauer algorithmisch, erfolgen. Infrage kommen hierfür verschiedene Vertreter aus dem Bereich des Natural Computings.

¹ Universität Oldenburg, Department für Informatik, Ammerländer Heerstr. 114-118, 26129 Oldenburg, b.rapp@uol.de, <https://orcid.org/0000-0003-3028-9565>

² Universität Oldenburg, Institut für Chemie und Biologie des Meeres, Ammerländer Heerstr. 114-118, 26129 Oldenburg, lea.christina.freese@uol.de

³ Universität Oldenburg, Department für Informatik, Ammerländer Heerstr. 114-118, 26129 Oldenburg, lukas.cremers@uol.de

Um die optimale Vernetzung der verschiedenen Akteure, ggf. unter Berücksichtigung jeweils eigener Interessen, zu ermöglichen, bietet sich der Einsatz eines Multi-Agenten Systems mit kontrollierter Selbstorganisation an.

In diesem Paper stehen der Aspekt der Kaskadennutzung, am Beispiel der Ölpflanze Raps (Abschnitt 2), sowie relevante Vertreter aus dem Bereich des Natural Computing im Vordergrund (Abschnitt 4). Die eigentliche Optimierung kann aus Platzgründen leider nicht im Detail behandelt werden; einige Aspekte daraus wurden aber bereits in verschiedenen anderen Publikationen aufgegriffen (z.B. [RS13] und [RBS11]).

2 Nutzung Biogener Rohstoffe

Biomasse selbst umfasst sämtliche Stoffe organischer Herkunft. Dabei kommt nicht alles an organischem Material für eine Nutzung in Frage. Doch auch der für die Nutzung geeignete Teil der Biomasse, die biogenen Rohstoffe, bietet ein großes Anwendungsspektrum, von Futter- und Nahrungsmittel über eine stoffliche Nutzung hin zum Energieträger. In Deutschland angebaute Pflanzen beispielsweise eignen sich zur Gewinnung von Stärke, Zucker, Ölen, Fasern, Arznei- und Färberstoffen sowie Holz. Hinzu kommt die Nutzungsmöglichkeit der Pflanzen zur Energiegewinnung [Fa14]. Das nicht jeder Rohstoff gleichermaßen als Energieträger oder als Futter- und Nahrungsmittel geeignet ist, ist bekannt. Doch auch Stärke ist nicht gleich Stärke, Öl nicht gleich Öl. Je nach Verarbeitungsgrad können auch einzelne Rohstoffe Eigenschaften aufweisen, wodurch sie für einen bestimmten Nutzungsbereich geeignet oder auch ungeeignet sind. Auch im Hinblick einer größeren Nachfrage an Biomasse in der stofflichen und energetischen Nutzung in Deutschland, können solche Eigenschaften für eine Optimierung von Nutzungspfaden relevant sein. Sie ermöglichen eine Entscheidungsfindung, ob ein Rohstoff für das vorgesehene Produkt, oder auch schon viel eher für einen bestimmten Verarbeitungsprozess in Frage kommt.

2.1 Nutzungspfade von Biomasse am Beispiel Raps

Um ein Verständnis über die Komplexität der Nutzungspfade biogener Rohstoffe zu gewinnen, kann es sinnvoll sein, solche einmal anhand eines Beispiels aufzuzeigen. Dabei eignen sich nicht jede Art von Biomasse für eine solche Darstellung, da es durchaus biogene Rohstoffe gibt, die nur für ganz spezifische Bereiche eingesetzt werden können und nicht für alle Anwendungsbereiche von der energetischen Nutzung über Nahrungs- und Futtermittel bis hin zum Rohstoff für eine stoffliche Nutzung in Betracht kommen. Für alle drei Marktsegmente geeignet ist beispielsweise der Raps [Fa14].

Abbildung 1 zeigt daher exemplarisch für biogene Rohstoffe die Verarbeitungsschritte des Raps mit Beispielen für die Endnutzung. Nach jedem Verarbeitungsprozess (braune Umrandung) steht ein Rohstoff (grüne Umrandung) mit veränderten Eigenschaften zur Verfügung, der entweder direkt in einem Anwendungsfeld für das Endprodukt (blaue Umrandung) genutzt oder einem weiteren Verarbeitungsprozess unterzogen werden kann.

Der Verarbeitungsprozess beginnt mit der Biomasse in Form der Rapspflanze, die dem ersten Verarbeitungsschritt, der Ernte, unterzogen wird. Bereits bei der Ernte werden Anforderungen an den Raps bzw. an die Biomasse gestellt, um einen späteren Transport, eine spätere Lagerung und Weiterverarbeitungsprozesse zu ermöglichen. Neben dem Hauptprodukt selbst, dem Öl, können bei den Verarbeitungsprozessen auch Nebenprodukte (gelbe Umrandung) entstehen. Sie können selbst auch für bestimmte Anwendungen bereitgestellt werden, teilweise haben auch sie wiederum eine Vielzahl an Anwendungsgebieten. Beim Raps beispielsweise entsteht bei der Ernte Rapsstroh, das für die Humusanreicherung verwertet werden kann. Bei der Ölpressung wiederum bleibt je nach Verfahren Rapskuchen oder Rapsextraktionsschrot übrig. Beide Nebenprodukte werden im Futtermittelbereich eingesetzt. Im späteren Verlauf der Rapsverarbeitung kann bei bestimmten Modifizierungen außerdem Glycerin entstehen. Glycerin ist für verschiedene Anwendungen einsetzbar, weshalb hier in der Grafik auf Beispiele verzichtet worden ist. Die nach der Ernte anfallenden Rapskörner können weiteren Verarbeitungsschritten unterzogen werden. Dazu zählen in diesem Beispiel Ölpressung, Raffination sowie eine große Auswahl an weiteren Modifikationen für die Oleochemie wie z.B. Veresterungen [Gr06]. Dabei werden die Rapskörner so verarbeitet, dass am Ende verschiedene Rohstoffe wie rohes Rapsöl, raffiniertes Öl, modifizierte Fettsäuregemische, Derivate und einzelne Fettsäuren vorliegen. Sie alle haben als Basis die Rapspflanze mit den Körnern, weisen aber nach den verschiedenen Arbeitsschritten in ihren Eigenschaften unterschiedliche Ausprägungen auf. Diese sind letztendlich entscheidend dafür, ob der Rohstoff sich für einen bestimmten Prozess bzw. für eine bestimmte Endnutzung eignen. Rohes Rapsöl (brauner Pfad) ist demnach im Nahrungsmittel- und Kraftstoffbereich einsetzbar, nicht aber in der Oleochemie. Hier muss raffiniertes Rapsöl (lila Pfad) vorliegen, das aber auch für Nahrungsmittel und als Kraftstoff eingesetzt werden kann. Einzelne Fettsäuren und Derivate (grüner Pfad) eignen sich vor allem in der Oleochemie, als Rapsmethylester aber auch im Kraftstoffbereich.

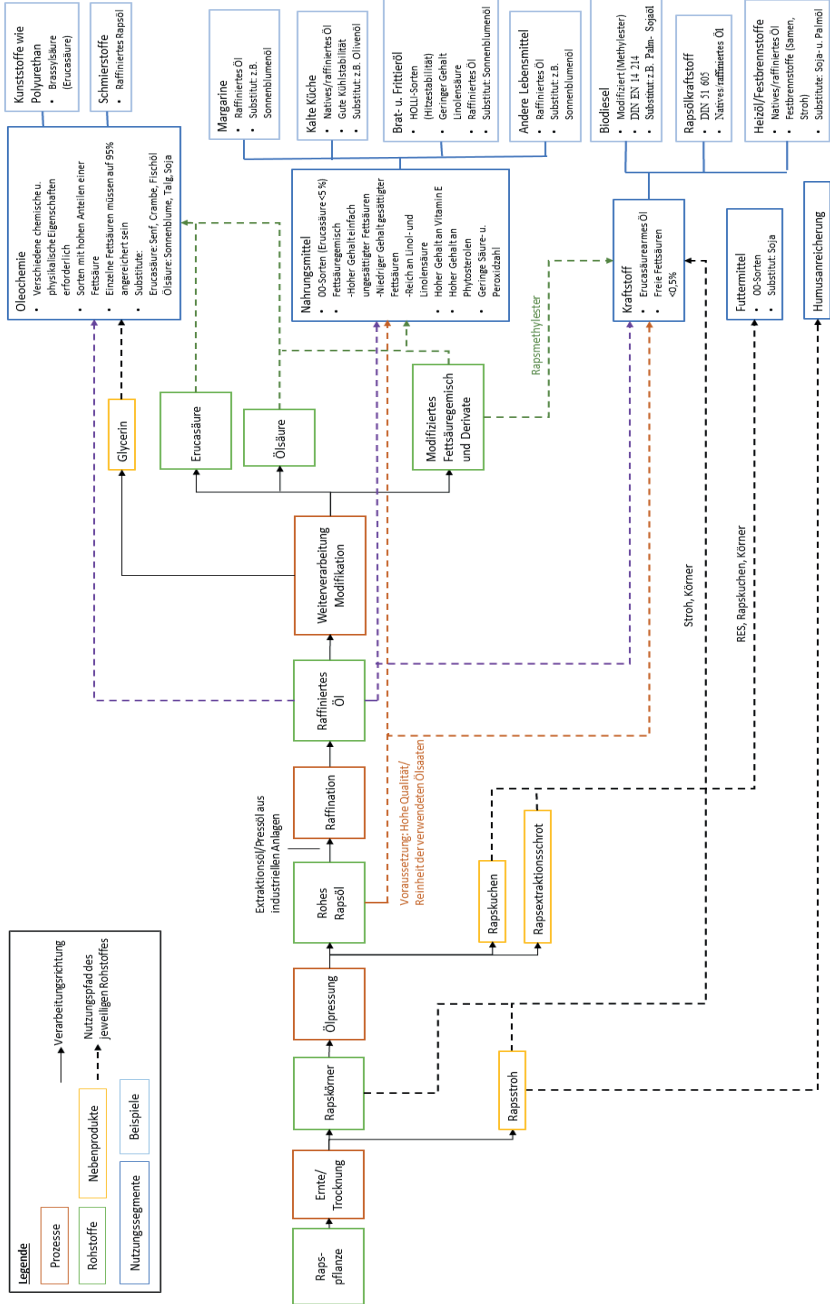


Abbildung 1: Ausgewählte Verarbeitungsschritte der Öl-pflanze Raps mit Beispielen für die Endnutzung.

2.2 Relevante Eigenschaften beim Raps

Nicht jeder Rohstoff eignet sich für jeden Verarbeitungsprozess oder für jede Nutzung. Aus Rapskörnern lässt sich nur schwer Stärke für den Nahrungsmittelbereich gewinnen, Kartoffeln eignen sich weniger für Arzneistoffe. Aber auch Öl ist nicht gleich Öl wie Abbildung 1 zeigt. Wie andere Produkte auch, weisen biogene Rohstoffe charakteristische Eigenschaften auf, durch die sie eindeutig identifiziert werden und die gleichzeitig auch ausschlaggebend sind, ob der Rohstoff für eine Nutzung oder einen Prozess in Betracht gezogen werden kann. Abbildung 1 zeigt neben den Nutzungspfaden auch beispielhaft solche Anforderungen an den Rohstoff im jeweiligen Nutzungssegment. Nicht alle Eigenschaften, die für eine Charakterisierung des Rohstoffes möglich sind, sind aber auch gleichzeitig sinnvoll. Es gibt Merkmale, die zwar zur Identifizierung des Rohstoffes beitragen, aber keinerlei Relevanz aufweisen, wenn es um die Eignung des Rohstoffes für einen Prozess geht [Di14]. Um eine automatisierte Vernetzung der Rohstoffe und Prozesse zu realisieren, sollten daher diese Eigenschaften bekannt sein, die für eine solche Zuordnung von Bedeutung sind.

Am Beispiel Raps zeigt sich, dass je nach Verarbeitungsschritt sich die relevanten Eigenschaften oder die Ausprägungen, die gewünscht sind, unterscheiden können. Bei dem Prozess der Ölpressung kommt es bei der Rapssaat vor allem auf die Eigenschaften Ölgehalt, Reinheit und Feuchtigkeit an [Mi14]. Nur so kann letztendlich Rapsöl mit gewünschter Qualität garantiert werden. Eigenschaften wie die Morphologie der Körner oder der Name des Rohstoffes sind für den Prozess selbst irrelevant. Nach der Pressung in den Ölmühlen liegt rohes Rapsöl vor. Sofern dieses direkt genutzt und nicht weiterverarbeitet werden soll, werden je nach Anwendungssektor verschiedene Eigenschaften gefordert. Im Nahrungsmittelbereich sind vor allem der Geruch und der Geschmack entscheidend [CF11]. Aber auch die Kenntnis über sekundäre Inhaltsstoffe wie z. B. Vitamine oder die Fettsäurezusammensetzung des Öls sind ernährungstechnisch relevant. Im Nahrungs- und Futtermittelbereich sind überwiegend 00-Sorten gewünscht, die wenig Erucasäure und einen geringen Anteil an Glucosinolate aufweisen. Wird das Rapsöl als Energieträger genutzt, ist vor allem ein geringer Anteil an Erucasäure und ein hoher Energiegehalt entscheidend [Mi14]. Eine stoffliche Nutzung in der Oleochemie setzt wiederum andere Eigenschaften bzw. andere Ausprägungen voraus. Der Verschmutzungsgrad des Öls sollte möglichst gering sein, weshalb hier nur raffiniertes Rapsöl zum Einsatz kommt. Auch entscheidend ist die Fettsäurezusammensetzung. Sind einzelne Fettsäuren wie Eruca- oder Ölsäure mit hohen Anteilen vorhanden, eignet sich das Öl für die Gewinnung dieser einzelnen Fettsäuren, aber auch Fettsäuregemische sind möglich. In der Oleochemie sind aber auch chemisch-physikalische Eigenschaften von Bedeutung, beispielsweise Viskosität, Schmierfähigkeit und Eigenschaften, die den Rohstoff als Bindemittel beschreiben [Gr06]. Neben den genannten Eigenschaften, die sich überwiegend auf die Qualität des Rapsöls beziehen, können auch Eigenschaften wie Lagerfähigkeit, Preis und Menge relevant sein, insbesondere wenn es um die Logistik und Wirtschaftlichkeit des Prozesses geht.

Somit zeigt sich bereits bei den Nutzungspfaden des Raps, dass je nach Prozess, je nach Nutzungsart, verschiedene Eigenschaften des Rohstoffes relevant und bekannt sein sollten, um mögliche Nutzungspfade zu identifizieren.

3 Anforderungen an die Optimalität der Kaskadennutzung

Der vorangegangene Abschnitt gibt einen ersten Überblick über ausgewählte Nutzungsmöglichkeiten der Ölpflanze Raps. Dass die in Abbildung 1 dargestellten Möglichkeiten nicht abschließend sind, versteht sich von selbst. Allerdings geben sie einen ersten guten Einblick in die Komplexität einer orchestrierten Nutzung. Soll das vollständige Potential verschiedener biogener Rohstoffe ausgenutzt werden, so kommen sie jeweils u.a. für Lebens-, Futter- und Arzneimittel sowie bio-basierte Produkte wie Biomaterial, Biochemikalien oder Kraftstoffe und Energie in Betracht. Zusätzlich zu den mannigfaltigen Nutzungsmöglichkeiten jeder einzelnen Ressource ergeben sich unüberschaubar viele Kombinationsmöglichkeiten, sobald mehr als eine Ressource gleichzeitig betrachtet wird. Hier ist die Identifikation einer optimalen Nutzungskaskade unabdingbar.

Optimalität bezeichnet in diesem Zusammenhang die richtige Kombination aus mindestens den Aspekten Bedürfnisbefriedigung, Nachhaltigkeit sowie übergeordneter Ziele. Bedürfnisbefriedigung meint, inwieweit die gewünschte Nutzungsmöglichkeit (Prozess) mit allen benötigten Inputs in der notwendigen Menge und Qualität versorgt werden kann. Nachhaltigkeit bedeutet, wie eine konkrete Nutzungsmöglichkeit unter ökologischen, ökonomischen und sozialen Gesichtspunkten zu bewerten ist. Übergeordnete Ziele adressieren primär den Optimierungszweck und gewichten vor allem die Aspekte Flexibilität und Robustheit entsprechend; es macht einen Unterschied, ob ein autarkes Netzwerk oder ein auf einen bestimmten Output fokussiertes Netzwerk gefunden werden soll. Eindeutig handelt es sich bei dem hier vorliegenden Problem um eine Mehrzieloptimierung.

Neben den bereits benannten Herausforderungen müssen auch eine Reihe induzierter Eigenschaften berücksichtigt werden. Biogene Rohstoffe sind bzgl. Menge, Qualität und Zeit ein nur begrenzt verfügbares Gut, welches sich unterschiedlich für bestimmte Prozesse oder Prozessschritte eignet. Maßgeblich wird die Eignung durch Form und Stadium des Rohstoffs bestimmt und kann nur begrenzt durch Vorverarbeitung, Vorbehandlung oder Konditionierung modifiziert werden. Ebenso können die Verwertungspfade innerhalb des Netzwerks nicht als fix angesehen werden, da beteiligte Akteure und/oder auch die Reihenfolge variabel ist. Ferner ist die immanent vorhandene zeitliche und räumliche Fluktuation eine große Herausforderung; die Verfügbarkeit und Qualität der Ressourcen sind zeit- und ortsabhängig.

4 Algorithmen zum Aufbau einer optimalen Kaskadennutzung

Wie bereits ausgeführt, kann die optimale Vernetzung - aufgrund der ihr innewohnenden Komplexität - nur automatisiert erfolgen. Zur Erreichung eines gewissen Grades an Automatisierung muss eine angemessene Algorithmik zum Einsatz kommen, von besonderem Interesse sind hier die Vertreter aus dem Forschungsgebiet des Natural Computing.

Grundsätzlich sind viele Prozesse aus der Natur rechnerbasierten Handlungsweisen nicht unähnlich [BOM15]. Diese beobachtete Parallelität veranlasst Forschende bereits fast ein Jahrhundert, die Fähigkeit der Natur, seit jeher widrigen Bedingungen und komplexen Problemen zu trotzen, als Inspirationsquelle für problemlösende Algorithmen zu nutzen [CP17, KR08]. Eigenschaften, die eine Vielzahl derartiger Systeme oder Prozesse der Natur vereinen, und sie als Vorbild für die technische Adaption attraktiv machen, sind nach [BOM15] bspw. die Folgenden: Die Populationsbasiertheit - also die Existenz einer Mehrzahl an potenzieller Lösungen - ermöglicht die breite Abdeckung einer Suchumgebung und darüber hinaus eine parallele Suche. Ein hoher Grad an Dispersion (Verteiltheit) der Population trägt dazu bei, eine Konvergenz der Suche in lokalen Optima zu verhindern. Eine dritte wichtige Eigenschaft ist die Fähigkeit der Individuen einer Population zur Kommunikation. Auf verschiedenste Arten kann miteinander kommuniziert werden, ob direkt oder indirekt. Durch den Informationsaustausch kann der Prozess der Lösungsfindung gemeinschaftlich in eine bestimmte Richtung gelenkt werden.

Ein System, welches die Eignung ausgewählter Algorithmen des Natural Computing zur Lösung der zuvor angesprochenen multikriteriellen Optimierungsprobleme ermöglicht, kann wie folgt aufgebaut sein: Angebote sowie Nachfragen sind räumlich verteilt und weisen neben ihren Koordinaten weitere wichtige Eigenschaften, wie sie z.B. in Abschnitt 2.2 aufgeführt sind, auf. Es empfiehlt sich, dass entweder Angebote oder Nachfragen als handelnde Akteure in Aktion treten. Das Problem wird dementsprechend angebots- bzw. nachfrageseitig gelöst. Im Lösungsfindungsprozess wird dann z.B. eine Nachfrage durch einen Agenten in einem Multi-Agenten System verkörpert. Maßgebliches Charakteristikum eines solchen ist die Komplexitätsreduktion des Problems durch Verteilung der Aufgaben auf mehrere Agenten. Ziel der Agenten ist es, durch Anwendung eines der hier vorgestellten Algorithmen, eine optimale Vernetzung von Angeboten und Nachfragen sukzessive zu erarbeiten. Dies geschieht durch Kommunikation und Verarbeitung der von ihnen erzielten Beiträge zur gemeinschaftlichen Lösung des Gesamtproblems [BT07]. Die Definition der Optimalität ist dabei von besonderer Bedeutung; mit der globalen Zielfunktion als zentralem Element. In sie fließen mehrere Kriterien ein, welche - je nach Gewichtung - die Güte der Lösung (Vernetzung) beeinflussen. Oftmals wird das betrachtete Problem vereinfachend als Problem der Kostenminimierung bzw. Gewinnmaximierung dargestellt. Optimal heißt demnach bspw., die wenigsten Kosten verursachend. Verschiedene der bereits eingeführten Optimalitätskriterien wie Merkmale oder Kennzahlen der Nachhaltigkeit, der Bedürfnisbefriedigung sowie die Struktur der Verwertungspfade können über Strafkosten

modelliert werden. Neben der globalen Zielfunktion können auch lokale Restriktionen bestehen [CT18]. Diese können das Zustandekommen eines Matchings von Angebot und Nachfrage, bzw. die Güte dessen, ebenso beeinflussen, jedoch auf lokaler Ebene, also einzelne spezielle Kombinationen betreffend.

Im Rahmen einer kleinen Simulationsstudie wurde ein solches System prototypisch umgesetzt. Zufällig gezogene Koordinaten für Angebote und Nachfragen (Verhältnis 1:3) wurden auf einen Kartenausschnitt projiziert und die Distanzen mithilfe der GraphHopper API auf Basis von OpenStreetMap-Kartendaten bestimmt. Handelnde Akteure waren stets Nachfragen, die nach “passenden” Angeboten suchten. Angebote sowie Nachfragen enthielten u.a. Eigenschaften wie Angebots- bzw. Nachfragemenge, Reinheit (Qualität) der Mengeneinheiten und lokale Restriktionen. Letztere konnten etwa derart sein, dass ein Angebot nur exklusiv genutzt werden kann oder eine Nachfrage nur eine bestimmte Distanz zurücklegen kann. Wird eine dieser Restriktionen verletzt, konnten Angebot und Nachfrage nicht zusammengeführt (gematcht) werden. Aus der Teilmenge der gültigen Kombinationen erarbeitete der jeweils untersuchte Algorithmus eine bestmögliche Vernetzung unter Verwendung der globalen Zielfunktion, deren Funktionswert maßgeblich durch die zurückzulegende Distanz und Strafkosten für ungleiche Reinheit sowie für übrigbleibende Nachfrageeinheiten bestimmt wurde.

Es kamen dabei folgende Algorithmen zum Einsatz, welche den Bereichen der evolutionären Algorithmen, der Schwarmintelligenz und weiteren naturanalogen Prozessen zuzuordnen sind: Evolutionsstrategien (ES), die Partikelschwarmoptimierung (PSO), das künstliche Bienenvolk (ABC) und das Simulated Annealing (SA), deren Funktionsweise im Einzelnen in [BOM15] nachvollzogen werden kann. Bewertet wurden die Ergebnisse der Algorithmen allen voran unter den Aspekten Berechnungsaufwand, Lösungsgüte und Konsistenz der Lösungen sowie zusätzlich nach Distanz pro befriedigter Einheit und Anteil der befriedigten Nachfrageeinheiten. Werden dann für mehrere Simulationsläufe die Anzahl der Angebote und Nachfragen sowie die Anzahl der lokalen Restriktionen variiert, lassen sich mit Hilfe der Bewertungskriterien und des Wissens über die Funktionsweise der Algorithmen einige Ergebnisse auf bestimmte Charakteristika der Algorithmen zurückführen [SVH17].

Als von einem Agenten generierter Lösungskandidat zur Befriedigung der durch ihn repräsentierten Nachfrage ist im Folgenden ein n -dimensionaler Vektor zu verstehen, wobei n die Anzahl der Angebote ist. Die zu befriedigenden Einheiten einer Nachfrage müssen also über die Menge der Angebote verteilt werden. Dabei überzeugt die PSO durch einen stets geringen Berechnungsaufwand verglichen mit den anderen Algorithmen. Zurückzuführen ist dies einerseits auf das Gedächtnis eines Individuums, das die bisher beste gefundene Lösung abspeichert sowie auf den Einfluss des Gesamtschwarms bzw. der Nachbarschaft, die ihrerseits die besten gefundenen Lösungen kommunizieren. So ist der Suchprozess als stark zielgerichtet zu beschreiben. Bei ES hingegen, welche der biologischen Evolution nachempfunden sind, ist dem nicht so. Die Mutationen, die neue Lösungskandidaten erzeugen, basieren auf dem Zufallsprinzip und dem Konzept der Erzeugung einer größtmöglichen Diversität. Allein die Logik gebietet, dass hier in der Regel signifikant mehr neue Kandidaten zu erzeugen sind, um positive Veränderungen

herbeizuführen. Dies schlägt sich negativ im Berechnungsaufwand nieder. In kleineren Szenarien (Lösungsvektor mit wenigen Dimensionen) wird durch diese sehr breite Abdeckung des Suchraums daher jedoch sehr sicher das globale Optimum gefunden und dies auch sehr konsistent. Hier ergibt sich ein Vorteil gegenüber der PSO, welche diese "guten" Lösungen zwar auch erarbeiten kann, jedoch nicht zuverlässig in jedem Anlauf. Bei größeren Lösungsvektoren stoßen die Algorithmen mit zufallsbasierten Mutationen jedoch an ihre Grenzen, so auch das SA. Durch die kleinen Suchschritte, derer es für diese Art von Algorithmen bedarf, damit sie sich Optima zuverlässig annähern können, werden zu Anfang des Suchprozesses sehr viele Kanten aufgebaut (50 Einheiten einer Nachfrage werden bspw. von 40 verschiedenen Angeboten abgerufen). Dies wird in der späteren Suchphase zum Problem, da Kanten, auch wenn sie sich als nicht vorteilhaft erwiesen haben, durch die zufälligen Mutationen nicht zielgerichtet abgebaut werden können. Mit einer etwas andersartigen Funktionsweise wartet der ABC-Algorithmus auf. Neue Lösungskandidaten werden hier mit der Hilfe einer stochastischen Variable erzeugt, die auf der Differenz zweier bestehender Lösungen beruht. Mit Voranschreiten des Suchprozesses werden die Differenzen zwischen den Individuen und somit auch die Stärke der Mutationen immer geringer. So wird sich an Optima angenähert. In diesem speziellen System von Angebot und Nachfrage kann der ABC-Algorithmus in keiner der Bewertungsdimensionen über die verschiedenen Szenarien hinweg besonders überzeugen. Geschuldet ist dies vermutlich dem Verwerfen und Ersetzen von Lösungskandidaten, aus denen längere Zeit keine Verbesserung hervorgegangen ist. Diese Eigenschaft wirkt sich in der betrachteten statischen Umwelt eindeutig negativ aus. Sowohl der natürliche Lebensraum der Honigbiene als auch die Nutzung biogener Rohstoffe lassen sich hingegen eher als hochgradig dynamische Umwelt beschreiben. Laut [BOM15] darf angenommen werden, dass die Vorgehensweise der Honigbienen in diesem Kontext einen Vorteil gegenüber den anderen betrachteten Algorithmen darstellt.

5 Fazit und Ausblick

In diesem Paper wurden die Aspekte der Kaskadennutzung am Beispiel der Ölpflanze Raps erläutert sowie relevante Vertreter aus dem Bereich des Natural Computing betrachtet und deren Eignung für den Einsatz im Kontext einer multikriteriellen Optimierung zusammenfassend dargestellt. Beide Aspekte halten gerade Einzug in ein Multi-Agentensystem mit kontrollierter Selbstorganisation. Erste technische Vorüberlegungen dazu finden sich bereits in [Rap19]. Ein Paper, das die Variabilität der Zielfunktion in Abhängigkeit vom gewünschten Modellzweck untersucht, ist in Planung.

Neben der hier beleuchteten Ölpflanze Raps sollten weitere Rohstoffe betrachtet werden, um die Relevanz bestimmter Eigenschaften zu bestätigen. Dazu bietet es sich insbesondere an, nicht nur Öle zu beschreiben, sondern auch Rohstoffe aus anderen Bereichen wie Arzneimittel, Holz, Zucker und Stärke.

- [BOM15] Brabazon, A.; O'Neill, M.; McGarraghy, S.: Natural Computing Algorithms, 1st ed. 2015 Aufl., Springer Berlin Heidelberg, 2015
- [BT07] Bullinger, H.-J.; ten Hompel, M.: Internet der Dinge, Springer Berlin Heidelberg, 2007
- [CF11] Christen, O.; Friedt, W.: Winterraps. Das Handbuch für Profis. DLG-Verlag, Frankfurt Main, 2011.
- [CP17] Chawda, B.; Patel, J. M.: Investigating performance of various natural computing algorithms, International Journal of Intelligent Systems and Applications (IJISA), 1, 46-59, 2017
- [CT18] Chopard, B.; Tomassini, M.: An Introduction to Metaheuristics for Optimization, 1st ed. 2018 Aufl., Springer International Publishing, 2018
- [Di14] Diepenbrock, W.: Nachwachsende Rohstoffe - 44 Tabellen. Ulmer-Verlag, Stuttgart, 2014.
- [Fa14] Fachagentur Nachwachsende Rohstoffe e.V.: Marktanalyse nachwachsende Rohstoffe. Gülzow-Prüzen, 2014.
- [Gr06] Grimm, E.: Industrierohstoffe. In (Heyland, K.-U.; Hanus, H.; Keller, E., Hrsg.): Ölf Früchte, Faserpflanzen, Arzneipflanzen und Sonderkulturen. Ulmer-Verlag, Stuttgart, S. 141-148, 2006.
- [KR08] Kari, L.; Rozenberg, G.: The many facets of natural computing, Communications of the ACM, 51, 72-83, 2008
- [Mi14] Miedaner, T.: Raps - Vom Leuchtöl zum Lebensmittel. In (Miedaner, T. (Hrsg.): Kulturpflanzen: Botanik - Geschichte - Perspektiven. Springer Science and Business Media, S. 183- 200, 2014.
- [Rap19] Rapp, B.: Fitnesslandschaftsanalyse als Vorstufe zur Algorithmenentwicklung. In (Schaldach, R.; Simon, K.; Weismüller, J.; Wohlgemuth, V., Hrsg.): Environmental Informatics: Computational Sustainability. Düren : Shaker Verlag, S. 160-167, 2019.
- [RBS11] Rapp, B., Bremer, J., Sonnenschein, M.: Sustainable, Multi-Criteria Biomass Procurement: A Game Theoretical Approach, Information Technologies in Environmental Engineering, 2011.
- [RS13] Rapp, B.; Sonnenschein, M.: Identification of Optimal Biomass Utilization – Characteristics and Challenges, Shaker Verlag, 2013.
- [SVH17] Siebertz, K.; van Bebber, D.; Hochkirchen, T.: Statistische Versuchsplanung: Design of Experiments (DoE), Springer Berlin Heidelberg, 2017

BUIS in der Lehre: Theorie und Praxis Hand in Hand

Barbara Rapp,¹ Andreas Solsbach²

Abstract: In diesem Paper berichten wir über die Veranstaltungen Betriebliche Umweltinformationssysteme I und II wie sie an der Carl von Ossietzky Universität Oldenburg angeboten werden. Ursprünglich stand die Vermittlung verschiedenster theoretischer Methoden, Modelle usw. im Vordergrund. Mit der Zeit hat sich der Fokus in Richtung Forschendem Lernen verschoben. So kann es gelingen, die vielfältigen und ohne Frage wichtigen theoretischen Grundlagen aus dem Bereich BUIS direkt mit einer praktischen Anwendung zu verbinden.

Keywords: Lehre; BUIS; Google Design Sprint; Urban Farming; Nachhaltige Mobilität

1 Grobkonzept

Im Rahmen des Masterstudiengangs Wirtschaftsinformatik an der Carl von Ossietzky Universität Oldenburg haben Studierende die Möglichkeit, verschiedene Veranstaltungen zum Themenkomplex Betriebliche Umweltinformationssysteme (BUIS) zu belegen. In der Veranstaltung BUIS I wird die Grundlage für weitere Veranstaltungen gelegt und hier werden im Theorieteil beispielsweise die Themen Materialflussanalyse, Bilanzbewertung und Schwachstellenanalyse, produktionsnahe BUIS, Umweltmanagementsysteme, Ökobilanzen und Nachhaltigkeitskommunikation behandelt. Begleitet wird der Theorieteil durch eine Projektaufgabe, in der eine Gruppe von i. d. R. drei Personen eine IKT-basierte Nachhaltigkeitskommunikation in einem produzierenden, mittelständischen Unternehmen prototypisch umsetzt. Die Studierenden haben bei der Wahl des (fiktiven) Unternehmens und dessen Kernprodukt freie Hand, sie müssen die bisherigen Produktionsabläufe für das gewählte Produkt sowie dessen überarbeitete nachhaltiger ausgerichtete Fertigung modellieren und beide Abläufe gegenüberstellen. Ebenso ist es die Aufgabe der Gruppe, ein grobes Konzept für die Einführung eines strukturierten Umweltmanagementsystems im Wahlunternehmen zu erstellen. Die verschiedenen Bausteine werden in einem Konzept für einen unternehmensweiten, online verfügbaren und zielgruppenorientierten Nachhaltigkeitsbericht zusammengeführt und am Ende der Vorlesungszeit in Form eines Pitches präsentiert. Neben den Studierenden sind Mitarbeiter:innen und gelegentlich auch Gäste der Abteilung Very Large Business Applications, wie z. B. Vertreter:innen von Unternehmen oder der Stadt Oldenburg, bei den Pitches vertreten.

¹ Universität Oldenburg, Department für Informatik, Ammerländer Heerstr. 114-118, 26129 Oldenburg, b.rapp@uol.de, <https://orcid.org/0000-0003-3028-9565>

² Universität Oldenburg, Department für Informatik, Ammerländer Heerstr. 114-118, 26129 Oldenburg, andreas.solsbach@uol.de, <https://orcid.org/0000-0002-8799-0478>

So ist es möglich, die im Rahmen der Vorlesung kennengelernten theoretischen Inhalte unmittelbar praktisch anzuwenden und entsprechende Fach- und Methodenkompetenzen auf- und auszubauen. Die Teilnehmenden der Veranstaltung kommen aus verschiedenen Masterstudiengängen: Wirtschaftsinformatik, Informatik sowie Sustainability Economics and Management. Da die Studierenden in heterogenen Gruppen arbeiten und ihre eigenen (Teil-) Ergebnisse präsentieren und diskutieren müssen, verbessert sich ebenfalls ihre Sozialkompetenz. Das erhaltene Feedback und ggf. vorgebrachte Kritik wird als Vorschlag für die Weiterentwicklung des eigenen Handelns aufgenommen, was auf eine hohe Selbstkompetenz schließen lässt.

Die Veranstaltung Betriebliche Umweltinformationssysteme II (BUIIS II) findet als zweiwöchiger Blockkurs im Wintersemester statt und baut auf dem vorhandenen Wissen auf. Hier wird im Projektformat ein funktionsfähiger Hard- und/oder Softwareprototyp von den Studierenden realisiert. In der ersten Woche des Blockkurses wird die theoretische und methodische Grundlage gelegt. Klassischerweise finden am Montag der ersten Woche Vorträge zum, bereits im Vorfeld bekannten, Dachthema statt; die Szenarien aus den vergangenen beiden Jahren werden in den Abschnitten 3.1 und 3.2 im Detail beschrieben. Dienstag bis Donnerstag der ersten Woche widmen sich dem Google Design Sprint (siehe Abschnitt 2) und dem Sensorworkshop, am Freitag der ersten Woche wird die eigentliche Aufgabe vorgestellt und besprochen, ferner findet, wenn immer möglich, eine Exkursion statt. An allen Tagen der ersten Woche werden die Studierenden aktiv in die Erarbeitung des Contents eingebunden, der Anteil des Dozentenvortrags ist auf das absolute Minimum reduziert; andere Formen genießen Vorrang.

In Abhängigkeit vom Dachthema haben wir bislang zwei Varianten für die Vorträge am Montag der ersten Woche ausprobiert. Bei eher breit aufgestellten Dachthemen wie z. B. Urban Farming (siehe Abschnitt 3.1) konnten sich die teilnehmenden Studierenden im Vorfeld für verschiedene Konkretisierungen bzw. Fragestellungen entscheiden und haben dazu ein Referat zu Beginn des Blockkurses gehalten. So konnten verschiedene Aspekte, die die Studierenden interessierten, beleuchtet und für die eigentliche Aufgabenstellung berücksichtigt werden. Bei eher abgeschlossenen Dachthemen, wie z. B. der Fahrradinfrastruktur (siehe Abschnitt 3.2), wurde eine Vortragsreihe organisiert, in der verschiedene Forschungsprojekte (u. a. unseres Lehrstuhls und der Stadt Oldenburg) und studentische Projektgruppen jeweils ihre Forschungsfragen, erste Lösungsansätze und (Teil-) Ergebnisse vorgestellt haben. Dadurch ist es gelungen, die große Diversität innerhalb eines eher abgegrenzten Forschungsgebiets für die Studierenden sichtbar zu machen und ihnen zu ermöglichen, aktuelle Forschungsergebnisse bei der Bearbeitung der eigentlichen Aufgabenstellung zu berücksichtigen.

Erst am Freitag der ersten Blockkurswoche wird die eigentliche Aufgabenstellung bekanntgegeben. Auf der einen Seite ist es so recht einfach möglich, die Startbedingungen für die einzelnen Gruppen zu harmonisieren und zum anderen kann sich so noch kein festes Meinungsbild oder eine konkrete Lösungsidee festsetzen. Unsere Erfahrung zeigt, dass aufgrund der späten Veröffentlichung der Aufgabenstellung die in der zweiten Woche

erarbeiteten Ideen und Lösungskandidaten sehr mannigfaltig sind. Trotz exakt gleicher Aufgabenstellung für alle Gruppen gab es bislang keine identischen Lösungen, sondern jede Gruppe hat sich für einen jeweils anderen Schwerpunkt entschieden.

Da die BUIS II-Teilnehmer:innen, wie schon bei BUIS I, aus verschiedenen Masterstudiengängen stammen, bringen sie verschiedene Kompetenzen und unterschiedliches Vorwissen ein. Die wenigsten verfügen über Erfahrung bei der Entwicklung von Hardwareprototypen. Dieser Aspekt wird im Sensorworkshop adressiert. Für die Studierenden stehen Hardwarebaukästen (Arduino, verschiedene Sensoren, etc.) zur Verfügung und im Rahmen eines etwa vierstündigen Workshops werden verschiedene Prototypen gebaut. Begonnen wird damit die Helligkeit von LEDs mittels verschiedener Widerstände zu beeinflussen. In der nächsten Aufgabe steht das Farbspektrum einer RGB-LED im Mittelpunkt und der Prototyp soll stetig zwischen den Zuständen rot, grün, blau wechseln und in den Übergängen die dazwischenliegenden Farben anzeigen. Diese beiden (sehr einfachen) Aufgaben sind bewusst gewählt und der Erfolg der Studierenden ist nahezu garantiert. Interessant zu beobachten ist, dass es bei vielen Studierenden zunächst Vorbehalte bzgl. des eigenen Könnens gibt, diese aber schrittweise abgelegt werden. In fast jedem Workshop werden von den Studierenden Fotos und Videos der Prototypen gemacht und mit Freunden und/oder der Familie geteilt. Als dritte Aufgabe soll ein thermischer Widerstand in Verbindung mit einem LCD genutzt werden, um die Temperatur anzuzeigen. Diese Aufgabe ist deutlich schwieriger als die vorangegangenen Aufgaben, allerdings ist zu diesem Zeitpunkt i. d. R. das Interesse der Studierenden geweckt und sie arbeiten gruppenübergreifend zusammen, um die gestellte Aufgabe zu lösen. In erster Linie dient der Workshop dazu, bestehende Vorurteile und Ängste gegenüber der Hardware abzubauen und den Studierenden aufzuzeigen, dass ggf. auch ein Hardwareprototyp als Lösungskandidat in Frage kommt.

In der ersten Woche des Blockkurses sind die einzelnen Tage fest strukturiert, montags Vorträge, Dienstag bis Donnerstagmittag der Google Design Sprint, Donnerstagnachmittag der Sensorworkshop und am Freitag die Vorstellung der Aufgabenstellung und nach Möglichkeit eine Exkursion. Die Uhrzeit zu denen an den einzelnen Tagen begonnen wird sowie die Anzahl und Länge der Pausen wird mit den Studierenden abgestimmt. In der zweiten Woche müssen sich die einzelnen Gruppen vollständig selbst organisieren, was auch der Idee des Forschenden Lernens ([Le17]) sehr zugute kommt. Lediglich der Termin für das tägliche Stand-up sowie für die Pitches am Freitag sind fix. Die Dozierenden stehen während des gesamten Kurses als Ansprechpartner:innen zur Verfügung.

Bis einschließlich Wintersemester 2019/2020 konnte der Kurs in Präsenzform durchgeführt werden, inklusive Exkursion ins benachbarte Ausland. Im Wintersemester 2020/2021 musste der Kurs, bedingt durch die COVID-19-Pandemie, erstmals online umgesetzt werden. Hierzu wurden verschiedene BigBlueButton-Räume für die einzelnen Gruppen sowie ein Vortragsraum genutzt. Parallel dazu wurde auf verschiedene Kollaborationssoftware (z. B. Conceptboard, Miro, Mural) gesetzt. Die Hardware wurde unter strenger Einhaltung eines Hygienekonzeptes an die Studierenden ausgehändigt, der Sensorworkshop fand online statt. Generell waren wir sehr positiv überrascht, wie problemlos sich das etablierte Präsenzformat

in ein Onlineformat überführen lies; dies mag u. a. darin begründet sein, dass alle Beteiligten bereits verschiedene Erfahrungen aus dem vorangegangenen Onlinesemester einbringen und so neue Wege beschritten werden konnten. Zukünftig könnte man beide Ansätze kombinieren, wobei das Präsenzformat insbesondere die Zusammenarbeit der Studierenden in der Gruppe deutlich vereinfacht.

2 Google Design Sprint

In seiner ursprünglichen Ausprägung erstreckt sich der Google Design Sprint über fünf Tage, wobei je Tag etwa acht Stunden Arbeitszeit (inklusive 2 Pausen) vorgesehen sind. Jeden Tag steht eine andere, auf die vorangegangenen Arbeiten aufbauende, Aktivität im Vordergrund. Abbildung 1 zeigt diesen Ablauf.

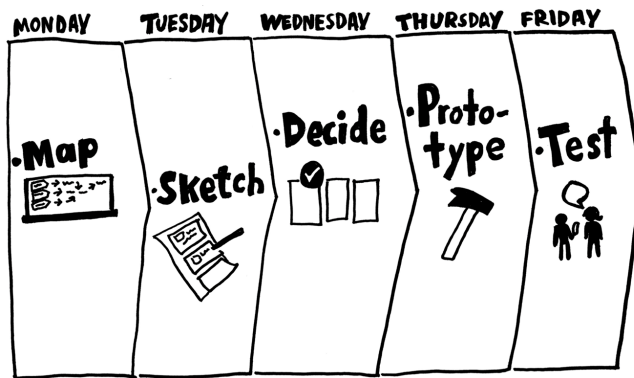


Abb. 1: Ablauf des Google Design Sprints (nach [KZK16]).

Die Design Sprint-Woche beginnt mit der Phase Map. Im Fokus steht hierbei der Wissensaustausch, das genaue Verständnis des Problems sowie die Festlegung des Ziels. In einem ersten Schritt (Start at the End) werden das langfristige Ziel festgelegt sowie mögliche Risiken auf dem Weg dorthin identifiziert. Von diesem Ziel ausgehend werden rückwärts die Schritte zur Erreichung dieses Ziels identifiziert. Im Anschluss (Make a Map) wird erarbeitet, wie die Kunden:innen mit dem (geplanten) Produkt interagieren. Dies wird bereichert durch die Befragung von Experten:innen, indem deren Visionen aber auch bisherige Erfahrungen und Ansätze in Erfahrung gebracht werden. Ferner wird versucht bisherige Probleme in Möglichkeiten umzuwandeln (How might we?). Ein klassisches Beispiel ist hier die Aussage einer Kundin, dass es bei technischen Problemen immer Ewigkeiten dauert, bis ihr geholfen wird. In der How might we-Form wird daraus: Wie können wir unseren Kunden:innen schneller bei technischen Problemen helfen? Das Wissen aus allen Schritten wird zusammengetragen und am Ende des Tages werden die wichtigsten Kunden:innen und Ziele identifiziert (pick a target).

Am Dienstag steht das Sketching im Vordergrund; der Fokus liegt somit auf Lösungen. Zunächst werden bestehende Lösungen (eigene oder Lösungen von Marktbegleitern) kritisch betrachtet (remix and improve). Im Anschluss werden Lösungen aus abstrakten Ideen generiert und bewertet. Dafür wird in Einzelarbeit zunächst notiert was wichtig sein könnte und es werden grobe Ideen skizziert. Eine Idee wird achtmal variiert, sodass innerhalb von acht Minuten acht Varianten derselben Idee entstehen (crazy 8s). Im Anschluss wird aus allen Vorarbeiten des Tages der Solution Sketch entwickelt. Dafür werden die besten Ideen durch das Team ausgewählt und verfeinert.

Mittwoch ist der Tag der Entscheidungen. Mithilfe eines fünfstufigen Entscheidungsprozesses werden die besten Ideen (von Dienstag) ausgewählt. Zunächst werden alle Sketches nebeneinander aufgehängt (art museum) und jedes Teammitglied nominiert ein bis maximal drei Elemente, die ihm:r am besten gefallen (heat map). Dies geschieht ohne Worte. Sobald das Team fertig ist, werden in der Gruppe pro Sketch die Highlights besprochen (Zeitansatz 3 Minuten, speed critique). Anschließend ermittelt jedes Mitglied seinen bzw. ihren Sieger ohne dies zu kommentieren, alle stimmen gleichzeitig ab (straw poll). Der:Die Entscheider:in bekommt drei Stimmen und kann diese entsprechend der eigenen Präferenz verteilen. Die Gewinner der Entscheidung werden in ein Storyboard übernommen, welches in der Regel aus fünf bis fünfzehn Szenen besteht.

Am Donnerstag wird der Prototyp erstellt, um diesen am Freitag mit Kunden:innen testen zu können. Daher liegt der Fokus meist auf der Benutzungsschnittstelle. Klassischerweise fährt man hierbei sehr gut mit einer Mischform aus experimentellem und explorativem Prototyping. Es können verschiedene Möglichkeiten ausgetestet und auf Realisierbarkeit geprüft werden und es sollte Einblick in eine Anwendungssituation gegeben werden, weshalb eine konkrete Funktion gut ausdefiniert werden sollte. Die Erstellung des Prototyps kann an einem einzigen Tag erfolgen, da bereits alle Entscheidungen getroffen und im Storyboard festgehalten wurden. Allerdings bietet es sich an, das Storyboard in kleinere Szenen aufzuteilen und einzelnen Teammitgliedern zuzuordnen (divide and conquer). Bezüglich der technischen Umsetzung empfehlen wir den Studierenden Prototypingtools wie Adobe XD, Figma oder SAP Build zu nutzen.

Freitag stehen der Test und somit die Präsentation gegenüber den Kunden:innen im Vordergrund. [KZK16] empfehlen dazu fünf Personen zu interviewen und die Interviews jeweils in fünf Blöcke zu unterteilen. Im Prinzip kann man sich unter dem Interview in fünf Akten eine begleitete Nutzung des Prototyps vorstellen; die Nutzerin soll den Prototyp selbst erkunden und die Nutzung kommentieren. Um das laute Denken zu forcieren, stellt der Interviewführende verschiedene Fragen. Gegen Ende des Interviews wird der:die Interviewte beispielsweise gebeten seine:ihre Erfahrungen zusammenzufassen.

In seiner ursprünglichen 5-Tage-Form ist der Google Design Sprint nicht perfekt geeignet für den Einsatz in unserer Veranstaltung. Wir empfehlen unseren Studierenden deshalb wie folgt vorzugehen: Am Montag gilt es das Problem zu verstehen, eine Kontextanalyse durchzuführen, die Ziele festzulegen und Nutzer:innen für die Evaluation zu rekrutieren. Am

Dienstag sollen Ideen gesammelt werden und aus den vielen verschiedenen Lösungen die eine Lösung identifiziert und verfeinert werden. Gegebenenfalls kann auch schon der Prototyp vorbereitet werden (zum Beispiel indem sich die Studierenden mit den Prototypingtools vertraut machen). Am Mittwoch wird der Prototyp erstellt und gegen Abend die Evaluation geplant. Am Donnerstag findet die Evaluation statt und die Präsentation für Freitag wird vorbereitet. Am Freitag wird das Ergebnis der gesamten Woche präsentiert. Hier bietet sich eine Unterteilung in Produktpräsentation und Präsentation der spezifischen Entscheidungen und angewandten Methoden an. Der Freitag schließt mit einem dedizierten Feedback für jede einzelne Gruppe ab.

3 Szenarien

In den zwei nachfolgenden Unterkapiteln werden die in den Wintersemestern 2019/2020 und 2020/2021 genutzten Szenarien vorgestellt. Das Konzept der Lehrveranstaltung erfordert von den Studierenden sich in vielfältige Fragestellungen zu einem Szenario einzuarbeiten. Die Szenarien sollen hierbei für die Studierenden anschaulich und übertragbar auf die persönliche Lebenssituation sein.

3.1 Urban Farming - Am Beispiel des digitalen Gewächshaus

Die stetige Digitalisierung und das Fortschreiten des Einflusses der Technik zeigt sich nicht nur bei Unternehmen oder dem Bildungssektor, sondern auch in anderen Branchen wie z. B. der Landwirtschaft [Ta16]. Der Trend zur Selbstversorgung hat Konzepte wie „Urban Farming“, „Vertical Farms“ oder „Digital Farming und Smart Greenhouses“ hervorgebracht bzw. zu einer Weiterentwicklung veranlasst [En11] [Eg15]. Urban Farming als Oberbegriff umfasst unterschiedlichste Aspekte landwirtschaftlicher Konzepte, die sich in einer städtischen Umgebung umsetzen lassen.

Das Smart Greenhouse oftmals auch „digitales Gewächshaus“ genannt, eignet sich als ein Aspekt von Urban Farming für die Vermittlung unterschiedlichster Fragestellungen mit BUIS-Bezug, da hierbei theoretische Methoden direkt in die Praxis umgesetzt werden können - wie z.B. Sensorik und die Verarbeitung von Umweltdaten bzw. umweltrelevanter Daten. Die Komplexität und Größe von digitalen Gewächshäusern ist beliebig skalierbar, so dass sich Studierende mit unterschiedlichem Vorwissen aus ihren Studienfächern in Kleingruppen das Wissen in den vorher genannten Aspekten aneignen und die Veranstaltung erfolgreich besuchen können. Das digitale Gewächshaus kann hierbei von den Studierenden prototypisch mit einer kostengünstigen Sensorik aufgebaut werden, um die Praktikabilität der im Rahmen des Google Design Sprint entwickelten Lösungen (z. B. automatische Belüftung und Beschattung durch Aktuatoren oder eine automatische Bewässerung) im städtischen Umfeld (z.B. des eigenen Balkons) zu testen. Die Möglichkeit der Umsetzung im eigenen Umfeld im Nachgang der Veranstaltung kann die Motivation der Studierenden

weiter erhöhen und einen nachhaltigen Effekt über die Lehrveranstaltung hinaus bewirken. Die Bedeutung von digitalen Gewächshäusern als ein Aspekt im Rahmen des Urban Farming wird auch durch das weltweite Marktvolumen für „Smart Greenhouses“ deutlich, das 1,37 Milliarden US-Dollar (in 2019) betragen hat und sich geschätzt bis 2027 auf 3,23 Milliarden US-Dollar erhöhen wird [BKK20].

3.2 Nachhaltige Mobilität am Beispiel des Radverkehrs

Die aktuelle Diskussion zur Klimaneutralität Deutschlands bis 2045 zeigt auf, dass alle Möglichkeiten der Einsparungen von CO₂ von der Wissenschaft, Politik und Gesellschaft in Erwägung gezogen werden müssen. Die persönliche Mobilität wird sich verändern: „Die Menschen fahren deutlich mehr mit öffentlichen Verkehrsmitteln oder mit dem Rad, und sie gehen mehr zu Fuß“ [Ta21]. Eine nachhaltige Mobilität beinhaltet somit auch den Radverkehr, wobei Städte mit einem hohen Anteil an Radverkehr in der Mobilität oftmals von Studierenden geprägt sind wie z. B. Münster mit einem Radverkehrsanteil von 39 % [Ta21].

Aus diesem Grund eignet sich im Rahmen des Konzeptes der Lehrveranstaltung das Szenario „Nachhaltige Mobilität am Beispiel des Radverkehrs“ besonders für eine anschauliche und auf den Alltag der Studierenden übertragbare Lebenssituation. Durch die persönlichen Erfahrungen der Studierenden kann insbesondere die Phasen des Verstehens und der Findung von Lösungsansätzen (Tag 1 und 2 im Google Design Sprint) besonders unterstützt werden. Eigens erlebte Alltagsituationen zeigen den Studierenden auf, an welcher Stelle Verbesserungen am Rad oder der Radinfrastruktur notwendig wären, um entweder die Akzeptanz und Nutzung des Radverkehrs zu erhöhen oder andererseits die Sicherheit zu erhöhen bzw. die Emissionen des Radverkehrs auf die Umwelt zu verringern. Dieses können Ansätze zur Optimierung der Routenplanung inklusive des Abstellens der Fahrräder sein oder Veränderungen am Rad (z.B. Toter-Winkel-Assistent oder eines Schalter für einen Blinker am Lenkrad, damit durchgehend zwei Hände am Lenker sind für eine bessere Kontrolle des Fahrrads) sein. Da viele Studierende ein eigenes Rad besitzen, kann auch eine Hardwareumsetzung der Lösungsansätze oftmals ohne großen Aufwand erfolgen und selbst erprobt werden, welches sich positiv auf die Motivation der Studierenden im Kurs auswirkt. Die rechtlichen Voraussetzungen zur Realisierung der von den Studierenden entwickelten Ideen werden in dem Kurs nicht weiter betrachtet, da die Methodiken mit dem Design Sprint sowie die Innovation über Brainstorming etc. im Vordergrund der Veranstaltung stehen.

4 Ausblick

Das aktuelle Konzept der Lehrveranstaltung wird von den Studierenden der unterschiedlichen Fachrichtungen gut angenommen. Besonders gelobt wird die Interdisziplinarität sowie die Umsetzung des Forschenden Lernens. Sehr gut gefällt den Studierenden, dass die Setzung eines eigenen Schwerpunkts innerhalb eines Dachthemas möglich ist.

Zukünftig sollen weitere Themen mit BUIS-Bezug bearbeitet werden. Ferner ist es vorstellbar alternative Vorgehensweisen auszuprobieren oder den aktuell genutzten Design Sprint um Aspekte des Design Thinkings zu erweitern.

Literaturverzeichnis

- [BKK20] Borasi, Pramod; Khan, Shadaab; Kumar, Vineet: Smart Greenhouse Market Size, Share & Growth | Analysis - 2027. Bericht, Allied Market Research, 2020.
- [Eg15] Egnolff, Mareike: Publikationen der Uds: Die Sehnsucht nach dem Ideal: Landlust und Urban Gardening in Deutschland. Dissertation, Universität Saarland, 2015.
- [En11] Engelhardt, Marc: Selbstversorger: Die Urban Farmers liegen im Trend. ZEIT ONLINE, 2011.
- [KZK16] Knapp, Jake; Zeratsky, John; Kowitz, Braden: Sprint: How to Solve Big Problems and Test New Ideas in Just Five Days. Simon & Schuster, 2016.
- [Le17] Leitungsteam des Projektes Forschungsbasiertes Lernen im Fokus (FLiF): Forschungsbasiertes Lehren und Lernen an der Universität Oldenburg. Bericht, Carl von Ossietzky Universität Oldenburg, 2017.
- [Ta16] Taenzer, Miriam: Positionspapier - Digitalisierung in der Landwirtschaft. Bericht, Bundesverband Informationswirtschaft, Telekommunikation und Neue Medien e.V., Berlin, Deutschland, 2016.
- [Ta21] Tagesschau: Laut Studie: Klimaneutralität schon 2045 möglich | tagesschau.de. Bericht, Norddeutscher Rundfunk Anstalt des öffentlichen Rechts, 2021.

Kennzahlen zur nachhaltigen Gestaltung von Mobilität in Unternehmen durch Umstellung auf Elektromobilität

Bettina Steden,¹ Jorge Marx Gómez

Abstract: In diesem Paper werden Kennzahlen zur Betrachtung der Mobilitätssituation eines Unternehmens vorgestellt. Das Ziel ist es, mit Hilfe dieser Kennzahlen zu untersuchen, ob die Fahrzeugflotte des Unternehmens durch Umstellung auf Elektromobilität nachhaltig gestaltet werden kann. Die Bedürfnisse des Standortes sollen vollständig erfasst werden, daher werden Anforderungen aus den Bereichen Ökologie, Ökonomie, Technik und Organisation abgedeckt. Die Kennzahlen wurden im Rahmen des Forschungsprojekts ELogZ identifiziert. Sie ermöglichen eine individuelle Ableitung von Anforderungen an ein Konzept zur Einführung von Elektromobilität.

Keywords: Nachhaltigkeit, Elektromobilität, Unternehmen, Mobilität

1 Einleitung

Elektromobilität wird als Möglichkeit nachhaltiger Antriebsalternativen angesehen [La19][AH19]. Im Unternehmenskontext bietet sich daher die Untersuchung an, ob eine Umstellung der Fahrzeugflotte auf Elektromobilität möglich sowie ökologisch und ökonomisch vorteilhaft ist. Dies erfordert jedoch vorangehende, sorgfältige Planung.

Um die Planung zu ermöglichen, werden in diesem Paper zentrale Kennzahlen vorgestellt, anhand derer die Mobilitätssituation in Unternehmen im Hinblick auf das Elektrifizierungspotential bewertet werden kann. Mithilfe dieser Kennzahlen soll die Ableitung von Anforderungen an ein Konzept zur Einführung von Elektromobilität ermöglicht werden. Die Kennzahlen wurden im Rahmen des Forschungsprojekts ELogZ - Energieversorgungskonzepte für klimaneutrale Logistikzentren in Zusammenarbeit mit Partnern aus Industrie und Forschung identifiziert. In diesem Rahmen wurde die Fahrzeugflotte eines Logistikzentrums untersucht. Die Übertragbarkeit der Kennzahlen auf Unternehmen in anderen Bereichen ist jedoch sichergestellt.

Kennzahlen stellen im Mobilitätsbereich ein wichtiges Mittel zur Strukturierung dar und wurden in diesem Bereich bereits vielfältig eingesetzt. So betrachtet [BB14] Mobilitätskennzahlen für den motorisierten Individualverkehr. Die Kriterien sollen für den Vergleich zwischen Elektrofahrzeug und herkömmlichen Fahrzeug herangezogen werden können. Er unterscheidet in quantitative sowie qualitative Kriterien in den für den Privatanutzer relevanten Bereichen Informationsbeschaffung, Nutzungsphase und Verkauf.

¹ Universität Oldenburg, VLBA, Ammerländer Heerstraße 114-118, 26129 Oldenburg, Deutschland bettina.steden@uni-oldenburg.de

In [JG10] werden verschiedene Definitionen von Kennzahlen angegeben, die in der Literatur verwendet werden, insbesondere im Bereich Umwelt und Nachhaltigkeit.

Technische Kennzahlen zum Vergleich zwischen Elektrofahrzeug und Verbrenner werden in [Fü17] betrachtet. Er unterscheidet in Bewertungs- und Effektkriterien und untersucht die Wechselwirkungen zwischen diesen Kriterien. Untersucht wurden zum Beispiel Reichweite und Wirkungsgrad.

In [Do09] wurde ein Indikatorenkatalog aufgestellt, um die Nachhaltigkeit des Transportsektors in den EU Staaten zu evaluieren. Der Katalog umfasst 55 technische, soziale, institutionelle, ökonomische und ökologische Kennzahlen.

[Kö13] stellt ein Informations- und Planungssystem zur Förderung nachhaltiger Mobilität vor. Um dieses System zu entwickeln, werden ökonomische, ökologische und soziale Kennzahlen genutzt, mit denen Nachhaltigkeit messbar gemacht werden soll.

2 Kennzahlen zur Bewertung der Mobilität

Damit sich die Kennzahlen zur Bewertung der Mobilitätssituation und der Ableitung möglicher Anforderungen an ein Einführungskonzept eignen, sind verschiedene Perspektiven abzudecken. Es muss aus ökonomischer Sicht auf die Situation geschaut werden, um die Kosten einzuschätzen, aus ökologischer, um die Nachhaltigkeit der Umstellung beurteilen zu können, sowie aus technischer und organisatorischer Sicht, um strukturelle Änderungen zu erfassen.

Darüber hinaus sollen die Kennzahlen unabhängig voneinander ermittelbar sein. Sie sollen einen Vergleich zwischen verschiedenen Antriebsarten ermöglichen sowie in Verhältnis zueinander gesetzt werden können. Im Folgenden werden die ermittelten Kennzahlen sowie die Evaluation an einem Beispielunternehmen vorgestellt.

Betrachtungsgegenstand der Kennzahlen ist die Mobilität, die durch die Mitarbeiter eines Unternehmens verursacht wird. Dies umfasst die Fahrzeugflotte bestehend aus Fahrzeugen mit Verbrennungsmotor sowie die Privatfahrzeuge der Mitarbeiter, die zur An- und Abfahrt an den Arbeitsplatz genutzt werden.

2.1 Kennzahlen

Eingesparte Kraftstoffe

Die Kennzahl eingesparte Kraftstoffe erfasst die Kosten und Treibhausgasemissionen, die durch die Nutzung von Kraftstoffen beziehungsweise Strom während des Betriebes der Fahrzeuge anfallen.

Zur Erfassung der Kosten in € werden die jährlichen Kosten, die durch Tanken bzw. Laden der Fahrzeuge entstehen, berechnet. Siehe Formel (1).

$$\text{Kraftstoffkosten/Jahr [€]} = \text{Verbrauch [l/100km]} \cdot \frac{\text{Kosten Kraftstoff [€/l]}}{100} \cdot \text{Jahresfahrleistung [km]} \quad (1)$$

Zur Berechnung der jährlichen Treibhausgasemissionen werden die Well to Wheel Emissionen (WTW) betrachtet. So werden die Emissionen, die während der Herstellung des Kraftstoffs entstehen, sowie die Emissionen, die das Fahrzeug bei der Nutzung des Kraftstoffs ausstößt, mit einbezogen [AH19]. Siehe Formel (2). Die Treibhausgasemissionen werden in CO_2 -Äquivalenten (CO_2e) angegeben. Diese umfassen die klimarelevanten Gase CO_2 , CH_4 und N_2O [Ch20].

$$\text{Em. Betrieb/Jahr [kg} CO_2e] = \text{Verbrauch [l/100km]} \cdot \frac{\text{Emissionsfaktor Kraftstoff [kg } CO_2e/l]}{100} \cdot \text{Jahresfahrleistung [km]} \quad (2)$$

Durch diese Kennzahl wird ein direkter Vergleich zwischen Kosten und Treibhausgasemissionen, die ein Fahrzeug durch die Nutzung von Kraftstoffen verursacht, ermöglicht. Neben dem Vergleich zwischen Kosten und Treibhausgasemissionen ist der Vergleich zwischen Elektrofahrzeug und Verbrennerfahrzeug möglich, um Einsparmöglichkeiten während des Betriebes zu erfassen.

Lebenszyklusanalyse

Zur Beurteilung, ob die Umstellung der Mobilität auf elektrisch angetriebene Fahrzeuge zu Einsparungen in den Treibhausgasemissionen führt, ist eine alleinige Betrachtung der eingesparten Kraftstoffe nicht ausreichend. So tragen die Herstellung und die Nutzung der Fahrzeuge bei den verschiedenen Antriebsmöglichkeiten unterschiedlich stark zur Treibhausgasbilanz bei [Ag19].

Um diese Unterschiede zu erfassen, stellt die Lebenszyklusanalyse (LCA) eine weitere Kennzahl dar. Durch die LCA werden neben den Emissionen während des Betriebes die Emissionen, die bei der Herstellung sowie der Entsorgung der Fahrzeuge anfallen, berücksichtigt. Siehe dazu auch [AH19]. Die Emissionen der Fahrzeugherstellung und Entsorgung werden dabei auf die Nutzungsjahre verteilt, siehe Formel (3), vgl. [La19].

$$\text{Em./Jahr [kg } CO_2e] = \frac{\text{Em. Herstellung} + \text{Em. Entsorgung}}{\text{Nutzungsjahre}} + \text{Em. Betrieb/Jahr} \quad (3)$$

Im Kontext der Kennzahlen ist die LCA als Leitindikator bei der Erfassung der Unterschiede in den Treibhausgasemissionen zwischen Verbrennerfahrzeug und Elektrofahrzeug anzusehen, da alle anfallenden Emissionen abgedeckt werden. Geringere Emissionen werden als positiv bewertet.

Total Cost of Ownership

Ähnlich wie bei den Treibhausgasemissionen ist eine alleinige Betrachtung der Kraftstoffkosten nicht ausreichend, um Kosten von Elektrofahrzeugen und Verbrennerfahrzeugen

zu vergleichen. Da Elektrofahrzeuge in den Anschaffungskosten häufig teurer sind, im Betrieb jedoch Kostenvorteile entstehen [La19][WIB15], wird die Betrachtung der Kosten eines Fahrzeugs über die Besitzdauer hinweg (Total Cost of Ownership – TCO) als weitere Kennzahl gewählt.

In der TCO werden neben den Anschaffungskosten und dem Restwert am Ende der Halte-dauer die weiteren Kosten für Versicherung, Haupt- und Abgasuntersuchung, Kfz-Steuer und Reparaturkosten berücksichtigt. Hinzu kommen die jährlichen Kraftstoffkosten aus Formel (1). Vergleiche dazu [HvWM15]. Da es sich bei den Anschaffungskosten und dem Restwert um einmalige Kosten handelt, die übrigen jedoch regelmäßig anfallende Kosten sind, werden Erstere durch die Anzahl der Nutzungsjahre geteilt. So kann die TCO in jährlichen Kosten angegeben werden, siehe (4).

$$\text{TCO/Jahr [€]} = \frac{\text{Anschaffungskosten-Restwert}}{\text{Nutzungsdauer}} + \text{Kraftstoffkosten/Jahr} + \text{weitere jährliche Kosten} \quad (4)$$

Eine Möglichkeit zur Berechnung der TCO mit Diskontierung zukünftiger Kosten ist ebenfalls in [WIB15] nachzulesen. Die TCO ist als Leitindikator bei der Erfassung der Kostenunterschiede zwischen Verbrennerfahrzeug und Elektrofahrzeug anzusehen. Geringere Kosten werden als positiv bewertet. Ein Vergleich zwischen TCO und LCA eines Fahrzeuges ist ebenfalls möglich.

Zufriedenheit der Mitarbeiter

Das Verhalten und die Zufriedenheit der Mitarbeiter spielen eine wichtige Rolle bei der Umgestaltung der Mobilität, da diese unmittelbar von Veränderungen betroffen sind. Die Kennzahl Zufriedenheit der Mitarbeiter zielt daher darauf ab, die Mitarbeiter und deren Bedürfnisse durch eine Mitarbeiterbefragung mit einzubinden.

Die Emissionen, die die Privatfahrzeuge der Mitarbeiter während der An- und Abreise zum Standort verursachen, sind den Emissionen des Standortes zuzurechnen. Daher zählt zur Einbindung der Mitarbeiter die Kenntnis über die Distanzen zwischen Wohn- und Arbeitsort. Mit der Kenntnis über die Distanzen zwischen Wohn- und Arbeitsort erfolgt die Berechnung der Emissionen, wie in Formel (2) unter eingesparte Kraftstoffe vorgestellt.

Zur Einbindung der Mitarbeiter zählt ebenfalls die Bereitschaft, privat Elektrofahrzeuge zu nutzen und diese während der Arbeitszeit am Standort zu laden. Um daraus resultierende Energiebedarfe und Bedarfe an der Nutzung von Ladesäulen einzuschätzen, ist die Erfassung der Arbeitszeiten und -tage erforderlich. An Tagen, an denen die Mitarbeiter vor Ort sind, lässt sich der Energiebedarf wie folgt abschätzen.

$$\text{Energiebedarf [kWh]} = \frac{\text{Verbrauch [kWh/100km]}}{100} \cdot \text{zurückgelegte Strecke [km]} \quad (5)$$

Insgesamt erhöht die Kennzahl die Auswertungsqualität der vorgestellten Kennzahlen und damit den Erfolg der Maßnahmen.

Technische Unterschiede

Die Erfassung der technischen Unterschiede zwischen Verbrennerfahrzeugen und Elektrofahrzeugen wird als weitere Kennzahl aufgenommen. Es sind die Unterschiede zu betrachten, die eine Auswirkung auf die tägliche Nutzung sowie die Infrastruktur des Standorts haben. Relevant sind daher die eingeschränkte Reichweite von Elektrofahrzeugen und die unterschiedlichen Anforderungen an die Kraftstoffbereitstellung.

Die Reichweitenbeschränkung erfordert eine Analyse der regelmäßig zurückgelegten Strecken und einen Vergleich mit der Reichweite verschiedener Elektrofahrzeuge (Formel (6)). Ein Elektrofahrzeug kommt als Ersatz infrage, falls alle regelmäßig zurückgelegten Strecken innerhalb der Reichweite des Fahrzeugs liegen. Da die angegebenen Reichweiten von Elektrofahrzeugen je nach Nutzung und Jahreszeit schwanken [Si21], wird ein Puffer von 30% einkalkuliert.

$$\text{Zurückgelegte Strecke [km]} \leq \text{Reichweite Elektrofahrzeug [km]} \cdot 70\% \quad (6)$$

Diese Kalkulation nimmt an, dass unterwegs keine Ladung des Fahrzeugs erfolgt, jedoch zwischen zwei Fahrten eine Ladung möglich ist.

Mit den unterschiedlichen Anforderungen an die Kraftstoffbereitstellung werden die Ladeleistung der Elektrofahrzeuge und die zu errichtenden Ladesäulen betrachtet. Diese sind aufeinander abzustimmen, da einige Fahrzeuge über eine beschränkte Ladeleistung verfügen. Im Zusammenhang mit der Reichweite kann eine Analyse der unterwegs verfügbaren Ladestationen sowie der Ladeleistung des Fahrzeugs zur Aufweichung der Reichweitenrestriktion beitragen.

Insgesamt bestimmt diese Kennzahl, inwieweit die Umstellung aus technischer Sicht möglich ist. Sie kann die Ergebnisse der Kennzahlen LCA und TCO einschränken.

Auslastung und Verfügbarkeit

Durch Umstellung der Fahrzeugflotte auf Elektrofahrzeuge ändert sich die Auslastung der Fahrzeuge. Zu erwarten ist, dass sie aufgrund der längeren Ladedauern von Elektrofahrzeugen steigt. Bei einem Fahrzeug, das bereits eine hohe Auslastung aufweist, kann die Ersetzung durch ein Elektrofahrzeug dazu führen, dass durch die zusätzlichen Ladezeiten nicht mehr alle Fahrten bedient werden können. Daher ist eine Abschätzung der Ladezeiten notwendig. Diese erfolgt mithilfe der Ladeleistung und des Energiebedarfs [Br20], siehe (7).

$$\text{Ladedauer [h]} = \frac{1}{\alpha} \cdot \text{Energiebedarf [kWh]} \cdot \text{Ladeverlust [\%]} \quad (7)$$

mit $\alpha = \min(\text{Ladeleistung Ladesäule [kW]}, \text{Ladeleistung Fahrzeug [kW]})$

Die Betrachtung der Auslastung ist somit als Kennzahl notwendig, um Anforderungen an die Ladeleistung, Reichweite und Anzahl der Elektrofahrzeuge sowie an die Ladesäuleninfrastruktur zu identifizieren.

Eine Auslastung, die die Fixkosten der Fahrzeugflotte minimiert und den Mobilitätsbedarf deckt, wird als positiv angesehen. Des Weiteren kann diese Kennzahl zu Einschränkungen der Ergebnisse der Kennzahlen LCA und TCO führen und stellt eine Ergänzung der Betrachtung der technischen Unterschiede dar.

2.2 Evaluation an einem Beispielunternehmen

Die Evaluation der vorgestellten Kennzahlen erfolgt im Rahmen des Projektes ELogZ. Die dort untersuchte Fahrzeugflotte besteht aus vornehmlich dieselbetriebenen Dienstwagen. Im Laufe dieses Projektes soll die Fahrzeugflotte auf Elektromobilität umgestellt sowie die passende Ladeinfrastruktur am betrachteten Unternehmen errichtet werden. Dazu ist die Auswertung der Fahrzeugflotte anhand der Kennzahlen erforderlich.

Die Fahrzeuge zeichnen sich überwiegend durch lange Standzeiten während der Arbeitszeit auf dem Gelände des Unternehmens aus. Dies bietet eine hohe Flexibilität bei der Auswahl der Ladesäulen. Die regelmäßig zurückgelegten Strecken dieser Fahrzeuge liegen je nach Fahrzeug zwischen 15 und 500 km. Inklusive eines Puffers von 30 % ergeben sich Anforderungen an die Reichweite Elektrofahrzeugen von bis zu 650 km. Da eine derartige Reichweite bisher nur von wenigen Elektrofahrzeugen erreicht werden kann, werden Plug-In Hybrid Fahrzeuge als weitere Option mitberücksichtigt.

Das Verhalten der Mitarbeiter wurde mithilfe einer Umfrage erfasst. Das Ergebnis zeigt, dass nur wenige Mitarbeiter dazu bereit sind, sich Elektrofahrzeuge anzuschaffen. Genannte Gründe gegen die Anschaffung waren vor allem der Preis. Daher ist ein geringer Einfluss der Mitarbeiter auf den Energiebedarf und die Nutzung der Ladesäulen zu erwarten. Insgesamt beläuft sich der erwartete Energiebedarf durch Privatfahrzeuge der Mitarbeiter auf 9 kWh pro Arbeitstag.

Zur Analyse der Fahrzeugflotte anhand ökologischer und ökonomischer Kennzahlen wurde ein Vergleich zwischen den aktuellen Fahrzeugen und möglichen Elektrofahrzeugen durchgeführt. Die zum Vergleich genutzten Elektrofahrzeuge sind so gewählt, dass eine möglichst breite Spanne an Reichweiten abgedeckt ist (Tab. 1). Der Vergleich zwischen Elektrofahrzeug und konventionellem Fahrzeug erfolgte jeweils unter der Annahme, dass die Elektrofahrzeuge mit Strom von 21 ct/kWh und 463 g CO_2e/kWh geladen werden. Um die Produktions- und Entsorgungsphase der LCA mit in die Analyse einzubeziehen, wurden Werte aus [Ag19] herangezogen.

Die Fahrzeugflotte weist im Mittel eine Jahresfahrleistung von 30.000 km und eine Nutzungsdauer von 5 Jahren auf. Die Ergebnisse, die der Vergleich eines Flottenfahrzeugs mit eben diesen Parametern sowie einem Verbrauch von 4,8 l und Anschaffungskosten von 32.000 € ergibt, sind in Tab. 2 dargestellt. Erkennbar ist, dass vor allem Elektrofahrzeuge mit hohen Reichweiten aufgrund der meist höheren Batteriekapazität und des höheren Anschaffungspreises im Vergleich mit dem bestehenden Fahrzeug schlechter abschneidet.

Der Stand der Untersuchungen an dem Beispielunternehmen zeigt auf, dass die Betrachtung der Kennzahlen aus allen vorgestellten Bereichen zur Einschätzung der Mobilitätssituation nötig ist.

Fahrzeug	Art	Reichweite [km]	Batteriekapazität [kWh]	Verbrauch [kWh/100km]	Verbrauch [l/100km]	Anschaffungskosten [€]
A1	E-Fahrzeug	100	16	17	0	21.250
A2	E-Fahrzeug	230	33	18	0	34.707
A3	E-Fahrzeug	258	36,8	14,5	0	21.421
A4	E-Fahrzeug	305	42	15	0	35.336
A5	E-Fahrzeug	320	50	17,7	0	37.715
A6	E-Fahrzeug	385	62	18,5	0	37.237
A7	E-Fahrzeug	411	85	22,6	0	69.484
A8	E-Fahrzeug	455	67	15,9	0	38.105
A9	E-Fahrzeug	560	80,5	16	0	52.980
A10	E-Fahrzeug	652	100	18,9	0	77.970
A11	Plug-In-Hybrid	73	15,6	17,1	1	37.765

Tab. 1: Parameter der zum Vergleich ausgewählten Elektrofahrzeuge mit Angaben aus [AD]

Fahrzeug	Reichweite [km]	LCA Differenz [kg CO ₂ e/Jahr]	TCO Differenz [€/Jahr]
A1	100	1538	3157
A2	230	881	1616
A3	258	1251	3713
A4	305	1023	891
A5	320	403	1094
A6	385	-74	383
A7	411	-1346	-5547
A8	455	135	256
A9	560	-291	-2757
A10	652	-1289	-6906
A11	73	718	-123

Tab. 2: Ergebnisse des Vergleichs der LCA und TCO eines Dieselfahrzeugs mit Elektrofahrzeugen am Beispiel eines Fahrzeugs mit 30.000 km Jahresfahrleistung

3 Ergebnisse

Es wurden Kennzahlen vorgestellt, anhand derer untersucht werden kann, ob die Mobilität in Unternehmen durch Umstellung auf Elektromobilität nachhaltiger gestaltet werden kann. Die Kennzahlen decken ökonomische, ökologische sowie technische und organisatorische Gesichtspunkte ab, so entsteht ein umfassendes Bild der Mobilitätssituation.

Mit den Kennzahlen werden unterschiedliche Ziele verfolgt. Die Kennzahl eingesparte Kraftstoffe dient dem Vergleich zwischen Kosten und Emissionen der Fahrzeuge im Betrieb. Die LCA und TCO ermöglichen einen ökologischen und ökonomischen Vergleich zwischen Verbrennern und Elektrofahrzeugen. Mithilfe der Kennzahlen Auslastung, technische Unterschiede und Zufriedenheit der Mitarbeiter werden Anforderungen an Fahrzeuge und Ladeinfrastruktur identifiziert. Weiterhin können diese zur Einschränkung der Ergebnisse der ökonomischen und ökologischen Kennzahlen führen.

Zwischen einigen Kennzahlen besteht somit ein Zielkonflikt. Gleichzeitig ist beim Vergleich verschiedener Kennzahlen untereinander Vorsicht geboten. So ist ein Vergleich zwischen eingesparten Kraftstoffen in CO_2e und der TCO aufgrund der unterschiedlichen Komponenten, die in die Analyse mit eingegangen sind, nicht sinnvoll.

Um aus den Kennzahlen ein Einführungskonzept ableiten zu können, ist daher eine Gewichtung der Kennzahlen vorzunehmen. Die Gewichtung erfolgt standortindividuell. Dies kann beispielsweise mithilfe von Promethee geschehen. Promethee wurde unter anderem in [BVM86] dargestellt und dient zur multikriteriellen Entscheidungsunterstützung, wie hier vorliegend.

Dieses Paper beschränkt sich auf die Einführung von Elektromobilität. Nachhaltige Mobilität umfasst jedoch mehr als den Umstieg auf Elektromobilität. Andere Kriterien, die zur nachhaltigen Gestaltung führen, erfordern weitere oder veränderte Kennzahlen.

Die Arbeit ist Teil des Projektes ELogZ. Das Projekt wird vom Bundesministerium für Wirtschaft und Energie (BMWi) gefördert.

Literaturverzeichnis

- [AD] ADAC: , ADAC Autokatalog. <https://www.adac.de/rund-ums-fahrzeug/autokatalog/marken-modelle/>. Aufgerufen am: 24.07.2021.
- [Ag19] Agora Verkehrswende: Klimabilanz von Elektroautos. Einflussfaktoren und Verbesserungspotenzial. 2019.
- [AH19] Ajanovic, Amela; Haas, Reinhard: On the Environmental Benignity of Electric Vehicles. *J. sustain. dev. energy water environ. syst.*, 7 (3):416–431, 2019.

- [BB14] Bertram, Mathias; Bongard, Stefan: Elektromobilität im motorisierten Individualverkehr Grundlagen, Einflussfaktoren und Wirtschaftlichkeitsvergleich. Springer Vieweg, 2014.
- [Br20] Bruckmüller, Thomas: Entwicklung einer Methodik zur Bewertung des Ladeinfrastrukturbedarfs für Elektrofahrzeuge hinsichtlich Anzahl, Kosten und Auswirkung auf die Energieversorgung am Beispiel Österreichs bis 2030. Dissertation, Technische Universität Wien, 2020.
- [BVM86] Brans, J.P.; Vinke, Ph.; Mareschal, B.: How to select and how to rank projects: The Promethee method. *European Journal of Operational Research*, 24:228–238, 1986.
- [Ch20] Choia, Wonjae; Yoo, Eunji; Seol, Eunsu; Kim, Myoungsoo; Song, Han Ho: Greenhouse gas emissions of conventional and alternative vehicles: Predictions based on energy policy analysis in South Korea. *Applied Energy*, 265:1–17, 2020.
- [Do09] Dobranskyte-Niskota, A; Perujo, A; Jesinghaus, J; Jensen, P. , Indicators to Assess Sustainability of Transport Activities, 2009.
- [Fü17] Füßel, Andreas: Technische Potenzialanalyse der Elektromobilität Stand der Technik, Forschungsausblick und Projektion auf das Jahr 2025. Springer Vieweg, 2017.
- [HvWM15] Hacker, Florian; von Waldenfels, Rut; Mottschall, Moritz: Wirtschaftlichkeit von Elektromobilität in gewerblichen Anwendungen Betrachtung von Gesamtnutzungskosten, ökonomischen Potenzialen und möglicher CO₂-Minderung. Abschlussbericht, 2015.
- [JG10] Joumard, Robert; Gudmundsson, Henrik: Indicators of environmental sustainability in transport: An interdisciplinary approach to methods. *INRETS*. INRETS, S. 422, 2010.
- [Kö13] Kölpin, Sven: Informations- und Planungssystem für nachhaltige Mobilität. In (Marx Gómez, Jorge; Lang, Corinna; Wohlgemuth, Volker, Hrsg.): *IT-gestütztes Ressourcen- und Energiemanagement*, S. 259–270. Springer Vieweg, 2013.
- [La19] Labode, Jonas: Evaluation des Einsatzes von Elektrofahrzeugen in betrieblichen Fahrzeugflotten mit spezifischen Methoden des Carsharing; Untersuchungen von Nachhaltigkeit und Wirtschaftlichkeit sowie Ansätzen zur Steigerung dieser. Dissertation, Universität Hildesheim, 2019.
- [Si21] Siebenpfeiffer, Wolfgang: *Mobilität der Zukunft Intermodale Verkehrskonzepte*. Springer-Verlag GmbH Deutschland, 2021.
- [WIB15] Wu, Geng; Inderbitzin, Alessandro; Bening, Catharina: Total cost of ownership of electric vehicles compared to conventional vehicles: A probabilistic analysis and projection across market segments. *Energy Policy*, 80:196–214, 2015.

Parametrizing Complex Co-Simulations to Support Decision Making in Mobility

Leonard Stepien ¹, Frank Köster²


Abstract: The radical change in mobility calls for the simulative support of decision making for large-scale projects which asks for the simultaneous consideration of multiple stakeholders additionally. The major challenge in such simulation systems is nowadays the consistent and modifiable tool-independent parameterization to enable scenario analysis by systematic parameter variation and reach reliable support for those extensive decisions. Taking the example of public charging infrastructure, a holistic co-simulation framework is proposed, in which an approach for the central parameterization as well as the belonging scenario analysis tool take part. Necessary functionalities as well as the belonging interfaces are revealed. Proposing a loop of mainly getting and setting parameters automatically, a domain ontology is required to reach standardization of naming conventions. With this work, a research gap in the filled of co-simulation is addressed by proposing a framework which allows the central parametrization of such systems to enable decision support.

Keywords: Co-Simulation, Central Parameterization, Decision support, Electric Mobility, Charging Infrastructure

1 Introduction

1.1 Motivation and Aim

Currently the mobility undergoes changes in manifold directions: Intermodal connections, the change of user behaviour as well as technological changes due to political decisions and technical progress. Regarding the high required invest and the long-term implications to evolve mobility systems, decisions have to be taken carefully and under consideration of relevant interdependencies, although an uncertainty will remain. As future developments cannot be experienced economically in real-life on a large-scale base, simulative analysis to support decisions and regard foreseeable scenarios is a suitable approach.

¹ ITK Engineering GmbH, Frankfurter Straße 5, 38122 Braunschweig leonard.stepien@itk-engineering.de 
<https://orcid.org/0000-0001-7392-5977>

² DLR Institut für KI-Sicherheit in Sankt Augustin & Ulm, Lilienthalplatz 7, 38108 Braunschweig,
frank.koester@dlr.de

Those foreseeable scenarios shall cover the realistic range of future developments, which is technically a parameter space consisting of the modelled behaviour of the stakeholders and their interdependencies. From this parameter space, the relevant combinations have to be derived as scenarios for the simulation. The consideration of different scenarios within the same simulation can be described as scenario variation and calls for the ability of the used simulation environment for systematic parameter variation. Only this systematic parameter variation enables the aimed decision support by analyzing and comparing different scenarios.

Simulation is widely established in the mobility domain, e. g. regarding traffic [Lo18, MDC19] or vehicle relocation in car sharing [Ke09], which allows the visualization and analysis of certain developments and changes.

Within this domain, the simulative approach is challenging due to multiple stakeholders and required adaptations to local circumstances. Therefore, the approach must be flexible, adoptable and calls for a high reliability of the results to meet the demand of such cyber-physical systems.

Co-simulation meets the described challenges: First, it allows the domain-specific modelling for the stakeholders' perspectives. Second, models can be exchanged within reasonable effort to make local adaptations or to later add a further model. Moreover, models can be reused for different purposes: E. g. a traffic simulation can be not only used for traffic light optimization but as well to identify the demand for car parks time- and location-specifically.

The challenge arises when a co-simulation shall be enabled for scenario variation and therefore, parameter variation respectively. By the example of planning public charging infrastructure for electric vehicles (EV), the needs as well as the challenges are revealed. To overcome EV drivers' concerns such as range anxiety and necessary detours for charging [KKS13, AG20], public charging infrastructure is regarded as a crucial point for the market penetration of EVs and is therefore in political focus as an urgent issue [Bu19, Be19, KKS13]. Only with sufficient public charging infrastructure, the emissions associated with individual mobility as established nowadays can decrease with the desired speed.

In general, the framework enables careful decision making in the mobility sector under consideration of multiple relevant aspects. This facilitates necessary adaptations in the mobility sector to meet the demands of climate change and changing users' request. Consequently, it contributes to the field of green software. The concrete benefits of the presented approach depend on the concrete application. Therefore, the benefits regarding environmental and sustainable aspects are explained by the example of planning public charging infrastructure in the following.

The efficient location of charging infrastructure ensures several point: The charging points are frequently used, as they are placed according to the driver's behavior and under consideration of specific local circumstances. In contrast to private or semi-public charging points, the accessibility is not restricted. Without restrictions, the achievable capacity utilization is higher and the overall amount of charging points can be reduced. Moreover, the simulation allows an enhanced precision of necessary charging stations, which prevents the authorities from installing too many charging stations. As the production and installation process of such charging systems asks for significant use of resources and investments, the improved planning procedure helps to decrease the environmental impact caused.

1.2 Status Quo and Challenges in Co-Simulation

Cyberphysical systems call for the simultaneous analysis of relevant subsystems. Due to the heterogeneity a unified modelling or monolithic approach is not suitable. Instead, the use of specialized software packages and modelling approaches is promising. The necessary merge of those distributes models can then be reached by co-simulation. This is generally defined as a coupled simulation of independent black boxes with an orchestrator – hereafter master algorithm – to coordinate the simulation. [Pa17]

A co-simulation scenario is regarded as the obtaining of a correct co-simulation by providing the necessary information. This necessary information mainly comprises the inputs, outputs, and their experimental frames beside others. There are four main types of co-simulation which have to be considered within the domain: discrete-event (DE), continuous-time (CT) and hybrid approaches combining those two. Depending on the major coordination of the overall simulation, there are either called hybrid DE or hybrid CT simulations. [Go18] As time is the major unit in the mobility domain and there are both, discrete as well as continuous systems, the type of hybrid CT simulation is the most relevant here.

Standardization of co-simulation is aimed in multiple projects which results in several co-existing standards. The High Level Architecture (HLA) focuses on simulation interoperability by providing a runtime infrastructure and is also applicable on cyber-physical systems [Da97, Na19]. The functional mock-up interface (FMI) standard concentrates on the standardization of the model's interface by defining a model description standard to enable the exchange of simulation models as a container [Ma20]. Related to the FMI standard, the System Structure and Parameterization (SSP) standard aims to provide a tool independent standard to enable the parameterization of composed simulation systems. The SSP standard provides an opportunity to provide a container of several pre-defined FMUs to be used as a container. [Ma19] Next to the technical challenges to be overcome for co-simulation, the consideration of different sub-domains calls for the standardization of used terms to address semantic challenges in parameterization. Exemplary this is conducted in [Te20] in the related domain of power and energy systems.

The issue of systematic parameter variation for scenario analysis has been addressed in the context of automated driving among others. The field of safety assessment for automated driving calls for the analysis of a variety of defined scenarios which has to be conducted also in simulation. Therefore, in [Na19] an approach is presented, in which the parameter setting is conducted via a specific tool, from which the further application are supplied.

A related co-simulation approach towards the planning of charging infrastructure is proposed in [MBN16]. Montori et al. integrate “Simulation of Urban Mobility” (SUMO) as an urban traffic simulator with an electric power system simulator ElectroMagnetic Transient program (EMTP-rv) via the open-source simulation environment OMNeT++. The parameterization itself is not conducted centrally but rather model-individually, completed by a synchronization mechanism. Binder et al. pursues a bottom-up approach towards scenarios in co-simulation by creating comprehensive scenario descriptions from co-simulation runs using the smart grid co-simulation framework Mosaik [Bi20]. In the domain of planning charging infrastructure, known co-simulation frameworks build on decentralized approaches for parameterization [MBN16, Na19], which limits their value for scenario analysis and decision support.

Having summarized the state of the art, the challenges to be overcome are introduced. There is mainly the reuse of existing models, the manifold synchronization of models and finally, the systematic parameter variation to enable the scenario analysis and the subsequent optimization for decision support.

Model Reuse

Models have been developed once to serve in a specific environment and a defined purpose. In the context of mobility, this might be a grid model originally developed to plan a factory location which should be now used for charging infrastructure. Therefore, it is necessary to find ways which allow the reuse with only small adaption which can be ideally conducted without expert-knowledge on the model. Regarding the parametrization several issues might occur:

Variable only part in a specific model: If a variable can be only found within one model which is not supported by the other included models, this variable must be handled via the parameterization. Either by setting a constant value or by a function which uses further variables from the participating models. If a parameter is only part of a specific model this can be directly handled via the central parameterization without regarding any interdependencies.

Variable names different: Either exchanged variables as well as parameters might occur in different models with different names. Furthermore, the variables should possibly appear with a more user-friendly name than it is handled internally in the simulation analysis tool. Those naming differences have to be recognized and solved for the actual simulation runs.

Model Exchangeability

The model exchangeability is closely related to the model reuse. When it is aimed to reuse existing models, e. g. for local adaptations or the consideration of a specific stakeholder, those models shall be integrated with limited effort required. Therefore, the framework should deal with different degrees of models' accuracy, scope, and interfaces. Models' accuracy means the outputs' resolution in terms of details, e. g. spatially, as well as the calculation frequency or step size. With the models' scope, its included content is meant. E. g. if a model only considers technical or technical and financial aspects. The models' interfaces are relevant for the connection of inputs and outputs between the participating models within the framework.

Models' Synchronization

In recent research the simultaneous time progress has been widely addressed. Regarding the parameterization, it is important to ensure that all models have the same values for identical parameters (see above). In the given context the pricing for electric energy have to be consistent for all models. In particular with to regard to the systematic scenario variation it has to be ensured that all variables and parameters are adopted consistently to the scenario. Furthermore, during simulation runs the variables have to be exchanged to all relevant parts of the simulation framework according to a determined order.

Systematic Parameter Variation for Scenario Analysis

This hypothesis is closely related to the models' synchronization above. The systematic parameter variation has two major challenges: The technical aspect of setting and getting the parameter values and the content-wise related perspective towards the parameter values and realistic combinations of them. Therefore, the values' variation between the parameters in distinct simulation runs have to be defined for the parameters or a group of parameters depending on the domain. Afterwards, the decision on how to vary the combination of parameters efficiently has to be taken. Due to the high number of parameters within mobility issues full-factorial plans are unlikely to be suitable.

1.3 Main Contribution and Paper Structure

This paper reveals the required functionalities to overcome the aforementioned challenges in parameterization and embed them in a general, modular, and tool-independent co-simulation framework that serves the decision support in current mobility issues. Exemplary explained by the field of planning public charging infrastructure, the required functionalities, their interaction and remaining challenges regarding implementation and adaptation are addressed.

Hereafter, the framework for decision support by co-simulation introduced. This includes the presentation of its aim, components, and workflow.

Thereafter, the concrete functionalities divided into the simulation preparation to ensure a consistent parameterization and the simulation and analysis section are elaborated. In parallel the belonging interactions and support concepts are introduced as they appear. This contribution concludes with a discussion and conclusion of the work followed by an outlook on future work.

2 Framework for Decision Support by Co-Simulation

The presented work is summarized in figure 1. Overall components and the major division are shown in blue, the concrete required functionalities and components in grey. The details on this framework are given in the following. The main information and workflow are highlighted with bold arrows, whereas the others represent supporting functionalities.

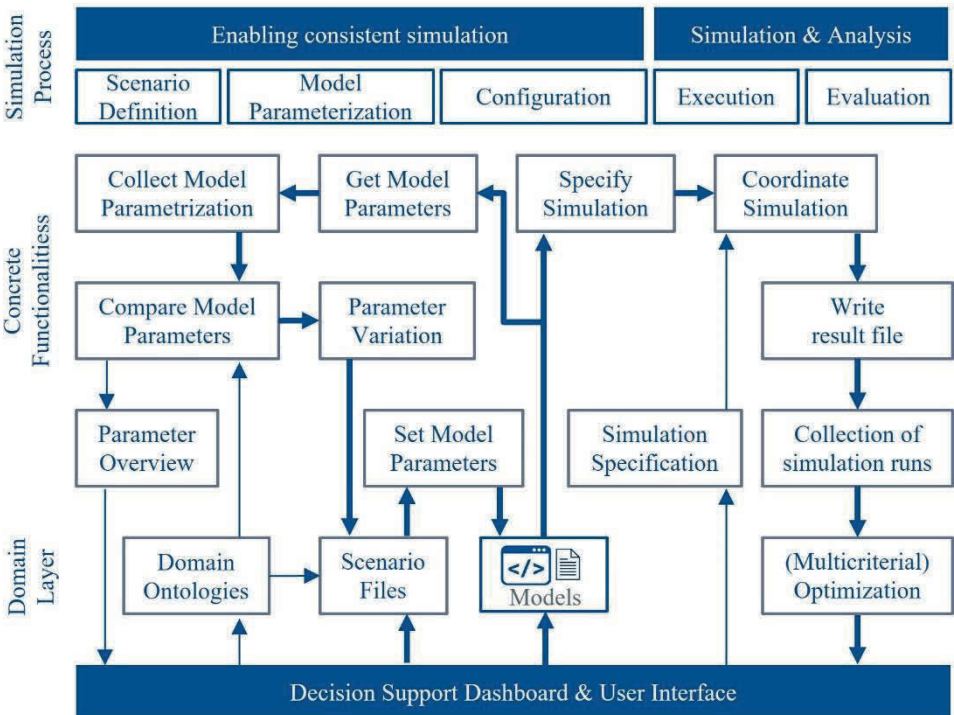


Fig. 1: Framework for Decision Support by Co-Simulation

The proposed framework should overall support decisions in the mobility sector which do require the consideration of multiple stakeholders. Their interests, relevant behaviour and causalities usually are modelled separately from each other due to the heterogeneity in

tooling and content-wise focus. To overcome existing differences in terms of parameters, such as default-values, naming and required interconnections, this framework provides the required functionalities and additionally analyses the user's interaction.

The framework is organized as a matrix with the procedure horizontally and the layers vertically clustering the contents for each step: First, there are the generally applicable layers with the simulation process and the concrete functionalities addressing the arising challenges. Second, it follows the domain layer which is context-specific and includes the user interface as well as the decision support dashboard.

Horizontally, the framework has two major segments spanning the range from starting with models and scenario ideas until reaching optimized results and comparison between different simulation runs: At first, the pre-simulation section to ensure the consistent parameterization needs to be handled. The scenario definition, the model parameterization and partly, the simulation configuration are belonging steps. It follows the simulation and analysis part which mainly contains the simulation execution and the evaluation. Details on the framework's components are elaborated in the following sections, separated in the pre-simulation side "Enabling Consistent Simulation" and its (post-)simulation counterpart "Simulation and Analysis for Decision Support".

3 Enabling Consistent Simulation

Beginning with the consistent parameterization, the required functionalities are presented which incorporate the process steps of scenario definition, model parameterization and partly the simulation configuration. Required inputs from user's side compromise the models to be included, the first approach of the scenario files and the domain ontology.

In the beginning, the types of variables to be handled must be defined. Within the framework, there are three main types of variables to be regarded: Constants, Model inputs and outputs as well as variables for analysis.

Constants are those simulation variables, which are not changed during a simulation run. An example for this type in the context of charging infrastructure is the price for energy. To reach a consistent parameterization, a defined price must be set in all participating models which include that parameter. Consequently, this type of simulation variable must be modifiable during the initialization of the simulation run and fixed afterwards.

The models' inputs and outputs must be served and routed during the entire simulation run. An example in this context, is the current power flow at a specific charging point. For those variables, it is important that the exchanged values match all the same time stamp at any time during the simulation run between the models. Regarding the central parameterization, the initial values must be set accordingly. In parallel to the constants, the aim of consistent parameterization calls for concise initial values between all models.

Moreover, it might be necessary to define an order, in which the models are executed and exchange those variables. This task is conducted by the master algorithm.

As a third type of simulation variables, the class of analysis variables must be considered. Those variables are not directly used within the simulation run, but are required for sophisticated analysis towards decision support. An example could be the average daily degree of capacity utilisation of a charging point. Whereas during the simulation run, the status quo of the charging point might be exchanged, the decision support can be further enhanced by investigating the daily situation based on weekdays. Therefore, the desired analysis variables have to be defined in advance, then the necessary information has to be made available in the individual model and last, this information must be prepared for analysis.

For the consistent parameterization, the variable types constants and the initial values for the models' input are most relevant. Therefore, those parameters have to be identified with the relevant information. This comprises their name, initial or fixed value and, if applicable, the unit. The difficulty of this step lies in the potential limited accessibility of black-box-models' variables. Therefore, the successful application of the proposed framework requires the visibility of all relevant parameters as well as their accessibility to overwrite values for consistency across models. The implementation of this function depends on the models' standard.

The second required functionality is the collection of the models' parameters in a structured manner with the aforementioned scope. This does not necessarily require the format of an external standard as this is internal within the simulation framework. The domain ontology supports this function by revealing potential overlapping parameters. From this status, a user overview is provided which enables the user to further intervene or analyse overlapping and similar parameter. Optionally, the naming of the parameters can be standardized via the application of the domain ontology.

Having created and cleared the parameters' overview, the preparation for the scenario analysis is required. Besides the parameter overview, the scenario files are necessary inputs. At first, a comparison and following adjustment of the prepared scenario files has to be conducted. Potentially, parameters were defined in the scenario files and are not included in the models or vice versa. An adoption of units and related conversion might be necessary. After balancing possible inconsistencies, the user can adjust the scenarios to be simulated according to the aim.

At this point, the required preparation for the simulation runs is mainly finished. For each simulation run, a comprehensive scenario file has to be selected and the parameter values as well as initial values have to be automatically set in the participating models. Therefore, the steps of scenario definition and model parameterization have been completed.

4 Simulation and Analysis for Decision Support

The simulation and analysis part starts with the final steps of simulation configuration and further comprises the simulation execution and its evaluation. Besides the models and their parameterization, the simulation specification is a necessary input for a simulation run. In contrast to the models' parameter variation by the scenario files, the simulation specification should be fixed for all simulation runs within a given context to ensure consistent results. It should comprise the model execution order, the models' step size, the time to be simulated and general specifications such as the participating models. With the twofold simulation input of parameterized models and the simulation specification, the actual simulation can be executed. The coordination functionality is mainly provided by a master algorithm which can be chosen considering the application scope and the amount of data to be exchanged. Generally, the requirements for a master algorithm in this decision support context are limited.

From the simulation execution the results are stored. They comprise the models output and optionally, analysis variables. In the following the result files from each simulation execution are collected and an optimization according to the domain and the user's desire is conducted. This information is then provided to the user in a decision support dashboard.

Besides the technical aspects within the framework, the users' interaction as well as the aimed decision support itself shall be discussed here. A major issue to reach decision support is the definition of a suitable objective function. Regarding the manifold stakeholders, a multicriteria objective function seems adequate. These objectives have to be defined specifically for an application in advance, but the weight of individual components might also be a field of study. Therefore, sensitivity analysis is relevant, equally for parameters in simulation and the weight of objectives regarding the simulation results. Therefore, the result can be either a specifically optimized solution or a set of similar optimal configurations.

5 Conclusion

In this paper an approach towards the central parametrization of distributed simulations for decision support in future mobility was presented. Decision support by simulation is required due to the high uncertainty of the upcoming development, the large time horizon and the high invest. Due to multiple stakeholders from different domain, a co-simulation approach is suitable to enable the model reuse and take advantage of varied modelling approaches. This work contributes to fill a gap in the state of the art in co-simulations as it ensures consistency with the included simulation models and enables the framework for efficient parameter studies in the context of scenario variation.

Limits of the proposed idea lay in the necessary information regarding the included parameters in black-box models. Furthermore, the presented approach highly benefits from a common domain ontology for the included simulation models, but shortages have

to be resolved here. For the specific domain application of charging infrastructure this would include at least electrical engineering, economics and individual mobility. Moreover, a concise definition has to be taken of which models to include. Moreover, those models must provide a certain fit to each other to enable the analysis of interdependencies between stakeholders.

Implementation-wise, there is no general solution for central parametrization in simulation. This can be only achieved in a defined context, as different standards require different implementations. A further issue is the fix setting of parameters in models when they are not tuneable. In case such a parameter needs to be edited for a specific scenario, it is required to change the model settings and then to rebuild the model for simulation. Therefore, the models to be used do not require content-wise adaptations to be handled by the framework but it is mandatory to set some model properties according to the requirements of the proposed framework.

As the planning of charging infrastructure has a long-time horizon and no high-frequently changing simulation parts, numerical stability is not necessarily in scope for such a decision-support-tool. Nonetheless, this might be relevant when aiming for a detailed grid model with the physical modelling of plugging an EV. Whereas the idea of a central parameterization is not related to this, the tooling of the simulation framework might need further adaptations then.

Further fields of application within the mobility domain could be the positioning of car-sharing stations, the placement of mobile charging stations according to local events and the placement of battery swapping stations, e. g. for micro mobility offers like E-scooters.

References

- [AG20] Bedarfsgerechte und wirtschaftliche öffentliche Ladeinfrastruktur. Plädoyer für ein dynamisches NPM-Modell. Arbeitsgruppe 5 "Verknüpfung der Verkehrs- und Energienetze, Sektorkopplung", Berlin, 2020.
- [Be19] Beckers, T. et al.: Die Bereitstellung der Schnellladeinfrastruktur für die Elektromobilität in Deutschland. Eine ökonomisch-juristische Analyse zentraler Fragestellungen und alternativer Organisationsmodelle, Berlin, 2019.
- [Bi20] Binder, C. et al.: Towards a Tool-Based Approach for Dynamically Generating Co-Simulation Scenarios based on complex Smart Grid System Architectures: 2020 IEEE 15th International Conference of System of Systems Engineering (SoSE). IEEE, pp. 199–204, 62020.
- [Bu19] Masterplan Ladeinfrastruktur der Bundesregierung. Ziele und Maßnahmen für den Ladeinfrastrukturaufbau bis 2030, Berlin, 2019.
- [Da97] Dahmann, J. S.: High Level Architecture for simulation: Proceedings First International Workshop on Distributed Interactive Simulation and Real Time Applications. IEEE Comput. Soc. Press, pp. 9–14, 1997.

-
- [Go18] Gomes, C. et al.: Co-Simulation. *ACM Computing Surveys* 3/51, pp. 1–33, 2018.
- [KE09] Kek, A. G. et al.: A decision support system for vehicle relocation operations in carsharing systems. *Transportation Research Part E: Logistics and Transportation Review* 1/45, pp. 149–158, 2009.
- [KKS13] Kuby, M. J.; Kelley, S. B.; Schoenemann, J.: Spatial refueling patterns of alternative-fuel and gasoline vehicle drivers in Los Angeles. *Transportation Research Part D: Transport and Environment* 25, pp. 84–92, 2013.
- [LO18] Pablo Alvarez Lopez et al.: 2018 IEEE Intelligent Transportation Systems Conference. November 4-7, Maui, Hawaii. IEEE, Piscataway, NJ, 2018.
- [MBN16] Montori, F.; Borghetti, A.; Napolitano, F.: A co-simulation platform for the analysis of the impact of electromobility scenarios on the urban distribution network: 2016 IEEE 2nd International Forum on Research and Technologies for Society and Industry Leveraging a better tomorrow (RTSI). IEEE, pp. 1–6, 92016.
- [MDC19] Mollier, S.; Delle Monache, M. L.; Canudas-de-Wit, C.: A decision support and planning mobility method for large scale traffic networks: 2019 18th European Control Conference (ECC). IEEE, pp. 1–6, 62019.
- [Mo19] Modelica Association: System Structure and Parameterization, 2019.
- [MO20] Modelica Association: Functional Mock-up Interface for Model Exchange and Co-Simulation 2.0.2, 2020.
- [Na19] Nalic, D. et al.: Development of a Co-Simulation Framework for Systematic Generation of Scenarios for Testing and Validation of Automated Driving Systems*: 2019 IEEE Intelligent Transportation Systems Conference (ITSC). IEEE, pp. 1895–1901, 102019.
- [NH17] Nagele, T.; Hooman, J.: Co-simulation of cyber-physical systems using HLA: 2017 IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC). IEEE, pp. 1–6, 12017.
- [Pa17] Palensky, P. et al.: Cosimulation of Intelligent Power Systems: Fundamentals, Software Architecture, Numerics, and Coupling. *IEEE Industrial Electronics Magazine* 1/11, pp. 34–50, 2017.
- [Te20] Teixeira, B. et al.: Application Ontology for Multi-Agent and Web-Services’ Co-Simulation in Power and Energy Systems. *IEEE Access* 8, pp. 81129–81141, 2020.

Global Environmental Website Disclosure of Large Production Companies and their CDP Scores

Heiko Thimm ¹, Karsten Boye Rasmussen ², Milan Tanik¹


Abstract: This work contributes to the exposure of the current state of environmental disclosure through an empirical transnational study of 344 websites of large production companies from China, Europe, Japan, and the USA. The focus of the investigation is on four formative indices based on variables covering observation of disclosure extent of environmental management and environmental compliance management. All countries show low levels of disclosure. However, some variation on the extent of disclosure was found. The lowest level of disclosure was found for the Chinese websites. Higher - though not remarkable higher - levels of disclosure extents were observed for the companies from Europe, Japan, and the USA. Also, the correlation between disclosure scores based on these observational data and the CDP carbon disclosure scores of the year 2020 did only reveal insignificant correlations. However, websites having webpages containing a higher number of terms of the corporate environmental management terminology also achieve higher CDP scores.


Keywords: corporate environmental management, corporate environmental compliance management, website disclosure, sustainability ranking, Carbon Disclosure Project, CDP

1 Introduction

In recent years website publishing has evolved to a major component of corporate environmental communication. Up-to-date information about environmental concerns at various levels of granularity can be disclosed to the stakeholders and the greater public with relatively low costs. In addition to market conditions and climate change awareness, the ever-growing global environmental legislation is a main driver for environmental website disclosure.

The evaluation of the business world by rating companies which originally focussed on financial performance has been extended in the last decade. Based on complex scoring methodologies, now the companies are also evaluated in terms of their environmental performance and environmental disclosure. In general, evidence-based scoring methods will only yield high scores for companies that through verifiable facts can demonstrate

¹ Pforzheim University, School of Engineering, Tiefenbronner Str. 65, Pforzheim, D-75175,
heiko.thimm@hs-pforzheim.de,  <https://orcid.org/0000-0002-6200-9655>

² University of Southern Denmark, Dept. of Business & Management, Campusvej 55, Odense, DK-5230,
kbr@sam.sdu.dk,  <https://orcid.org/0000-0002-0545-3098>

environmental efforts. Hence, regular sustainability rankings such as published by Corporate Knights, Standard and Poor's, and the Carbon Disclosure Project [Di20], [SA11] have to be regarded as a further driver of corporate environmental disclosure.

Environmental website disclosure has been addressed by several scientific disciplines including environmental accounting, environmental communication, and environmental informatics. Of the available studies many focussed on a single country and/or a specific business sector. Because of the dynamics of the above described disclosure drivers as well as the permanently evolving digital world, the findings of these earlier works may not accurately describe the current situation of environmental website disclosure. The primary purpose of the research described in this article is to reveal an up-to-date status of environmental website disclosure in China, countries of the European Union (EU), Japan, and the United States of America (USA). The production sector was chosen for analysis because production is often considered to be a major pillar for the respective economies and production companies are believed to often leave a heavy environmental burden. A large data set for the production companies was obtained by a thorough website inspection identifying specific features and disclosure subjects. Four formative indices were developed in order to explore disclosure extent of corporate environmental management and corporate environmental compliance management (CECM).

Corporate websites play a major role in image building, reputation management, and investor attraction of today's business world. Hence, it can be assumed that companies are intrinsically motivated to publish information about their environmental achievements such as good scores in sustainability rankings. Conversely, they are less motivated to disclose information that may have a negative impact on the environmental reputation of the company such as a bad ecological footprint of a product, hazardous substances of product components, environmental incidents, and violations of environmental legislation. These considerations lead to the following research hypothesis (H1): There is a positive correlation between the level of environmental website disclosure and the company's score in fair and objective sustainability rankings. Test of this hypothesis is using the 'Climate Change' scores of the company from the 2020 CDP Rating [CD21b] and their corresponding disclosure scores obtained from observations at the company website.

The literature overview contained in the next section is followed by a description of the research method in Section 3. The study results are described in Section 4 and Section 5 concludes the article with a summary and an outlook to our future research.

2 Literature Overview

Environmental disclosure at corporate websites has been investigated empirically earlier. Cho and Roberts [CR10] investigated the environmental disclosure of US firms. The researchers found that 'worse environmental performing firms' provide more extensive disclosure. The categorization of companies is formed by official pollution data from the EPA's Toxics Release Inventory. From this inventory the top-100 highest toxic scoring

companies were considered as environmentally worse performing.

Jose and Lee [JL07] presented a global cross-industry study which focused on the website disclosure policies and practices of the 200 largest multinational companies. The article reports that 29% of the study sample disclosed details of their Environmental Management System (EMS) and that approximately one-third of the companies addressed office and site practices. The researchers found that when companies disclose, 31% report compliance information regarding legal standards.

The investigation results of a Canadian research group who looked at oil firms [BCT13] suggest that the larger the firm and the greater its media exposure, the more likely the firm is to include environmental management disclosures on its website. Furthermore, Suttipun and Stanton [SS12] found a relationship between the amount of disclosures and type of industry, ownership status, and audit firm in their study of Thai companies.

A study by Brazilian researchers used website observation data and website sustainability reports for a transnational comparison of environmental website disclosure [PB20]. It was revealed that companies from the USA stood out compared to Brazilian companies concerning environmental disclosure. The researchers concluded from the study results that the country of origin of the company has to be regarded as a variable that explains the environmental disclosure in corporate website.

3 Research Method

The target population of the study is implied by our research focus on production companies. We focus on large companies with more than 10,000 employees registered in China, Europe including Great Britain (in the following referred to as EU), Japan, and the USA. A production company was defined as a company with the particular codes 13 to 16 and 22 to 31 among its primary NACE codes [EU]. The selection of the population was performed in the Global Company Database ORBIS [Bu19] which lead to a random sample of 344 Companies. The distribution of the sample by country and by industry sector is shown in Table 1.

The data collection was performed by about 250 well-trained student research assistants during three periods: 1. EU companies: 10/15/-11/30/2019, 2. USA: 03/15/-04/30/2020, 3. China & Japan: 10/15/-11/30/2020. For each website three steps were completed (see Figure 1) in order to observe features (e.g. environmental menu items, environmental news sections, options to contact the environmental management team) and disclosure subjects (e.g. environmental awards, certificates, audits, compliance violations) at the websites as described in [TR20]. In the first step the start page was inspected. In step two followed an observation of lower-level pages according to a restricted top-down traversal of 'environmental links'. The 'environmental links' can be supplied by menu items or hypertext links in the content area of the start page. A link established a new traversal starting point for a further potential traversal path. As opposed to 'ordinary links' the

Industry Sector using NACE Classification	China N	EU N	Japan N	USA N	All N
13 textiles	3	.	1	.	7
14 wearing apparel	2	2	1	3	8
15 leather and related products	.	1	.	3	4
16 wood and products of wood and	.	.	.	1	1
17 paper and paper products	.	2	.	1	3
22 rubber and plastic products	2	4	5	2	13
23 other non-metallic mineral products	3	3	4	1	11
24 basic metals	12	6	8	1	27
25 fabricated metal products	.	4	8	7	19
26 computer, electronic, optical	16	14	25	34	89
27 electrical equipment	11	8	5	7	31
28 machinery and equipment	10	11	20	11	52
29 motor vehicles, trailers	12	18	22	10	62
30 other transport equipment	1	6	4	9	20
All	72	79	103	90	344

Tab. 1: Distribution of sample by countries and manufacturing sector

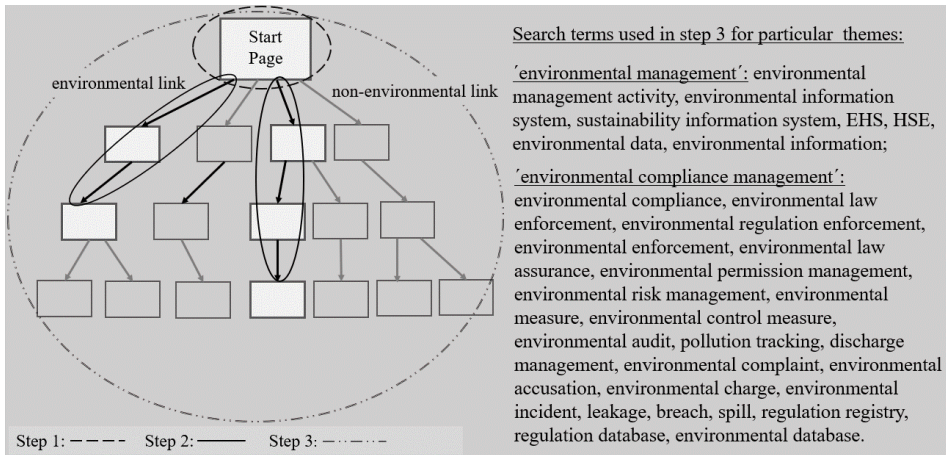


Fig. 1: The three-step website observation approach and the search terms of step three

'environmental links' are defined as links where relevant content is implied by the link name or the context. Consequently, the traversal was terminated when reaching a page without 'environmental links'. Step three used Google Advanced Search full text searches for the 29 particular terms displayed at the right side of Figure 1. The number of hits and the top five urls of the result list that referred to a html page were recorded excluding advertising links.

Drawing on theoretical foundations for the construction of multiscale measures four formative disclosure indices were developed similar to indices used in other disclosure studies [BE02], [CI08]. In general, formative indices are based on formative indicators as described in more detail by Diamantopoulos and Winklhofer [DW01]. Following the guidelines distinctive sets of weighted variables were chosen [TR20] to construct the following formative indices: the *Environmental Management Overview Disclosure Index (EODI)* measures extent of overview information of the company's general environment management engagement; the *Compliance Management Disclosure Index (CMDI)* measures extent of information focused on the company's environmental compliance management efforts; the *Environmental Presentation Features Index (EPFI)* measures extent of dialogical communication capabilities through website features such as contact options and specific information about environmental management; the *Environmental Disclosure and Presentation Index (EDPI)* is a summary index thus measuring websites from all of the perspectives of the three former indices.

The indices were built through a pragmatic scheme to assign weights of value 0.5, 1.0, 1.5, and 2.5 to the indices' variables [TR20]. Features and disclosure subjects at the start page are assigned a higher value than their counterparts at lower-level pages of the website. The higher weight reflects that an environmental news section at the start page is more likely being noticed by ordinary website visitors than when the section is at a lower-level web page. Also, the frequencies of the search terms listed in Figure 1 are reflected in the indices. The frequencies of the terms that are addressing the theme 'environmental management' were used for the EODI index and the other terms were used for the CMDI index. A weigh factor of 0.1 was uniformly used for all frequencies of search terms in order to reflect that the great majority of the obtained hits did not lead to web pages with extensive environmental management disclosure.

For every website the index score for the EODI, the CMDI, and the EPFI was computed by an accumulation of the weighted values of the observation variables associated with the index. The results were used to compute the websites' EDPI scores through the formula $EDPI = EODI + CMDI + EPFI$.

4 Study Results

This section describes the results of investigations of the website data collection. Box-and-whisker plots are used to visualize the results: a box depicts the middle 50% of the observations, the middle line of a box signifies the median, and the calculated mean is depicted by a diamond.

Differences of environmental website disclosure at a regional level can have various reasons such as differences in the environmental legislation (including fines for breaches), differences in the environmental regulatory authorities, differences in general business practices, and cultural differences. However, this work does not address reasons for disclosure variations.

4.1 Disclosure of Environmental Management and CECM

The EDPI index is used to reveal the extent of disclosure in general. From the box plots in the left pane of Figure 2 two major characteristics of the general environmental disclosure practices can be derived. First, the companies fell far short of the maximum possible EDPI value of 49.9, as the mean is 10.1 and the highest EDPI score achieved was 37.4 (by a company of the USA). In fact, the majority of the companies disclose environmental information to a relatively small extent. The Chinese companies are slightly behind the companies of the other three regions with the Japanese companies at the top.

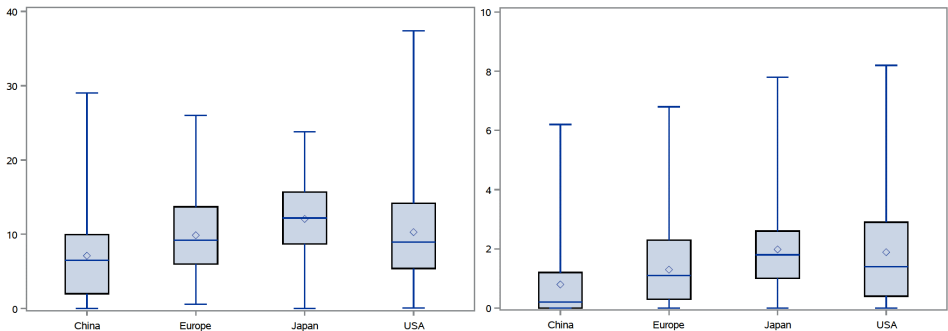


Fig. 2: EDPI (left) and CMDI (right)

As the box plots in the right half of Figure 2 show, only a small amount of CECM information was found at the websites. None of the investigated countries came close to the maximum possible CMDI index value of 14.2. CECM information found at websites often consisted of short general statements about the company's CECM commitment but did not cover details regarding the organizational concept, processes, measures, and use of IT. At some websites, visitors looking for CECM information were referred to corporate sustainability reports. In a small number of cases hyperlinks to online versions of the reports were supplied. A regional comparison shows that Chinese companies are again slightly behind the other three regions that achieve similar CMDI numbers with the USA companies at the top.

Building on the observational data an automated investigation was performed on the recorded webpages of the websites. For the set of company urls recorded during observation it was determined whether an English webpage successfully opens and a registration of the search terms (listed in Figure 1) was recorded. For each observed website the frequency for occurrence of each search term was registered. In the further analysis, only the occurrence or non-occurrence of the search term was registered. The proportion of companies in China, Europe, Japan, and the USA where the search terms occurred is displayed in Figure 3. Each bar displays the occurrence of a particular search term for the corresponding set of companies. For example, the term 'breach' was found at 60 companies of the set of 89 inspected US websites which results a mean of $60/89 = 0.674$. Chinese companies show a lower proportion for the search terms. For example, compare

the proportion for the term ‘carbon disclosure project’ and the corresponding acronym ‘cdp’. Some local differences stand out. For the term ‘environmental measure’ the Chinese companies are no different from the companies of Europe and the USA. For this term only Japan stands out with a higher proportion.

Generally, Europe, Japan, and the USA are quite similar compared to the Chinese companies that have a less frequent occurrence of the search terms.

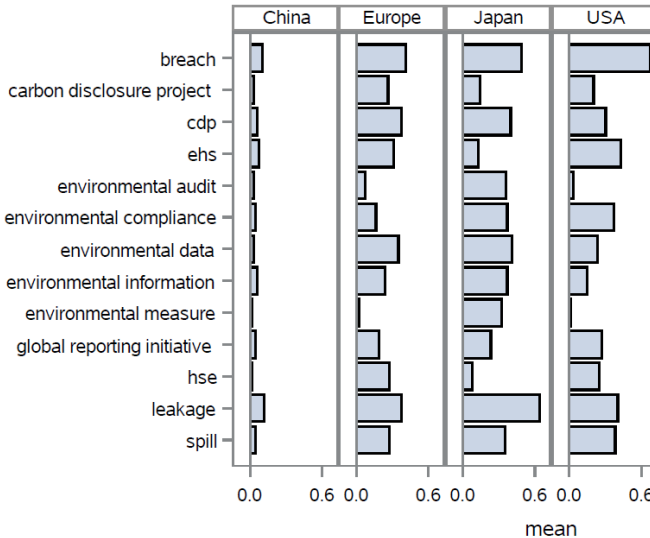


Fig. 3: Proportion of occurrence of selected search terms of companies in countries

4.2 Investigations Considering the CDP Ranking

In this work it is focused on CDP’s climate change sector scores of the sample companies. A description of the general questionnaire-based scoring approach and the method of the 2020 scoring round can be found in an article recently published by the CDP organization [CD21a]. The scores are based on a sector-specific questionnaire and a scoring methodology where responding companies are ‘assessed across four consecutive levels which represent the steps a company moves through as it progresses towards environmental stewardship’ ([CD21a], p. 6). The levels with ‘D’ being the lowest score and ‘A’ the highest scores are in ascending order: Disclosure (D-, D), Awareness (C-, C), Management (B-, B), Leadership (A-, A).

The results of the 2020 scoring have been published on the web as the so-called CDP 2020 ranking [CD21b]. Through a semi-automated procedure, the 2020 Climate Change scores for 180 companies of the sample were obtained from this online data source and are contained in Table 2. Because only 4 Chinese companies with a CDP Climate Change score were found China is excluded from the further investigations. The two leftmost

columns of Table 2 describe the recoding scheme used to transform the CDP scores into numeric scores to enable statistical investigations. From the mean CDP scores and further data of Table 2 it can be obtained that the EU companies achieved the highest CDP climate change scores. The EU companies are followed by the companies of Japan which still outperform the companies of the USA.

CDP Score	Transformed Score	China N	EU N	Japan N	USA N	All N
A	11	1	16	14	10	41
A-	10	1	15	17	10	43
B	9	1	15	14	10	40
B-	8	.	1	9	2	12
C	7	.	5	7	19	31
D	5	1	2	4	4	11
D-	4	.	.	.	2	2
All		4	54	65	57	180
Mean CDP score		8.75	9.52	9.09	8.37	8.98

Tab. 2: CDP 2020 Climate Change Scoring Results of the sample companies

In the following the research hypothesis (H1) described in the introduction is tested through correlation analyses focusing on companies with CDP scores from the EU, Japan, and the USA. Two correlation tests are performed with different proxies for environmental website disclosure. The EDPI scores of the companies serves as first proxy. The amount of different search terms found on the webpages described above serves as second proxy.

Pearson tests revealed highly insignificant correlations between the EDPI scores and the CDP scores of the companies. That is, the observational investigation using the EDPI score as proxy for environmental website disclosure did not lead to any evidence for the research hypothesis H1.

The plot of data points for the number of found search terms (i.e. second proxy for environmental website disclosure) and the CDP scores are displayed in Figure 4. From the left side to the right side the plots show the data points for the complete sample, for Europe, for Japan, and for the USA. Note that some circles in the diagrams are obscured due to overlapping data points. A medium and significant positive correlation was obtained for the USA ($r=0.323$, $p=0.014$) as well as for Japan ($r=0.294$, $p=0.036$). For Europe the result is insignificant. For all companies (Figure 4, most left plot) the result presents some significant correlation ($r=0.202$, $p=0.007$). These results can be viewed as a confirmation of the research hypothesis H1 mostly obtained by the companies of USA and Japan.

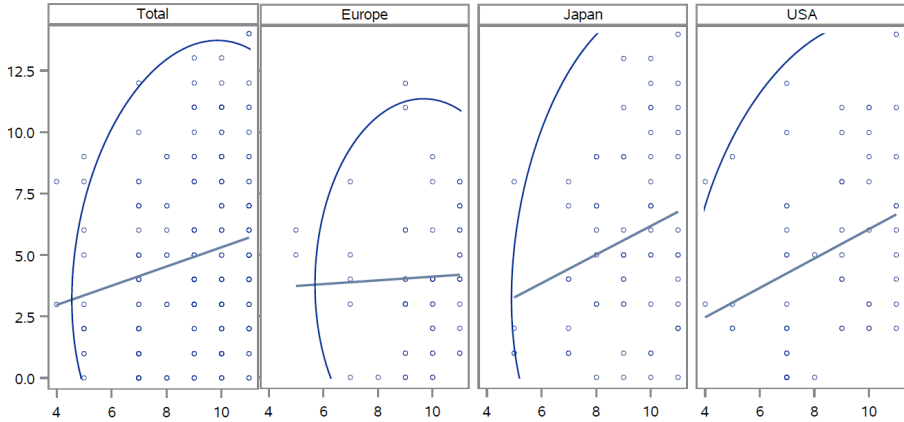


Fig. 4: Scatter plot to explore correlation between number of found search terms (vertical) and CDP scores for complete sample (left) and for countries of sample

5 Conclusions

The study results suggest that large production companies in China, Europe, Japan, and the USA today only to a small extent disclose environmental information on their websites. Many of the potential disclosure capabilities investigated are left unused. For the subtopic of environmental compliance management, the study revealed that the companies disclose information to an even smaller extent. One may argue that companies are assuming the CECM topic to be too narrow for being of large interest for the greater public and thus on purpose disclose only very little CECM information if at all. Another more extreme explanation of the low CECM disclosure is that extensive and detailed CECM information may bring critical attention to the company. Chinese companies are in terms of the two addressed disclosure extents somewhat behind the other countries. For companies of the USA the statistical tests revealed support of the research hypothesis that companies with a good carbon disclosure score disclose environmental information to a larger extent than low scoring companies. However, the proxies for disclosure have to be further investigated and additional sustainability rankings are of interest, too.

Bibliography

- [BCT13] Berthelot, S.; Coulmont, M.; Thibault, K.: Sustainability Content on Oil and Gas Company Websites. *Business and Management Research* Vol 2, No. 1, 94-103, 2013.
- [BE02] Bonson; Escobar: A Survey on Voluntary Disclosure on the Internet. Empirical Evidence from 300 European Union Companies. *Int. J. of Digital Accounting Research* 1/2, pp. 27–51, 2002.

- [Bu19] Bureau van Dijk: Untangling the world of private company information (White Paper). <https://www.bvdinfo.com/en-gb/knowledge-base/white-papers/>, accessed 20 May, 2019.
- [CD21a] CDP: Scoring Introduction 2020. An introduction to 2020 scoring. www.cdp.net, accessed March 19th, 2021.
- [CD21b] CDP: The A List 2020- Explore the full 2020 scores. <https://www.cdp.net/en/companies/companies-scores>, accessed February 25th, 2021.
- [CI08] Clarkson, P. M. et al.: Revisiting the relation between environmental performance and environmental disclosure. An empirical analysis. *Accounting, Organizations and Society* 4-5/33, pp. 303–327, 2008.
- [CR10] Cho, C. H.; Roberts, R. W.: Environmental reporting on the internet by America's Toxic 100. Legitimacy and self-presentation. *Int. J. of Accounting Information Systems* 1/11, pp. 1–16, 2010.
- [Di20] Diez-Cañamero, B. et al.: Measurement of Corporate Social Responsibility: A Review of Corporate Sustainability Indexes, Rankings and Ratings. *Sustainability* 5/12, p. 2153, 2020.
- [DW01] Diamantopoulos, A.; Winklhofer, H.: Index Construction with Formative Indicators: An Alternative to Scale Development. *J. of Marketing Research* 2/38, pp. 269–277, 2001.
- [EU] EUROSTAT: NACE rev. 2. Office for Official Publications of the European Communities, Luxembourg.
- [JL07] Jose, A.; Lee, S.-M.: Environmental Reporting of Global Corporations. A Content Analysis based on Website Disclosures. *Journal of Business Ethics* 4/72, pp. 307–321, 2007.
- [PB20] Portella, A. R.; Borba, J. A.: Environmental disclosure in corporate websites. A study in Brazil and USA companies. *RAUSP Management Journal* 3/55, pp. 309–324, 2020.
- [SA11] Sinreich, E.; Amer, M.: Green Ranking Systems: How They Really Work. Green Edge Research Report. <http://greenedgellc.com/wp-content/uploads/2010/07/Green-Ranking-Systems-How-They-Really-Work.pdf>, accessed February 16th, 2021.
- [SS12] Suttipun, M.; Stanton, P.: A study of Environmental Disclosures by Thai listed Companies on Websites. *Procedia Economics and Finance* 2, pp. 9–15, 2012.
- [TR20] Thimm, H.; Rasmussen, K. B.: Website disclosure of environmental compliance management—the case of European production companies. *J Environ Stud Sci* (2020), 2020.

Langfristige Wasserbedarfsprognosen in Theorie und Praxis

Marius Wybrands,¹ Anne Dyck²

Abstract: Langfristige Wasserbedarfsprognosen sind ein wichtiges Instrument für die strategische Planung in der Wasserwirtschaft. Es zeigt sich jedoch, dass die genutzten Methoden der Praxis sich stark von denen in der Literatur unterscheiden. In dieser Publikation wird daher betrachtet, welche Methoden in der Literatur diskutiert und in der Praxis genutzt werden, um ein besseres Verständnis für den Anwendungsfall der langfristigen Wasserbedarfsprognose im Kontext der Wasserwirtschaft 4.0 zu erlangen und den aktuellen IST-Zustand zu bestimmen.

Keywords: Langfristige Wasserbedarfsprognosen; Wasserversorgung 4.0; Systematische Mapping Studie

1 Motivation

In einer vom Bundesumweltamt aufgegebenen Studie [Ho20], sind Chancen und Herausforderungen der Verknüpfungen der Systeme in der Wasserwirtschaft systematisch dargestellt. Die Autoren kommen zur Erkenntnis, „dass eine 'Wasserwirtschaft 4.0' bisher nur in Ansätzen und / oder einzelnen 4.0-Anwendungen umgesetzt ist“ [Ho20] und dass das zentrale Thema bei der Weiterentwicklung der Digitalisierung der Wasserwirtschaft eine verstärkte Standardisierung und Harmonisierung der Datenerfassung-, -haltung und -nutzung sein muss [Ho20]. Differenzierter geben Sappl et al. [SHR19] einen Überblick, wie Methoden und Erkenntnisse aus dem Bereich des Maschinellen Lernens in den verschiedenen Anwendungsbereichen der Wasserwirtschaft eingesetzt werden. Sie treffen in ihrer Publikation die Aussage, dass realisierte Anwendungen mit Maschine Learning in der Praxis überschaubar sind. Dabei betrachten sie auch den Anwendungsbereich der Wasserbedarfsprognosen (WBP) und zeigen dort anhand von Metastudien [Gh17, HCC16], welche theoretischen Ansätze in der Literatur diskutiert werden.

WBP sind ein vielfältig anwendbares Instrument zur Sicherstellung einer kontinuierlichen und ausreichenden Trinkwasserversorgung. Wird der langfristige Wasserbedarf betrachtet, können die Erkenntnisse als Entscheidungsgrundlage verschiedenartiger strategischer Planungsprozesse genutzt werden. Das Ziel einer langfristigen WBP ist, ein differenziertes Zukunftsbild hinsichtlich der räumlichen und zeitlichen Verteilung des Wasserbedarfs zu erstellen [YY18].

¹ Carl von Ossietzky Universität Oldenburg, Wirtschaftsinformatik/VLBA, Ammerländer Heerstraße 114-118, 26129 Oldenburg, Land marius.wybrands@uol.de

² Carl von Ossietzky Universität Oldenburg, Wirtschaftsinformatik/VLBA, Ammerländer Heerstraße 114-118, 26129 Oldenburg, Land anne.dyck@uol.de

Es besteht jedoch kein allgemeiner Konsens darüber, welches detaillierte Vorgehen und welche Methodik zur Durchführung einer langfristigen WBP genutzt wird. Das DVGW-Arbeitsblatt W 410 [De08] ist bei der Berechnung des langfristigen Wasserbedarfs eine Orientierungshilfe und wird in der praxisnahen Fachliteratur, wie z. B. Bauer et al. [Ba19] benannt. Die in der aktuellen wissenschaftlichen Literatur beschriebenen Prognosemodelle sind theoretisch ausgelegt und untersuchen eher eine spezifische Fallstudie oder Prognosemethode, sodass die verwendeten Vorgehen in der Regel nicht direkt auf ein anderes Untersuchungsgebiet übertragen werden können [Ri15]. Die Folge davon ist, dass sich eine Vielzahl an Vorgehen und Methoden zur Durchführung von langfristigen WBP in der Praxis etabliert haben, die sich von den in der Literatur diskutierten Ansätzen unterscheiden.

Diese Publikation beschreibt eine Gegenüberstellung von in der Literatur behandelte Methoden und Praxisanwendungen für den Anwendungsbereich der langfristigen WBP in Wasserversorgungskonzepten und ähnlichen Veröffentlichungen. Durch eine systematische Mapping Studie [PVK15] ist eine Teilmenge der existierenden Literatur analysiert worden. Die Ergebnisse dieser Recherche werden anschließend mit den Untersuchungen aus der Praxis verglichen, um Unterschiede in den Vorgehensweisen der betrachteten Prognosemethoden, Einteilung der Verbrauchergruppen und Einflussfaktoren zu ermitteln und im Anwendungsfall der langfristigen WBP die Schlussfolgerungen der Studie des Bundesumweltamtes [Ho20] zu detaillieren.

2 Durchführung der systematischen Mapping Studie

Die systematische Mapping Studie ist nach der Methodik von Petersen et al. [PVK15] durchgeführt. Der Prozess beginnt mit der Erstellung von Leitfragen. In diesem Fall heißt dies zu erkennen, welche Methoden zur Durchführung von langfristigen WBP verwendet werden und welche Einflussfaktoren dabei relevant sein können. Daraus ergeben sich zwei Leitfragen: (1) Mit welchen Methoden kann der langfristige Wasserbedarf prognostiziert werden? (2) Durch welche Faktoren kann der langfristige Wasserbedarf beeinflusst werden?

In einer vorgeschalteten Pilotstudie wurden verschiedene Datenbanken und Schlüsselbegriffe getestet und für die Recherche gegebenenfalls ausgeschlossen. Initial sind deutsche Begriffe verwendet worden, um den Suchraum zu definieren. Jedoch stellte sich heraus, dass die Ergebnismenge nicht ausreichend ist und in der Vielzahl von deutschsprachigen Publikationen, praxisbezogene Projekt die Basis für die Publikationen sind. Beispiele hierfür sind die Publikationen von Liehr et al. [LSS18] oder Roth [Ro11], die jeweils begleitend zu Projekten aus der Praxis publiziert haben und somit auch Limitationen in der Wahl der Methodik und Daten unterlagen. Daher wurde der Suchraum durch die englischen Begriffe definiert.

Dieses führte jedoch zu einem Ungleichgewicht zwischen den zwei betrachteten Räumen. Die systematische Mapping Studie betrachtet internationale Publikationen und der Praxisanteil beschränkt sich auf Deutschland. Ebenso gibt die Literaturstudie keinen vollumfänglichen

Sachstand über die aktuellen Forschungen, sondern zeigt nur einen Ausschnitt als Benchmark für die in der Praxis genutzten Methoden. Dieses muss bei der Interpretation der Ergebnisse mit betrachtet werden. Aus den Erkenntnissen folgt die Auswahl über zwei verschiedene Suchbegriffe: *Long-term water demand* und *Water demand forecasting*.

Obwohl auch ähnliche Suchbegriffe wie Wassernachfragemanagement oder Wasserbedarfsplanung hätten ausgewählt werden können, werden diese ausgeschlossen. Der Grund hierfür ist, dass diese Begriffe vermehrt zu wissenschaftlichen Arbeiten führen, die sich auf die Angebotsseite beziehen oder betriebliche Managementaufgaben diskutieren, deren Inhalte zu relevanten Themen kurzfristiger WBP gehören. Als wissenschaftliche Datenbanken zur Suche wird der Bibliothekskatalog der Universität Oldenburg, Science Direct, Research Gate und ACM Digital Library ausgewählt. Weitere Datenbanken (PLOS one, Springer Link, IEEEExplore, Ebsco, Google Scholar) wurden nach der Pilotsuche ausgeschlossen. Ebenso wurden Publikationen ausgeschlossen, die vor dem Jahr 2010 veröffentlicht wurden, nicht auf englisch verfügbar waren und Bücher sowie graue Literatur wurde ebenfalls ausgeschlossen.

Die gefundenen Publikationen wurde zunächst nach den Titeln analysiert. Im Anschluss an die Auswahl aller relevanter Publikationen in Bezug auf den Titel werden in jedem einzelnen Abstract relevante Schlüsselbegriffe und Konzepte identifiziert, sowie anschließend zu übergeordneten Kategorien zusammengefasst.

2.1 Ergebnisse der systematischen Mapping Studie

Alle Titel sind im ersten Schritt mithilfe des Abstracts überprüft worden. Die hohe Anzahl der bei diesem Schritt ausgeschlossenen Publikationen ist, ebenso wie auch die Auswahl nach Titeln, dadurch begründet, dass sich diese Publikationen mit der Untersuchung kurzfristiger WBP beschäftigen. Im nächsten Schritt sind 29 Publikationen vollständig gelesen worden. Ausgeschlossen werden in diesem Schritt hauptsächlich Publikationen, die sich auf spezielle Einzelfälle fokussieren. Final ausgewählt wurden folgende 20 Publikationen, anhand derer die zuvor formulierten Leitfragen 1 und 2 im Folgenden beantwortet werden können: [Qi14, Ge18, DF14, Wa18, BMS14, Ha15, Fo12, Wa11, VB15, Sa14, SYN17, Yo17, AFG14, Zh12, MA10, Zi10, LWG10, FH16, Ha14, Ha18]

2.1.1 Leitfrage 1: Mit welchen Methoden kann der langfristige Wasserbedarf prognostiziert werden?

Zur Prognose des langfristigen Wasserbedarfs werden in den betrachteten Publikationen eine Vielzahl von Methoden (Tabelle 2) verwendet, die zum Ziel haben, anhand eines geeigneten mathematischen Modells mögliche Veränderungen des Wasserbedarfs zu ermitteln.

Vor der Erstellung der Prognose des Wasserbedarfs, werden in den Publikationen zunächst relevante Einflussfaktoren ermittelt. Von den insgesamt 20 Publikationen weisen 15 eine direkte Bestimmung dieser auf, ohne dass ein erkennbarer Auswahlprozess des Vorgehens beschrieben wird. In den verbleibenden 5 Publikationen werden zur Bewertung des Einflusses potenzieller Faktoren auf den Wasserbedarf drei unterschiedliche methodische Vorgehensweisen genutzt. Dieses sind die *Principal Component Analysis* [Ge18, ?, Ha18], *Sensitivity Analysis* [BMS14] und *Segmentation* [Fo12].

In den Publikationen [Ge18] und [Ha18] werden jeweils sieben verschiedene Methoden verglichen. Anhand dieser Methoden werden die Daten möglicher Einflussfaktoren so aufbereitet, dass die relevanten unabhängigen Variablen für eine spätere Regression identifiziert werden können. Verglichen werden in beiden Publikationen die Methoden Forward Selection, Backward Elimination, Stepwise Selection sowie Residual Mean Square Error, Mallow's CP, Akaike Information Criterion und Principal Component Analysis. Die Ergebnisse beider Publikationen zeigen, dass die entwickelten Wasserbedarfsmodelle bessere Simulationsergebnisse des Wasserbedarfs ergeben, wenn die Datensätze der unabhängigen Variablen mit Principal Component Analysis vorverarbeitet werden [Ge18, Ha18]. Unterstützt wird diese Aussage durch Haque et al. [Ha15] in der ebenfalls Principal Component Analysis zur Auswahl der unabhängigen Variablen genutzt wird. Sobald alle Einflussgrößen identifiziert sind, wird in allen Publikationen zunächst der aktuelle Wasserbedarf der untersuchten Region, gemessen am durchschnittlichen jährlichen oder monatlichen Bedarf, betrachtet. Dabei fokussieren zehn Publikationen [Ge18, Wa18, VB15, AFG14, Zh12, MA10, LWG10, FH16, Ha14, Ha18], unabhängig von der später verwendeten Prognosemethodik, auch die historische Entwicklung der Nutzungsdaten. Die Betrachtung der Entwicklung des Wasserverbrauchs erweist sich in den Publikationen als sinnvoll, um später die Ergebnisse der Prognosen mit der Trendentwicklung zu vergleichen. Die Publikationen [Sa14] und [LWG10] vergleichen die Ergebnisse mehrerer Methoden, um auf diesem Wege die effektivste Prognosemethode bestimmen zu können. Deutlich wird dabei vor allem, wie stark das Prognoseergebnis und somit die Genauigkeit einer Prognose, durch die gezielt gewählte Methode beeinflusst werden kann. Elf Publikationen [Qi14, DF14, Wa18, BMS14, Fo12, Wa11, AFG14, Zh12, MA10, LWG10, Ha18] nutzen für die Prognose mehrere Methoden, sodass die Prognosemethode dem jeweils betrachteten Einflussfaktor angepasst wird.

2.1.2 Leitfrage 2: Durch welche Faktoren kann der langfristige Wasserbedarf beeinflusst werden?

Die verschiedenen Einflussfaktoren wirken sich oftmals nicht auf den gesamten Wasserbedarf aus, sondern beeinflussen lediglich die Nachfrage einer bestimmten Verbrauchergruppe. Daher ist es von hoher Bedeutung zunächst die Verbrauchergruppen der betrachteten Publikationen zu untersuchen.

Tabelle 3 zeigt, dass der Fokus der überwiegenden Publikationen auf der Analyse des häuslichen Wasserbedarfs liegt, wobei oftmals diese Verbrauchergruppe auch Verbraucher enthält, die als Kleingewerbe anzusehen sind. Ein Grund dafür ist, dass die meisten der betrachteten Publikationen sich auf ein urbanes Verbrauchsgebiet beziehen, in dem Privathaushalte einen überwiegenden Teil der Nutzer darstellen. Dieser Fokus der Publikationen ist darüber hinaus dadurch zu erklären, dass die Nachfrage der Haushalte im Vergleich zu anderen Verbrauchergruppen deutlich stärkeren Veränderungen unterliegt. Alle Publikationen, die neben dem häuslichen Wasserbedarf ebenfalls andere Verbrauchergruppen betrachten, kommen zu dem Ergebnis, dass in diesen Verbraucherbereichen nicht mit vergleichbaren Veränderungen zu rechnen ist. Alle weiteren Verbrauchergruppen werden in deutlich weniger Publikationen betrachtet.

Der langfristige Wasserbedarf jeder einzelnen Verbrauchergruppe wird durch eine Vielzahl von Einflussfaktoren beeinflusst. Die Identifikation der wichtigsten und relevantesten Faktoren unter allen möglichen Variablen ist für die Entwicklung des Prognosemodells von entscheidender Bedeutung, da die Prognosegenauigkeit maßgeblich von der Auswahl eines geeigneten Satzes erklärender Variablen abhängt [Ha18].

Die in den betrachteten Publikationen verwendeten Einflussfaktoren auf den langfristigen Wasserbedarf sind in Tabelle 4 zusammengefasst. Nahezu alle der betrachteten Publikationen integrieren mehrere unabhängige Variablen in ihre Prognosemodelle, wodurch die Schätzung aussagekräftiger werden soll [Ha18]. Babel et al. [BMS14] argumentieren, dass der Erfolg einer langfristigen WBP maßgeblich von der Anzahl und Type der erklärenden Variablen abhängt.

Die Analyseergebnisse der Tabelle 4 zeigen, dass die meisten langfristigen WBP den Untersuchungsschwerpunkt auf klimatische sowie sozioökonomische und soziodemografische Faktoren legen. Knapp dreiviertel aller betrachteten Publikationen integrieren mindestens einen dieser Bereiche in ihre Prognose. Der Grund für diesen Schwerpunkt zeigt sich in den Ergebnissen der WBP. Diese verdeutlichen, dass diese Faktoren den langfristigen Wasserbedarf deutlich stärker beeinflussen als andere Variablen [BMS14, SYN17]. Hinzu kommt, dass für diese Variablen die benötigten Daten am häufigsten zur Verfügung stehen, beziehungsweise in vielen Fällen bereits Prognosen für einzelne Einflussfaktoren vorliegen, auf die zurückgegriffen werden kann [MA10].

Aus den Ergebnissen der Prognosen wird insgesamt deutlich, dass kein allgemein gültiger oder wichtigster Einflussfaktor benannt werden kann, da der Einfluss eines Faktors immer abhängig von den Gegebenheiten des jeweils betrachteten Versorgungsgebietes ist.

3 WBP in der Praxis

Ausgehend von den 50 größten deutschen Städten als initialer Suchraum, wurden folgende WBP ermittelt: Bei der Recherche konnten für 14 der zuvor ausgewählten Städte die

Ziel Stadt	Gefundene Region	Einordnung	Bezeichnung	Zeitraum
Berlin	Berlin	Direkt	Wasserversorgungskonzept	2008-2040
Hamburg	Hamburg	Direkt	Wasserbedarfsprognose	2014-2045
München	Oberbayern	Größere Region	Wasserversorgungsbilanz	2016-2025
Köln	Bergheim	Alternative	Wasserversorgungskonzept	2018-2028
Frankfurt	Rhein-Main-Region	Größere Region	Situationsanalyse zur Wasserversorgung	2016- 2030
Stuttgart	Tübingen	Alternative	Bedarf und Sicherstellung der Wasserversorgung	2017-2035
Düsseldorf	Düsseldorf	Direkt	Wasserversorgungskonzept	2018- 2030
Leipzig	Sachsen	Größere Region	Grundsatzkonzeption	2012-2020
Dortmund	Ense	Alternative	Wasserversorgungskonzept	2018-2041
Essen	Dorsten	Alternative	Wasserversorgungskonzept	2018-2035
Bremen	-	-	-	-
Dresden	Sachsen	Größere Region	Grundsatzkonzeption	2012-2020
Hannover	-	-	-	-
Nürnberg	Mittelfranken	Größere Region	Wasserversorgungsbilanz	2016-2025
Duisburg	Voerde	Alternative	Wasserversorgungskonzept	2017-2027
Bochum	Castrop-Rauxel	Alternative	Wasserversorgungskonzept	2017-2027
Wuppertal	Wuppertal	Direkt	Wasserversorgungskonzept	2018-2025
Bielefeld	Bielefeld	Direkt	Wasserversorgungskonzept	2016-2046
Bonn	Swistal	Alternative	Wasserversorgungskonzept	2018-2046
Münster	Münster	Direkt	Wasserversorgungskonzept	2018-2041
Karlsruhe	Karlsruhe	Direkt	Wasserbedarfsprognose	2006-2040
Mannheim	Frankenthal	Alternative	Wasserbedarfsprognose	2009-2025
Augsburg	-	-	-	-
Wiesbaden	Rhein-Main-Region	Größere Region	Situationsanalyse zur Wasserversorgung	2016-2030
Mönchengladbach	Mönchengladbach	Direkt	Wasserversorgungskonzept	2018-2030
Gelsenkirchen	Gelsenkirchen	Direkt	Wasserversorgungskonzept	2017-2027
Braunschweig	-	-	-	-
Kiel	-	-	-	-
Chemnitz	Sachsen	Größere Region	Grundsatzkonzeption	2012-2020
Aachen	Aachen	Direkt	Wasserbedarfsprognose	2018-2040
Halle (Saale)	-	-	-	-
Magdeburg	Sachsen	Größere Region	Grundsatzkonzeption	2012-2020
Freiburg	Oberer Neckar	Größere Region	Strukturanalyse	2014-2035
Krefeld	Krefeld	Direkt	Wasserversorgungskonzept	2018-2020
Lübeck	Zugehörig zu Hamburg	Alternative	-	-
Mainz	Teilgebiet 4 in RLP	Größere Region	Teilgebiet Wasserversorgungsplan	2005-2012
Erfurt	Thüringen	Größere Region	Trinkwasserbilanz	2009-2040
Oberhausen	Oberhausen	Direkt	Wasserversorgungskonzept	2018-2035
Rostock	-	-	-	-
Kassel	-	-	-	-
Hagen	Gevelsberg	Alternative	Wasserversorgungskonzept	2018-2022
Saarbrücken	-	-	-	-
Hamm	Sendenhorst	Alternative	Wasserversorgungskonzept	2017-2027
Potsdam	Brandenburg	Größere Region	Wasserversorgungsplan	2009-2030
Ludwigshafen am Rhein	Teilgebiet 6 in RLP	Größere Region	Wasserversorgungsplan	2005-2012
Mühlheim an der Ruhr	Mühlheim an der Ruhr	Direkt	Wasserversorgungskonzept	2018-2035
Oldenburg (Oldb)	-	-	-	-
Osnabrück	-	-	-	-
Leverkusen	Leverkusen	Direkt	Wasserversorgungskonzept	2018-2040
Heidelberg	Wiesloch	Alternative	Wasserertragsrechnung	2015-2045

Tab. 1: Übersicht über langfristige WBP der untersuchten deutschen Städte.

dazugehörigen langfristigen WBP direkt über eine Internetrecherche ermittelt werden. Für weitere 13 Städte konnten langfristige WBP gefunden werden, die eine größere regionalen Skala untersuchen, welche auch die jeweils gesuchte Stadt umfasst. Davon betreffen die Prognosen für sechs Städten das gesamte Bundesland. Für alle Städte, denen keine langfristige WBP zugeordnet werden konnte, wird nach einem alternativen sowie nahegelegenen Versorgungsgebiet mit einer zugehörigen langfristigen WBP gesucht. Dadurch konnten 12 weitere Prognosen ermittelt werden. Für die verbleibenden zwölf Städte wurden keine langfristige WBP gefunden. Die Entsprechende Zuordnung mit den Begriffen direkt, größere Region und Alternative ist in Tabelle 1 aufgelistet

3.1 Analyse der WBP

Die betrachteten langfristigen WBP aus der Praxis verfolgen das Ziel, ein möglichst umfassendes Gesamtbild über die Versorgungssituation zu liefern. Dieses Ziel wird bereits in den Titeln der Berichte deutlich. Begriffe wie *Wasserversorgungskonzept* oder *Wasserversorgungsplan* machen deutlich, dass der Fokus auf einer umfassenden Darstellung zum aktuellen Stand und den zukünftigen Entwicklungen der Wasserversorgung liegt. Die Prognose des langfristigen Wasserbedarfs ist in den Berichten daher nur ein Teilaspekt der gesamten Untersuchung.

Zunächst wird die Ist-Situation des Versorgungsgebietes dargestellt. Auf diesen Informationen aufbauend, wird in allen Berichten die räumliche und zeitliche Verteilung des Trinkwassers dargestellt und analysiert. Dadurch wird deutlich, welche Verbrauchergruppen welchen Bedarf an Trinkwasser benötigen. Die aktuelle Wassernachfrage wird als mittlerer jährlicher Bedarfswert für die jeweils betrachteten Verbrauchergruppen angegeben. In einigen Berichten werden darüber hinaus auch die Spitzenfaktoren ermittelt. Aufbauend auf der Betrachtung der historischen Bedarfsentwicklung erfolgt im nächsten Schritt die Prognose des Wasserbedarfs. Für eine Gegenüberstellung des Wasserangebots und der jährlichen Fördermenge nutzen nahezu alle vorliegenden Berichte eine Wasserbilanzierung. Mehrere Prognosen bewerten in diesem Schritt außerdem die Qualität der Wasserressourcen sowie des Rohwassers der Wasserversorgungsunternehmen und stellen Überwachungskonzepte vor. Abschließend kann dann die Versorgungssicherheit bewertet werden. Wenige Berichte diskutieren zusätzlich anhand einer Gefahrenanalyse.

Die überwiegende Anzahl der untersuchten WBP aus der Praxis machen keine direkten Angaben zu der verwendeten Prognosemethodik. In 15 Berichten (Berlin, Sachsen, Ense, Düsseldorf, Thüringen, Dorsten, Oberer Neckar, Karlsruhe, Krefeld, Frankenthal, Mönchengladbach, Münster, Brandenburg, RLP, Tübingen) werden die Ergebnisse der WBP direkt angegeben, ohne dass Rückschlüsse auf die verwendete Prognosemethodik oder Teile der Berechnungen gezogen werden können. In weiteren 15 Berichten (Aachen, Bielefeld, Castrop-Rauxel, Voerde, Rhein-Main Region, Gelsenkirchen, Gevelsberg, Wiesloch, Köln, Leverkusen, Mühlheim an der Ruhr, Oberbayern, Mittelfranken, Oberhausen, Wuppertal) werden zwar keine direkten Angaben zur verwendeten Methode gemacht, jedoch können indirekt Rückschlüsse auf die jeweils verwendete Methode geschlossen werden. In den verbleibenden Berichten wird die Prognosemethode direkt benannt. Die Ergebnisse der Untersuchung sind in Tabelle 2, 3 und 4 dargestellt.

4 Gegenüberstellung von Literatur und Praxis

Tabelle 2 listet die Methoden zur Modellierung des langfristigen Wasserbedarfs auf und zeigt, welche dieser Methoden in der Literatur betrachtet und in der Praxis angewendet werden. Der ökonomische Ansatz wird bereits seit den 60er Jahren für WBP angewendet und basiert

auf einer statistischen Schätzung historischer Zusammenhänge zwischen unabhängigen Erklärungsvariablen und der abhängigen Variable, welche in diesem Fall den langfristigen Wasserbedarf darstellt [Ha18]. Sowohl in der Praxis als auch in der Literatur wird diese Methode betrachtet. Sehr ähnlich verhält es sich mit Zeitreihenanalysen. Diese basieren auf der Annahme, dass der Wasserverbrauch zwar tendenziell schwankt, grundsätzlich aber auf einem gleichbleibenden Weg verlaufen wird, ohne dass strukturelle Veränderungen eintreten. Durch diese Annahme kann der Wasserbedarf direkt prognostiziert werden, ohne dass Einflussfaktoren berücksichtigt werden müssen [Zh12]. In der Praxis wird vor allem die lineare Trendextrapolation genutzt. Die anderen Methoden werden nur in der Literatur betrachtet. Methoden der Künstlichen Intelligenz, System Dynamics und Monte-Carlo-Simulationen sind in der Praxis nicht zu finden, werden aber in der Literatur diskutiert. Die Methode der Szenarien wird in der Praxis angewendet. Die Szenarien werden vor allem angewendet, um Unsicherheiten über einen großen Zeitraum zu betrachten. Es hat sich gezeigt, dass bei einem Ensemble von Methoden Szenarien genutzt werden.

In Tabelle 3 wird dargestellt, welche Verbrauchergruppen in den langfristigen WBP aus der Praxis und in der Literatur betrachtet werden. In den meisten Prognosen wird der Fokus auf den Bereich Haushalte und Kleingewerbe gelegt. Darüber hinaus werden in 22 Prognosen aus der Praxis die Verbräuche aus dem Bereich Industrie und Gewerbe analysiert. Dieser Bereich ist in der Literatur mit 4 Publikationen vertreten. Eine Prognose der Nachfrage durch die Land- und Forstwirtschaft wird in den Prognosen von Ense, Thüringen und Hamburg vorgenommen. In mehreren Berichten wird erklärt, dass das jeweilige Versorgungsgebiet auch land- und teilweise forstwirtschaftliche Flächen enthält, der Bedarf jedoch nicht prognostiziert wird. Eine weitere Verbrauchergruppe ist der Eigenbedarf der Wasserversorgungsunternehmen. Die Prognosen für Berlin, Bielefeld, Thüringen, Dorsten, Rhein-Main, Wiesloch, Frankenthal, Oberbayern und Hamburg betrachten hierbei auch

Prognosetyp	Methode	Praxis	Literatur
Expertenbefragung	Expertenbefragung	Berlin, Hamburg (2)	[Zi10]
Ökonometrischer Ansatz	Regression	Aachen, Bielefeld, Rhein-Main-Region, Hamburg (4)	[Qi14], [Ge18], [DF14], [Wa18], [Ha15], [Fo12], [Yo17], [LWG10], [Ha14], [Ha18]
Zeitreihenanalyse	Regressionsanalysen	Castrop-Rauxel, Voerde, Oberer Neckar, Gelsenkirchen, Gevelsberg, Wiesloch, Köln, Leverkusen, Mülheim an der Ruhr, Oberbayern, Mittelfranken, Oberhausen, Thüringen, Wuppertal	[Qi14], [DF14], [VB15], [Zh12]
	Seasonal Auto-Regressive Integrated Moving-Average	-	[Fo12], [Sa14]
	Random Walk Methode	-	[FH16]
	Box-Cox Transformation (BATS)	-	[Sa14]
Künstliche Intelligenz	Artificial Neural Networks (ANN)	-	[BMS14]
	Adaptive Neuro Fuzzy Inference System	-	[VB15]
	Support Vector Maschine (SVM)	-	[Sa14], [SYN17]
	Gene Expression Programming (GEP)	-	[Yo17]
	IWR-MAIN Software	-	[MAI10]
System Dynamics	Space-time Analyses	-	[LWG10]
System Dynamics	System Dynamics	-	[DF14], [Wa11]
Szenario	Szenario	Berlin, Thüringen, Hamburg, Frankenthal, Tübingen, Gelsenkirchen	[DF14], [Wa18], [BMS14], [Fo12], [AFG14], [Zh12], [MAI10], [FH16], [Ha14]
	Monte-Carlo-Simulation	-	[DF14], [FH16], [Ha14]

Tab. 2: Methoden aus der Literatur und Praxis zur Prognose des langfristigen Wasserbedarfs.

Nutzergruppen	Verbrauchergruppen	Praxis	Literatur
Primäre Industrie	Landwirtschaft, Forstwirtschaft, Tierhaltung und Fischerei	Ense, Thüringen und Hamburg	[Wa11], [AFG14], [Zh12]
Sekundäre Industrie	Generelle Industrie und Baugewerbe	Berlin, Bielefeld, Castrop-Rauxel, Ense, Voerde, Düsseldorf, Thüringen, Dorsten, Rhein-Main Region, Gelsenkirchen, Gevelsberg, Hamburg, Wiesloch, Leverkusen, Frankenthal, Mönchengladbach, Mühlheim an der Ruhr, Oberbayern, Mittelfranken, Oberhausen, RLP, Tübingen	[Qi14], [Wa11], [AFG14], [Zh12]
Tertiäre Industrie	Andere Branchen mit Ausnahme der Primär- und Sekundärbranche, einschließlich Dienstleistungen, Banken, Versicherungen, Transportwesen, medizinische Behandlung und Ausbildung	-	-
Häusliche Nutzer	Einwohner und Kleingewerbe	Alle	[Qi14], [Ge18], [DF14], [Wa18], [BMS14], [Ha15], [Fo12], [Wa11], [VB15], [Sa14], [SYN17], [AFG14], [Zh12], [MA10], [Zi10], [LWG10], [FH16], [Ha14], [Ha18]
Ökologische Nutzer	Natürliches und künstliches Ökosystem	-	[Qi14], [Wa11]
Eigenbedarf und Weiterleitung	Eigenbedarf der Wasserversorgungsunternehmen, Wasserverluste, Weiterleitung	Berlin, Bielefeld, Thüringen, Dorsten, Rhein-Main, Wiesloch, Frankenthal, Oberbayern und Hamburg	-

Tab. 3: Einteilung der Verbrauchergruppen in der Literatur und Praxis zur Prognose des langfristigen Wasserbedarfs.

Bereich	Einflussfaktor	Praxis	Literatur
Ökonomische & sozioökonomische Variable	Preis & Preiselastizität	Thüringen, Rhein-Main	[Ge18], [DF14], [BMS14], [Ha15], [Fo12], [Ha14], [Ha18]
	Einkommen Wirtschaftliche Entwicklung	- Thüringen, Hamburg	[DF14], [Ha15] [Qi14], [BMS14], [Ha15]
Soziodemografische Variable	Bevölkerungsentwicklung	Alle	[Qi14], [Ge18], [DF14], [Wa18], [BMS14], [Fo12], [Wa11], [AFG14], [Zh12], [MA10]
	Altersstruktur	Berlin, Düsseldorf, Frankenthal, Hamburg	[Fo12]
	Haushaltsgröße (Anzahl Personen oder qm2)	Rhein-Main, Hamburg	[DF14], [Fo12], [LWG10]
	Bildung	-	[Fo12]
Territoriale Variable	Bevölkerungsdichte Urbanisierung	Hamburg, Thüringen, Frankenthal Thüringen	[Fo12] [LWG10], [Qi14], [Wa18], [Zh12], [LWG10]
Kulturelle Variable	Kultureller Einfluss auf Konsumentenpräferenzen	-	[Fo12], [Wa11], [Zh12], [FH16], [Ha18]
Klimatische Variable	Temperatur	Berlin, Thüringen, Hamburg, Wiesloch	[Ge18], [Wa18], [BMS14], [Ha15], [Fo12], [Wa11], [SYN17], [Yo17], [AFG14], [MA10], [Zi10], [LWG10], [Ha14], [Ha18]
	Niederschlag	Berlin, Thüringern	[DF14], [BMS14], [Fo12], [Wa11], [Yo17], [AFG14], [Zh12], [MA10], [Zi10], [LWG10], [Ha14], [Ha18]
Technologische Variable	Wassersparende Technologie	Thüringen, Rhein-Main, Hamburg, Frankenthal	[Ge18], [Wa18], [Ha15], [Fo12], [Zh12], [Ha14], [Ha18]

Tab. 4: Einflussfaktoren auf den Wasserbedarf aus der Literatur und Praxis zur Prognose des langfristigen Wasserbedarfs.

die Wasserverluste. Durch den vergleichsweise geringen Anteil am gesamten Wasserbedarf wird diese Nutzergruppe vermutlich nicht in der Literatur betrachtet.

5 Fazit und Ausblick

Die vorangegangenen Analyseergebnisse zeigen, dass die betrachteten Studien eine Vielzahl an unterschiedlichen Vorgehensweisen für die Erstellung einer langfristigen WBP aufzeigen und der Bereich ein attraktives Forschungsgebiet darstellt. Auch in der deutschen Praxis ist die Thematik für Wasserversorgungsunternehmen und die öffentliche Verwaltung ein zentrales

Anliegen, um die Versorgung aller Verbrauchergruppen mit der benötigten Wassermenge bei ausreichender hydraulischer Kapazität langfristig gewährleisten zu können.

Der Vergleich zwischen den Methoden, die in den Wasserversorgungskonzepten genutzt werden und der internationalen Literatur muss mit Einschränkungen und Limitationen betrachtet werden. Es bleibt offen, welche Methoden in den weltweit größten Städten zur Berechnung des langfristigen Wasserbedarfs genutzt werden, so dass in dieser Studie ein Ungleichgewicht zwischen den Vergleichsobjekten in der Interpretation zu berücksichtigen ist. Ebenfalls werden spezifisch Wasserversorgungskonzepte und ähnliche Publikationen betrachtet. Es kann keine Aussage darüber getroffen werden, in welchem Detailgrad die Prognosen bei den Wasserversorgungsunternehmen vorliegt. Dies bedeutet, dass sich die Prognosen in einer räumlichen Granularität stark unterscheiden können, was die Wahl der Prognosemethode beeinflussen könnte. Ein weiterer Aspekt ist die Bewandnis der Bedarfsprognose und der damit verbundene methodische Ansatz. So dient die Wasserversorgungsbilanz der bayrischen Bezirke dazu, die unterschiedlichen Wasserversorgungsunternehmen in ihrer Gesamtheit zu erfassen und strukturelle Herausforderungen zu identifizieren. Dieses unterscheidet sich im Detailgrad und der Zielsetzung von z. B. der WBP für Hamburg, die einen höheren Detailgrad adressiert.

Die Erkenntnisse aus dem Vergleich machen aber auch deutlich, dass in den betrachteten internationalen Publikationen kein eindeutiger Konsens darüber besteht, wodurch sich eine erfolgssichere Prognose auszeichnet. Die jeweiligen Vorgehensweisen zur Durchführung langfristiger WBP unterscheiden sich stark in der Komplexität, den Datenanforderungen und der Menge an benötigtem Expertenwissen. Die unterschiedlichen Vorgehensweisen haben ihren Ursprung vor allem in dem jeweils gesetzten Ziel. In der untersuchten Literatur wird meist einem festgelegten Forschungsziel gefolgt und dabei hauptsächlich eine ausgewählte Methode oder ausgewählte Einflussfaktoren untersucht. Hingegen sind die betrachteten WBP aus der Praxis immer Teil eines Berichts, in welchem die gesamte Versorgungssituation beleuchtet wird. In welchem Ausmaß die Schätzung des zukünftigen Bedarfs dabei erfolgt, schwankt zwischen den Berichten stark.

Es zeigt sich, dass auch im Anwendungsfall der langfristigen WBP in der Praxis eher auf Regressionsanalysen und einfachere Berechnungsmodelle gesetzt wird. Das Methodenspektrum der Künstlichen Intelligenz wird in der Betrachteten Menge an WBP aus der Praxis noch nicht genutzt. Diese Erkenntnis bestärkt die Aussage der Studie des Bundesumweltamtes [Ho20] im Anwendungsfall der langfristigen WBP. In dieser Publikation wurde betrachtet, welche Methoden in der Literatur und in der Praxis eingesetzt werden, um langfristige WBP durchzuführen. Der nächste Schritt ist, zu betrachten wieso diese nicht eingesetzt werden und welche unterstützenden informationstechnischen Methoden, Prozesse oder Systeme entwickelt werden müssen.

Literaturverzeichnis

- [AFG14] Ahmed, Ayman A.; Fogg, Graham E.; Gameh, Mohsen A.: Water Use at Luxor, Egypt: Consumption Analysis and Future Demand Forecasting. Environmental Earth Sciences,

72(4):1041–1053, August 2014.

- [Ba19] Baur, Andreas; Fritsch, Peter; Hoch, Winfried; Merkl, Gerhard; Rautenberg, Joachim; Weiß, Matthias; Wricke, Burkhard: *Mutschmann/Stimmelmayer Taschenbuch der Wasserversorgung*. Springer Fachmedien Wiesbaden, Wiesbaden, 2019.
- [BMS14] Babel, Mukand S.; Maporn, Nisuchcha; Shinde, Victor R.: *Incorporating Future Climatic and Socioeconomic Variables in Water Demand Forecasting: A Case Study in Bangkok*. *Water Resources Management*, 28(7):2049–2062, Mai 2014.
- [De08] Deutscher Verein des Gas- und Wasserfaches: , *Arbeitsblatt W 410*, 2008.
- [DF14] Dhungel, Ramesh; Fiedler, Fritz: *Price Elasticity of Water Demand in a Small College Town: An Inclusion of System Dynamics Approach for Water Demand Forecast*. *Air, Soil and Water Research*, 7:ASWR.S15395, Januar 2014.
- [FH16] Ferdosian, Iman; Haie, Naim: *Random Walk Forecast of Urban Water in Iran under Uncertainty*. *Water Resources*, 43(1):200–206, Januar 2016.
- [Fo12] Fontdecaba, Sara; Grima, Pere; Marco, Lluís; Rodero, Lourdes; Sánchez-Espigares, José A.; Solé, Ignasi; Tort-Martorell, Xavier; Demessence, Dominique; Martínez De Pablo, Victor; Zubelzu, Jordi: *A Methodology to Model Water Demand Based on the Identification of Homogenous Client Segments. Application to the City of Barcelona*. *Water Resources Management*, 26(2):499–516, Januar 2012.
- [Ge18] Gedefaw, Mohammed; Hao, Wang; Denghua, Yan; Girma, Abel: *Variable Selection Methods for Water Demand Forecasting in Ethiopia: Case Study Gondar Town*. *Cogent Environmental Science*, 4(1):1537067, Januar 2018.
- [Gh17] Ghalekhondabi, Iman; Ardjmand, Ehsan; Young, William A.; Weckman, Gary R.: *Water Demand Forecasting: Review of Soft Computing Methods*. *Environmental Monitoring and Assessment*, 189(7):313, 2017.
- [Ha14] Haque, Md Mahmudul; Rahman, Ataur; Hagare, Dharma; Kibria, Golam: *Probabilistic Water Demand Forecasting Using Projected Climatic Data for Blue Mountains Water Supply System in Australia*. *Water Resources Management*, 28(7):1959–1971, Mai 2014.
- [Ha15] Haque, Md Mahmudul; Egodawatta, Prasanna; Rahman, Ataur; Goonetilleke, Ashantha: *Assessing the Significance of Climate and Community Factors on Urban Water Demand*. *International Journal of Sustainable Built Environment*, 4(2):222–230, Dezember 2015.
- [Ha18] Haque, Md; Rahman, Ataur; Hagare, Dharma; Chowdhury, Rezaul: *A Comparative Assessment of Variable Selection Methods in Urban Water Demand Forecasting*. *Water*, 10(4):419, April 2018.
- [HCC16] Hadjimichael, Antonia; Comas, Joaquim; Corominas, Lluís: *Do Machine Learning Methods Used in Data Mining Enhance the Potential of Decision Support Systems? A Review for the Urban Water Sector*. *AI Communications*, 29(6):747–756, 2016.
- [Ho20] Holländer, Robert; Lautenschläger, Sabine; Görlitz, Stefan; Pielow, Christian: *Chancen und Herausforderungen der Verknüpfungen der Systeme in der Wasserwirtschaft (Wasser 4.0)*. Bericht, Umweltbundesamt, 2020.
- [LSS18] Liehr, Von Stefan; Schramm, Engelbert; Schulz, Oliver: *Wasserbedarfsanalysen: Planungssicherheit für Kommunen*. 3R, 2018(9/2018):38–41, 2018.
- [LWG10] Lee, Seung-Jae; Wentz, Elizabeth A.; Gober, Patricia: *Space–Time Forecasting Using Soft Geostatistics: A Case Study in Forecasting Municipal Water Demand for Phoenix, Arizona*. *Stochastic Environmental Research and Risk Assessment*, 24(2):283–295, Februar 2010.

- [MA10] Mohamed, Mohamed M.; Al-Mualla, Aysha A.: Water Demand Forecasting in Umm Al-Quwain (UAE) Using the IWR-MAIN Specify Forecasting Model. *Water Resources Management*, 24(14):4093–4120, November 2010.
- [PVK15] Petersen, Kai; Vakkalanka, Sairam; Kuzniarz, Ludwik: Guidelines for Conducting Systematic Mapping Studies in Software Engineering: An Update. *Information and Software Technology*, 64:1–18, August 2015.
- [Qi14] Qin, T. L.; Yan, D. H.; Wang, G.; Yin, J.: Water Demand Forecast in the Baiyangdian Basin with the Extensive and Low-Carbon Economic Modes. *Journal of Applied Mathematics*, 2014:1–10, 2014.
- [Ri15] Rinaudo, Jean-Daniel: Long-Term Water Demand Forecasting. In (Grafton, Quentin; Daniell, Katherine A.; Nauges, Céline; Rinaudo, Jean-Daniel; Chan, Noel Wai Wah, Hrsg.): *Understanding and Managing Urban Water in Transition*, Jgg. 15, S. 239–268. Springer Netherlands, Dordrecht, 2015.
- [Ro11] Roth, Ulrich; Herber, Werner; Mikat, Hermann; Wagner, Holger: Die Entwicklung des Trinkwasserverbrauchs in Kommunen unterschiedlicher Größe. *gwf - Wasser|Abwasser*, 2011(02):162–169, 2011.
- [Sa14] Sampathirao, Ajay Kumar; Grosso, Juan Manuel; Sopsakis, Pantelis; Ocampo-Martinez, Carlos; Bemporad, Alberto; Puig, Vicenç: Water Demand Forecasting for the Optimal Operation of Large-Scale Drinking Water Networks: The Barcelona Case Study. *IFAC Proceedings Volumes*, 47(3):10457–10462, 2014.
- [SHR19] Sappl, Johannes; Harders, Matthias; Rauch, Wolfgang: Maschinelles Lernen in der Siedlungswasserwirtschaft. *Österreichische Wasser- und Abfallwirtschaft*, Mai 2019.
- [SYN17] Shabani, Sina; Yousefi, Peyman; Naser, Gholamreza: Support Vector Machines in Urban Water Demand Forecasting Using Phase Space Reconstruction. *Procedia Engineering*, 186:537–543, 2017.
- [VB15] Vijayalaksmi, D.P.; Babu, K.S. Jinesh: Water Supply System Demand Forecasting Using Adaptive Neuro-Fuzzy Inference System. *Aquatic Procedia*, 4:950–956, 2015.
- [Wa11] Wang, Xiao-Jun; Zhang, Jian-Yun; Amgad, Elmahdi; He, Rui-min; Zhang, Li-Ru; Chen, Feng: Water Demand Forecasting under Changing Environment: A System Dynamics Approach. In: *Risk in Water Resources Management*. IAHS Publication, International Association of Hydrological Sciences, Wallingford, Oxfordshire, UK, 2011.
- [Wa18] Wang, Xiao-Jun; Zhang, Jian-Yun; Shahid, Shamsuddin; Xie, Wei; Du, Chao-Yang; Shang, Xiao-Chuan; Zhang, Xu: Modeling Domestic Water Demand in Huaihe River Basin of China under Climate Change and Population Dynamics. *Environment, Development and Sustainability*, 20(2):911–924, April 2018.
- [Yo17] Yousefi, Peyman; Shabani, Sina; Mohammadi, Hadi; Naser, Gholamreza: Gene Expression Programming in Long Term Water Demand Forecasts Using Wavelet Decomposition. *Procedia Engineering*, 186:544–550, 2017.
- [YY18] Yan, Kun; Yang, Min-Zhi: Water Demand Forecast Model of Least Squares Support Vector Machine Based on Particle Swarm Optimization. *MATEC Web of Conferences*, 246:01029, 2018.
- [Zh12] Zhai, Yuanzheng; Wang, Jinsheng; Teng, Yanguo; Zuo, Rui: Water Demand Forecasting of Beijing Using the Time Series Forecasting Method. *Journal of Geographical Sciences*, 22(5):919–932, Oktober 2012.
- [Zi10] Ziervogel, Gina; Johnston, Peter; Matthew, Margaret; Mukheibir, Pierre: Using Climate Information for Supporting Climate Change Adaptation in Water Resource Management in South Africa. *Climatic Change*, 103(3):537–554, Dezember 2010.

Workshop: 2. Workshop Künstliche
Intelligenz in der Umweltinformatik
(KIUI-2021)

2. Workshop Künstliche Intelligenz in der Umweltinformatik

Andreas Abecker¹, Julian Bruns², Stefan Naumann³

Abstract: In den vergangenen Jahren konnte die Umweltforschung eine zunehmend datengesteuerte Perspektive einnehmen, weil immer leistungsfähigere In-situ-Sensorik mit telemetrischer Datenübertragung, riesige Sensornetze, satelliten- und drohnengestützte Erdbeobachtung und ein fast allgegenwärtiger Internetzugang technisch ganz neue Möglichkeiten eröffnen. Von einigen dieser datengestützten Ansätze wird erwartet, dass sie Visionen einer nachhaltigen Zukunft deutlich unterstützen können. Zum Beispiel, indem sie es ermöglichen, in nachhaltigen intelligenten Städten besser zu leben, die Menschheit mit intelligenterer Landwirtschaft besser zu ernähren, die Umweltverschmutzung oder weltweite Rodung von Wäldern mit besserer Erdbeobachtung zu bekämpfen, oder durch intelligente Steuerung von Prozessen und Vernetzung von Akteuren wichtige Menschheitsfragen wie die Umsetzung der Energiewende, die Realisierung der Kreislaufwirtschaft oder die Bekämpfung und den Umgang mit dem Klimawandel zu unterstützen. Im Workshop KIU werden in diesem großen Themenkomplex neue innovative Ideen vorgestellt, offene Forschungsfragen und Lösungsansätze diskutiert und Anwendungsbeispiele aus den Schnittfeldern von KI und Umwelthanwendungen präsentiert.

Keywords: Umweltinformatik, Künstliche Intelligenz, Maschinelles Lernen, Umweltschutz, Nachhaltigkeit

1 Ziele und Motivation des Workshops „Künstliche Intelligenz und Umweltinformatik“

Die Umweltinformatik befasst sich interdisziplinär mit der Analyse und Bewertung von Umweltsachverhalten. Aus informationstechnologischer Sicht spielen dabei beispielsweise Simulationen komplexer Systeme, Geographische Informationssysteme (GIS) und räumliche Datenanalyse, Messnetze und Sensordatenverarbeitung sowie Fernerkundung und Bildverarbeitung eine große Rolle. Außer in der Wissenschaft findet die Umweltinformatik ihre wichtigsten Anwendungen in der öffentlichen Verwaltung (Natur- und Umweltschutz, Umweltdatenportale, Katastrophenschutz, Verbraucherschutz, Wassermanagement etc.), aber auch in der Wirtschaft (betriebliche Umweltinformationssysteme). Enge Bezüge und teilweise Überlappungen ergeben sich auch zur Agrar-, zur Hydro-, zur Energie- und zur Gesundheitsinformatik sowie im Bereich Green IT.

¹ Disy Informationssysteme GmbH, Ludwig-Erhard-Allee 6, 76131 Karlsruhe andreas.abecker@disy.net

² Disy Informationssysteme GmbH, Ludwig-Erhard-Allee 6, 76131 Karlsruhe julian.bruns@disy.net

³ Hochschule Trier, Umwelt-Campus Birkenfeld, Institut für Softwaresysteme, Postfach 1380, 55761 Birkenfeld s.naumann@umwelt-campus.de

Die Umweltinformatik betrachtet in aller Regel sehr komplexe Prozesse in Ökosystemen, deren Verhalten (noch) nicht vollständig bekannt und verstanden ist, deren Verhalten nur approximativ oder vereinfachend modelliert, simuliert oder vorhergesagt werden kann und bei deren Beobachtung häufig relevante Größen unbekannt sind oder nur geschätzt werden können. Zusätzlich müssen bei der Betrachtung von Ökosystemen noch weitere Systeme, die aus Informatiksicht nicht einfach zu behandeln sind, wie z.B. Wetter, betrachtet werden, da diese ebenfalls direkte und indirekte Einflüsse ausüben. Entscheidungen in Anwendungsfällen (wie z.B. Planungsverfahren, Notfallmanagement, Politikgestaltung) sind in der Regel schwierige Abwägungen und erfordern Fach- und Erfahrungswissen. Die entsprechenden Fragestellungen haben vielfältige Wechselwirkungen zu hoch aktuellen und enorm wichtigen Themen wie z.B. Klimawandel, Energiewende, Biodiversität und Nachhaltigkeit, aber auch zu großen Technologietrends wie Smart Cities, Smart Agriculture oder Smart Grids.

Die Anwendung von Methoden und Technologien der KI drängt sich also auf. Im Rahmen der INFORMATIK 2020 fand der Workshops KIU das erste Mal statt, mit der Idee, eine neue Community zu dieser Thematik zu begründen. Teilnahme und Verlauf von KIU-2020 waren ermutigend. Es gab einen eingeladenen und 8 eingereichte und begutachtete Vorträge. Bei der Online-Durchführung waren permanent etwa 30 Zuhörer präsent, in Spitzenzeiten waren es sogar bis zu 50 – bei insgesamt über 100 unterschiedlichen Tagungsteilnehmern im Workshop. Nichtsdestotrotz waren die Breite der Themen (umwelt-fachlich und KI-technisch) und die Tiefe der Diskussionen durchaus noch ausbaufähig. Deshalb freuen wir uns für die zweite Durchführung KIU-2021 ebenfalls auf eine möglichst große und interdisziplinäre Teilnehmerschaft aus Forschung, Entwicklung und Anwendung eingeladen, um gemeinsam ihre Fragestellungen, Lösungsansätze und Ergebnisse zu intelligenten IT-Ansätzen für Umwelthanwendungen zu diskutieren.

Der Call for Papers betrachtete die Umwelt-KI mit der gesamten Breite intelligenter Software-Ansätze, also aus symbolischer und subsymbolischer KI, intelligenter Datenanalyse, maschinellem Lernen usw.

2 Thematische Schwerpunkte

Der Workshop war bewusst breit angelegt und umfasste einerseits KI-Technologien für Umwelt und Umweltschutz und andererseits Anwendungen aus dem Bereich Umwelt, die durch KI unterstützt werden können.

Mögliche KI-Technologien waren (nicht ausschließliche Liste):

- Big / Smart / Linked / Open Data
- Bildverarbeitung und Fernerkundung
- Data Mining, Machine Learning, Deep Learning

- Digital Twins
- ELSI-Aspekte zur Umwelt-KI, Responsible AI
- Explainable AI
- Human-Centered AI
- Multiagentensysteme
- Multimodale Interfaces, AR, VR und KI
- Robotik
- Semantische Technologien
- Soft Computing / Computational Intelligence
- Spatial Data Mining, Spatio-Temporal Data Analytics
- Sprachverarbeitung
- Unsicherheit und Vagheit
- Wissensbasierte Systeme

Mögliche Umwelthanwendungen mit KI-Unterstützung umfassen:

- Betriebliche und behördliche Umweltinformationssysteme
- Biodiversität
- Computational Sustainability
- Erneuerbare Energien und Energiewende
- Green IT und Energiemanagement
- Katastrophenschutz und -management (aus Umweltsicht)
- Klimawandel
- Nachhaltige Produktion, Kreislaufwirtschaft
- Natur- und Umweltschutz
- Ressourcenschutz und Landmanagement
- Smart Agriculture, Smart Forestry (aus Umweltsicht)
- Smart City, Smart Mobility (Umweltaspekte)
- Umweltbildung
- Verbraucherschutz (Umweltaspekte)

- Wasser 4.0

Auch übergreifende Aspekte wie die Energie- und Ressourcenverbräuche durch KI-Training und -Anwendung sind durch die Workshop-Themen angesprochen.

3 Eingereichte Fachbeiträge

Beim zweiten Workshop KIU-2021 freuen wir uns auf eine spannende Keynote von Frau Dr. Ludwig vom Exzellenzcluster Maschinelles Lernen der Universität Tübingen sowie wiederum 8 begutachtete Fachbeiträge zu unterschiedlichen KI-Technologien und Anwendungsgebieten. Diese betreffen aus Anwendungssicht Themen wie z.B. Energiesystem (Windkraft, Photovoltaik), Luftqualität und Kreislaufwirtschaft. Aus technischer Sicht sind weiterhin viele Ansätze des maschinellen Lernens vertreten, aber auch Fragen des intelligenten Sensordatenmanagements, der Geostatistik sowie allgemeine methodische Fragestellungen.

4 Programmkomitee

- Dr. Ansgar Bernardi; Deutsches Forschungszentrum für Künstliche Intelligenz, Kaiserslautern; <https://www.dfki.de/>
- Dr. Matthias Budde; Disy Informationssysteme GmbH, Karlsruhe; <http://www.disy.net/>
- Prof. Dr. Frank Fuchs-Kittowski; Hochschule für Technik und Wirtschaft, Berlin; <https://www.htw-berlin.de/>
- Dr. Desirée Hilbring; Fraunhofer IOSB, Karlsruhe; <https://www.iosb.fraunhofer.de/>
Julian Huber ; FZI Forschungszentrum Informatik, Karlsruhe; <https://www.fzi.de/>
- Dr. Christian Jolk; Ruhr-Universität Bochum, Bochum; <https://www.ruhr-uni-bochum.de>
- Prof. Dr. Gerlinde Knetsch; HTW, Berlin; <https://www.htw-berlin.de/>
- Dr. Sven Lautenbach; Universität Heidelberg, Heidelberg; <https://www.geog.uni-heidelberg.de/>
- Dr. Tanja Liesch; Karlsruhe Institut für Technologie, Karlsruhe; <http://www.kit.edu/>
- Dr. Martin Memmel; Deutsches Forschungszentrum für Künstliche Intelligenz, Kaiserslautern ; <https://www.dfki.de/>
- Prof. Dr. Jens Nimis; Hochschule Karlsruhe Technik und Wirtschaft, Karlsruhe; <https://www.hs-karlsruhe.de/>

- Dr. Steffen Thoma; FZI Forschungszentrum Informatik, Karlsruhe; <https://www.fzi.de/>

KIU-2021 Keynote: Uncertainty in Machine Learning for Environmental Research

Nicole Ludwig ¹

Abstract: Advances in metering infrastructure, data collection and storing have led to data-driven research approaches such as machine learning to become more prevalent in environmental and sustainability research. However, probabilistic approaches which handle uncertainty in the available data and models are still underrepresented. This underrepresentation is prevalent in environmental research, in addition there are different uncertainties influence the models we design and their output. For example, the data sets we have at hand are not always equally well distributed in space and time; some areas might have lots of data points (high resolution in space) while others might have more frequent data (high resolution in time). When working with a mixed data set, probabilistic measures can help clarify where the models are certain about the results and where they are uncertain. Additionally, many research questions we want to answer, such as the optimal location for a wind turbine, rely on weather and climate data input. As weather and climate are a chaotic system and thus only predictable to a certain extent, accounting for this uncertainty is crucial for reliable forecasts and any decision making based on them. In this talk, we will address some of the open issues with uncertainty in space and time in the input data and output of probabilistic machine learning models at the example of forecasting in sustainable energy systems.

Frau Dr. Ludwig leitet die Early Career Forschungsgruppe für Maschinelles Lernen in Nachhaltigen Energiesystemen am Exzellenzcluster Maschinelles Lernen der Universität Tübingen. Ihre Forschungsarbeiten umfassen die Entwicklung von Algorithmen des Maschinellen Lernens für Probleme im Zusammenhang mit zukünftigen, nachhaltigen Energiesystemen mit hohem Anteil erneuerbarer Energiequellen. Ihr besonderes Interesse gilt dem Probabilistischen Lernen auf Zeitreihen und dem Bestärkenden Lernen mit unsicheren Daten. Frau Ludwig hat am KIT in Informatik promoviert und besitzt einen MSc in Information Systems and Network Economics sowie einen BSc in Economics der Universität Freiburg.

Siehe auch: <https://nicoleludwig.github.io/>

¹ Exzellenzcluster Maschinelles Lernen, Maria-von-Linden Str. 6, 72076 Tübingen, nicole.ludwig@uni-tuebingen.de

Digitalisierungsworkflow zur Strukturierung und Standardisierung von Instandhaltungsinformationen von Windenergieanlagen

Marc-Alexander Lutz^{1,2}, Katharina Beckh³, Jörg Kindermann³, Juliane Schneider⁴, Julia Walgern⁴, Sebastian Pfaffel¹, Stefan Faulstich¹, Alisa Staack¹

Abstract: Die Instandhaltung von Windenergieanlagen erfolgt derzeit zumeist reaktiv und bietet damit Potenzial zur Kostensenkung und Steigerung der Anlagenverfügbarkeit. Die hierzu notwendige genaue Kenntnis der Anlagenzuverlässigkeit erfordert die detaillierte Analyse historischer Instandhaltungsberichte. Daten aus den Instandhaltungsberichten müssen hierzu strukturiert und einheitlich klassifiziert vorliegen. In der Praxis bestehen Einsatzbeschreibungen in Instandhaltungsberichte häufig aus Freitexten, existierende Standards bleiben ungenutzt und der vorhandene Datenbestand bleibt aufgrund des hohen Aufwands einer manuellen Aufbereitung unerschlossen. In dieser Arbeit wird ein Digitalisierungsworkflow skizziert, welcher bestehende Instandhaltungsinformationen nutzbar machen soll. Hierzu werden existierende Verfahren zur Digitalisierung, Informationsextraktion und Klassifizierung vorgestellt und hinsichtlich des Einsatzes im beschriebenen Anwendungsfall bewertet. Diese Verfahren können zu einem Digitalisierungsworkflow kombiniert werden, um die für eine Instandhaltungsoptimierung notwendige Datengrundlage zu erreichen. Die hier betrachteten Verfahren stammen aus dem Bereich der Optical Character Recognition sowie des Natural Language Processing und der Textklassifikation.

Keywords: Windenergie; Instandhaltung; Digitalisierung; Textklassifikation; Optical Character Recognition

1 Einleitung

Betrieb und Instandhaltung verursachen bis zu 25 Prozent der Gesamtkosten im Lebenszyklus einer Windenergieanlage (WEA) [Co21]. Nach der Inbetriebnahme der WEA sind die Instandhaltungskosten oft die einzigen variablen Kosten [In] und damit ein wesentlicher Ansatzpunkt für Maßnahmen zur Kostenreduktion. " Instandhaltung definiert und umfasst alle technischen und administrativen Maßnahmen, die zum Erhalt oder der Wiederherstellung der Funktionsfähigkeit eines Objektes (hier der WEA) notwendig sind "[DIb, DIa].

¹ Fraunhofer-Institut für Energiewirtschaft und Energiesystemtechnik IEE, Königstor 59, 34119 Kassel

² E-Mail korrespondierender Autor: marc-alexander.lutz@iee.fraunhofer.de

³ Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS, Schloss Birlinghoven, 53757 Sankt Augustin

⁴ Fraunhofer-Institut für Windenergiesysteme IWES, Am Seedeich 45, 27572 Bremerhaven

Instandhaltung von WEA ist zumeist mit einem Vor-Ort-Einsatz von Technikern⁵ an der WEA verbunden. Zur Dokumentation dieser Arbeiten werden in der Regel Serviceberichte erstellt. Ein Einsatz kann beispielsweise mit dem Austausch einer oder mehrerer Komponenten, einer turnusmäßigen Wartung oder dem Reset einer Software verbunden sein. Die verschiedenen Arbeiten an einer WEA werden oft durch unterschiedliche Unternehmen, Instandhaltungsteams und Servicetechniker ausgeführt. Da die Prozesse zum Austausch und der Erfassung von Instandhaltungsdaten nicht vereinheitlicht sind, resultiert hieraus eine Vielzahl unterschiedlicher Dokumententypen zu einer Vielzahl an Maßnahmen. All diese Dokumente liegen dem Betreiber bzw. Betriebsführer der WEA vor, der diese archiviert. Eine weitergehende Auswertung der in den Instandhaltungsberichten verfügbaren Informationen ist zum gegenwärtigen Zeitpunkt oft mit einem hohen manuellen Aufwand verbunden und findet daher meist nicht statt.

Zur Senkung der Betriebs- und Instandhaltungskosten bzw. der Erhöhung der Anlagenverfügbarkeit wird in einer Instandhaltungsoptimierung idealerweise zu jeder WEA bzw. jedem WEA-Typ die richtige Instandhaltungsstrategie aus reaktiven und präventiven Maßnahmen ermittelt. Zur Instandhaltungsoptimierung ist eine genaue Kenntnis des Betriebs- und Zuverlässigkeitsverhalten der WEA notwendig, die sich auf Key Performance Indicators (KPIs) wie die Fehlerhäufigkeit von Bauteilen stützt. Entsprechende KPIs können auch zur Verwendung im Rahmen des Ersatzteilmanagements berechnet werden. Wie zuvor beschrieben, ist eine Berechnung von Zuverlässigkeit-KPIs auf Basis der derzeit vorliegenden Dokumente ohne eine aufwendige manuelle Aufbereitung jedoch nicht möglich.

Eine Lösung für das skizzierte Problem kann durch einen umfassenden Workflow zur Digitalisierung von Instandhaltungsinformationen (Digitalisierungsworkflow) erreicht werden. Die Lösung dieser Problemstellung wird auch durch das vom BMWi geförderte Projekt "Digitalisierung von Instandhaltungsinformationen - DigMa"[IE21]⁶ angestrebt. Eines der Ziele des Projektes und Ziel des Digitalisierungsworkflows ist es, die vorliegenden Daten in den Dokumenten zu digitalisieren und zu extrahieren, durch eine Klassifizierung hinsichtlich der betroffenen Bauteile und ergriffenen Maßnahmen nutzbar zu machen und dadurch die Informationen aus den Instandhaltungsberichten verwerten zu können. Auf dieser Basis können Betreiber bspw. Ursachen für Anlagenstillstände oder kostentreibende Komponenten identifizieren und Zuverlässigkeits-KPIs bestimmen.

Derzeit gibt es nur einige Arbeiten zu dem Thema der Digitalisierung und Klassifizierung von Instandhaltungsinformationen, insbesondere in der Windbranche. Carchiolo et al. [Ca19] zeigen die Anwendung einer Ontologie um Fehlerursachen an Erzeugungseinheiten, die in verschiedenen Instandhaltungsdokumenten beschrieben sind, zu vereinheitlichen. Im Gegensatz dazu wird in der vorliegenden Arbeit keine Ontologie verwendet, um eine einheitliche Struktur abzubilden, sondern bereits existierende Branchenstandards dafür verwendet. Brundage et al. [Br21] schlagen einen neuen Bereich innerhalb der Domäne Natural Language Processing (NLP) vor: Technical Language Processing (TLP). TLP umfasst neben den NLP-Verfahren auch Expertenwissen sowie spezielle Fachtermini, um welche der Textkorpus angereichert werden kann. Die vorliegende Arbeit gliedert sich in den vorgeschlagenen Bereich des TLP ein. Gao et al. [Ga20] zeigen den Einsatz einer Pipeline zum maschinellen Lesen von unstrukturierten Daten von Instandhaltungsaufträgen. Der Einsatz der Pipeline ermöglicht ein besseres Verständnis der Daten und weiterführende Zuverlässigkeitsanalysen. Katterbauer [Ka] verwendet Methoden des NLP als Vorstufe einer Fehlerfrüherkennung in WEA. Dabei

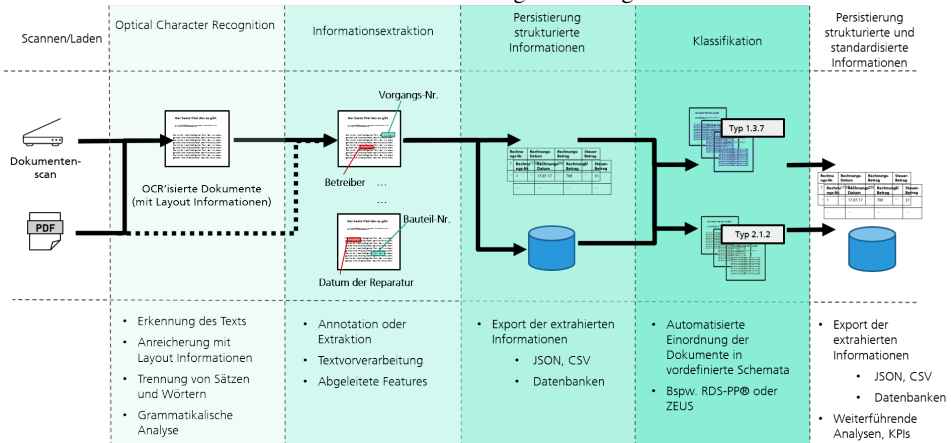
⁵ Folgend wird aus Gründen der besseren Lesbarkeit ausschließlich die männliche Form verwendet. Sie bezieht sich auf Personen beiderlei Geschlechts

⁶ Förderkennzeichen: 03EE2016A

werden zunächst Instandhaltungsberichte mittels NLP erschlossen, um Anlagenfehler zu identifizieren. Anschließend wird ein Modell anhand von WEA-Sensordaten und den auftretenden Fehlern trainiert, um diese Fehler frühzeitig zu erkennen. Sexton et al. [Se17] erläutern den Einsatz von Methoden der Künstliche Intelligenz (KI), um aus unstrukturierten, nicht einheitlichen Instandhaltungsinformationen strukturierte Semantiken abzuleiten. Weiterhin wird aufgezeigt, wie der aufwendige Prozess des Tagging, der oft manuell durch Experten erfolgt, durch den Einsatz der vorgestellten KI-Verfahren vereinfacht wird. Eine deutliche Zeitersparnis ist die Folge. Blanco-M. et al. [BI19] stellen ein Verfahren vor, das in der Historie von Instandhaltungsberichten diejenigen identifiziert, welche Fehler an WEA beschreiben. Ebenfalls relevant sind Arbeiten zur Extrahierung von Informationen aus PDF-Berichten [Oe18, CF04]. In weiteren Arbeiten [Ca21] wird der Einsatz von NLP diskutiert, um die Routenplanung für Instandhaltungsmaßnahmen an WEA zu optimieren.

Idealerweise ist eine Klassifizierung der Instandhaltungsinformationen nach der Vorgabe eines Branchenstandards beziehungsweise einer Richtlinie, wie z.B. Reference Designation System for Power Plants® (RDS-PP®) [VG14] oder Zustands-Ereignis-Ursachen-Schlüssel (ZEUS) [FG10a] nicht notwendig, da diese schon bei der Datenerfassung angewendet wurden. In der Praxis ist dies jedoch oft nicht der Fall. Eine Richtlinie zum einheitlichen Austausch von Daten im Instandhaltungsprozess ist beispielsweise das Globale Service Protokoll [FG10b]. Ob sich die Richtlinien zum Klassifizieren oder zum einheitlichen Austausch von Daten im Instandhaltungsprozess in der Windbranche durchsetzen, kann nicht eindeutig durch die Autoren beantwortet werden. Die Autoren erwarten jedoch, dass durch den Einsatz des Digitalisierungsworkflows die Anwendung von Standards in der Branche beschleunigt werden kann. In den bisher eingesetzten und hochgradig ausgereiften Document Management oder Enterprise Resource Planning Systemen besteht zwar die Möglichkeit Dokumente zu klassifizieren, dies erfolgt jedoch manuell und ist mit hohem Aufwand und entsprechenden Kosten verbunden. Ziel dieses Dokuments ist es, eine Übersicht über die verschiedenen Möglichkeiten innerhalb eines automatisierten Digitalisierungsworkflow zu geben. Der Digitalisierungsworkflow setzt sich aus unterschiedlichen Verfahren zusammen und ist in Abbildung 1 zu sehen. Die einzelnen Verfahren sowie die Ein- und Ausgangsdaten werden detaillierter in Kapitel 3 beschrieben.

Abb. 1: Übersicht über den Digitalisierungsworkflow



Als Bilder vorliegende Instandhaltungsinformationen werden mit Hilfe von Optical Character Recognition (OCR) in Text umgewandelt. Der Text ist jedoch noch unstrukturiert. Daher gilt es im Schritt der Informationsextraktion gleiche Informationen in gleichen Kategorien zu bündeln. Eine Vereinheitlichung der unterschiedlichen Dokumente in Tabellenform ist das Resultat. Ebenso werden in diesem Schritt die Informationen so vorbereitet, dass sie als Eingabe für den Klassifikator verwendet werden können. Darauf aufbauend und auch im abschließenden Schritt kann eine Persistierung der Daten in eine Datenbank erfolgen. Die Persistierung ist weitestgehend automatisiert und wird im Dokument nicht weiter beschrieben. Nach der Informationsextraktion liegen einige der Informationskategorien noch als Freitexte vor, beispielsweise die Beschreibung einer Instandhaltungstätigkeit an einer Komponente der WEA. Diese Freitexte gilt es im letzten Schritt mit Hilfe von Verfahren im Bereich der Textverarbeitung in gängige Richtlinien und Branchenstandards zu klassifizieren. Dadurch werden eindeutige Bezeichner für die Komponenten und Maßnahmen vergeben. Durch die so vorgenommene Strukturierung und Klassifizierung der Daten nach der Vorgabe von Branchenstandards können unter anderem Zuverlässigkeits-KPIs berechnet werden.

Die vorliegende Arbeit gliedert sich wie folgt: Zuerst werden Anforderungen aus der Praxis (Kapitel 2) aufgezeigt, die Problemstellung (Kapitel 2.1) und mögliche Klassifikationsschemata (Kapitel 2.2) zur Vereinfachung der Problemstellung werden diskutiert. Daran schließt sich eine detailliertere Beschreibung der einzelnen Schritte des Digitalisierungsworkflow, wie in Abbildung 1 zu sehen, an. Die Schritte die dabei beschrieben werden sind: OCR (Kapitel 3.1), die Informationsextraktion (Kapitel 3.2) und die Multi-Label-Klassifikation (Kapitel 3.3). Die Schritte werden abschließend in Kapitel 3.4 diskutiert. Zum Schluss des Dokuments wird eine Zusammenfassung und ein Ausblick gegeben (Kapitel 4).

2 Anforderungen aus der Praxis

Instandhaltungsdaten sind in der Praxis heterogen, unterschiedlich strukturiert, nur teilweise standardisiert und werden in unterschiedlichen Formaten und Sprachen erfasst. Für die Analyse von Instandhaltungsdaten müssen diese jedoch einheitlich strukturiert unter Verwendung einheitlicher Klassifizierungssysteme für Komponenten und Tätigkeiten vorliegen. Aus dieser Divergenz leiten sich zwingende Anforderungen an den Digitalisierungsworkflow ab.

In diesem Abschnitt wird zunächst auf unterschiedliche Herausforderungen eingegangen, die die Arbeit mit Instandhaltungsdaten von WEA erschweren. Im Anschluss daran werden mögliche Klassifikationsschemata zum Standardisieren der Daten erläutert.

2.1 Problemstellung

Instandhaltungsinformationen stehen häufig in unterschiedlichsten Instandhaltungsdokumenten zur Verfügung. Die Dateiaustauschformate reichen von digitalisierten Dokumenten, die bereits in strukturierter Form vorliegen (Excel-Tabellen, CSV-Format) bis zu Scans von Briefen im PDF-Format oder als Bild. Instandhaltungsdokumente können je nach Verwendungszweck in Dokumententypen wie beispielsweise Rechnungen, Materiallisten und Reparatur- bzw. Arbeitsprotokolle unterteilt werden. Diese unterscheiden sich maßgeblich in der Informationstiefe zu technischen Aspekten der

Instandhaltungsmaßnahme. So können Reparaturprotokolle eine detaillierte Fehler- und Maßnahmenbeschreibung enthalten, wohingegen Rechnungen für gewöhnlich einen deutlich reduzierten technischen Informationsgehalt aufweisen, da diese lediglich zur grundlegenden Nachvollziehbarkeit für den Auftraggeber dienen. Ein hohe Heterogenität an Dokumenten erschwert die Aufbereitung, da diese an den Aufbau des jeweiligen Dokumententyps angepasst werden muss. Gleichzeitig können nicht aus allen Dokumenten die gleichen Informationen extrahiert werden.

Auch innerhalb einer Dokumentenart kann die Informationstiefe je nach Datenbereinsteller unterschiedlich sein. Hierbei spielt insbesondere die Datenhoheit eine große Rolle: Abhängig davon, wer den Service an der WEA durchführt und welche Dokumentation für diesen Service vereinbart wurde, kann der für den Betreiber zugängliche Informationsgehalt stark variieren. Ein geringer Informationsgehalt erschwert es, die Maßnahmen, die auf der Anlage durchgeführt wurden, nachzuvollziehen. Insbesondere eine Klassifizierung von betroffenen Komponenten und den ergriffenen Maßnahmen kann eine große Herausforderung darstellen.

Die Vereinheitlichung der unterschiedlichen Berichte soll die Informationen in den Dokumenten für weitergehende Analysen nutzbar machen. Ein Teil dieser Informationen liegt jedoch oft als Freitext vor, in dem der Einsatz kurz beschrieben ist. Die Klassifikation von diesem Freitext nach anerkannten Standards soll den Zustand der Anlage und die betroffene Komponente einheitlich beschreiben und somit eine Vergleichbarkeit gewährleisten.

2.2 Klassifikationsschemata

Nachfolgend sollen gängige technische Richtlinien vorgestellt werden, die als Grundlage für eine Klassifikation von Freitexten in Instandhaltungsdaten verwendet werden können. Abschließend soll für eine Klassifikation ein Beispiel gegeben werden.

Die IEA Wind [Ha17] unterscheidet verschiedene Kategorien von Daten im Bezug auf WEA: Betriebsdaten und Messwerte, Anlagendaten, Fehlerdaten und Instandhaltungsdaten. Die nachfolgend zur Klassifizierung vorgestellten Richtlinien und Standards folgen den Vorschlägen der IEA Wind zur Datenkategorie „Instandhaltungsdaten“.

2.2.1 Klassifizierung der Anlagenstruktur

Um ein einheitliches branchenübergreifendes System zur eindeutigen Identifizierung und Strukturierung von WEA und deren Komponenten zu erhalten, ist es notwendig ein gemeinsames Schema für die Anlagenbestandteile zu verwenden. Diesen Zweck erfüllen diverse Normen: RDS-PP® [VG14], NERC-GADS [GA10], eine Ontologie für Objekte von Offshore-WEA basierend auf dem Kommunikationsstandard IEC 61400-25 für Betriebsdaten [Ng14] als auch eine von der IEA Wind aus der DIN ISO 14224 [De04] abgeleitete Taxonomie zur Gliederung von WEA.

Da RDS-PP® den höchsten Detaillierungsgrad und die bisher größte Marktdurchdringung aufweist, soll RDS-PP® im hier skizzierten Digitalisierungsworkflow zur Klassifizierung der Komponenten in Freitexten von Instandhaltungsdaten verwendet werden. Andere vergleichbare Standards sind für

dieses Vorgehen ebenso denkbar, dafür gilt es die Instandhaltungsdaten nach diesem alternativen Standard zu klassifizieren.

RDS-PP® bildet eine funktionale Struktur der WEA bis zur kleinsten austauschbaren Einheit, dem Ersatzteil, ab. Jeder Bestandteil der Anlagenstruktur hat einen eindeutigen Code und ist somit eindeutig beschrieben. RDS-PP® kann auf WEA unterschiedlicher Hersteller angewendet werden. Es bildet somit ein System zur Vergleichbarkeit von Anlagen unterschiedlicher Hersteller und unterschiedlichen Typs. Ebenso ist RDS-PP® unabhängig von der natürlichen Sprache [RRW14] der Beteiligten, was einen internationalen Einsatz ermöglicht. Eine Übersicht über das durch RDS-PP® vorgegebene Schema ist in Abbildung 2 im Anhang zu sehen.

Ebenso wird durch RDS-PP® eine klare Hierarchie zwischen Systemen und Sub-Systemen vorgegeben. So ist es möglich Maßnahmen, die an Sub-Systemen vorgenommen wurden, beispielsweise eine Reparatur der Planetenstufe des Getriebes, auch gleichzeitig dem Antriebsstrangsystem zuzuordnen. Ebenso kann ein wechselnder Informationsgehalt durch die Hierarchie der Systeme berücksichtigt werden.

2.2.2 Klassifizierung von Zuständen und Maßnahmen

Zur Klassifizierung und eindeutigen Beschreibung von Fehler- und Instandhaltungsdaten können laut IEA Wind [Ha17] folgende Normen und Richtlinien verwendet werden: Der ZEUS [FG10a], die ISO 14224 [De04] und NERC-GADS [GA10].

Von den genannten Normen und Richtlinien zur Klassifizierung von Fehler- und Instandhaltungsdaten bietet ZEUS den höchsten Grad der Genauigkeit und einen hohen Freiheitsgrad ohne zu abstrakt zu sein. Daher soll ZEUS im skizzierten Digitalisierungsworkflow für die Klassifizierung von Fehler- und Instandhaltungsdaten verwendet werden. Alternative Standards oder auch unternehmensinterne Klassifizierungsschemata lassen sich ebenfalls anwenden, erfordern jedoch ein individuelles Training.

ZEUS betrachtet sowohl den übergeordneten Zustand der WEA als auch im Detail den Zustand der Komponente im Rahmen eines Fehlers oder einer Instandhaltung. Jeder ZEUS-Block wirft dabei eine Frage auf, die durch einen spezifischen ZEUS-Schlüssel beantwortet werden kann. Durch die Kombinationen der verschiedenen Informationsblöcke bzw. der zugehörigen Schlüssel ist der Zustand der WEA eindeutig beschrieben. Wie RDS-PP® ist ZEUS sprachagnostisch. Das Schema von ZEUS ist exemplarisch in Abbildung 3 im Anhang zu sehen.

2.2.3 Beispielhafte Klassifikation

Nachfolgend soll ein Beispiel für eine Klassifikation von Instandhaltungstätigkeiten, welche als Freitexte vorliegen, gegeben werden. Wie in Abschnitt 2.1 schon beschrieben, liegt nicht innerhalb jedes Instandhaltungsdokuments eine Beschreibung der Instandhaltungstätigkeit (IHT) als Freitext vor. Liegt jedoch eine Beschreibung der IHT als Freitext vor, ist es oft notwendig die Informationen erst mit Hilfe von OCR, Annotation und Extraktion zugänglich zu machen. Ungeachtet dieser Einschränkung, soll hier exemplarisch darauf eingegangen werden, wie diesen Freitexten ein Schema zugeordnet werden kann. Drei exemplarische Freitexte sind in Tabelle 1 im Anhang in der ersten Zeile zu sehen.

Die Klassifikationen zum Schema sind ab der zweiten Zeile zu sehen. Ein Problem bei der Arbeit mit Freitexten wird bei dem Vergleich der ersten und der zweiten IHT offensichtlich. Der semantische Inhalt ist identisch, die Beschreibung jedoch unterschiedlich. Beide Tätigkeiten beschreiben die Inspektion von Rost an dem Azimut-System der WEA.

3 Entwurf des Digitalisierungsworkflows

Um eine Übersicht über die Digitalisierungsverfahren zu erhalten, erfolgt in diesem Abschnitt eine Vorstellung des Stands der Technik im Bereich von OCR, der Informationsextraktion und der Klassifizierung.

3.1 Optical Character Recognition

Die Instandhaltungsberichte liegen in unterschiedlichen Formaten vor. Teilweise sind die Berichte nur als eingescannte Bilder vorhanden. In dieser Form lassen sich die Dokumente nicht durchsuchen und die Instandhaltungsinformationen nicht extrahieren. Um die Texte inhaltlich zu erschließen wird OCR eingesetzt. Somit werden die Bilder in bearbeitbare und durchsuchbare Texte umgewandelt.

Bevor neuronale Netze weitreichend eingesetzt wurden, erfolgte die Texterkennung zeichenweise. Mittlerweile werden Varianten rekurrenter neuronaler Netze eingesetzt, die Text zeilenweise verarbeiten können (siehe z.B. [SJ15]). Die Forschungsentwicklungen der letzten Jahre beziehen sich auf die Verbesserung der Erkennung in Spezialfällen wie beispielsweise seltenerer Sprachen oder starker Verzerrungen, die wir an dieser Stelle nicht weiter behandeln.

Es existieren eine Vielzahl an Verfahren, die OCR ermöglichen. Zu den bekanntesten gehören: ABBYY FineReader [AB21], OmniPage [In21], Tesseract [Te21].

Um die Qualität der OCR-Ergebnisse zu erhöhen, ist häufig eine Qualitätsverbesserung der Eingangsbilder unabdinglich. Aus diesem Grund enthalten einige Produkte eine Komponente zur Bildoptimierung.

In dem beschriebenen Digitalisierungsworkflow findet OCR Anwendung, wenn die Instandhaltungsdokumente in Form von Bilddateien vorliegen. Ist das vorliegende Dokument eine lesbare PDF- oder eine CSV-Datei, so startet der Workflow direkt mit der Informationsextraktion.

Die Autoren können noch nicht abschließend bewerten, welches der vorgestellten OCR-Verfahren sich am besten innerhalb des Digitalisierungsworkflows eignet. Dies ist Aufgabe weiterführender Arbeiten.

3.2 Informationsextraktion

In der Computerlinguistik ist die Informationsextraktion eine zentrale Aufgabe, die in mehrere aufeinander folgende Arbeitsschritte unterteilt ist. Hier wird unterschieden zwischen der Erkennung von benannten Entitäten, der Identifikation von Beziehungen (Relationen) zwischen Entitäten und der

Identifikation von komplexeren Abläufen oder Fakten (Event Detection) (siehe [Ju07] und [JM20] als geplante Neuauflage). Ziel der Informationsextraktion ist es, die in un- oder semi-strukturierten Texten verborgenen Informationen in eine maschinenlesbare Form wie z.B. eine relationale Datenbank zu überführen.

Für den Digitalisierungsworkflow werden die Ergebnisse der Informationsextraktion nicht in einem standardisierten Format benötigt, sondern als Vorverarbeitungsstufe, die für jedes Dokument das Eingabeformat realisiert, das für die Anwendung eines Klassifikators notwendig ist. Die vorgenannten Teilschritte der Informationsextraktion werden daher nur in Kürze beschrieben. Im Kontext dieser Arbeit ist mit der Informationsextraktion derjenige Teilschritt des Workflows gemeint, der ein in der Form als OCR-Resultat vorliegendes Dokument oder PDF-Dokumente, die nicht als Bild vorliegen und somit selektierbaren Text enthalten, in eine Struktur überführen, die für das Training oder die Anwendung eines Klassifikators geeignet ist. Dieser Teilschritt kann in die Unterteilschritte Template Filling, Text Preprocessing und Feature Extraction aufgegliedert werden. Die vorgenannten Teilschritte der Informationsextraktion und gängige hierbei verwendete Verfahren werden nachfolgend vorgestellt.

Template Filling bezeichnet den Vorgang, Text aus einem oder mehreren Dokumenten in eine einheitliche Struktur zu bringen. Beispielsweise wird der Textinhalt von mehreren CSV-Dateien nach Informationen gleicher Kategorie durchsucht. Diese Kategorien werden gruppiert und als Ergebnis gespeichert. Das Ergebnis des Template Filling ist eine tabellarische Struktur, z.B. eine Excel-Datei oder relationale Datenbank, in der zeilenweise der Textinhalt der einzelnen CSV-Dateien zu sehen ist und in Spalten die Informationen gleicher Kategorie zu finden sind. Die Informationen sind übersichtlicher organisiert und leichter zugänglich.

Die Umsetzung dieses Vorgangs kann regelbasiert durch ein Extraktionstool erfolgen oder KI-basiert durch ein Annotationstool. Beide sollen im Nachfolgenden vorgestellt werden. Da diese Vorgänge sehr spezifisch sind und als Vorverarbeitung und Organisation der Daten zu verstehen sind, kann auf keine Quellen und weitere Verfahren verwiesen werden. In der Vergangenheit gab es zu dem Template Filling kein einheitliches Vorgehen [JLBF11], nach Auffassung der Autoren ist dies auch im Jahr 2021 noch so. Daher soll im Folgenden die Funktionsweise für das vorliegende Anwendungsfeld näher beschrieben werden.

Für die Anwendung eines regelbasierten Extraktionstools im skizzierten Digitalisierungsworkflow ist es notwendig für jeden Dokumententyp die Regeln, die zum Erkennen bestimmter Informationskategorien notwendig sind, im Voraus durch den Entwickler zu implementieren. So sind in einem Dokument unter anderen mehrere Kategorien vorhanden: Anschrift des Betreibers, Beschreibung der IHT als Freitext, Arbeitsdauer, etc.. Durch Regeln können diese Kategorien erkannt und gruppiert werden. Die Informationskategorie „Beschreibung der IHT als Freitext“ wird beispielsweise durch den Textbaustein „Tätigkeitsbeschreibung“ eingeleitet. Wird dieser Textbaustein im Text erkannt, wird der Text, welcher darauf bis zum nächsten Textbaustein folgt, dieser Kategorie zugewiesen.

Ein Annotationstool ist hingegen KI-basiert. Hier werden die Regeln nicht durch den Entwickler implementiert, sondern von einem Algorithmus gelernt. Dafür gilt es in einer Benutzeroberfläche gleiche Informationskategorien in mehreren Texten zu annotieren. Sind ausreichend viele Texte markiert, so wird dieser Zusammenhang zwischen annotiertem Text und Informationskategorie erkannt und kann auf weitere Dokumente gleichen Typs angewandt werden.

Im Digitalisierungsworkflow soll ein regelbasiertes Extraktionstool eingesetzt werden, da hier der zu erwartende Aufwand am geringsten eingeschätzt wird. Das Annotationstool benötigt ca. hundert annotierte Texte pro Informationskategorie um die Zuordnung der Informationskategorien zu erlernen. Der Aufwand dieser Annotation wird höher eingeschätzt, als eine softwaretechnische Implementierung dieser Regeln. Ein Vergleich der beiden Tools ist Gegenstand zukünftiger Arbeiten.

Text Preprocessing Im Folgenden werden die typischen Schritte für eine Text-Vorverarbeitung beschrieben. Diese sind Kowsari et. al [Ko19] entnommen. Ausgangspunkt ist eine Textrepräsentation als String bestehend aus einzelnen Zeichen, heute im Allgemeinen in der UTF-8 Codierung. Wichtig ist die Voraussetzung, dass alle Teilschritte vollautomatisch durchgeführt werden können. Wir werden in dieser Aufzählung jedoch nicht auf die technischen Realisierungsmöglichkeiten eingehen, da dies den Rahmen sprengen würde, sondern die Schritte in Hinblick auf den Anwendungsfall, der Nutzbarmachung von Instandhaltungsinformationen, diskutieren. Einige Schritte sind durch einfache Programme realisierbar (wie z.B. die Stopwortentfernung), andere erfordern bereits die Anwendung von maschinellen Lernverfahren.

- **Tokenisierung** [Ko19] bezeichnet die sinnvolle Zerlegung einer Zeichenfolge in einzelne Teilstrings. In der Regel handelt es sich hier um die Zerlegung des Satzes in einzelne Worte oder auch Abkürzungen.
- **Entfernung von Stopworten** [Ko19] Hier handelt es sich um die Löschung bestimmter Worte aus der Textrepräsentation, die für die im Folgenden auszuführende Aufgabe (z.B. Textklassifikation) keinen oder störenden Einfluss haben können, wie z.B. häufig unabhängig von der Textklasse vorkommende Worte, die keine sinngebende Rolle bei der Klassifizierung spielen. Im skizzierten Digitalisierungsworkflow ist es schwierig, die Stopworte in den zur Verfügung stehenden kurzen Texten der Instandhaltungsberichte zu identifizieren. Daher wird dieser Verarbeitungsschritt nicht genutzt.
- **Angleichung von Groß-/Kleinschreibung** [Ko19] Besonders für Sprachen, die außer am Satzanfang keine Regeln zur Großschreibung haben, werden zur weiteren Verarbeitung häufig alle Worte in Kleinschreibung überführt, damit identische Begriffe, die zufällig am Satzanfang stehen, nicht doppelt verarbeitet werden.
- **Auflösung von Abkürzungen** [Ko19] Dies wurde in älteren Arbeiten häufig genutzt. Es bestehen hier jedoch Schwierigkeiten, dies konsistent und vor allem vollautomatisch umzusetzen. Es werden Listen von verwendeten Abkürzungen benötigt, die in der Regel nicht vollständig sein können. Außerdem existieren viele mehrfach verwendete Abkürzungen, die nur im jeweiligen Kontext eindeutig sind und die daher nur mit menschlichem Textverständnis entschlüsselt werden können.
- **Entfernung von unerwünschten Zeichen** [Ko19] Hierzu zählen je nach Zweck (mehrfache) Leerzeichen, Tabulatoren, und Interpunktionszeichen.
- **Rechtschreibkorrektur** [Ko19] In letzter Zeit hat die Rechtschreibkorrektur wesentlich an Bedeutung verloren, denn neuere maschinelle Lernverfahren sind in der Lage, auch fehlerhaft geschriebene Begriffe meistens richtig zuzuordnen. Die gängigen einfachen Methoden zur Korrektur sind sprachabhängig und nicht kontext-spezifisch und können daher dazu führen, dass Begriffe gerade durch die Korrektur verfälscht werden. Im vorliegenden Fall kommt

erschwerend hinzu, dass häufig Texte mit deutschen und englischen Abschnitten gemischt bearbeitet werden müssen.

- **Überführung in Stammformen** [Ko19] Hierunter ist die Entfernung von Wortteilen, die sich unter Deklination oder Konjugation ändern, zu verstehen. So wird beispielsweise „Achsen“ zu „Achs“ und „untersucht“ zu „untersuch“.
- **Überführung in Grundformen** [Ko19] Dies bezeichnet die Umwandlung von deklinierten oder konjugierten Worten in ihre Grundformen, also Infinitiv für Verben und Nominativ für Nomen und Adjektive.

Einige dieser Verarbeitungsschritte wurden bereits in den Erläuterungen als in diesem Falle ungeeignet gekennzeichnet. Die übrigen Schritte sollen im praktischen Test des Digitalisierungsworkflows darauf überprüft werden, ob sie zu einer Verbesserung der Ergebnisse führen oder nicht.

Feature Extraction oder die Extraktion von Merkmalen bezeichnet die an die Vorverarbeitung anschließende Umwandlung des Textes in ein standardisiertes Eingabeformat, das für das zu trainierende maschinelle Lernverfahren zweckmäßig ist. Für die Zwecke des skizzierten Digitalisierungsworkflows sind die folgenden beiden Ansätze relevant.

- **Bag of Words** bezeichnet eine Textrepräsentation, bei der der Text in Form von hochdimensionalen Vektoren gespeichert wird [Ko19]. Zunächst wird anhand einer Referenzmenge von Texten (im Zusammenhang mit maschinellen Lernverfahren ist dies meistens die Trainingsmenge von Texten) eine Liste aller vorkommenden Worte erstellt. Wenn in der Vorverarbeitung der Texte keine Stammformen oder Grundformen gebildet werden, beinhaltet diese Liste auch alle deklinierten und konjugierten Formen. Ein Text wird nun durch einen Vektor codiert, dessen Dimension gleich der Anzahl der Worte der Wortliste ist. Die Komponenten des Vektors werden anhand der Wortliste benannt. Die Komponenten, deren Worte nicht im Text vorkommen, erhalten den Wert Null. Die Komponenten der vorkommenden Worte erhalten Werte größer als Null. Hier gibt es unterschiedliche Möglichkeiten der Wertzuweisung, wie z.B. die Worthäufigkeit im Text oder deren Logarithmus. Des Weiteren kann der Komponentenwert noch mit einem global errechneten Einflussgewicht wie TF-IDF [Ko19, LK02] multipliziert werden. Bei dieser Art der Merkmalsextraktion geht die Reihenfolge der Worte im Text als Information verloren. Wenn in einem neuen Text Worte vorkommen, die nicht in der Trainingskollektion vorhanden waren, so können diese nicht codiert werden und müssen in der weiteren Verarbeitung ignoriert werden.
- **Worteinbettungsvektoren** Nicht nur Texte, sondern auch Worte werden seit langem in Form von mittel- oder hochdimensionalen Vektoren dargestellt [CH90, Ko19]. Neue Dynamik erhielt die Entwicklung durch die Arbeiten von Mikolov und anderen [Mi13, Bo17]. Obwohl die Berechnung von Wortvektoren durch diese neuen Entwicklungen effizient möglich wurde, bedeutet die Erstellung einer großen Zahl von Einbettungsvektoren einen beträchtlichen Aufwand. Inzwischen wurden solche vortrainierten Einbettungsvektoren auch zur freien Verwendung bereitgestellt, wie z.B. [PSM14]. Allerdings handelt es sich bei diesen Sammlungen in der Regel um Verzeichnisse, die auf Basis von Corpora mit breitem Themenbereich wie Wikipedia Dumps erzeugt wurden. Sie sind für unsere Zwecke eher ungeeignet, da viele

technische Begriffe nicht enthalten sind oder eine abweichende Verwendung in den Corpus Texten hatten. Worten in neuen Texten, die nicht in der Liste der Einbettungen vorhanden sind (sog. OOV = Out of Vocabulary Words), wird meist ein spezieller Vektor, dessen Komponenten zufällig besetzt sind, oder der Nullvektor zugeordnet. Darüber hinaus existieren Verfahren, die jedem OOV-Wort einen speziell berechneten Vektor zuordnen, wie z.B. [Bo17].

3.3 Klassifikation

Die Klassifikation von Datensätzen strukturierter Daten wie z.B. wirtschaftliche Kennzahlen oder unstrukturierter Daten wie insbesondere Freitexten ist seit jeher ein zentrales Anwendungsthema für künstliche Intelligenz [Ko19]. Unterschieden wird zwischen der Aufgabe einer eindeutigen Klassifikation (Multi-Class-Klassifikation) und einer Zuweisung von mehreren relevanten Klassen für einen Datensatz (Multi-Label-Klassifikation). Die beiden in Abschnitt 2.2 beschriebenen Klassifikations-Schemata RDS-PP® und ZEUS sind Multi-Label-Schemata und werden für die Klassifizierung der Freitexte in den Instandhaltungsberichten verwendet.

Maschinelle Lernverfahren codieren zunächst die aus einer Lernmenge von Daten (Trainingsdaten) abgeleiteten Regularitäten in einer Parametrisierung - auch Modell genannt. Dieses Modell kann daraufhin automatisch ohne menschlichen Eingriff auf neue Daten angewendet werden, um diese zu klassifizieren. Die von einem Lernverfahren erzeugten Modelle sollten vor Nutzung in Bezug auf die erwartbare Klassifikationsgüte auf einer unabhängig von den Trainingsdaten gewählten Testdatenmenge evaluiert werden.

Für den skizzierten Digitalisierungsworkflow gilt es das beste maschinelle Lernverfahren für die automatische Klassifikation der Instandhaltungsberichte auszuwählen. Es ist daher essentiell, eine möglichst große Anzahl von maschinenlesbaren Instandhaltungsberichten zur Verfügung zu stellen, um Modelle aus Trainingsdaten zu generieren und diese zur Auswahl des besten Verfahrens auf Testdaten zu evaluieren.

Im Anwendungsfall dieser Veröffentlichung handelt es sich um Kategorienanzahlen von 100 bis maximal ca. 1000 Kategorien. Daher wird sich auf entsprechende Verfahren konzentriert, die für diesen Bereich veröffentlicht wurden. Verfahren, die Kategorienanzahlen im Millionenbereich verarbeiten können, sind für den Digitalisierungsworkflow nicht relevant.

- **One-Vs-All (OVA)** Ansätze aus diesem Bereich nutzen einen trainierbaren Klassifikator für jede der zu unterscheidenden Klassen. Aus Gründen der Performanz wurde hierfür seit längerer Zeit die Support Vector Machine (SVM) [Jo98, LK02] bevorzugt. Die SVM ist ursprünglich für die Single-Label-Klassifikation gedacht. Durch eine nachgeschaltete Softmax-Funktion kann eine Kombination mehrerer SVMs auch für Multi-Label-Klassifikation verwendet werden. Allerdings skaliert der Trainingsaufwand linear mit der Anzahl der zu unterscheidenden Kategorien, da für jede Kategorie ein eigenes SVM-Modell trainiert werden muss. Neuere Arbeiten schlagen eine Effizienzsteigerung durch parallelisiertes Training [BS17] sowie durch eine vorab durchgeführte statistische Analyse der Häufigkeitsverteilung der Kategorien in den Trainingsdaten [Ye16] vor.
- **Partitionierungs-Methoden** Für Klassifikationsaufgaben - sowohl Multi-Class als auch Multi-Label - ist die Instanz-Label-Matrix definiert als $Y \in \{0, 1\}^{N \times L}$. Hier ist N die Anzahl der

Trainingsdaten und L die Anzahl der Kategorien. Bei großen L ist die Instanz-Label Matrix tendenziell eher spärlich besetzt. Daher kann eine Partitionierung der Label, die als duales Optimierungsproblem formuliert wird, zu einer sublinearen Trainingszeit bei gleichzeitig guter Fehlermarge führen. Diesen Ansatz verfolgen z.B. [Ye16] und [Pr18]. Es handelt sich hier im Vergleich zu SVM noch nicht um etablierte Verfahren, sondern um Beiträge aus der aktuellen Forschung.

- **Einbettungsbasierte Methoden** Diese Klassifikatoren basieren auf dem Prinzip der Worteinbettungsvektoren. Diesem Ansatz liegt das Verfahren zugrunde, die hochdimensionale und spärliche Kookkurrenz-Matrix der in den Trainingsdaten vorkommenden Worte auf einen niedrigdimensionalen latenten Vektorraum abzubilden (siehe auch Abschnitt Feature Extraction). Die erste effiziente Implementierung dieses Ansatzes ist Word2Vec [Mi13]. Ein erster Erfolg von einbettungsbasierten Methoden konnte im letzten Jahrzehnt mit Convolutional Neural Networks (CNNs) erzielt werden, die zunächst für die Klassifikation von Bildern entwickelt wurden. Sie wurden dann auch für die Textklassifikation eingesetzt ([Wa12]). Die Beziehung von Worteinbettungsvektoren zu Kookkurrenz-Matrizen durch Faktorisierung wurde von [LG14] bewiesen. Das StarSpace Programm [Wu17] nutzt den Zusammenhang mit Kookkurrenz-Matrizen für das Training der Multi-Label-Klassifikation, indem nicht nur die Worte eines Textes, sondern auch die Kategorien in einen latenten Vektorraum abgebildet werden. Beckh und Kindermann [KB20] nutzen StarSpace zur Klassifikation auf großen Textsammlungen von Abstracts von Patenten und naturwissenschaftlichen Publikationen.
- **Deep Learning Methoden** Auf Basis von vortrainierten Worteinbettungsvektoren wie FastText [Jo16] und ebenfalls vortrainierten Sprachmodellen wie BERT [De18] versuchen neuere Arbeiten [Ch19] [ZLR20], durch sogenanntes Refinement die Sprachmodelle auch für Klassifikation zu nutzen. Für den Anwendungsfall der Instandhaltungsberichte ergeben sich generelle Nachteile: Die verfügbaren vortrainierten Sprachmodelle sind mit Textdaten erzeugt worden, die sehr weit von der Thematik der Instandhaltungsberichte entfernt sind. Daher ist zu erwarten, dass für das Refinement Training ebenfalls große Mengen von einschlägigen deutschen und englischen Textdaten benötigt werden, die uns nicht zur Verfügung stehen. Die Erfolgsaussichten für das Refinement werden zudem kritisch gesehen (siehe [PKC21]).
- **Ensemble Methoden** Es handelt sich hier um Ansätze, die die Berechnungen von mehr als einem Klassifikator zusammenfassen [Ro19] mit dem Ziel, die Klassifikationsgenauigkeit zu erhöhen. Eine Anzahl von Klassifikatoren kann durch eine Aufteilung der vorhandenen Trainingsdaten in mehrere Teilmengen, oder durch Zusammenfassung von Ergebnissen unterschiedlicher Klassifikationsverfahren, oder eine Kombination beider Techniken erreicht werden. In diesem Sinne ergibt sich eine Überschneidung mit den unter Partitionierungs-Methoden genannten Verfahren.

3.4 Diskussion und Auswahlprozess der Verfahren

Die Verfahren OCR, Informationsextraktion und Klassifizierung sollen auf Basis von unterschiedlichen Kriterien bewertet und so hinsichtlich ihrer Eignung für den Digitalisierungsworkflow untersucht werden. Die Kriterien sind: Genauigkeit, Ressourceneinsatz, Komplexität und Fehlermöglichkeiten. Die Kriterien sollen im Folgenden kurz erläutert werden. Die Genauigkeit beschreibt die Qualität des Ergebnisses eines Verfahrens. Die Genauigkeit für die Verfahren Extraktion und Annotation

unterscheidet sich von der Genauigkeit der Klassifizierung. Ersteres bezeichnet den Anteil von vorliegenden Dokumenten, die in ein gemeinsames Schema überführt werden konnten. Zweiteres bezeichnet geläufige Machine Learning-Metriken wie z.B. die "Accuracy" eines Klassifikators. Der Aufwand der bei dem Einsatz der Verfahren notwendig ist, wird unter dem Kriterium Ressourceneinsatz diskutiert. Die Komplexität und die Fehlermöglichkeiten beschreiben die Verfahren hinsichtlich selbiger Kriterien.

Erste Experimente mit SVM-Klassifikation erzielten vielversprechende Resultate. Es ist zu erwarten, dass weitere Arbeiten ein gutes Ergebnis erzielen, da bei ähnlichen kurzen Texten aus Nachrichten Edda und Kindermann [LK02] gute Resultate erhielten. Weiterführende Arbeiten werden außerdem CNN Architekturen [Wa12] und das StarSpace Programm [Wu17] anhand der projektspezifischen Daten auf Performanz prüfen. Nachdem die Verfahren implementiert und anhand der vorgenannten Kriterien untersucht wurden, sollen die am besten bewerteten Algorithmen in den Digitalisierungsworkflow eingesetzt werden. Diese Schritte sind Teil weiterführender Arbeiten.

Mit den Verfahren können nur dann gute Ergebnisse erzielt werden, wenn ausreichend Daten als Eingabe für die Verfahren vorhanden sind. Diese Daten werden durch die Praxispartner im Rahmen des DigMa-Projektes [IE21] bereitgestellt. Für die Anwendung des Klassifikators ist es notwendig, dass die Instandhaltungsdaten nach der Vorgabe des Klassifikationsschemas (siehe Kapitel 2.2) klassifiziert vorliegen. Wie in Kapitel 2.1 aufgezeigt, ist dies in der Praxis oft nicht der Fall. Um dennoch die erforderliche Datengrundlage zu erhalten, werden nicht-klassifizierte Instandhaltungsdaten durch einen Experten mit den entsprechenden Kategorien versehen.

Wichtige Anwendungsfälle, die sich an eine Strukturierung und Standardisierung von Instandhaltungsdaten von WEA anschließen sind die Erstellung von KPIs [Na21] und Zuverlässigkeitsanalysen. Es ist davon auszugehen, dass der Digitalisierungsworkflow viele weitere Anwendungszwecke ermöglicht.

4 Zusammenfassung und Ausblick

Das vorliegende Paper skizziert die notwendigen Schritte eines Digitalisierungsworkflows um die in der Windbranche häufig noch ungenutzten Berichte aus der Instandhaltung von WEA in eine zur Auswertung und Zuverlässigkeitsanalyse nutzbare Form zu überführen. Die zu Beginn vorgestellte Analyse der Problemstellungen aus der Praxis zeigt, dass die größte Hürde in der unstrukturierten und heterogenen Form der vorliegenden Dokumente liegt, welche u.a. stark durch die Vielzahl der Akteure in der Branche verursacht wird. Die oftmals in Instandhaltungsberichten verwendeten Freitexte lassen sich zum heutigen Stand nur manuell auswerten, sollen zukünftig jedoch automatisiert für die Datenanalyse erschlossen werden. Die zur Klassifizierung der Instandhaltungsobjekte, Zustände und Maßnahmen erforderlichen Richtlinien sind mit RDS-PP® und ZEUS in der Windbranche verfügbar. Nur durch eine automatisierte Aufbereitung und Klassifizierung von Berichten wird die Analyse ganzer Instandhaltungshistorien denkbar.

Der skizzierte Digitalisierungsworkflow besteht aus den Schritten der Datenbereitstellung, der Vorverarbeitung und ggf. notwendigen Anwendung von OCR, der Datenextraktion bzw. Annotation, der Persistierung und der Klassifikation. Die Herausforderung des Workflows liegt insbesondere in der Datenextraktion und der Klassifikation. Durch die in der kürzeren Vergangenheit schnellen Fortschritte im Feld des Natural Language Processing stehen für die Textklassifikation eine Vielzahl

an Verfahren zur Verfügung, die auf ihre Eignung im konkreten Anwendungsfall hin überprüft werden müssen.

Die nächsten Schritte liegen in der prototypischen Implementierung des skizzierten Digitalisierungsworkflows und der Erprobung der verschiedenen Verfahren der Textklassifikation im beschriebenen Anwendungsfall. Basis hierzu sind Instandhaltungsberichte, die im Rahmen des Forschungsprojektes DigMa derzeit als Trainingsdaten aufbereitet werden. Der anschließende Praxistest des Prototypen soll helfen den Mehrwert der Nutzbarmachung der bisher verborgenen Informationen zu quantifizieren.

A Anhang

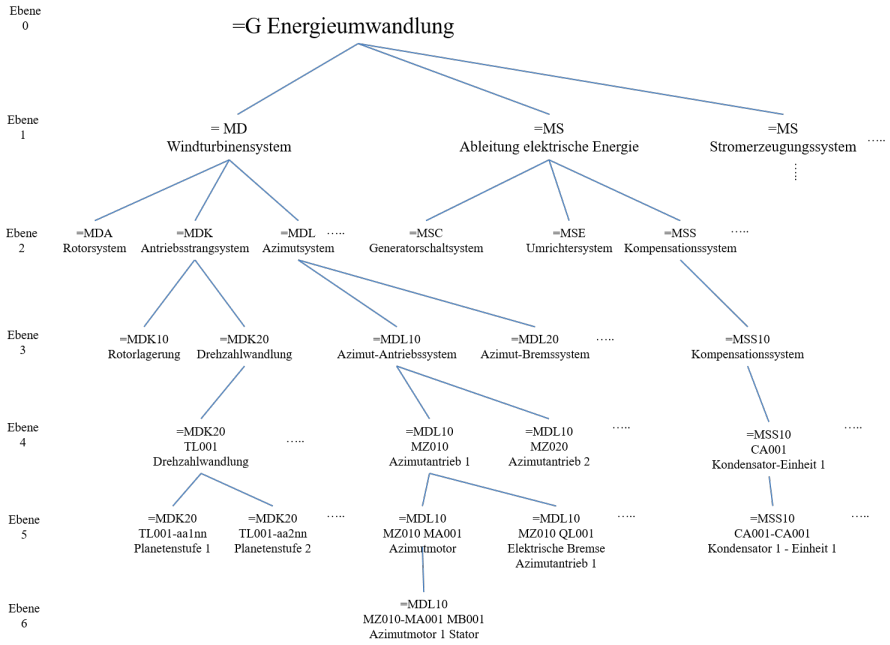


Abb. 2: Funktionale Anlagenstruktur des RDS-PP®, Eigene Darstellung der Autoren auf Basis von [VG14]

	IHT 1	IHT 2	IHT 3
Beschreibung der IHT als Freitext	Inspektion Oberflächenschaden am Azimut System	Überprüfung Rost Windrichtungsnachführung	Fehler aufgrund von Eisansatz am Rotorblatt, Anlage abgeschaltet
RDS-PP® Bezeichnung	Azimut System	Azimut System	Rotor Blattsystem 1
RDS-PP® ID	=MDL	=MDL	=MDA11
ZEUS 01-01 Bezeichnung	Betriebsbedingter manueller Stopp	Betriebsbedingter manueller Stopp	Außer Betrieb störungsbedingt
ZEUS 01-01 ID	01-01-02-02-01	01-01-02-02-01	01-01-02-02-04
ZEUS 01-02 Bezeichnung	Funktionsfähiger Zustand	Funktionsfähiger Zustand	Anlagenbedingte Störung
ZEUS 01-02 ID	01-02-01	01-02-01	01-02-03
ZEUS 01-04 Bezeichnung	Eingriff durch Personen innerhalb Anlage	Eingriff durch Personen innerhalb Anlage	Ereignis mit externer Störung
ZEUS 01-04 ID	01-04-01-01	01-04-01-01	01-04-03
ZEUS 02-01 Bezeichnung	Abweichung vom Sollzustand ohne Handlungsbedarf	Abweichung vom Sollzustand ohne Handlungsbedarf	Fehler
ZEUS 02-01 ID	02-01-02-01	02-01-02-01	02-01-03-02
ZEUS 02-05 Bezeichnung	Merkmal nicht relevant	Merkmal nicht relevant	Eis
ZEUS 02-05 ID	02-05-XX	02-05-XX	02-05-02-05
ZEUS 02-08 Bezeichnung	Merkmal nicht relevant	Merkmal nicht relevant	Sofortige korrektive IH
ZEUS 02-08 ID	02-08-XX	02-08-XX	02-08-01-02
ZEUS 02-09 Bezeichnung	Merkmal nicht relevant	Merkmal nicht relevant	Wiederherstellung
ZEUS 02-09 ID	02-09-XX	02-09-XX	02-09-08
ZEUS 02-12 Bezeichnung	Verkrustung / Verschlackung / Versalzung / Ansinterung	Verkrustung / Verschlackung / Versalzung / Ansinterung	Vereisung
ZEUS 02-12 ID	02-12-01-03	02-12-01-03	02-12-01-04

Tab. 1: Beispielhafte Klassifikation zum Schema

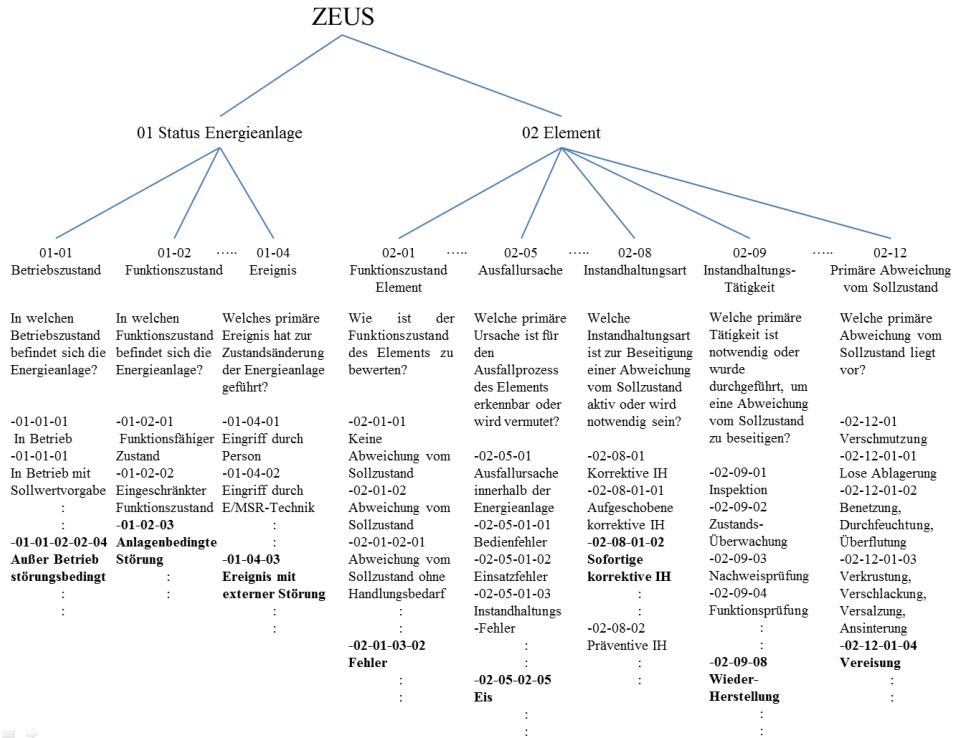


Abb. 3: Struktur des ZEUS, Eigene Darstellung der Autoren auf Basis von [FG10a]

Literaturverzeichnis

- [AB21] ABBYY: , <https://www.abbyy.com>, 11.03.2021.
- [Bl19] Blanco-M., Alejandro; Marti-Puig, Pere; Gibert, Karina; Cusidó, Jordi; Solé-Casals, Jordi: A Text-Mining Approach to Assess the Failure Condition of Wind Turbines Using Maintenance Service History. *Energies*, 12(10):1982, 2019.
- [Bo17] Bojanowski, Piotr; Grave, Edouard; Joulin, Armand; Mikolov, Tomas: Enriching Word Vectors with Subword Information. *Transactions of the Association for Computational Linguistics*, 5:135–146, 2017.
- [Br21] Brundage, Michael P.; Sexton, Thurston; Hodkiewicz, Melinda; Dima, Alden; Lukens, Sarah: Technical language processing: Unlocking maintenance knowledge. *Manufacturing Letters*, 27:42–46, 2021.
- [BS17] Babbar, Rohit; Schölkopf, Bernhard: Dismec: Distributed sparse machines for extreme multi-label classification. In: *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*. S. 721–729, 2017.
- [Ca19] Carchiolo, Vincenza; Longheu, Alessandro; Di Martino, Vincenzo; Consoli, Niccolo: Power Plants Failure Reports Analysis for Predictive Maintenance. In: *Proceedings of the 15th International Conference on Web Information Systems and Technologies*. SCITEPRESS - Science and Technology Publications, S. 404–410, 18.09.2019 - 20.09.2019.
- [Ca21] Carchiolo, Vincenza; Longheu, Alessandro; Malgeri, Michele; Mangioni, Giuseppe; Trapani, Natalia: Wind Farms Maintenance Optimization Using a Pickup and Delivery VRP Algorithm. In (Ziemba, Ewa; Chmielarz, Witold, Hrsg.): *Information Technology for Management: Towards Business Excellence*, Jgg. 413 in *Lecture Notes in Business Information Processing*, S. 67–86. Springer International Publishing, Cham, 2021.
- [CF04] Chao, Hui; Fan, Jian: Layout and Content Extraction for PDF Documents. In (Hutchison, David; Kanade, Takeo; Kittler, Josef; Kleinberg, Jon M.; Mattern, Friedemann; Mitchell, John C.; Naor, Moni; Nierstrasz, Oscar; Pandu Rangan, C.; Steffen, Bernhard; Sudan, Madhu; Terzopoulos, Demetri; Tygar, Dough; Vardi, Moshe Y.; Weikum, Gerhard; Marinai, Simone; Dengel, Andreas R., Hrsg.): *Document Analysis Systems VI*, Jgg. 3163 in *Lecture Notes in Computer Science*, S. 213–224. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004.
- [CH90] Church, Kenneth; Hanks, Patrick: Word association norms, mutual information, and lexicography. *Computational linguistics*, 16(1):22–29, 1990.
- [Ch19] Chang, Wei-Cheng; Yu, Hsiang-Fu; Zhong, Kai; Yang, Yiming; Dhillon, Inderjit: X-BERT: eXtreme Multi-label Text Classification with using Bidirectional Encoder Representations from Transformers. *arXiv preprint arXiv:1905.02331*, 2019.
- [Co21] Costa, Ángel M.; Orosa, José A.; Vergara, Diego; Fernández-Arias, Pablo: New Tendencies in Wind Energy Operation and Maintenance. *Applied Sciences*, 11(4):1386, 2021.
- [De18] Devlin, Jacob; Chang, Ming-Wei; Lee, Kenton; Toutanova, Kristina: Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.

- [De04] Deutsches Institut für Normung: , Erdöl-, petrochemische und Erdgasindustrie - Sammlung und Austausch von Zuverlässigkeits- und Wartungsdaten für Ausrüstungen (ISO 14224:2016, korrigierte Fassung 2016-10-01);, 2017-04.
- [Dia] DIN 31051:2019-06, Grundlagen der Instandhaltung.
- [Dib] DIN EN 13306:2018-02, Instandhaltung_ - Begriffe der Instandhaltung; Dreisprachige Fassung EN_13306:2017.
- [FG10a] FGW e.V. - Fördergesellschaft Windenergie und andere Erneuerbare Energien: , Technische Richtlinie für Erzeugungseinheiten - Teil 7: Instandhaltung von Kraftwerken für Erneuerbare Energien - Rubrik D2: Zustands-Ereignis-Ursachen-Schlüssel für Erzeugungseinheiten (ZEUS), 2013-10.
- [FG10b] FGW e.V. - Fördergesellschaft Windenergie und andere Erneuerbare Energien: , Technische Richtlinie für Erzeugungseinheiten - Teil 7: Instandhaltung von Kraftwerken für Erneuerbare Energien - Rubrik D3: "Globales Service Protokoll (GSP)" Einheitliches Datenformat für den elektronischen Austausch von Daten im IH-Prozess, 2013-10.
- [GA10] GADS Wind Turbine Generation: Data Reporting Instructions: Effective January 2010, 2010.
- [Ga20] Gao, Yiyang; Woods, Caitlin; Liu, Wei; French, Tim; Hodkiewicz, Melinda: Pipeline for Machine Reading of Unstructured Maintenance Work Order Records. In (Baraldi, Piero; Di Maio, Francesco; Zio, Enrico, Hrsg.): Proceedings of the 30th European Safety and Reliability Conference and 15th Probabilistic Safety Assessment and Management Conference. Research Publishing Services, Singapore, S. 1401–1408, 01.11.2020 - 05.11.2020.
- [Ha17] Hahn, Berthold: , 17. Wind Farm Data Collection and Reliability Assessment for O&M Optimization: Expert Group Report on Recommended Practices: First Edition, 2017, 2017.
- [IE21] IEE, Fraunhofer: , https://www.iee.fraunhofer.de/de/presse-infothek/Presse-Medien/Pressemitteilungen/2020/PM_DigMa.html, 28.04.2021.
- [In] International Renewable Energy Agency: , Renewable Energy Cost Analysis - Wind Power: Volume 1: Power Sector.
- [In21] Inc, Kofax: , <https://www.kofax.com/Products/omnipage>, 11.03.2021.
- [JLBF11] Jean-Louis, Ludovic; Besançon, Romaric; Ferret, Olivier: Text segmentation and graph-based method for template filling in information extraction. In: Proceedings of 5th International Joint Conference on Natural Language Processing. S. 723–731, 2011.
- [JM20] Jurafsky, Daniel; Martin, James H: , Speech and language processing (draft), 2020.
- [Jo98] Joachims, Thorsten: Text categorization with support vector machines: Learning with many relevant features. In: European conference on machine learning. Springer, S. 137–142, 1998.
- [Jo16] Joulin, Armand; Grave, Edouard; Bojanowski, Piotr; Mikolov, Tomas: Bag of tricks for efficient text classification. arXiv preprint arXiv:1607.01759, 2016.
- [Ju07] Jurafsky, Dan: Speech and language processing. Prentice Hall, New York, 2rev. ed.. Auflage, 2007.

- [Ka] Katterbauer, Klemens: , An integrated NLP Big Data framework for enhancing wind turbine reliability.
- [KB20] Kindermann, Jörg; Beckh, Katharina: Fusing Multi-label Classification and Semantic Tagging. 2020.
- [Ko19] Kowsari; Meimandi, Jafari; Heidarysafa; Mendu; Barnes; Brown: Text Classification Algorithms: A Survey. *Information*, 10(4):150, 2019.
- [LG14] Levy, Omer; Goldberg, Yoav: Neural word embedding as implicit matrix factorization. In: *Advances in neural information processing systems*. S. 2177–2185, 2014.
- [LK02] Leopold, Edda; Kindermann, Jörg: Text categorization with support vector machines. How to represent texts in input space? *Machine Learning*, 46(1-3):423–444, 2002.
- [Mi13] Mikolov, Tomas; Sutskever, Ilya; Chen, Kai; Corrado, Greg S; Dean, Jeff: Distributed representations of words and phrases and their compositionality. In: *Advances in neural information processing systems*. S. 3111–3119, 2013.
- [Na21] Navinchandran, Madhusudanan; Sharp, Michael E.; Brundage, Michael P.; Sexton, Thurston B.: Discovering critical KPI factors from natural language in maintenance work orders. *Journal of Intelligent Manufacturing*, 2021.
- [Ng14] Nguyen, Trinh Hoang; Dunlap, Rocky; Mark, Leo; Prinz, Andreas; Østgren, Bjørn Mo; Friisø, Trond: Offshore wind metadata management. *International Journal of Metadata, Semantics and Ontologies*, 9(4):333, 2014.
- [Oe18] Oevermann, Jan: Semantic PDF Segmentation for Legacy Documents in Technical Documentation. *Procedia Computer Science*, 137:55–65, 2018.
- [PKC21] Perez, Ethan; Kiela, Douwe; Cho, Kyunghyun: True Few-Shot Learning with Language Models. *arXiv preprint arXiv:2105.11447*, 2021.
- [Pr18] Prabhu, Yashoteja; Kag, Anil; Harsola, Shrutendra; Agrawal, Rahul; Varma, Manik: Parabel: Partitioned label trees for extreme classification with application to dynamic search advertising. In: *Proceedings of the 2018 World Wide Web Conference*. S. 993–1002, 2018.
- [PSM14] Pennington, Jeffrey; Socher, Richard; Manning, Christopher D: GloVe: Global vectors for word representation. In: *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*. S. 1532–1543, 2014.
- [Ro19] Rokach, Lior: *Ensemble learning: Pattern classification using ensemble methods*, Jgg. 85. World Scientific, 2019.
- [RRW14] Richnow, Jörg; Rossi, Clemens; Wank, Helmut: Kennzeichnung von Windkraftwerken mit dem Reference Designation System for Power Plants – RDS-PP®. *VGB PowerTech*, 94(7), 2014.
- [Se17] Sexton, Thurston; Brundage, Michael P.; Hoffman, Michael; Morris, K. C.: Hybrid datafication of maintenance logs from AI-assisted human tags. In: *2017 IEEE International Conference on Big Data (Big Data)*. IEEE, S. 1769–1777, 11.12.2017 - 14.12.2017.
- [SJ15] Sahu, Devendra K; Jawahar, CV: Unsupervised feature learning for optical character recognition. In: *2015 13th International Conference on Document Analysis and Recognition (ICDAR)*. IEEE, S. 1041–1045, 2015.

- [Te21] Tesseract: , <https://github.com/tesseract-ocr/tesseract>, 11.03.2021.
- [VG14] VGB PowerTech: , RDS-PP Anwendungsrichtlinie Teil 32: Windkraftwerke, 2014.
- [Wa12] Wang, Tao; Wu, David J; Coates, Adam; Ng, Andrew Y: End-to-end text recognition with convolutional neural networks. In: Proceedings of the 21st international conference on pattern recognition (ICPR2012). IEEE, S. 3304–3308, 2012.
- [Wu17] Wu, Ledell; Fisch, Adam; Chopra, Sumit; Adams, Keith; Bordes, Antoine; Weston, Jason: Starspace: Embed all the things! arXiv preprint arXiv:1709.03856, 2017.
- [Ye16] Yen, Ian En-Hsu; Huang, Xiangru; Ravikumar, Pradeep; Zhong, Kai; Dhillon, Inderjit: Pd-sparse: A primal and dual sparse approach to extreme multiclass and multilabel classification. In: International Conference on Machine Learning. S. 3069–3077, 2016.
- [ZLR20] Zhang, Zachariah; Liu, Jingshu; Razavian, Narges: BERT-XML: Large Scale Automated ICD Coding Using BERT Pretraining. arXiv preprint arXiv:2006.03685, 2020.

Modellierung eines Photovoltaik-Fehlererkennungsansatzes unter Berücksichtigung von maschinellem Lernen

Philipp Akharath¹, Jaqueline Altkrüger¹, Harkiran Sahota¹, Volker Herbort¹, Henrik te Heesen²

Abstract: Die Photovoltaik spielt eine zentrale Rolle bei der Transformation des globalen Energiesystems zu einer emissionsfreien Energieversorgung. In Deutschland tragen rund 1,8 Millionen installierte Photovoltaikdachanlagen mit einer Nennleistung bis zu 30 kWp zur elektrischen Energieerzeugung bei. Dennoch zeigt sich insbesondere in dieser Anlagenklasse, dass durch eine fehlende Fernüberwachung der Stromproduktion und ein fehlendes Qualitätssicherungskonzept technische Störungen auftreten können und es damit zu einer Minderung des Ertrags dieser Photovoltaikanlagen kommt [HHR18; HHR19; SH21]. Bislang wurden zur Erkennung von solchen Anomalien manuelle Verfahren verwendet; *Machine-Learning*-Konzepte können eine automatisierte und adaptive Alternative darstellen. Hierfür wird die Implementierung eines *Isolation Forests* mit dem Ansatz von Leloux [Le20] hinsichtlich des methodischen Aufbaus verglichen. Zur Bewertung der verschiedenen Ergebnisse werden exemplarische Ertragsdaten von einzelnen Anlagen analysiert. Der vorliegende Datensatz besteht aus den jeweiligen Erträgen der Anlagen in fünfminütigen Intervallen sowie den nötigen Stammdaten von Photovoltaikanlagen im Südwesten Deutschlands. Die Gegenüberstellung der Verfahren zeigt, dass die Anomalieerkennung durch *Isolation Forests* Betriebsstörungen von Photovoltaikanlagen automatisch identifizieren kann.

Keywords: Photovoltaik; Datenanalyse; Maschinelles Lernen; Fehlererkennung; Anomalie; Isolation Tree; Isolation Forest

1 Einleitung

Die weltweite Anzahl von Photovoltaik (PV)-Installationen ist in den letzten Jahren durch die globale Energiewende stetig gestiegen. Mit dem Ziel, den Klimawandel zu stoppen, leisten insbesondere PV-Dachanlagen einen großen Beitrag zur Minderung von CO₂-Emissionen in Deutschland, durch die Erzeugung von elektrischer Energie [HHR19]. In Deutschland wurden bis April 2021 knapp 1,8 Millionen Dachanlagen errichtet [Bu20]. Die durch Photovoltaikanlagen erzeugte elektrische Energie hat sich in den letzten zehn Jahren mehr als verdoppelt, wie in Abbildung 1 zu sehen ist. Im Jahr 2020 betrug die Energieerzeugung 50,7 TWh [Bu21].

¹ Technische Hochschule Ulm, Fakultät Informatik, Prittwitzstraße 10, 89075 Ulm, Deutschland
volker.herbort@thu.de

² Hochschule Trier, Umwelt-Campus Birkenfeld, Institut für Betriebs- und Technologiemanagement, Campusallee,
55768 Hoppstädten-Weiersbach, Deutschland
h.teheesen@umwelt-campus.de

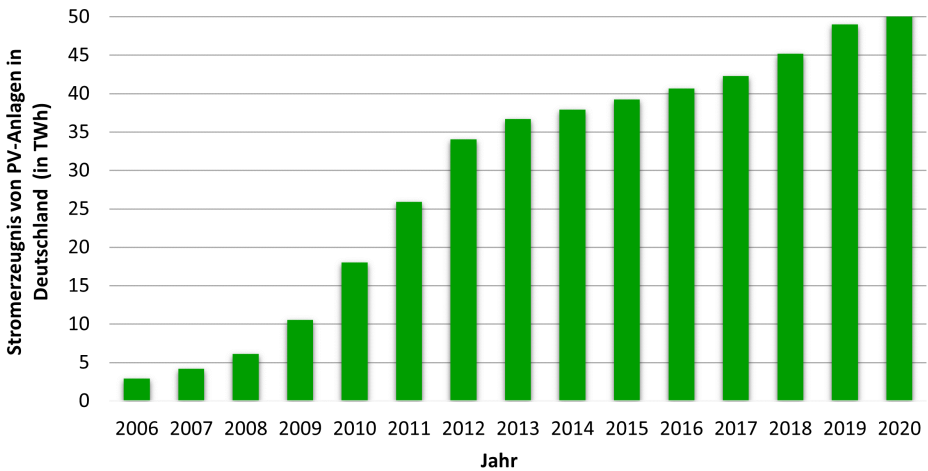


Abb. 1: Verlauf der produzierten Energie durch PV-Anlagen in Deutschland (Datenquelle [Bu21]).

Das Auftreten von technischen Störungen bei Photovoltaikanlagen kann jedoch die Energieproduktion und damit die Effizienz einer Anlage stark beeinträchtigen [Ga17; HH16; HHR18; SH21]. Dies kann durch verschiedene Arten von Fehlern verursacht werden. Einer der häufigsten Fehler ist eine technische Störung des Wechselrichters. Störungen bei den Photovoltaikmodulen treten seltener auf, können aber auch langfristig zu einer Ertragsminderung führen. Wird ein Fehler nicht rechtzeitig erkannt und behoben, beeinträchtigt dies die Produktivität und damit die Profitabilität einer PV-Anlage erheblich, weshalb ein verlässliches Fehlererkennungsverfahren unerlässlich ist [Ga17].

Solche Fehler lassen sich mithilfe von Anomalieerkennungsverfahren identifizieren, wobei abnorme Datenpunkte innerhalb eines Datensatzes ermittelt werden. Verfahren wie diese sind in unterschiedlichen Anwendungsdomänen einsetzbar, um seltene, aber signifikante Ereignisse zu erkennen [AMH16]. Ein automatisierter Ansatz zur Erkennung solcher Anomalien reduziert den händischen Aufwand zur Kalibrierung und ist adaptiv. Alternativ hierzu wurden bislang manuelle Verfahren zur Fehlererkennung eingesetzt, die mittels statistischer Methoden ein Profil regulärer Daten erstellen. Zur Ermittlung von Anomalien werden dabei Grenzwerte benötigt, welche von Domänenexperten festgelegt werden [St21].

Im Folgenden wird ein manueller Ansatz mit einer automatisierten Methode zur Anomalieerkennung im Kontext von Photovoltaikanlagen verglichen. Dabei soll die Applikabilität von *Machine-Learning*-Ansätzen zur Anomalieerkennung auf Basis exemplarischer Resultate nachgewiesen werden.

2 Grundlagen

2.1 State of the Art

Ein manuelles Anomalieerkennungsverfahren wird von Leloux [Le20] vorgestellt. In der Arbeit wird ein neuartiger Leistungsindikator - *Performance to Peer (P2P)* - zum Zweck der Fehlererkennung einer Solaranlage eingeführt, welcher allein auf dem Leistungsvergleich zwischen der betrachteten Anlage, im Folgenden als Fokusanlage bezeichnet, und den umliegenden Anlagen, im Folgenden als Nachbaranlagen bezeichnet, basiert. Dabei wird davon ausgegangen, dass Anlagen, die nah beieinander liegen (und damit vergleichbaren Einstrahlungsbedingungen ausgesetzt sind), und deren technische Parameter (Nennleistung, Ausrichtung, Neigung, etc.) ähnlich sind, vergleichbare spezifische Erträge produzieren.

Das P2P-Verfahren kann in zwei Phasen unterteilt werden. Zunächst erfordert die obige Annahme eine Einschätzung der Nachbaranlagen hinsichtlich ihrer Vergleichbarkeit mit der betrachteten Anlage. Hierbei beschränkt sich Leloux auf die umliegenden Nachbarn innerhalb eines Radius von 15 km. Der Ablauf der Bewertung der in Frage kommenden Nachbarn einer ausgewählten Fokusanlage ist in Abbildung 2 ersichtlich. Dabei wird die beschriebene Sequenz für jede der Nachbaranlagen ausgeführt. Um die Energieproduktion unterschiedlicher Anlagen effektiv vergleichen zu können, werden die Ertragswerte im Bezug auf die installierte Nennleistung normalisiert. Hierzu wird der Kapazitätsauslastungsfaktor (*Capacity Utilisation Factor CUF*)

$$CUF = \frac{E_{PV}}{P \cdot T} \quad (1)$$

aus der Energieproduktion E_{PV} über einen Zeitraum T und der Nennleistung P der Anlage berechnet. Zur Bestimmung des Ähnlichkeitsgrades einer Fokusanlage und einer ihrer Nachbaranlagen wird das sogenannte Kapazitätsauslastungsverhältnis (*Capacity Utilisation Ratio CUR*) zwischen zwei CUF -Werten zu einem bestimmten Zeitpunkt ermittelt als

$$CUR = \frac{CUF_{focus}}{CUF_{peer}} \quad (2)$$

Das Kapazitätsauslastungsverhältnis wird für jedes Fokus-Nachbar-Paar berechnet. Die Energieproduktion einer guten Nachbaranlage korreliert dabei stark mit der Fokusanlage, resultierend in einer geringen Schwankung des CUR -Wertes. Um diese Variabilität in den Werten untersuchen zu können ist die Auswahl einer geeigneten Zeitspanne wichtig. Diese sollte einerseits hinreichend lang sein, um statistisch repräsentativ zu sein und alle eventuellen Betriebszustände zu berücksichtigen, andererseits aber hinreichend kurz, um wesentliche Veränderungen in den Betriebszuständen wahrzunehmen. Leloux legt sich dabei auf einen effektiven Zeitraum von einem Monat fest. Als robustes Maß der Variabilität

in den berechneten *CUR*-Werten über den erwähnten Zeitraum eines Fokus-Nachbar-Paares wird schließlich die mittlere absolute Abweichung vom Median ($MAD(CUR)$) angewendet. Mittels

$$MAD(CUR) = med(|CUR_i - med(CUR)|) \tag{3}$$

lässt sich so die mittlere absolute Abweichung vom Median, *MAD*, berechnen, welche zur Ermittlung des Gewichts *w* durch

$$w = \frac{1}{MAD^4} \tag{4}$$

verwendet wird. Der normalisierte Wert λ

$$\lambda_i = \frac{w_i}{\sum w_i} \tag{5}$$

dient schließlich dazu, die einzelnen Nachbarn einer Anlage bei einem Vergleich zu gewichten. Diese Gewichtungen, welche exemplarisch über einen erwähnten Zeitraum berechnet werden, können anschließend zu jedem Zeitpunkt einer generellen Einschätzung der Nachbaranlagen dienen und sind nicht nur auf den angegebenen Zeitabschnitt zu beschränken.

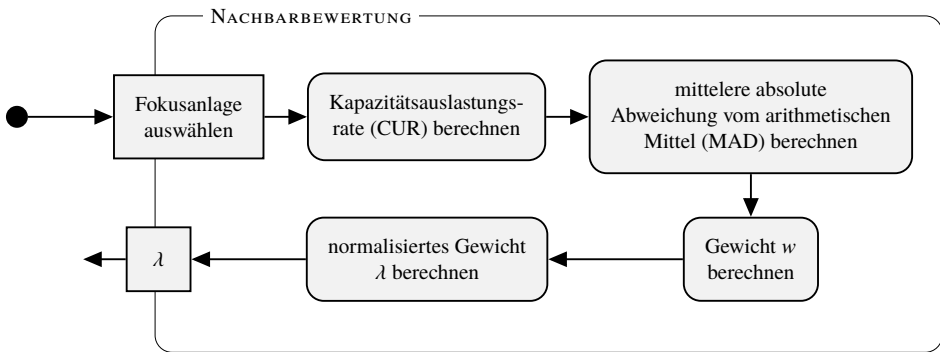


Abb. 2: Ablauf zur Bewertung der Nachbaranlagen gemäß Leloux [Le20].

Nachdem die Gewichtungen λ aller Nachbarn einer betrachteten Fokusanlage ermittelt wurden, kann in der zweiten Phase des Verfahrens nun der *P2P*-Wert zur Schätzung der Leistung eben dieser Anlage, durch den Vergleich mit den Nachbaranlagen zu einem beliebigen Zeitpunkt ermittelt werden. Über

$$P2P = \frac{CUF_{\text{focus}}}{CUF_{\text{ref}}} \quad (6)$$

lässt sich der $P2P$ -Wert berechnen, indem der CUF -Wert der Anlage zur fraglichen Zeit mit den CUF -Werten der umliegenden Anlagen verglichen wird. Hierfür wird ein Referenz- CUF als gewichtetes 50. Perzentil, Q_{50} , aller CUF Werte der Nachbaranlagen berechnet über

$$CUF_{\text{ref}} = Q_{50}(CUF, \lambda) \quad . \quad (7)$$

Ein $P2P$ -Wert von 1 indiziert, dass die Anlage im Vergleich zu den Nachbaranlagen einen vergleichbaren Ertrag erzielt. Aufgrund der Art, wie der $P2P$ -Wert berechnet wird, zeigt ein Wert größer eins, dass der Ertrag dieser Anlage den relativen Ertrag seiner Nachbarn übertrifft. Ebenso ist es möglich, dass der $P2P$ -Wert kleiner eins ist und der Ertrag damit kleiner als der Referenzertrag der Nachbaranlagen ist.

Da es sich hierbei um eine manuelle Methode zur Erkennung von Anomalien handelt, besteht der letzte Schritt darin, einen geeigneten Schwellwert für den berechneten $P2P$ -Wert festzulegen, um die Leistung der betrachteten Anlage zu kategorisieren. Fällt der Wert einer Anlage zu einem bestimmten Zeitpunkt unter den Schwellwert, so kann davon ausgegangen werden, dass ein Fehlverhalten vorliegt. Hierfür präsentiert Leloux ein Verfahren zur dynamischen Festlegung eines solchen Grenzwertes, welches im Weiteren jedoch keine Anwendung findet.

Dieser *State of the Art*-Ansatz wird im Folgenden als *Ground Truth* verwendet, um die Ergebnisse des entwickelten *Machine-Learning*-Ansatzes zur Anomalieerkennung zu verifizieren [Le20].

2.2 Isolation Forest

Der in dieser Ausarbeitung verfolgte automatisierte Ansatz zur Fehleridentifikation basiert auf der vorgeschlagenen Methodik des *Isolation Forest* nach Liu [LTZ08]. Der bislang verfolgte Prozess zur Anomalieerkennung bestand darin, intakte Datenpunkte zu erkennen und nicht übereinstimmende Fälle als Anomalien zu kennzeichnen. *Isolation Forests* hingegen isolieren Anomalien vorrangig durch iterative Partitionierung, anstatt zunächst normale Datenpunkte über Gruppierungen zu identifizieren [Ch20]. Dieser Ansatz unterscheidet sich von anderen Verfahren, wie dem *Local Outlier Factor*, durch die Tatsache, dass keine Berechnungen auf Basis der Distanz oder Dichte von Datenpunkten verwendet werden. Daher weist das Verfahren des *Isolation Forests* einen reduzierten Rechenaufwand und Speicherbedarf, sowie eine lineare Zeitkomplexität auf [JN19]. Somit lässt sich dieser

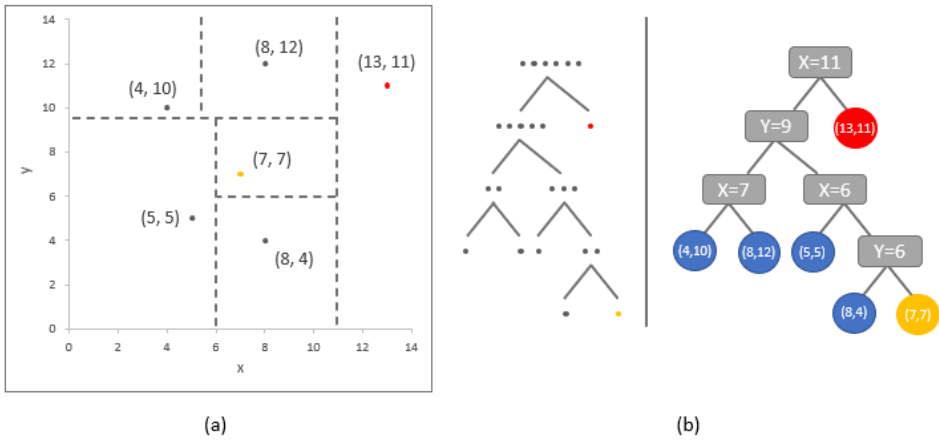


Abb. 3: Schematische Darstellung der Funktionsweise von *Isolation Forest* in der zweidimensionalen Ebene. Sechs Elementen werden jeweils ein Wertepaar (X, Y) zugewiesen (a). Über den binären Suchbaum (b) erfolgt eine Aufteilung der Instanzen und eine Isolation von Elementen (Datenquelle [Ch20])

Ansatz auch bei extrem großen Datenmengen effizient anwenden, liefert aber auch bei weniger Informationen zuverlässige Ergebnisse [LTZ08]. Dazu erfolgt zunächst die Erstellung zufälliger binärer Suchbäume, welche die Datenpunkte rekursiv partitionieren, bis alle Datenpunkte isoliert sind [Ch20].

Die Isolierung der Datenpunkte geschieht mittels eines zweistufigen Prozesses. Zunächst ist ein Training des Modells notwendig, bei dem die Baumstruktur gebildet wird. Bei der Erstellung eines Baumes wird ein zufälliges Merkmal der Daten selektiert und anschließend ein zufälliger Teilungswert zwischen dem minimalen und maximalen Wert dieses Attributs ausgewählt. Daraufhin werden die betrachteten Daten in zwei Gruppen entsprechend des Wertes dieses Attributs aufgeteilt. Dabei besteht die Annahme, dass Anomalien seltener sind als normale Daten und diese sich daher schneller vom restlichen Datensatz isolieren lassen. Dies zeigt sich durch eine kürzere Pfadlänge vom Wurzelknoten zum isolierten Datenpunkt. Dieser Vorgang wird in Abbildung 3 anhand eines zweidimensionalen Beispiels dargestellt. Die Verteilung der Datenpunkte (a) zeigt, dass die Instanz $(13,11)$ abseits der restlichen Punkte liegt und direkt isoliert werden kann. Der so entstandene binäre Suchbaum wird in (b) dargestellt. Zur Erstellung eines *Isolation Forests* werden mehrere solcher zufällig verteilten binären Suchbäume generiert und deren Ergebnisse zusammengefasst interpretiert.

Anschließend erfolgt die Evaluierung der Daten, wobei alle Testinstanzen an jeden *Isolation Tree* innerhalb des *Isolation Forests* gereicht werden. Dabei werden *Anomaly Scores* mittels der durchschnittlichen Pfadlänge der einzelnen Datenpunkte berechnet und zurückgegeben.

Die *Anomaly Scores* kennzeichnen das Maß der Anomalie, wobei Werte nahe 1 mit großer Sicherheit als Anomalie interpretiert werden. Werte, die kleiner als 0,5 sind, werden als regulär angesehen. Falls alle Ergebnisse eines Datensatzes *Anomaly Scores* um 0,5 aufzeigen, wurden keine Anomalien durch den *Isolation Forest* identifiziert [LTZ08].

Ein zentraler Vorteil bei der Verwendung von *Isolation Forests* ist, dass keine Anomalien im Trainingsdatensatz vorhanden sein müssen, um diese später auch zu identifizieren. Des Weiteren liefert dieses Verfahren schon bei vergleichsweise kleinen Datensätzen zuverlässige Ergebnisse. Da beide Aspekte für die vorliegenden Daten zutreffen, wird diese Methode im weiteren Verlauf angewandt [Ch20].

3 Methoden

Im Folgenden wird zunächst die praktische Umsetzung des manuellen Verfahrens zur Anomalieerkennung beschrieben, welche konzeptionell bereits in Abschnitt 2.1 vorgestellt wurde. Daraufhin folgt die Beschreibung der Implementierung des *Isolation Forests* basierend auf der Beschreibung in Abschnitt 2.2. Aufgrund der hohen Effektivität und Effizienz bei der Anomalieerkennung ist dies das Verfahren der Wahl in der vorliegenden Arbeit. Alternative Methoden wurden im Rahmen dieser Ausarbeitung nicht weiter beleuchtet. Die Umsetzung des *Isolation Forests* erfolgte in zwei Phasen. Aufgrund der nicht gekennzeichneten Daten wurden zuerst Zeitpunkte bestimmt, in denen Anomalien auftraten. Daraufhin fand eine Identifikation der fehlerhaften Anlagen bezogen auf die entsprechenden Zeiten statt. Diese Unterscheidung spiegelt sich in der Datenbeschaffung nicht wider, bei der Datenvorbereitung wurden aus diesem Grund zwei Verfahren unterschieden. Abschließend erfolgt die Interpretation der Ergebnisse durch den *Isolation Forest*.

3.1 Datenbeschaffung

Damit eine möglichst zuverlässige Aussage über Anomalien getroffen werden kann, sind vergleichbare Daten zu möglichst vielen PV-Anlagen in unmittelbarer Umgebung nötig. Zu diesem Zweck wurden frei verfügbare, webbasierte Ertrags- und Stammdaten des Fernüberwachungsanbieters Solare Datensysteme GmbH [So21] verwendet, um Informationen wie Postleitzahl, Nennleistung, Ausrichtung und Neigung zu beziehen [SH21]. Die Standortangaben verweisen dabei auf die Koordinaten der entsprechenden Postleitzahl und nicht dem tatsächlichen Standort der Anlage. Die benötigten Erträge in 5-Minutenintervallen wurden zusätzlich mit ein Scraping-Skript gesammelt.

Um die PV-Anlagen als Peers vergleichbar zu halten, müssen die technischen Eigenschaften (Ausrichtung, Neigung etc.) und der Standort der Peers hinreichend ähnlich sein, um zuverlässige Ergebnisse zu gewährleisten [Le20]. Aus diesem Grund wurden nur Daten von Aufdachanlagen berücksichtigt, die nach Süden ausgerichtet sind, eine Nennleistung

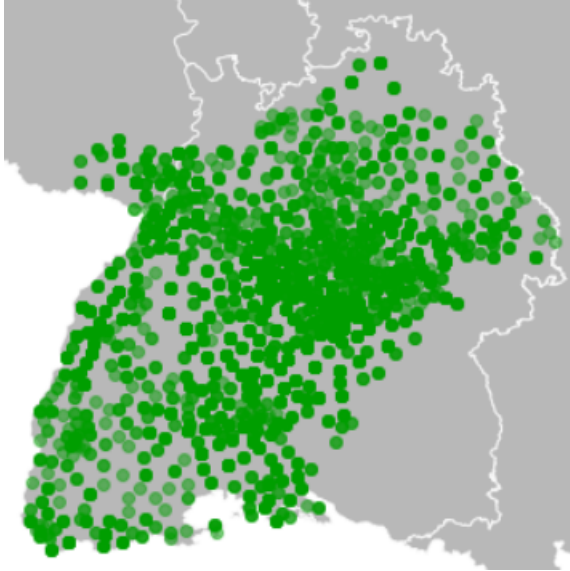


Abb. 4: Standorte der betrachteten Photovoltaikanlagen mit der Postleitzahl 7xxxx in den Bundesländern Baden-Württemberg und Rheinland-Pfalz.

bis 30 kW_p aufweisen und in der Postleitzahlregion beginnend mit der 7xxxx liegen (siehe Abbildung 4). Insgesamt besteht die Datenbasis aus 4 479 Photovoltaikanlagen, woraus die Daten für den exemplarischen Vergleich bezogen wurden, welche in Abschnitt 4 weiter beschrieben werden.

3.2 Datenaufbereitung

Um die Daten dem *Isolation Forest* zur Verfügung zu stellen und eine korrekte Anomalieerkennung zu gewährleisten, ist es notwendig, verschiedene Datenaufbereitungsschritte zu implementieren. Um einen Vergleich zwischen den Ergebnissen der manuellen Anomalieerkennung nach Leloux [Le20] zu ermöglichen, wurden die folgenden Schritte im Wesentlichen aus diesem Prozess nachgebildet. Demzufolge wurden die Nachbarn der betrachteten Anlage bestimmt, indem sämtliche PV-Anlagen identifiziert werden, die innerhalb eines 15 km-Radius liegen. Daraufhin wurden für jeden dieser Nachbarn die Daten im angegebenen Zeitintervall zusammengetragen. Dafür wurden sämtliche Datenpunkte für die untersuchten Anlagen auf ihre summierten Erträge von 5-Minutenintervallen auf eine Stunde aggregiert. Falls Anlagen nicht zwölf Datenpunkte innerhalb des Stundenintervalls aufwiesen, wurden diesen Anlagen in der Datenanalyse nicht weiter berücksichtigt. Anschließend werden die Stromerträge in Bezug auf die installierte Nennleistung normiert. Dieses Vorgehen ist in Abbildung 5 dargestellt.

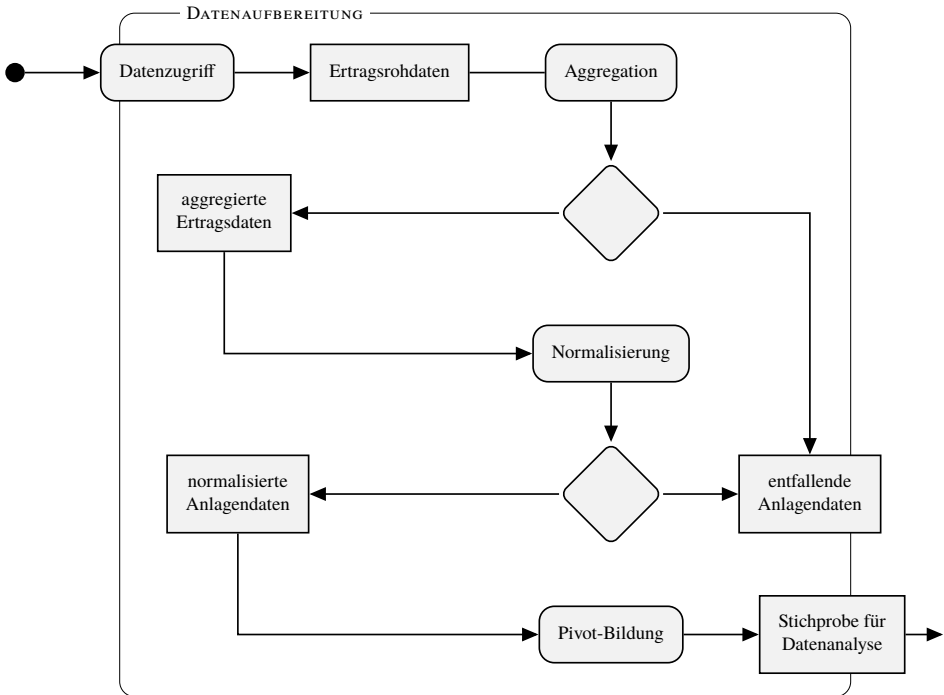


Abb. 5: Datenaufbereitung für die Datenanalyse. Aus den Rohertragsdaten werden Stundenenertragswerte aggregiert. PV-Anlagen mit Datenlücken werden aus der Datenauswertung ausgeschlossen. Die Erträge werden normalisiert und pivotiert.

Die Ertragsdaten auf stündlicher Basis werden nach der Datenbereinigung pivotiert, um die aggregierten Zeitstempel in Zeilen und die zugehörigen Anlagen in Spalten zu erhalten. Die Zellen werden anschließend mit den spezifischen Erträgen gefüllt. Dieser Datensatz diente als Basis zur Bestimmung des Zeitpunktes, zu welchem eine Anomalie auftritt.

Für den zweiten Schritt wird bestimmt, welche der betrachteten PV-Anlagen zu dem eben bestimmten Zeitpunkt Ertragsanomalien erzeugen. Hierfür werden die Nachbarn der betrachteten Anlage bestimmt und deren Daten mithilfe der Nennleistung normalisiert. Eine Aggregation nach der Zeit entfällt. Die Pivotierung der Daten zeigt die Anlagen als Zeilen und 5-Minutenintervalle für die betreffende Stunde als Spalten. Dieses Vorgehen ist in Abbildung 6 veranschaulicht.

Die räumliche Nähe zweier Peers ist ein entscheidender Faktor für ihre Ähnlichkeit. Da bereits kleinere regionale Wetteränderungen die spezifischen Erträge einer PV-Anlage beeinflussen können, muss der Standort eine übergeordnete Rolle spielen. Um sich auf lokal nahe gelegene Peers zu konzentrieren, schlägt Leloux vor, nur benachbarte Anlagen

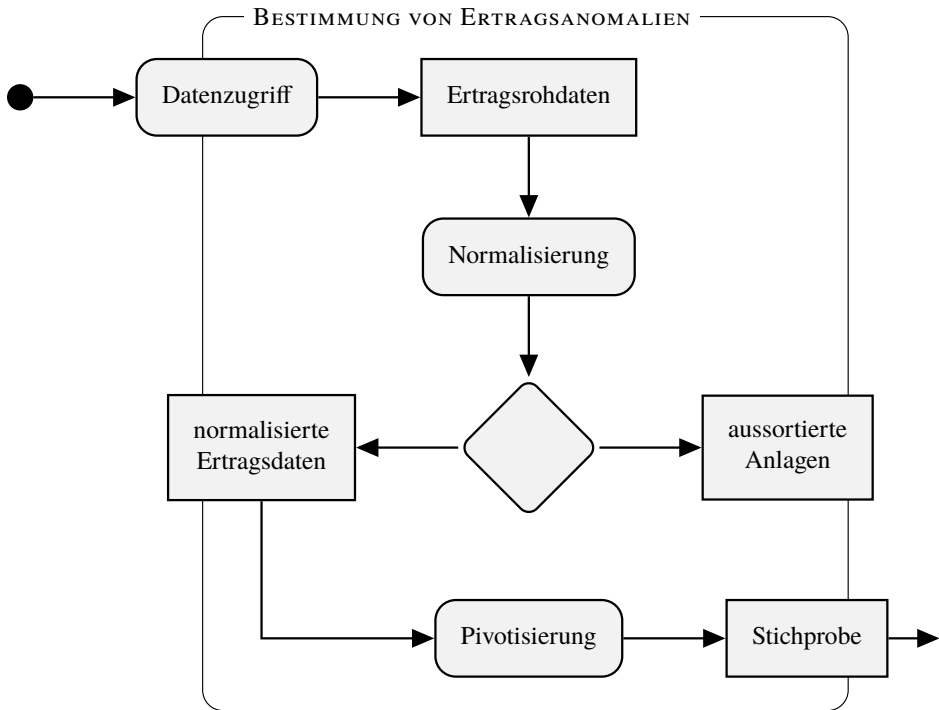


Abb. 6: Identifikation von Ertragsanomalien durch Bestimmung von Nachbaranlagen.

innerhalb eines Bereichs von 15 km zu berücksichtigen [Le20]. Um dies für den weiteren maschinellen Lernansatz umsetzen zu können, liefern die Datenvorbereitungsfunktionen die notwendigen Datenmatrizen für einen bestimmten gegebenen Fokus-Peer und filtert alle Anlagen, die nicht in diesem räumlichen Bereich liegen. Dies schränkt auch die Interpretation der Ergebnisse auf den oder die Fokus-Peer(s) ein, da für Anlagen, die am Rande dieses 15 km Radius liegen, Nachbarinformationen fehlen und die Ergebnisse nicht als vergleichbar angesehen werden können.

3.3 State-of-the-Art-Implementierung

Um einen *Ground Truth* für die folgende Analyse zu schaffen, wurde der in Abschnitt 2.1 beschriebene Ansatz von Leloux [Le20] implementiert. Hierbei soll es möglich sein, die Stromertragsproduktion einer Fokusanlage unter Einbeziehung ihrer Nachbarn zu bewerten. Dies bedarf zunächst einer Einschätzung aller Nachbarn in Bezug auf ihre Vergleichbarkeit zur Fokusanlage. Für die Berechnungen zur Gewichtung der Nachbaranlagen werden exemplarisch die Erzeugnisse der PV-Anlagen im Sommermonat August 2019 in Betracht

gezogen. Für jede Fokusanlage werden dabei ausschließlich die Nachbarn innerhalb eines 15 km-Radius beachtet. In der hier realisierten Variante wird zusätzlich die Restriktion eingeführt, dass lediglich Anlagen berücksichtigt werden, deren Standort in der Postleitzahlregion beginnend mit einer 7 liegt, um die Anzahl der PV-Anlagen zu reduzieren und die Machbarkeit des Vorgehens zu demonstrieren.

Für die somit infrage kommenden Nachbarn einer Fokusanlage werden für jeden der 31 Tagen 12 *CUF* zwischen 8 und 20 Uhr berechnet. Dazu werden die 5-Minutenwerte auf Stundenwerte aggregiert. Anlagen, welche nicht durchgehend Werte im festgelegten Zeitintervall liefern, werden verworfen. Anschließend können die stündlichen *CUR*-Werte für den betrachteten Monat berechnet werden, welche zur Ermittlung des einzelnen *MAD* Wertes pro Nachbar verwendet werden. Daraus werden schließlich die Gewichtung w und das normalisierte Gewicht λ berechnet.

Anschließend können in der zweiten Phase die *P2P*-Werte einer Fokusanlage berechnet und die Funktionsfähigkeit an einem bestimmten Tag eingeschätzt werden. Hierzu werden zunächst die stündlichen *CUF*-Werte dieser Anlage berechnet und für die Nachbaranlagen über die dazugehörigen λ -Werte die stündlichen Referenz *CUF*-Werte ermittelt. Somit können zwölf *P2P*-Werte ermittelt werden, um eine stündliche Einschätzung der Fokusanlage zu ermöglichen. Zur statischen Einschätzung der Anlagen muss ein Grenzwert definiert werden, welcher zur Kategorisierung der Anlagen dient. Dieser wird auf 0,85 festgelegt, um sämtliche Anlagen mit einem geringeren *P2P*-Wert als nicht funktionsfähig zu kategorisieren. Der Wert ergibt sich durch die Angabe von Leloux [Le20], dass eine funktionsfähige Anlage auch einen *P2P*-Wert von 90 Prozent annehmen kann. Ein zusätzlicher Interpretationsspielraum soll kleinere Abweichungen in den ersten Analysen zu Gunsten des *Isolation Forests* ausgleichen, weshalb dieser Grenzwert etwas verringert wird.

3.4 Dateninterpretation

An dieser Stelle werden die vorbereiteten Daten dem *Machine-Learnign*-Algorithmus übergeben. Dabei ist die Handhabung dieses Prozesses abhängig von den verwendeten Werkzeugen. Für den Einsatz des *Isolation-Forest*-Verfahrens wird die Python3-Bibliothek *Scikit-Learn* verwendet. Da die Anzahl der eingesetzten Bäume erheblichen Einfluss auf das Ergebnis hat, wurde der Wert 100 verwendet, welcher nach Liu [LTZ08] als optimal angesehen wird. Alle weiteren Parameter des Modells wurden entsprechend der Standardwerte übernommen.

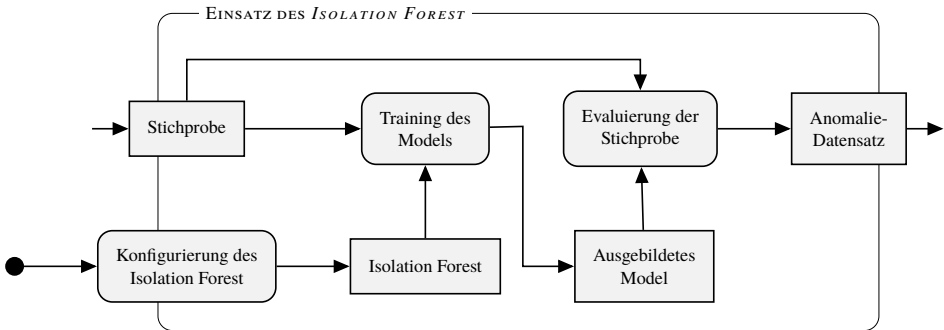


Abb. 7: Ermittlung von Anomalien durch den Einsatz des *Isolation Forest*.

In Abbildung 7 wird der *Isolation-Forest*-Prozess dargestellt. Dabei wird der Datensatz jeweils zum Training und zur Evaluation verwendet und anschließend den ermittelten Anomalien zugeordnet.

Für das Ermitteln einer Anomalie bezogen auf eine bestimmte Anlage und in Relation zu den dazugehörigen Nachbarn war ein iteratives Vorgehen mit den jeweiligen Datenvorbereitungen erforderlich. Aufgrund der fehlenden Kennzeichnung der vorliegenden Daten hinsichtlich ihrer Funktionstüchtigkeit wurden zunächst Tage ermittelt, welche eine Anomalie aufwiesen. Hierfür wurde die Datenvorbereitung nach Zeitpunkten verwendet. Der daraus resultierende anomalische Datensatz wurde herangezogen, um eine weitere Eingrenzung zu treffen. Dabei wurde der zu untersuchende Tag extrahiert, um daraufhin die Datenvorbereitung nach Anlagen durchzuführen. Nach Identifikation der anomalischen Anlagen mithilfe des *Isolation Forests* konnte der Datensatz der Anomalien zu den jeweiligen Anlagen zugeordnet werden. Das Format dieses Resultats ermöglicht einen Vergleich mit den jeweiligen *P2P*-Werten bezogen auf eine Anlage.

4 Ergebnisse und Diskussion

Ersichtlich ist, dass sowohl das Verfahren von Leloux [Le20] als auch die hier dargestellte Methode mittels eines *Isolation Forests* vielversprechende Ergebnisse liefern kann, um die Funktionsfähigkeit von Photovoltaikanlagen zu bewerten. Die Ergebnisse der exemplarischen Datenanalyse sind in Tabelle 1 dargestellt. Hierbei werden einzelne Anlagen an einem bestimmten Tag und zu einer bestimmten Stunde analysiert und die Ergebnisse der beiden Verfahren gegenübergestellt. Das Verfahren von Leloux [Le20] liefert den in der Studie entwickelten *P2P*-Wert. Dieser kennzeichnet die Leistung einer PV-Anlage im Vergleich zu den umliegenden Nachbarn. Werte, die dabei unter dem in Abschnitt 3.3 festgelegten

Tab. 1: Gegenüberstellung exemplarischer Resultate aus der Datenanalyse. Der *P2P*-Wert wurde nach Leloux [Le20] bestimmt. Die Klassifizierung nach dem *Isolation Forest* beträgt 1 für eine fehlerfreie Photovoltaikanlage und -1 für eine Anlage mit einer technischen Störung.

Anlagen-ID	Datum	Zeitspanne	<i>P2P</i>	<i>Isolation Forest</i>	
				Klassifizierung	<i>Anomaly Score</i>
12246	10.08.2019	14:00-15:00	1,033	1	0,372
15863	01.09.2020	11:00-12:00	1,311	-1	0,580
24874	10.08.2019	14:00-15:00	0,126	-1	0,697
25731	11.06.2020	09:00-10:00	0,860	1	0,366
25731	20.06.2020	09:00-10:00	0,857	1	0,386

Grenzwert von 0,85 liegen, kennzeichnen eine mangelhafte Anlage. Auch eine starke Überperformance bei Werten über 1,15 kann ein Kennzeichen für ein Fehlverhalten sein. Der *Isolation Forest* liefert eine automatisierte Identifizierung der Anomalien; eine -1 für abnorme und eine 1 für reguläre Werte. Diese Einschätzung geschieht anhand der *Anomaly Scores*, welche bei einem Wert nahe 1 eine Diskrepanz in den Daten feststellen.

Durch die Übereinstimmung der Ergebnisse in Tabelle 1 des *Isolation Forests* mit den eingangs als *Ground Truth* angenommenen *P2P*-Werten kann ersteres als plausible Alternative zu manuellen Verfahren angenommen werden. Die Zusammenhänge der verschiedenen Werte können bei näherer Betrachtung einzelner Zeilen erkannt werden. Die Anlage mit der ID 25731 hat beispielsweise am 11.06.2020 einen *P2P*-Wert von 0,860; somit lässt sie sich mit dem dazugehörigen statischen Grenzwert von 0,85 als funktionsfähig einstufen. Die respektiven Resultate des *Isolation Forest* stufen diese Anlage ebenso als funktionstüchtig ein, ersichtlich in der Klassifizierung von 1 und dem dazugehörigen *Anomaly Score* von 0,366. Dabei kennzeichnet der *Anomaly Score*, dass das Modell des *Isolation Forests* diese Stichprobe nur mit einer Wahrscheinlichkeit von 36 Prozent als Anomalie klassifizieren konnte. Anhand dieser exemplarischen Resultate können erste Annahmen getroffen werden, dass ein solcher *Machine-Learning*-Ansatz als funktionsfähig betrachtet werden kann.

Im Laufe der Implementierung haben sich besondere Charakteristiken des *Isolation Forests* gezeigt, welche für die Bestimmung von Anomalien bei den vorliegenden Daten zu Photovoltaikanlagen besonders von Nutzen sind. Wesentlich ist eine umfassende Datenaufbereitung der Rohdaten, um PV-Anlagen mit offensichtliche Fehlkonfigurationen, die zu Artefakten bei der Datenanalyse führen können, auszuschließen. Für den Vergleich von Anlagen zu deren Nachbarn ist des weiteren die Tatsache relevant, dass *Isolation Forests* gut auf kleineren Datensets anwendbar sind. Dies ist vor allem für Regionen mit einer geringeren Anlagendichte relevant.

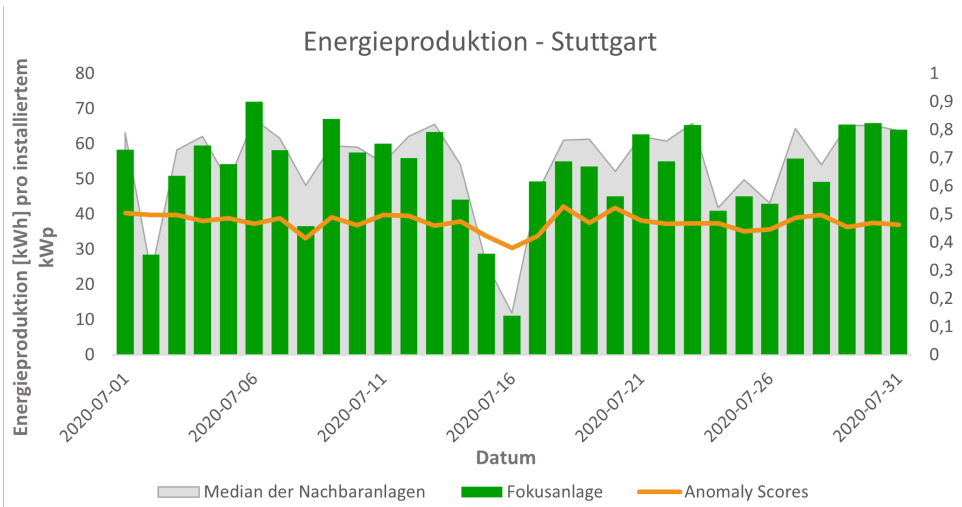


Abb. 8: Spezifischer Ertrag (grüne Balken) und *Anomaly Score* (orange Linie) der Photovoltaikanlage mit der ID 12246 (Standort Stuttgart) im Juli 2020.

Da die beiden Anlagen in Ostfildern (Anlagen-ID 24874) und Stuttgart (Anlagen-ID 12246) mit einer Distanz von circa 9 km als direkte Nachbarn angesehen werden, folgt eine detaillierte Interpretation der Ergebnisse dieser beiden Installationen.

Die Anlage in Stuttgart wird vom *Isolation Forest* als regulär eingestuft. Zur Validierung dieser Einschätzung werden die entsprechenden Daten manuell analysiert. In Abbildung 8 sind die aggregierten Messwerte der einzelnen Tage vom Juli 2020 als grüne Balken dargestellt. Dabei bildet die x-Achse die Zeit in Tagen ab. Die linke y-Achse zeigt die Energieproduktion in kWh, die rechte y-Achse veranschaulicht den *Anomaly Score*. Die graue Fläche im Hintergrund zeigt die summierten Erträge der Nachbaranlagen im Verhältnis zu deren Nennleistung. Der *Anomaly Score* der einzelnen Tage wird durch den Verlauf der orangenen Linie gekennzeichnet.

Auffällig ist, dass sowohl der Ertrag der Fokusanlage als auch die der Nachbaranlagen deutlich schwanken. Dies lässt sich durch die lokalen Wetter- und Einstrahlungsbedingungen erklären [De21]. Besonders die Tage am 02.07.2020, 15.07.2020 und 16.07.2020 sind hiervon betroffen. Ersichtlich ist dies auch in Abbildung 8, da sowohl die Fokusanlage als auch die Nachbaranlagen an diesen Tagen besonders geringe Erträge aufweisen. Durch den *Isolation Forest* werden in dem gesamten Monat keinerlei Anomalien verzeichnet, was durch durchgängige Werte um 0,5 ersichtlich ist [LTZ08].

Die Anlage in Ostfildern hingegen weist für alle Tage im Juli 2020 Anomalien auf (Abbildung 9). Die Abbildung ist ebenso zu interpretieren wie Abbildung 8. Hier ist vor

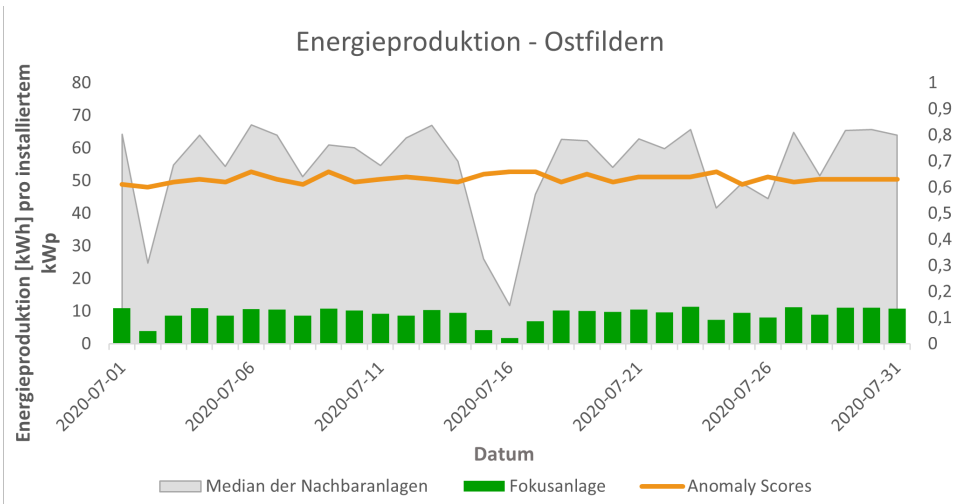


Abb. 9: Spezifischer Ertrag (grüne Balken) und *Anomaly Score* (orange Linie) der Photovoltaikanlage mit der ID 24874 (Standort Ostfildern) im Juli 2020.

allem auffällig, dass die Erträge der Fokusanlage im Verhältnis zu deren Nennleistung deutlich geringer sind als die normalisierten Erträge der Nachbarn. Grund hierfür kann entweder eine falsche Angabe der Nennleistung oder ein technischer Ausfall eines oder mehrerer Wechselrichter sein.

5 Zusammenfassung

Eine automatisierte Identifikation von technischen Störungen bei Photovoltaikanlagen, insbesondere von Dachanlagen mit einer Nennleistung bis 30 kW_p , dient einer schnellen Fehlererkennung, um eine Beseitigung von technischen Störungen einzuleiten. Hierzu wird die Tauglichkeit automatisierter Anomalieidentifikationsverfahren zur effektiven Einschätzung der Ertragsqualität einzelner Photovoltaikanlagen geprüft.

Zusammenfassend lässt sich festhalten, dass ein *Isolation Forest* als ein solches automatisiertes Verfahren mittels maschinellem Lernen anwendbar ist und eine gute Alternative zu den herkömmlichen, manuellen Verfahren (wie zum Beispiel das *P2P*-Verfahren nach Leloux [Le20]) bietet. Die Anwendung des *Isolation Forests* kommt im Gegensatz zum *P2P*-Ansatz ohne vorherige Gewichtung von Nachbaranlagen und Parametrisierung von Grenzen zur Fehlererkennung aus. Anhand eines exemplarischen Datensatzes aus den 4 479 PV-Anlagen im Südwesten Deutschlands konnte gezeigt werden, dass technische Störungen und damit verbundene Stromproduktionsausfälle durch den *Isolation Forest* erkannt werden können.

Eine ausführliche Analyse des Ansatzes anhand einer quantitativen Evaluierung ist Umfang der zukünftigen Forschungsarbeit. Auch die mögliche Parametrisierung des *Isolation Forests* zur verbesserten Anomalieerkennung ist Gegenstand weiterer Ausarbeitungen. Die Klassifizierung der identifizierten Anomalien nach verschiedenen Fehlerarten bietet ebenfalls Grundlage zur Anwendung weiterer *Machine-Learning*-Ansätze.

Zukünftig könnte die Umsetzung von Anomalieerkennung mithilfe von automatisierten Verfahren innerhalb privater oder gewerblich genutzter Photovoltaikanlagen eine effektive Möglichkeit für Anwender bieten, Ausfälle zu erkennen und diesen gegenwirken zu können. Eine zentrale Implementierung des präsentierten Verfahrens ermöglicht den Zugriff auf die notwendigen Daten benachbarter Anlagen und das Bereitstellen eines Services zum Abfragen der Funktionsfähigkeit einzelner Anlagen. Durch die Realisierung einer Fehlerfrüherkennung könnten so die Gesamtstörungszeiten von Photovoltaikanlagen reduziert werden. Infolgedessen ließe sich der Stromertrag von Photovoltaikanlagen verbessern und somit ein nachhaltiger Beitrag zur Energiewende erreichen.

Literatur

- [AMH16] Ahmed, M.; Mahmood, A. N.; Hu, J.: A survey of network anomaly detection techniques. *Journal of Network and Computer Applications* 60/, S. 19–31, 2016.
- [Bu20] Bundesnetzagentur: Marktstammdatenregister, 2020.
- [Bu21] Burger, B.: Burger, Bruno: Nettostromerzeugnis in Deutschland im Jahr 2020, 2021.
- [Ch20] Chen, H.; Ma, H.; Chu, X.; Xue, D.: Anomaly detection and critical attributes identification for products with multiple operating conditions based on isolation forest. *Advanced Engineering Informatics* 46/, S. 101139, 2020.
- [De21] Deutscher Wetterdienst: Deutscher Wetterdienst: Climate Data Center, 2021.
- [Ga17] Garoudja, E.; Harrou, F.; Sun, Y.; Kara, K.; Chouder, A.; Silvestre, S.: Statistical fault detection in photovoltaic systems. *Solar Energy* 150/, S. 485–499, 2017.
- [HH16] te Heesen, H.; Herbort, V.: Development of an algorithm to analyze the yield of photovoltaic systems. *Renewable Energy* 87/, 2016.
- [HHR18] te Heesen, H.; Herbort, V.; Rumpler, M.: Untersuchung des Ertrags von Photovoltaikdachanlagen bis 30 kWp in Deutschland im Zeitraum 2014 bis 2017. In: *Workshops der INFORMATIK 2018-Architekturen, Prozesse, Sicherheit und Nachhaltigkeit*. Köllen Druck+ Verlag GmbH, 2018.
- [HHR19] te Heesen, H.; Herbort, V.; Rumpler, M.: Performance of roof-top PV systems in Germany from 2012 to 2018. *Solar Energy* 194/June 2021, S. 128–135, Dez. 2019.

- [JN19] John, H.; Naaz, S.: Credit card fraud detection using local outlier factor and isolation forest. *Int. J. Comput. Sci. Eng.* 7/4, S. 1060–1064, 2019.
- [Le20] Leloux, J.; Narvarte, L.; Desportes, A.; Trebosc, D.: Performance to Peers (P2P): A benchmark approach to fault detections applied to photovoltaic system fleets. *Solar Energy* 202/, S. 522–539, 2020.
- [LTZ08] Liu, F. T.; Ting, K. M.; Zhou, Z.-H.: Isolation forest. In: 2008 eighth iee international conference on data mining. *IEEE*, S. 413–422, 2008.
- [SH21] Schardt, J.; te Heesen, H.: Performance of roof-top PV systems in Germany from 2012 to 2019. *Solar Energy* 217/June 2021, S. 235–244, 2021.
- [So21] Solar Datensysteme GmbH: Solar Datensysteme GmbH: Energy Management System Solar-Log, 2021.
- [St21] Steenwinckel, B.; De Paepe, D.; Hautte, S. V.; Heyvaert, P.; Bentefrit, M.; Moens, P.; Dimou, A.; Van Den Bossche, B.; De Turck, F.; Van Hoecke, S. et al.: FLAGS: A methodology for adaptive anomaly detection and root cause analysis on sensor data streams by fusing expert knowledge with machine learning. *Future Generation Computer Systems* 116/, S. 30–48, 2021.

Spatial Interpolation of Air Quality Data with Multidimensional Gaussian Processes

Paul Tremper¹, Till Riedel¹, Matthias Budde^{1,2}

Abstract: The central question of this paper is whether interpolation techniques applied to a distributed sensor network can indeed provide more information than using the constant background of an urban reference station to measure air pollution. We compare different interpolation techniques based on temporal-spatial machine learning in terms of their applicability for correctly predicting personal exposure. Using a dataset of stationary low-cost sensors, we estimate exposure on a route through the city and compare it to mobile measurements. The results show that while different machine learning-based interpolation methods yield quite different results, validation of machine learning-based approaches is still challenging.

Keywords: Air Quality; Gaussian Process Regression; Validation

1 Introduction

New fine-grained measurement networks promise to better quantify individual exposure to pollutants such as particulate matter. While official measurements look at large spatial and temporal integrals (daily averages for entire cities), the urban foreground can vary, and individual risk can vary widely based on daily routines. Since it seems impractical for individuals to carry dosimeters and simulations rely on typically incomplete information, low-cost sensor networks are becoming increasingly popular and have even been investigated by government agencies.

However, even with cheap sensors, there are limits to the density of such networks, and even dense networks have limited readings, typically spaced a few hundred meters apart in space and minutes apart in time. Therefore, measurement information must be interpolated, which means that even measurement systems always rely on data-driven predictions.

In this paper, we discuss how to evaluate basic temporal-spatial prediction methods against real data to improve the quality of machine learning (ML) algorithms in this domain and better understand relevant metrics.

¹ Karlsruhe Institute of Technology (KIT), TECO / Pervasive Computing Systems, Vincenz-Prießnitz-Straße 1, 76131 Karlsruhe, Deutschland tremper@teco.edu

² Disy Informationssysteme GmbH, Ludwig-Erhard-Allee 6, 76131 Karlsruhe, Deutschland matthias.budde@disy.net

To do this, we follow a simple methodology. The basis for the analysis is a large-scale open dataset collected in Augsburg over a three-year period. After calibrating a set of low-cost sensor readings against a monthly background station mean, we validate different interpolations of low-cost sensor readings with calibrated reference stations. We show how the different algorithms perform based on typical regression metrics. Our analysis shows that the difference in prediction is not covered by the metrics.

To represent a realistic scenario in exposure quantification, we then predict the expected exposure on a bicycle route through the city and compare it to actual mobile in situ measurements. While the bike route measurements are difficult to calibrate and are not accurate, we hypothesize that a better interpolation methodology will nonetheless reduce the mean square error against this imperfect ground truth and serve as a better metric for machine learning algorithms that reconstruct the structure of propagating emissions.

The paper is structured as follows: In section 2, we present some related work before describing the dataset and our calibration procedure in section 3. In section 4, we then describe our first analysis, in which we use data of the low cost network to predict ground truth values. In section 5, we describe our second analysis, in which we attempt to predict a scenario of personal exposure along a path through the city. The latter two sections are split into data, methods and results. Finally, we conclude our paper in section 6 with an overall discussion followed by a conclusion.

2 Related Work

The basic assumption underlying any spatial interpolation is succinctly expressed in *Tobler's First Law of Geography* [To70]: “*Everything is related to everything else, but near things are more related than distant things*”. The most simple approach that incorporates this idea is *Inverse Distance Weighting* (IDW) [Sh68].

Kriging [Kr51] is another popular approach in geoinformatics. Its mathematical basis are Gaussian Processes (GPs) which have recently received much attention as a general framework to deal with predictions under uncertainty and their close relation to unbounded neural networks. Kilibarda et al. e.g. introduced an automated mapping framework for predictions of daily air temperatures (mean, min and max) using regression-kriging for a resolution of 1km [Ki14]. Pebesma et al. [PH16] use copulas to enable spatio-temporal kriging: They show the application and benefit of their approach with a prediction of daily mean PM₁₀ concentration. In [Br18], the residual error of Gaussian Process has been reduced by coupling them with a calibration scheme for low cost sensors within a black box optimization approach.

Landuse regression (LUR) is another popular approach for modeling spatial variation in various domains, among them also air quality [Ho08]. Such a modelling approach has been successfully applied to the same dataset that was used in this study [Sh20]. LUR predicts

spatial variations in air pollution based on correlations with the measurement context instead of the geospatial one. LUR modeling requires air pollution measurements at multiple locations across the study area, stationary monitoring used by LUR is typically at 20 to 100 locations, spread over the study areas. To summarize the predictor variables used in the LUR models, frequently used data include: area of land-use, road network or traffic information, physical geography such as elevation and slope, and meteorological data. In this sense LUR has the same issues as simulation as it actually relies heavily on the availability of external information. In addition, machine-learning approaches such as ensemble regression methods have been utilized to handle complex and nonlinear relationships that exist within data and produce forecasting models with comparable performance in practice. Based on our review of papers [Yu16, ZLH13, Li17, Su16] from the domain of spatial data analysis, since the prediction accuracy follows algorithm design, the machine-learning algorithms are crucial for building air quality prediction models, whereas statistical models have not been heavily used recently. Moreover, the random forest based approach is a prominent technique in selecting variables and inferring air pollution values.

3 Dataset

3.1 Description

We selected our data from the SmartAQnet³ [Bu17] dataset. We focus on PM₁₀ (Particulate Matter of diameter < 10 $\mu\text{g}/\text{m}^3$ as pollutant and chose the month of September because of data availability. The SDS011 Crowdsensor sub-dataset consists of 1,479,343 data points across 35 devices, the EDM80 Scientific Scout sub-dataset consists of 253,220 data points across 35 devices. As references, we chose the official measurement stations within Augsburg, which give 2,756 data points across 4 stations.

SDS011 Crowdsensors are $\sim 30\text{€}$ sensor nodes, consisting of a Nova Fitness SDS011 fine dust sensor, an esp32 or esp8266 microcontroller and a BME280 or DHT22 humidity/temperature/pressure sensor, depending on the build⁴. In the context of this analysis, these differences in build are considered equivalent. These sensor nodes are considered in the ultra low cost range and well suited for use in large numbers and thus in various builds popular in citizen science contexts. While they have weaknesses in measuring precision, they perform reasonably well in measuring relative changes in air quality, which makes them an interesting component in large sensor networks.

EDM80 Scientific Scouts are considered low to mid cost sensors of $\mathcal{O}(1000\text{€})$, and come in two variants which we consider equivalent in the context of this analysis. The older

³ <https://www.smartaq.net>

⁴ A construction manual can be found, e.g., here: <https://www.smartaq.net/en/participate/>

variant (EDM80 NEPH) contains a nephelometer measuring cell, while the newer variant (EDM80 OPC) contains an Alphasense OPC (optical particle counter) measuring cell as fine dust sensor component. They have been developed, maintained and calibrated in project SmartAQnet by GRIMM⁵, a company that produces and maintains air quality sensors. With regard to the SDS011 Crowdsensors, the scientific scouts are considered as a higher tier in quality.

Reference Stations are the official measuring stations of the Bayerisches Landesamt für Umwelt (Bavarian state agency for environment). There are four such stations located in Augsburg⁶ which collect hourly values of various air quality parameters, including PM₁₀. In the context of this analysis, these stations are considered the ground truth.

3.2 Calibration Procedure

Off-the-shelf, low-cost light-scattering PM sensors often exhibit systematic deviation from ambient particle mass concentrations [Bu18] and therefore need to be calibrated.

First we took the monthly median of the official urban background station at Augsburg Bourges-Platz ($15.0 \mu\text{g}/\text{m}^3$) as our baseline to calibrate the sensors on. This choice is strictly speaking only valid in case of sensors which are also located in urban background environments. In the case of traffic environments, this will certainly under-calibrate the relevant sensors. Since we need to establish some kind of calibration on the sensors, we still proceed in this way and compare our interpolation results with the non-calibrated predictions later on to get information about the validity of the calibration. We note that for this reason, we do expect the calibration to be not 100% and thus leave an offset in the predictions. This method could be improved by taking into account the respective environment of each sensor and calibrating it against a reference station which is classified as the same environment.

Second we computed the monthly median for each single sensor (Crowdsensors and Scouts) and used the difference to the aforementioned baseline to shift the individual measurements. The effects of the calibration can be seen in subsection 4.3. For a discussion on why we used the median instead of the arithmetic mean, please see the discussion section.

⁵ <https://grimm-aerosol.com>

⁶ for a description of the official stations, see <https://www.lfu.bayern.de/luft/immissionsmessungen/dokumentation/index.htm> (in German)

4 Validation of Sensor Network Interpolation

4.1 Data

To validate the sensor network interpolation, we predicted the values of the four reference stations. Since the reference stations only give hourly values, we first had to aggregate the data to hourly values to be able to make predictions. We used the median instead of the arithmetic mean for the aggregation (again, see the discussion section). We then took data from the 21st and 22nd of September (48 hours) since the bike route, which we want to predict in the next section, falls into this interval. In that time, we have 31-33 active SDS011 Crowdsensors with a total of 1404 hourly values after aggregation in an area of 4.5 km x 7.9 km [lat x lon] (36 km²) and 6-27 active EDM80 OPC Scientific Scouts with a total of 839 hourly values in an area of 10.2 km x 9.2 km [lat x lon] (94 km²).

4.2 Methods

For the validation of the sensor network interpolation, we used two methods:

- One is the naive inverse distance weighted interpolation (IDW), which gives a prediction value v_p for each point by $v_p = N \sum_i \frac{v_i}{d_i^2}$ with the normalization constant N given by the condition $N \sum_i \frac{1}{d_i^2} = 1$. The d_i and v_i are the distances to each sensor and their values.
- The other method is Gaussian Process Regression (GP), which uses multidimensional Gaussian distributions to interpolate. We used an RBF Kernel (Radial Basis Function, $K(x, x') = \sigma^2 \exp(-(x - x')^2 / 2\ell^2)$) for the Gaussian Process Regression in every dimension (two spatial, one temporal). This kernel has one relevant parameter, the length scale ℓ , which is gauges the scale on which correlations between points are measured.

In case of the IDW, which is oblivious of the time coordinate, we performed an independent spatial interpolation for each set of hourly values. The GP, however, was trained on the full data of the two days. For that time frame, the GP finds RBF length scales of [0.00582, 0.00309, 1.38] (SDS011 Crowdsensor dataset) and [0.00785, 0.0149, 1.06] (Scientific Scouts dataset) in units of [lat,lon,hours]. To have better comparability, we adjust the length scales slightly and fix them at [0.00450,0.00680,1], which corresponds to ~500m in lat and lon dimensions, as well as one hour in the time dimension.

While we coded the inverse distance weighted interpolation ourselves (in python) using the formula above, we used the scikit-learn [Pe11] implementation of the Gaussian Process Regression.

4.3 Results

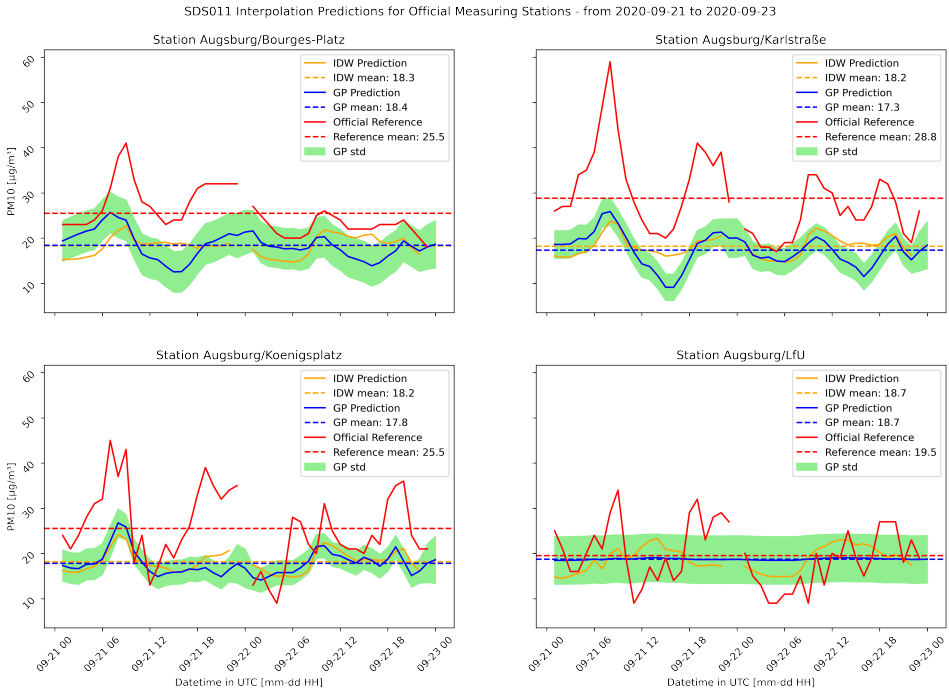


Fig. 1: SDS011 Crowdsensor dataset PM₁₀ predictions for each of the four reference stations. The inverse distance weighted interpolation is in orange, the (2+1)D Gaussian Process in blue, the measurements of the official reference station in red, the standard deviation of the Gaussian Process Regression in green.

The results of the prediction of the four reference stations are shown in Figure 1 (Figure 2) for the Crowdsensors (Scouts). We can see that in both cases, Crowdsensors and Scouts, the interpolations were able to capture the changes very well in case of Bourges-Platz and Karlstraße. The varying offsets of the prediction graphs to the ground truth graphs are systematic errors we will discuss later.

In case of the Crowdsensors (Figure 1), this is encouraging since the closest sensors to these stations are located 589m (Bourges-Platz) and 490m (Karlstraße) away. The length scale of the GP was set into that range to be sensitive to correlations of that length scale ($\ell = 500\text{m}$, see above). We thus interpret the similar shapes of the prediction and the ground truth graphs as a sign that:

- a) on the hourly level there are relevant correlations at the $O(500\text{m})$ scale in the data.

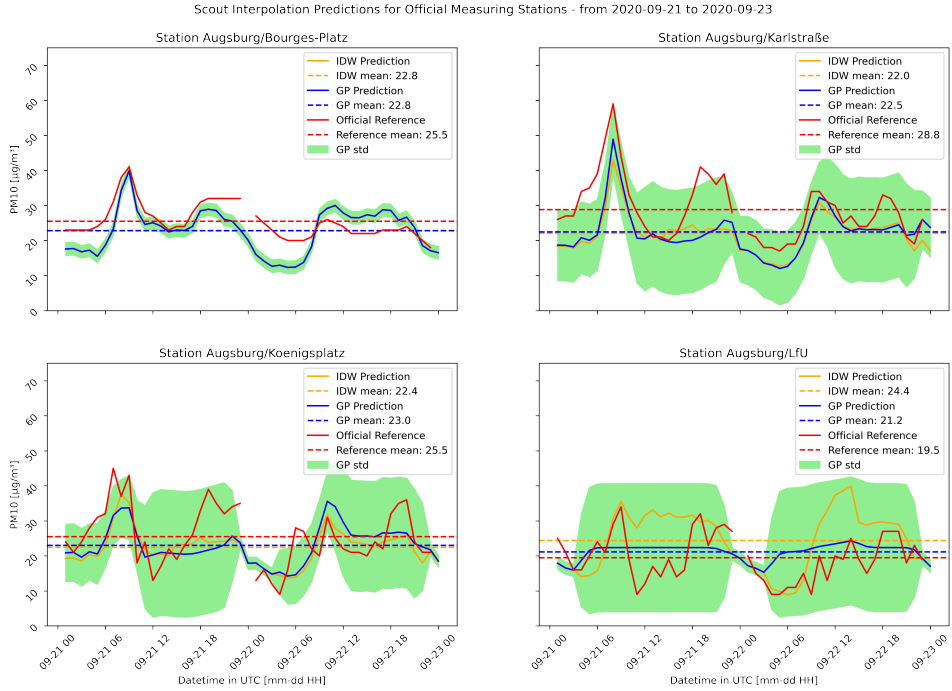


Fig. 2: Scientific Scout dataset PM_{10} predictions for each of the four reference stations. The inverse distance weighted interpolation is in orange, the (2+1)D Gaussian Process in blue, the measurements of the official reference station in red, the standard deviation of the Gaussian Process Regression in green.

- b) a network of low cost sensors, like the SDS011 Sensors at hand, can produce useful predictions, if supported by interpolation methods. That is, if it is meshed tightly enough to be sensitive to that length scale.
- c) The GP, as a representative of an AI based method that makes use of the temporal information, performs a lot better in tracing the graph of the ground truth than the naive IDW.

The Scientific Scouts are harder to interpret here. There is a Scientific Scout located closely to each of the reference stations ($<10\text{m}$). The results in Figure 2 show, however, that at least those at Karlstraße, Königsplatz and LfU seem not to produce any data. This is most notable at the size of the standard deviation band of the GP (in green), which represents a measure of confidence of the method in its own prediction. While the result at Bourges-Platz seems plausible for a sensor located rather close to the actual station (the IDW lying on top of it seems to back that), the predictions at the other three stations exhibit such large uncertainties, that it seems unlikely that the prediction is supported by a sensor very close.

Nevertheless, the Scouts were able to capture the shape of the ground truth very well in some parts for Karlstraße and Königsplatz.

For the LfU station, the behaviour of the GP for both, Crowdsensors and Scouts, shows that there is no sensor close enough to make a reliable prediction. This can be seen on the one hand by the large band of uncertainty, but most notably by the GP prediction dropping to the prior most of the time. And indeed, the closest Crowdsensor to the LfU station is located 1.163m away, which is more than twice the length scale of the GP. For the Scouts, the closest sensor is within 10 meters of the station, but the results suggests, that this Scout did not produce any measurements for most of the relevant timeframe.

The aforementioned offsets of most of the predictions are closely related to the calibration of the sensors. We indicated in subsection 3.2, that we expect the calibration to leave an offset because it underestimates foreground sensors. Therefore, a part of the systematic error that produces the remaining offset can be attributed to the under-calibrated foreground sensors in the Crowdsensor and Scout datasets.

In case of the Crowdsensors (Figure 1) this seems a plausible explanation as they underestimate Bourges-Platz, Karlstraße and Königsplatz by a somewhat comparable amount. LfU, being located far away in an suburban background, seems to match the means of the rather crude predictions by pure coincidence.

In case of the Scouts (Figure 2), the same argument applies and the pattern seems to be the same as for the Crowdsensors. The offset is considerably less, but that may well be due to them being higher quality sensors and individually calibrated during the duration of SmartAQnet a year ago. Also, if the ratio of background to foreground sensors is different in the dataset, this will also produce offsets of different sizes in the predictions.

Ref. Name	Ref. mean	IDW (nc)	IDW (c)	(2+1)D GP (nc)	(2+1)D GP (c)
Bourges-Platz	25.5	14.2	18.3	11.5	18.4
Karlstraße	28.8	12.8	18.2	5.6	17.3
Koenigsplatz	25.5	13.0	18.2	9.5	17.8
LfU	19.5	16.7	18.7	15.9	18.7

Tab. 1: Prediction means of the SDS011 Crowdsensor dataset for the reference stations. We give calibrated (c) and not-calibrated (nc) values.

Ref. Name	Ref. mean	IDW (nc)	IDW (c)	(2+1)D GP (nc)	(2+1)D GP (c)
Bourges-Platz	25.5	18.9	22.8	18.9	22.8
Karlstraße	28.8	18.8	22.0	19.6	22.5
Koenigsplatz	25.5	17.4	22.4	18.0	23.0
LfU	19.5	20.1	24.4	16.4	21.2

Tab. 2: Prediction means of the Scientific Scout dataset for the reference stations. We give calibrated (c) and not-calibrated (nc) values.

Table 1 (Table 2) shows the prediction means of each interpolation method for the SDS011 Crowdsensors (Scientific Scouts) at the locations of the reference stations. We see that in most cases, the calibration shift is substantial. We may safely disregard the reference at LfU, since it is a suburban background station, which we expect to be lower than both, the Crowdsensor and the Scout dataset, which are both located mostly in urban environment. We also see that (excluding LfU), not-calibrated (calibrated) predictions range between 19.4% - 55.7% (60.1% - 72.2%) for the Crowdsensor dataset and 65.3% - 74.1% (76.4% - 90.2%) for the Scientific Scouts dataset. We can see that the calibration, while not adjusting the sensors 100% accurate, substantially improved the predictions. We also observe, that the Crowdsensors improve more than the Scouts, which matches our expectation, since the scouts have been previously calibrated in project SmartAQnet.

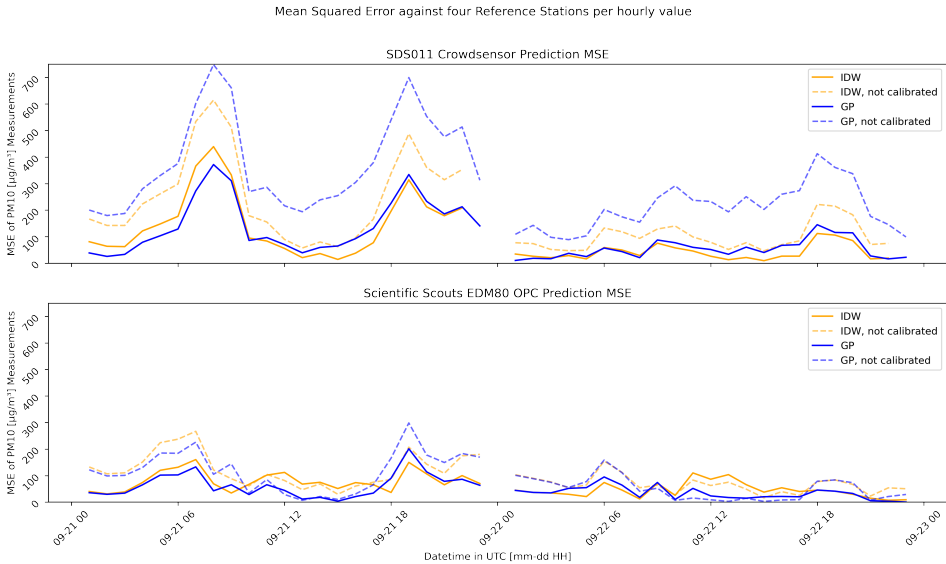


Fig. 3: MSE without (dashed) and with (solid) Calibration. Top (bottom) is the prediction of the SDS011 Crowdsensor (EDM80 Scientific Scout) dataset. We predicted each hourly value for 21st and 22nd of September 2020 (48hrs) for each reference station. Squaring the difference to the measurement and averaging over all four reference stations produced the shown plot. The hole at midnight of the 22nd is caused by missing data. In blue is the (2+1)D Gaussian Process Regression; in orange the inverse distance weighted interpolation.

Figure 3 shows the mean squared errors for Crowdsensors and Scouts with and without calibration. The Scouts perform overall better than the Crowdsensors, while the applied methods - inverse distance weighted interpolation and the (2+1)D Gaussian Process - perform similarly. This figure basically summarizes the essential information from this section.

5 Bike Route Prediction

5.1 Data

To tackle a prediction of a personal exposure scenario, we chose a bike route from 2020-09-22, which collected data from 6:48:20 to 7:10:56 over a length of 4.121 km for a total of 465 data points. The sensor used on the bike was a SDS011 sensor. The path is shown in Figure 4, along with the locations of the SDS011 Crowdsensors, EDM80 OPC Scientific Scouts and the official reference stations.

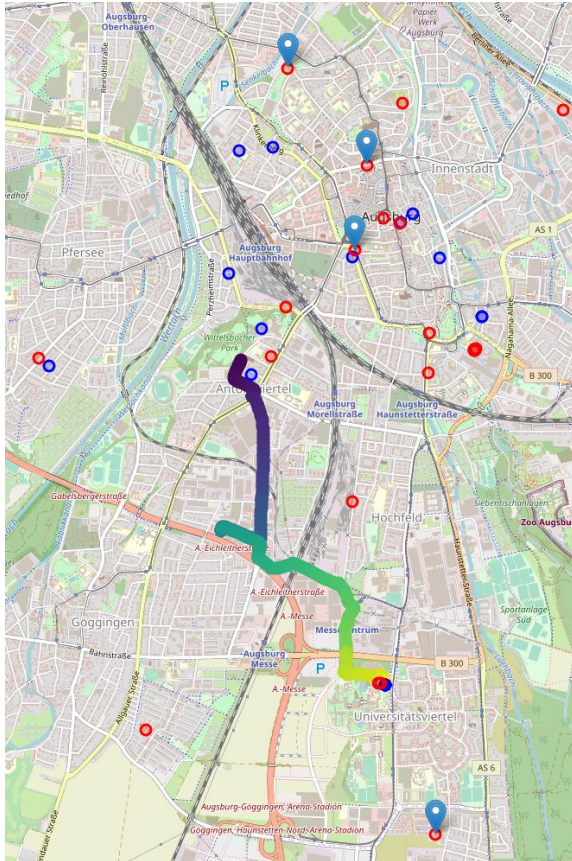


Fig. 4: The Location of the bike route within the city of Augsburg. Positions of the SDS011 Crowdsensors (Scientific Scouts) are marked in blue (red). The positions of the official reference stations are marked by the blue pins. The color coding of the bike path refers to the travel time with purple being the start and yellow the end. The map has been created using folium for python.

We used a total of 3454 (331) data points in the window from 6:00:00 to 7:15:00 to train the (2+1)D Gaussian Process for the Crowdsensors (Scouts). In contrast to the sample used in

the interpolation described in section 4, these were not further aggregated and correspond to 30 second to 1 minute values for the Crowdsensors and 5 minute values for the Scouts. We chose the RBF length scale as $[0.00450, 0.00680, 10]$, where the spatial dimensions again refer to lat, lon and correspond to $\sim 500\text{m}$ in each direction. The time direction is now encoded in seconds and fixed at 10 seconds to be sensitive to the frequency of the not aggregated sensor values.

5.2 Methods

We used the same two methods as in subsection 4.2, inverse distance weighted interpolation (IDW) and a (2+1)-dimensional Gaussian Process Regression ((2+1)D GP). Additionally, we performed a simple spatial Random Forest Regression (RF) using the scikit-learn implementation and a 2D (spatial) Gaussian Process Regression (2D GP) on the relevant time slice corresponding to each datapoint on the bike path.

To account for the timestamps of the bike path measurements and the sensor measurements not aligning, we used linear interpolation between values of the sensors to obtain values for the timestamps of each bike path datapoint for each sensor. The (2+1)D Gaussian Process was trained on the original data without interpolation and learned the interpolation.

We then subsequently aggregated the bike path measurements and calculated the predictions of each interpolation method for each aggregation interval.

For comparison, we calculate an IDW of the four reference stations. To that end, we split the values of the hours in question in the following way: Since the bike path ran from 6:48 to 7:10, we took 55% of the 6:00-7:00 hourly value and 45% of the 7:00-8:00 hourly value.

5.3 Results

Figure 5 shows the mean squared errors for each interpolation method as a plot over the aggregation time. There are two effects, which drew our attention: a) The interpolation methods all seem to perform vastly better than the background station and b) the (2+1)D GP performs way better than the other methods in case of the Scientific Scouts. However, in our subsequent analysis, we found that both effects stem from other sources and can not be attributed to any interpolation method performing any better than any other, nor the sensor network having an advantage over the reference stations, given the data at hand. We will first discuss the two effects that lead to Figure 5 and then discuss possible issues and improvements of our study.

a) The difference in MSE between the interpolation methods used on the Crowdsensors and Scouts and the prediction of the reference station depicted in Figure 5 is most likely purely

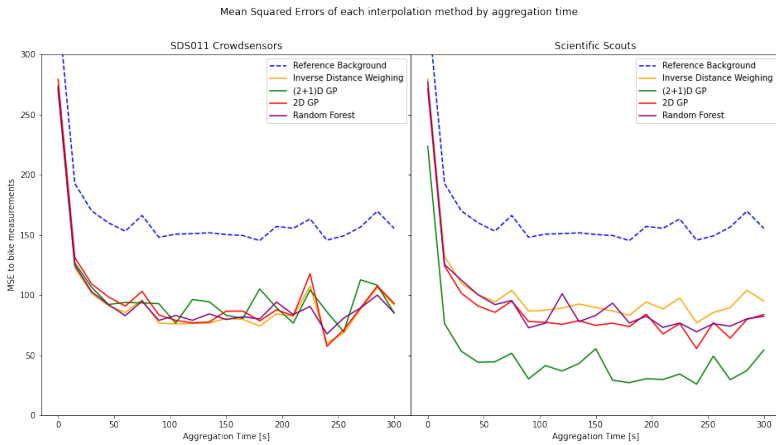


Fig. 5: Mean Squared Error for the SDS011 Crowdsensor dataset (left) and the Scientific Scout dataset (right) over aggregation time. The dashed blue line is the MSE against the mean of the reference stations for the time span of the bike path. MSE of inverse distance weighting, (2+1)D Gaussian Process, 2D Gaussian Process and 2D Random Forest in orange, green, red, and purple, respectively.

due to calibration. On the one hand, as stated earlier, our calibration of the Crowdsensors and Scouts improved their predictions a bit, but still left room for improvement. A proper calibration, where each single sensor is calibrated according to the specific environment it is deployed in would likely shift the interpolations of the Crowdsensors and Scouts in Figure 5 towards the prediction of the background stations. The large offset of the background station prediction MSE likely stems from the bike sensor not being well calibrated. The low mean of the bike sensor measurements, as shown in Figure 6, when compared to the Crowdsensors, Scouts and reference stations, supports that argument. Thus, any offset in Figure 5 likely stems from imperfect calibration of the relevant sensors and cannot be used to support our hypothesis.

b) The (2+1)D GP prediction MSE in the right side of Figure 5 seems to suggest, that the method outperforms the other methods. However, on further analysis, we found that this effect is just coincidence: When the temporal length scale is very low compared to the interval between measurements, the (2+1)D GP drops to the prior most of the time. This is because the GP interprets the temporal distances between the measurements as too large to be correlated. In case of the Scouts, which have a 5min measuring frequency, this is the case when training the GP with a temporal length scale of 10 seconds. The prior itself happens to be closer to the bike path measurements than the other interpolation methods by pure coincidence. This can be seen in Figure 6. We conducted further calculations with increasing temporal length scale (30s, 60s, 90s) and found that the (2+1)D GP starts to align

more with the shapes of the other methods and the MSE in Figure 5 also moves up to join the other methods. Therefore, the lower MSE of the (2+1)D GP also cannot support a claim of the method being superior in predicting personal exposure.

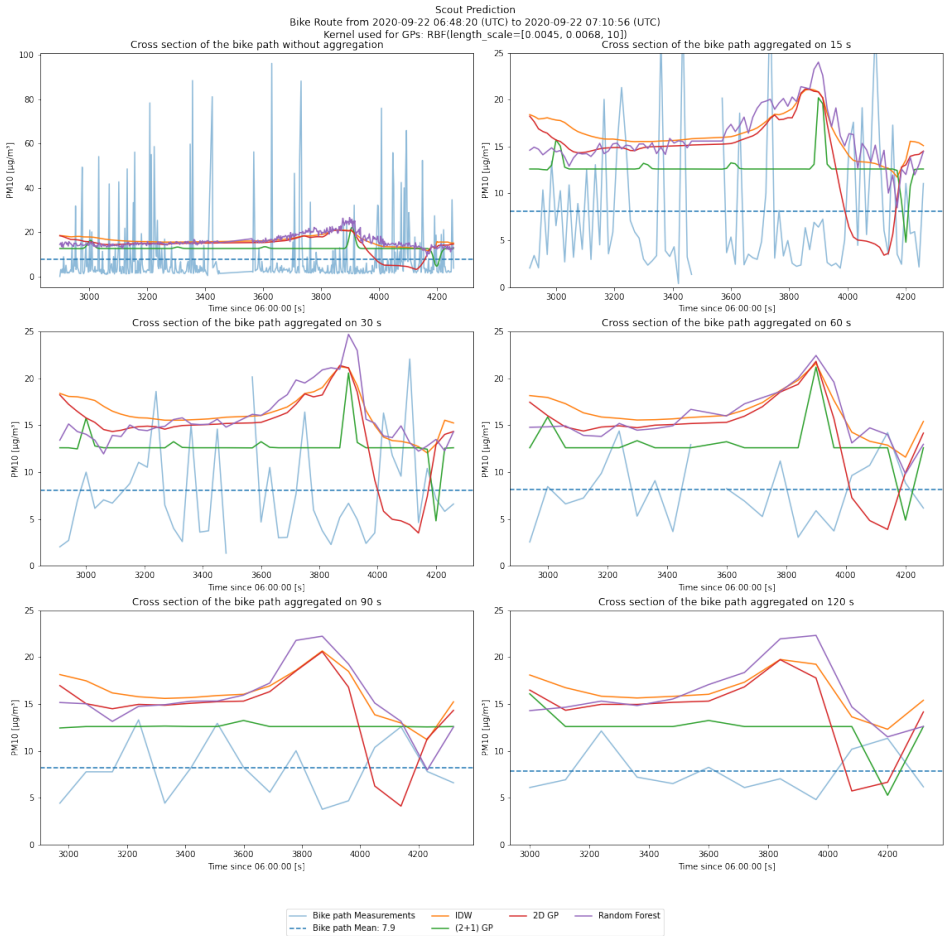


Fig. 6: Comparison of the various interpolation methods with the measurements along the bike path for different aggregation intervals.

We suggest, that the slope of a linear regression for each line in Figure 5 can show, that the corresponding method starts to pick up some predictive power as higher aggregation times wash out some of the fast fluctuations. In the case of the Crowdsensors, however, no reasonable slope is detectable. In case of the Scientific Scouts, the three AI supported methods (GPs and RF) seem to have a slope which, however, is still small compared to the fluctuations of the lines. This would be an interesting metric to follow for future work on a different dataset.

6 Discussion

We conclude, that for the given data, we were not able to show that a AI based interpolation of the network of Scientific Scouts or SDS011 Crowdsensors provides a significantly better prediction of personal exposure than an IDW of the reference stations. One reason why this part of our analysis led to no result is likely the geographical location of the route. We see from Figure 4, that the route most of the time is far away from both types of sensors, making a prediction by pure interpolation very unstable. An interesting question for future work would therefore be repeating the same analysis on a better suited dataset.

We were able to show in section 4, however, that accurate predictions are within reach of sensor networks comparable to those of the Crowdsensors and Scouts, and depend less on the sensor quality, as one might naively think. Our results suggest, that an individual calibration of the sensors based on the classification of their environment is the most important ingredient in obtaining accurate predictions (c.f. Figure 1). At the same time, the spatio-temporal (2+1)D Gaussian Process performs better than the non-AI based naive inverse distance weighted interpolation. This can be further bolstered by increasing the amount of training data and/or modifying the kernel function, where we used the most basic choice of an RBF kernel.

6.1 Limitations

The biggest limitations on the analyses we presented are:

- a) The location of the bike path relative to the locations of the sensors as well as the overall density of the sensor network with $O(1)$ sensor per km^2

Denser networks promise to better resolve the fast fluctuations seen on the bike path measurements, thereby allowing interpolation methods to capture these patterns. These high fluctuations likely stem from the bike sensor being in traffic and thereby occasionally very closely exposed to car emissions. We suggest that, while it may not be practicable to fully resolve the second-by-second fluctuations of the measurements, a denser network with sufficient time resolution might be able to resolve the changes after minimal aggregation. This translates into the question which spatial and temporal network densities lead to accurate predictions at which aggregation times. While it is clear, that it would require unrealistically high temporal and spatial resolutions to resolve the most highly localized, second-by-second fluctuations, the interesting question is which temporal accuracy of a prediction would be realistic to target and what network density would be required to achieve this. This is a task for future work to map out.

- b) Computing power to train spatio-temporal machine learning models on more data.

In our analysis, we restricted ourselves to two days in September, which we considered the relevant time frame for the bike route. For the prediction of the reference stations, we also did one run on which we trained the (2+1)D Gaussian Process using all September data (hourly aggregates, ~18.000 data points for the Crowdsensors and ~15.000 data points for Scouts) and predicted the four positions of the reference stations for each hour for the whole of September, which amounts to $4 \times 24 \times 30 \sim 3000$ predictions for each method. Predicting both, calibrated and not calibrated, for the evaluation makes roughly 12.000 predictions. On a local machine, using jupyter notebooks and scikit-learn as library for the GP, the process of training and predicting took more than two hours. Using several days of the unaggregated dataset described in section 3, would exceed the capabilities of a local machine. This run already excluded the most time consuming step: finding the optimal hyper-parameters for the Gaussian process. We argue, that the effects a network with an average distance of 1km between stations can estimate are in the ballpark of 500m correlation length. We confirmed that the GP can find lengths of that order with much smaller datasets (max. 2000 data points), and then fixed the GP for training with larger datasets on those 500m in each spatial direction. Increasing the input for finding the optimal hyper-parameters significantly increases the computing time needed, thus using the entire dataset of September (even if aggregated) for a training that tries to find the hyper-parameters is out of question on a local machine, let alone working with the raw, not-aggregated data. Yet, there may be phenomena and patterns (e.g. daily periods), which escape us when working with a restricted or pre-aggregated dataset.

6.2 On the use of the median in calibration and aggregation

The Crowdsensors occasionally have corrupt values which spike as high as nearly 1.000.000. These values extremely distort any calculation if the arithmetic mean is used and these values have not been discarded. This leaves us with picking the threshold by hand which induces a degree of arbitrariness. We chose to resolve this more elegantly by taking the median instead of the arithmetic mean for the calibration. Note that for the treatment of these spikes, it is secondary if they map real effects or are pure electronic errors within the device. Since high spikes can only occur in one direction, they expose the Poisson nature of the data to a degree, that any approximation as a normal distribution breaks down. The degree to which the dataset loses its normal shape, the arithmetic mean becomes erroneous as a measure of centrality. When calibrating on a monthly mean where huge spikes are present, the spikes will pull the arithmetic mean so far up, that they dominate the calibration procedure and render the result useless, e.g. by leading to a correction that pushes 95% of the sensors values into the negative range. This could be avoided by introducing a cutoff. The value of the cutoff, however, needs to be justified since it unavoidably excludes some high but likely real, physical measurements. Its purpose is only to shape the data to an approximate normal distribution so that the arithmetic mean can be used as a measure of centrality again. We chose to use the median as a measure of centrality to avoid having to pick an arbitrary threshold. While possibly the choice of the median for a given dataset can

be mapped on a mean with a fixed cutoff, the choice to us seemed less biased than choosing a threshold to reshape the data until we consider it sufficient for our methods.

Interestingly, in case of the Scientific Scouts, where no such high spikes were present, it did not matter whether we picked the mean or the median and Figure 2 and the respective parts of Figure 3 are nearly identical when using the median or the arithmetic mean for calibration and aggregation. This leads to the question how the scouts avoid such spikes. While it is possible, that unwanted effects like pure electronic errors or coarse dust entering the device may be prevented by appropriate measures in higher quality sensors, there is still the possibility for real high exposure, e.g. through a person carrying a cigarette walking by, which the sensor should correctly detect. As mentioned, it does not matter whether a true or false measurement destroys the approximate normal distribution, any spike will do. This poses a challenge when considering low cost sensors in urban areas and scenarios like mobile sensors, where high exposure is expected and the use of high-cost-high-quality sensors may be restricted.

In the case of aggregation, the same argument applies to a lesser degree. Here, one might argue that we are not primarily after a measure of centrality but a division of an integrated exposure. This would want to include real, valid measurements in the calculation. But since the mean in the presence of high spikes is dominated by them, it becomes crucial that the sensor in question still measures precisely in that high concentration range and that we are sure that the measured effect is indeed a real effect. Otherwise, errors will dominate our calculation. This sensitivity to the spikes being well measured and true effects thus effectively poses the same problem for low cost sensors, since they won't be able to guarantee the precision and truthfulness of a single measurement.

7 Conclusion

We conclude, that interpolation techniques applied on a distributed sensor network offer a lot of potential. While there are cases, in which they can provide additional value over the traditional, high-cost-high-quality measuring stations, the applications and their validation remains very challenging. We argue, that, in general, AI based interpolation - like any interpolation method - work better on more sensors than on less, and thus go hand in hand with sensor networks as an alternative approach to the traditional, high-cost-high-quality measuring stations. Especially when trying to tackle the intricate problem of accurately predicting real, personal exposure. We thus argue for both, sensor networks and AI based interpolation methods, as they can be seen as integrated system - neither can develop their full potential without the other. In the scope of future smart city applications, our analysis touches questions of urban decision makers, where the scenario is finding the best solution for monitoring air quality given a finite resources. The choice is basically to invest in very few, very accurate measuring stations – which is the traditional and standard way – or to invest into more, yet individually less accurate sensors. With machine learning methods

improving and starting to tap the potential of distributed sensor networks, they become an integral part of future smart cities.

8 Acknowledgement

This work was partially funded by the German Federal Ministry of Transport and Digital Infrastructure (BMVI) as part of project *SmartAQnet* [Bu17] (funding no. 19F2003B) and by the German Federal Ministry for Research as part of a *SDI-C* microproject.

References

- [Br18] Bruns, Julian; Riesterer, Johannes; Wang, Bowen; Riedel, Till; Beigl, Micheal: Automated quality assessment of (citizen) weather stations. arXiv preprint arXiv:1802.06018, 2018.
- [Bu17] Budde, Matthias; Riedel, Till; Beigl, Michael; Schäfer, Klaus; Emeis, Stefan; Cyrus, Josef; Schnelle-Kreis, Jürgen; Philipp, Andreas; Ziegler, Volker; Grimm, Hans et al.: SmartAQnet: remote and in-situ sensing of urban air quality. In: Remote Sensing of Clouds and the Atmosphere XXII. volume 10424. International Society for Optics and Photonics, p. 104240C, 2017.
- [Bu18] Budde, Matthias; Schwarz, Almuth D; Müller, Thomas; Laquai, Bernd; Streibl, Norbert; Schindler, Gregor; Köpke, Marcel; Riedel, Till; Dittler, Achim; Beigl, Michael et al.: Potential and limitations of the low-cost SDS011 particle sensor for monitoring urban air quality. ProScience, 5:6–12, 2018.
- [Ho08] Hoek, Gerard; Beelen, Rob; De Hoogh, Kees; Vienneau, Danielle; Gulliver, John; Fischer, Paul; Briggs, David: A review of land-use regression models to assess spatial variation of outdoor air pollution. Atmospheric environment, 42(33):7561–7578, 2008.
- [Ki14] Kilibarda, Milan; Hengl, Tomislav; Heuvelink, Gerard BM; Gräler, Benedikt; Pebesma, Edzer; Perčec Tadić, Melita; Bajat, Branislav: Spatio-temporal interpolation of daily temperatures for global land areas at 1 km resolution. Journal of Geophysical Research: Atmospheres, 119(5):2294–2313, 2014.
- [Kr51] Krige, Daniel G: A statistical approach to some basic mine valuation problems on the Witwatersrand. Journal of the Southern African Institute of Mining and Metallurgy, 52(6):119–139, 1951.
- [Li17] Lin, Yijun; Chiang, Yao-Yi; Pan, Fan; Stripelis, Dimitrios; Ambite, José Luis; Eckel, Sandrah P; Habre, Rima: Mining public datasets for modeling intra-city PM2. 5 concentrations at a fine spatial resolution. In: Proceedings of the 25th ACM SIGSPATIAL international conference on advances in geographic information systems. pp. 1–10, 2017.
- [Pe11] Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, E.: Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, 12:2825–2830, 2011.

- [PH16] Pebesma, Edzer; Heuvelink, Gerard: Spatio-temporal interpolation using gstat. *RFID Journal*, 8(1):204–218, 2016.
- [Sh68] Shepard, Donald: A two-dimensional interpolation function for irregularly-spaced data. In: *Proceedings of the 1968 23rd ACM national conference*. pp. 517–524, 1968.
- [Sh20] Shen, Yao; Lehmler, Stephan; Murshed, Syed Monjur; Riedel, Till: Characterizing Air Quality in Urban Areas with Mobile Measurement and High Resolution Open Spatial Data: Comparison of Different Machine-Learning Approaches Using a Visual Interface. In (Santos, Henrique; Pereira, Gabriela Viale; Budde, Matthias; Lopes, Sérgio F.; Nikolic, Predrag, eds): *Science and Technologies for Smart Cities*. Springer International Publishing, Cham, pp. 115–126, 2020.
- [Su16] Sun, L; Wei, J; Duan, DH; Guo, YM; Yang, DX; Jia, C; Mi, XT: Impact of Land-Use and Land-Cover Change on urban air quality in representative cities of China. *Journal of Atmospheric and Solar-Terrestrial Physics*, 142:43–54, 2016.
- [To70] Tobler, Waldo R: A computer movie simulating urban growth in the Detroit region. *Economic geography*, 46(sup1):234–240, 1970.
- [Yu16] Yu, Ruiyun; Yang, Yu; Yang, Leyou; Han, Guangjie; Move, Oguti Ann: RAQ—A random forest approach for predicting air quality in urban sensing systems. *Sensors*, 16(1):86, 2016.
- [ZLH13] Zheng, Yu; Liu, Furui; Hsieh, Hsun-Ping: U-air: When urban air quality inference meets big data. In: *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*. pp. 1436–1444, 2013.

Exploration and systematic assessment of the resource efficiency of Machine Learning

Achim Guldner¹, Sandro Kreten¹, Stefan Naumann¹

Abstract: Estimations of today's energy consumption of information and communication technologies (ICT) range from 2 to 9 % of the total produced energy and forecasts for the year 2030 predict an increase up to 21 %. Even though these numbers are controversial, it cannot be denied that the consumption growth of large impact factors, like data centers, networks, consumer devices, and the production of ICT needs to be reduced. In addition to Green IT, which is primarily focused on hardware, software is increasingly seen as an energy consumer with considerable savings potential. In this paper, we take a look at software for artificial intelligence (AI) and especially machine learning (ML). We describe a method for in-depth measurement and analyses of the energy consumption and hardware usage of ML algorithms and a series of experiments where we use the method on convolutional neural networks (CNN). We also compare existing estimation methods with our own. As outlook, we propose a holistic approach along the AI life cycle and additional experiments and assessments that could show potential efficiency improvements and consumption savings in AI.

Keywords: resource efficiency; energy and hardware assessment; measurement and analysis method

1 Introduction

In the last decade, AI and ML gained significant momentum, both in research and society. New technologies, like general-purpose computing on GPUs², new optimization algorithms, like ADAM (introduced by Kingma et al. in 2015 [KB15]), and new libraries, like TensorFlow (first released in 2015), brought the field to a level never seen before. Of course, this trend brings many advantages, especially, because of the many areas it can be applied to and the usability improvements that enable a more widespread usage. Companies such as Google, Amazon, and Microsoft already provide networks in their ML development environments, which can be applied in various productive areas. Thus, the combination of different pre-trained networks is often faster and more successful than the development of new processes. Especially in speech, text, and image recognition, rapid success can be achieved with pre-trained networks and ready-to-use data sets.

Recently, the implications of AI on a sustainable society shifted into research focus. Many works seek to implement new solutions to environmental issues, using machine learning.

¹ Hochschule Trier, Umwelt-Campus Birkenfeld, Postfach 1380, 55761 Birkenfeld, Germany {a.guldner|s.kreten|s.naumann}@umwelt-campus.de

² like Nvidia's Compute Unified Device Architecture (CUDA), released in 2007

Khakurel et al. [Kh18] find significant impacts of AI, both positive and negative, on all five dimensions of sustainability. Vinuesa et al. [Vi20] state that AI can enable the accomplishment of 134 targets across all of the United Nations 17 Sustainable Development Goals (SDGs), but it may also inhibit 59 targets. AI and ML can also be used in efforts to reduce the energy- and resource consumption across many subtopics, as evidenced, e.g. in the new journal "Energy and AI" (Jin et al. [JOJ]). The evaluation of the sustainability of AI and ML was first addressed by Strubell et al. [SGM19], estimating the CO_2 emissions of Natural Language Processing (NLP). They use Intel's RAPL tool³ and nvidia-smi⁴ while training the networks and approximate the energy usage and CO_2 emissions. They report, e.g. that a neural architecture search for translating English to German used 415409 kilowatt-hours (kWh), or 656347 kWh, factoring in Power Usage Effectiveness (this equals the annual energy consumption of ca. 200 households). Henderson et al. [He20] propose an *experiment-impact-tracker* framework for Python that also uses RAPL and nvidia-smi and encapsulates the assessment of the energy consumption. Schmidt et al. [Sc21] also provide a Python package based upon RAPL and nvidia-smi, called *CodeCarbon* that takes into account the computing infrastructure, location, usage, and running time. Both tools estimate how much CO_2 is produced by the algorithm. Canilang et al. [Ca21] take a look at the consumption in edge AI applications. This is especially advisable because of the limited hardware of those devices.

Since 2009, our research focuses on developing and assessing green and sustainable software systems. We defined *Sustainable Software* as software whose development, deployment, and usage results in minimal direct and indirect negative impacts or even positive impacts on the economy, society, human beings, and the environment (Naumann et al. [Na11]). From this definition, we derived criteria to evaluate a software product's sustainability in Kern et al. [Ke18], which includes our measurement method. In Naumann et al. [NGK21], we describe the process of the creation of a sustainability label (Blue Angel) for software products.

In this paper, we base the assessment of the sustainability of AI and ML upon the approach devised in these previous works and set three goals. We (1) assess, if the implemented measurement and analysis approach can be applied to the area of AI and ML and modify it where necessary, (2) perform two experiments and assess the energy consumption and hardware usage of a PC while training a CNN and compare it with the existing estimation approaches, and (3) provide an outlook on further experiments, assessments, and a holistic approach along the whole AI/ML life cycle.

2 Method

In this section, we describe the measurement setup and how we applied it to ML. The setup (depicted in fig. 1) is based upon Kern et al. [Ke18] and follows ISO/IEC 14756, as

³ <https://01.org/blogs/2014/running-average-power-limit-%E2%80%93-rapl> [accessed 2021-04-04]

⁴ <https://developer.nvidia.com/nvidia-system-management-interface> [accessed 2021-04-04]

introduced by Dirlwanger [Di06]. It consists of a system under test (SUT) on which the software that is to be assessed, is executed. In this case, the “software” is a Python script that trains CNNs. The power draw of the SUT is measured by a power meter (PM). A workload generator (WG) puts load on the SUT. With desktop or client-server software, this usually is a script that repeatedly performs and logs user inputs via an automation tool. Here, we used a python-script that executes and logs the training process. The recorded data is aggregated and evaluated (DAE) with a script, written in R, to produce the efficiency report.

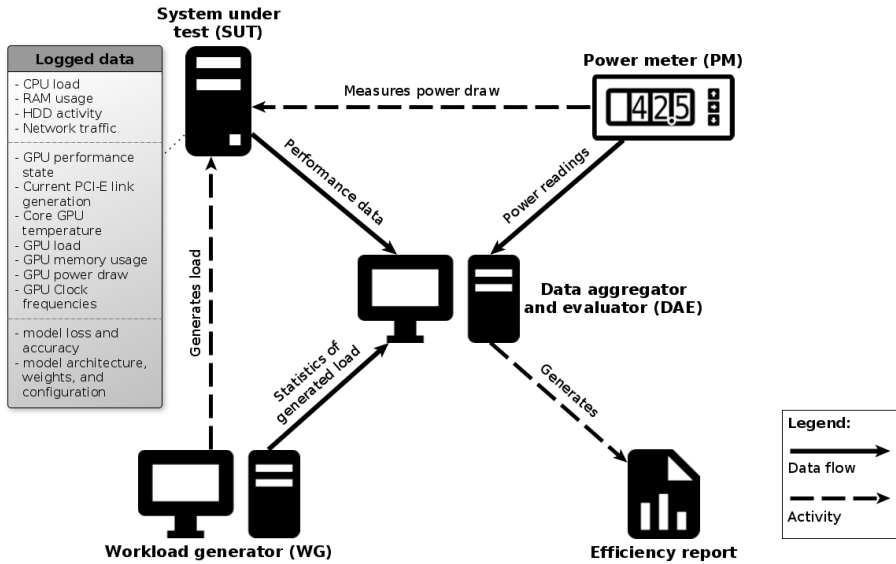


Fig. 1: Measurement setup

The PM saves the per second average power draw of the SUT. The WG script logs timestamps for the start and end of the test runs and training phases (epochs). The SUT collects its own performance data (CPU load, RAM usage, etc.) using `collectl`⁵. In addition to the previous setups, we integrated graphics card usage logs, collected with `nvidia-smi` and training and validation loss and accuracy in the analysis process. For later reference, the finished model architecture is also exported (using `model.save`).

The components of the used SUT are detailed in table 1. For the experiments, we set up the described software stack and measured the power draw and hardware usage for a baseline measurement (SUT runs only the operating system) and 3 usage scenarios: *idle* (Python script is running with no net being trained), *CNN* (training a CNN to classify pictures) and *transfer learning* (using a pre-trained CNN and training only the last layer). The scenarios are described in more detail in section 3. To combine and compare the method with the approaches from Henderson et al. [He20] and Schmidt et al. [Sc21], we set up additional

⁵ <http://collectl.sourceforge.net/> [accessed 2021-04-09]

Tab. 1: SUT hardware specifications and software stack

Component	Specifications
CPU	AMD Ryzen 7, 1700
RAM	16GB G.Skill RipJaws V DDR4-3200 DIMM
Main Board	MSI B350 PC MATE
Power supply	Cooler Master Silent Pro Gold, 600 W
SSD	256GB WD Black M.2 2280 NVMe SSD (WDS256G1X0C)
Graphics card	ZOTAC GeForce GTX 1060 AMP! Edition, 3 GB 192 Bit GDDR5
Operating System	Ubuntu 20.04.2 LTS
Python	Version 3.8.5
TensorFlow	Version 2.4.1
CUDA	Version 10.1.243

scenarios, where we measured the SUT's power draw, while simultaneously running the experiment-impact-tracker and CodeCarbon, respectively.

Finally, to get a feeling for the accuracy of `nvidia-smi`'s gpu power draw readings, we compared the logging results with voltage and current measurements at the power plug of the graphics card while running GPU benchmarks. This is by no means as accurate as the measurement with the PM, but the results showed that the `nvidia-smi` logs are believable. The whole measurement setup is shown in fig. 2.

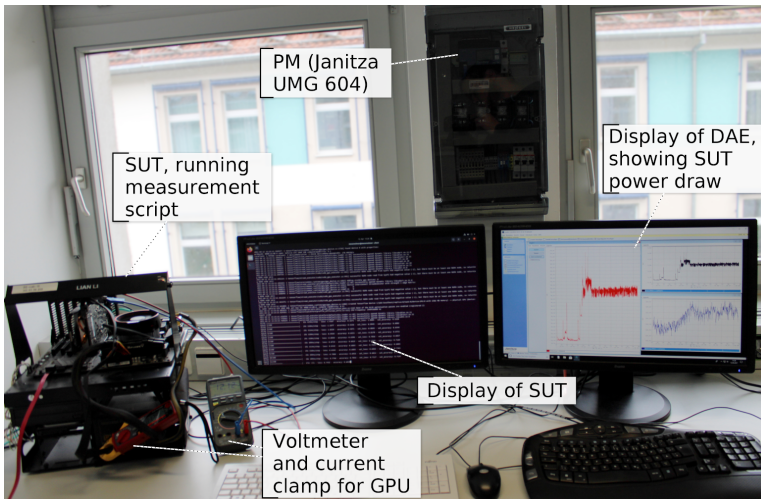


Fig. 2: Photo of measurement setup

The analysis method is largely identical to the one described in [Ke18]. We average each logged hardware component over the test runs and calculate the average energy consumption of the scenario in watt-hours [Wh], as the sum over all per-second averages of the power

draw (recorded by the PM during the execution of the scenarios, resulting in watt-seconds) and dividing the result by $3600 \frac{s}{h}$. Subtracting the mean baseline values, adjusted for the scenario duration, gives us the part of the energy consumption and hardware usage that is *induced by the software*. Additionally, we calculate these indicators also for each training epoch.

In order to make the experiments and analysis reproducible, we provide all source code files, measurement data, analysis scripts, and efficiency reports in a replication package as git repository at <https://gitlab.rlp.net/a.guldner855830/exploration-and-systematic-assessment-of-the-sustainability-of-machine-learning/-/tree/master>⁶.

3 Experiments

As a first set of experiments, we used the method to assess the energy consumption and hardware usage of the SUT while training CNNs to classify images from the `tf_flowers` data set⁷. The WG script first groups the images into folders according to their labels (roses, daisies, dandelions, sunflowers, and tulips) and then randomly augments them (re-scale, rotate, change width- and height, flip horizontally, and zoom) to enhance the training process and have access to more training and validation data sets. Figure 3 depicts some examples. Finally, the images are fed to the `fit` function using an `ImageDataGenerator` object.

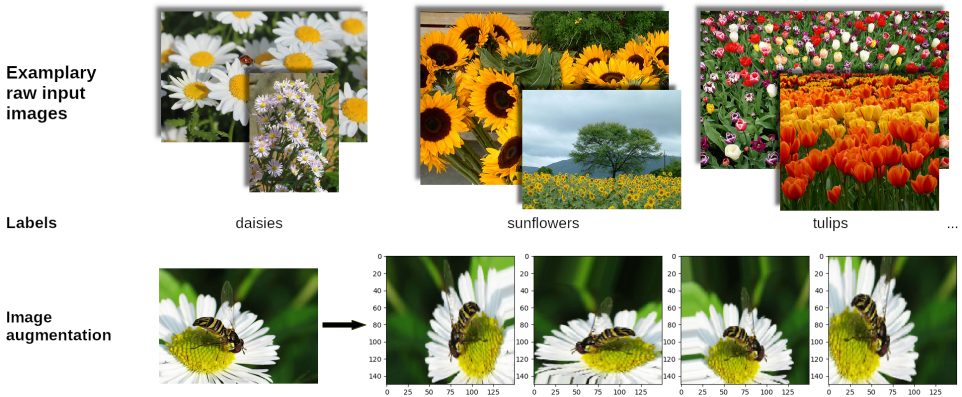


Fig. 3: Exemplary training images from the `tf_flowers` dataset (top) and image augmentation (bottom)

The scripts are based upon examples available from the TensorFlow git⁸. We ran each scenario (*idle*, *CNN*, and *transfer learning*) 30 times (=“test runs”), so as to produce normally distributed samples [Ke18]. The scripts execute the scenarios described below, and log the time for the beginning and end of each test run. In addition to the scenarios, we

⁶ the model architectures are available at <https://seafilerlp.net/d/6001c483b42342a0bdb5/>

⁷ available at https://www.tensorflow.org/datasets/catalog/tf_flowers [accessed 2021-04-09]

⁸ <https://github.com/tensorflow/examples> [accessed 2021-04-09]

also recorded the baseline for the SUT, with only the operating system installed and without any scripts running. The idle scenario consists of the same script as the CNN and transfer learning scenario, except that instead of a model being trained, the execution is suspended for 10 minutes, using Python's `time.sleep()` function. This was included e.g. to simulate developers running an interactive session when implementing algorithms.

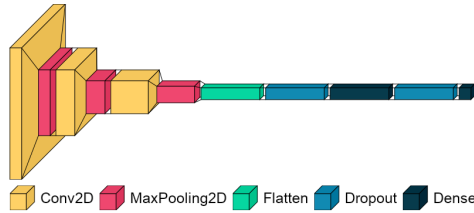
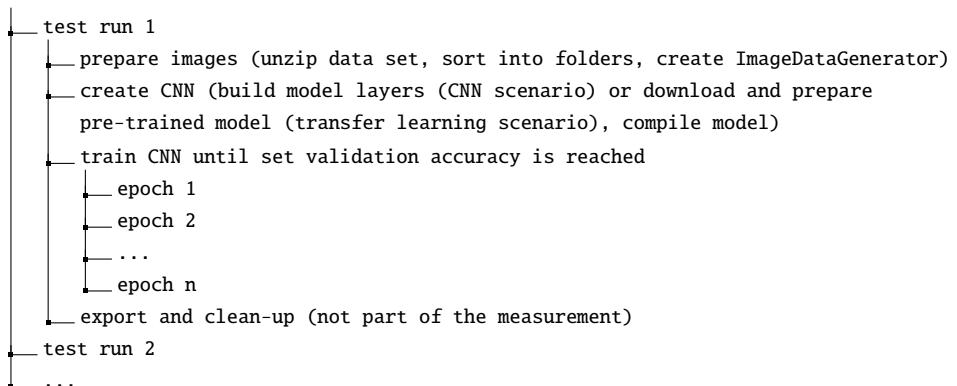


Fig. 4: CNN scenario model visualization

The CNN scenario consists of a `Sequential` model with 3 sets of `Conv2D` and `MaxPooling2D` layers, followed by a `Dense` hidden layer with 512 nodes and an output layer with 5 nodes (see fig. 4)⁹. For the transfer learning scenario, *MobileNet V2* from TensorFlow Hub is used¹⁰. Here, only the last layer is exchanged for an output layer with 5 nodes and only this layer is trained (this method is called "fine-tuning"). Both nets use the `adam` optimizer for a maximum of 100 training epochs, or until a validation accuracy of 80 % (CNN scenario) or 85 % (transfer learning scenario) is reached. This usually results in a different number of epochs for each test run. Therefore, the scripts also log timestamps for each *action* of the training process: image preparation, creation of the net, training process up to 80 % or 85 % validation accuracy, and for each epoch. After the test run, the script exports the trained model architecture and history (training- and validation-accuracy and loss) and resets the conditions as they were before the test run (deleting folder structures and pre-trained models). The actions within each test run are structured as follows:



⁹ Created with `visualkeras`, available at <https://github.com/paulgavrikov/visualkeras/> [accessed 2021-04-13]

¹⁰ https://tfhub.dev/google/tf2-preview/mobilenet_v2/feature_vector/4 [accessed 2021-04-09]

4 Analysis results

We now present an overview of the analysis and some explorative aspects. Table 2 shows the recorded data. From it, we calculated the energy consumption for the SUT and the graphics card, as well as the hardware usage averages, as described in section 2. Because the test run durations vary, we also calculate the averages for the whole training process and for each epoch. The data is visualised and aggregated in the report documents in the replication package.

Tab. 2: Recorded data, units, and recording method

Indicator	Unit	Method
SUT power draw	W	Power meter
SUT total CPU load	%	collectl
SUT RAM usage	kB	collectl
SUT disk activity (reading and writing)	kB	collectl
SUT network traffic (sending und receiving)	kB	collectl
GPU performance state	no unit	nvidia-smi
GPU PCIe link generation	no unit	nvidia-smi
GPU core temperature	°C	nvidia-smi
GPU load	%	nvidia-smi
GPU memory load	%	nvidia-smi
GPU memory allocated	MiB	nvidia-smi
GPU power draw	W	nvidia-smi
GPU SM, memory and graphics clock frequency	MHz	nvidia-smi

As an overview, fig. 5 shows the per second and overall average SUT power draw for the two usage scenarios. The results show the expected large benefits of using a pre-trained net and fine tuning it: The energy consumption induced by the transfer scenario amounts only to 0.374 Wh, compared to 10.032 Wh for the model that is trained from scratch (see fig. 6). Of course, this does not take into account the energy that was used to train the MobileNet in the first place. Thus, the two scenarios are not directly comparable, but that was also not the intention in this paper.

Looking at epoch-level, it is interesting to note that the energy consumption for the training process in the transfer scenario used 0.448 Wh per epoch with an average duration of 29.95 seconds, whereas for the CNN scenario, each epoch used only 0.423 Wh and, on average, took 28.304 seconds. This is even more surprising, because the complexity of the self-trained CNN is much larger (total and trainable parameters according to `model.summary()`: 25,716,773) than that of the adapted MobileNet (total parameters: 2,264,389; trainable parameters: 6,405).

Because the reports provide an in-depth analysis down to epoch-level for all recorded indicators, depending on the use case, it is possible, e.g. to analyse the energy consumption and hardware usage up to a certain epoch or until the validation accuracy of the net reaches

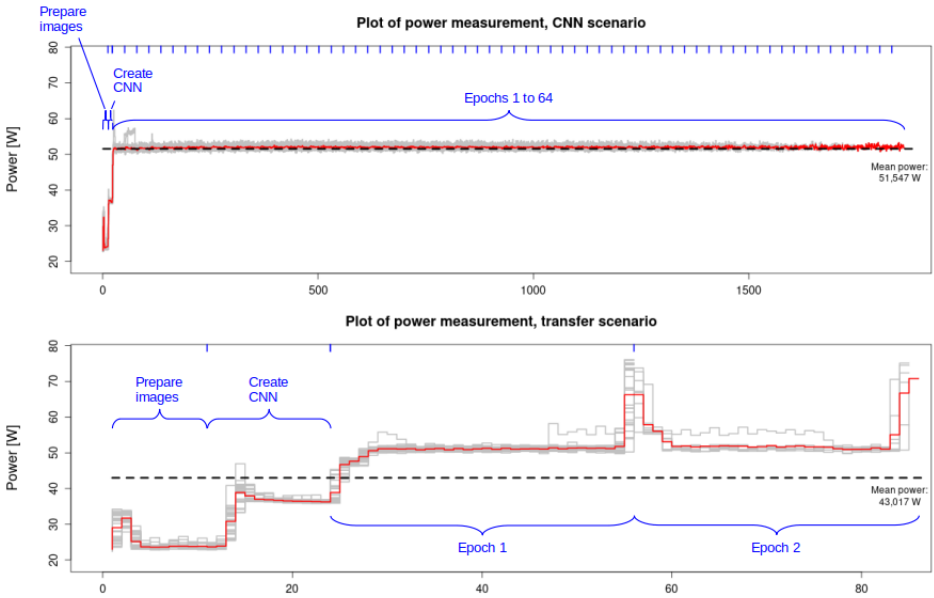


Fig. 5: Power draw measured with the PM for the CNN- (top) and transfer scenario (bottom).

a certain percentage. This allows detailed insights into how adjustments to the net relate to the energy and hardware consumption. E.g. it caught our eye that as soon as the model is initialized with `model = Sequential()`, the GPU memory was allocated even though no layers had been added or trained. Here, more research is necessary (see also section 5).

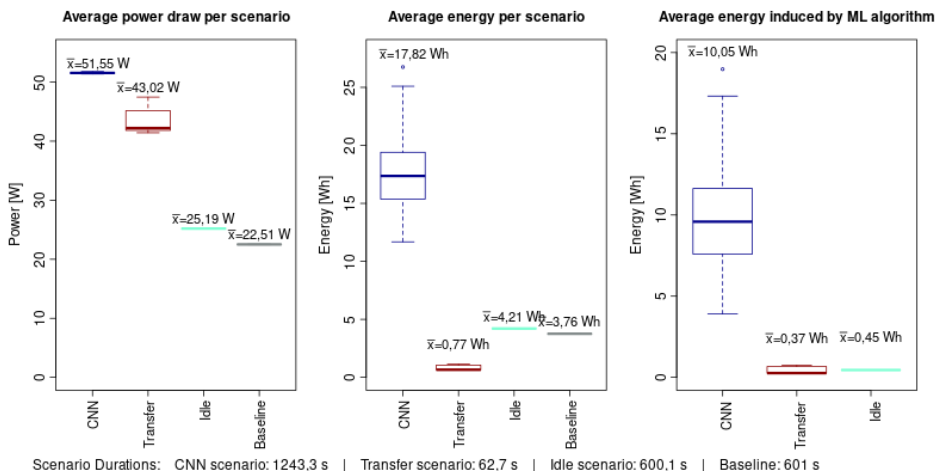


Fig. 6: Power draw and energy usage overview for all scenarios.

Fig. 6 depicts the average power draw, the average energy consumption, and the average energy consumption induced by the ML-algorithm (scenario energy minus adjusted baseline energy), per scenario. From this overview we can gain some insights: (1) considering energy consumption, using a pre-trained CNN seems to be by far the better choice for image classification, (2) the GPU’s energy consumption rises from about 35 % in the baseline to about 46 % in the idle scenario and to about 70 % in the scenarios that train CNNs.

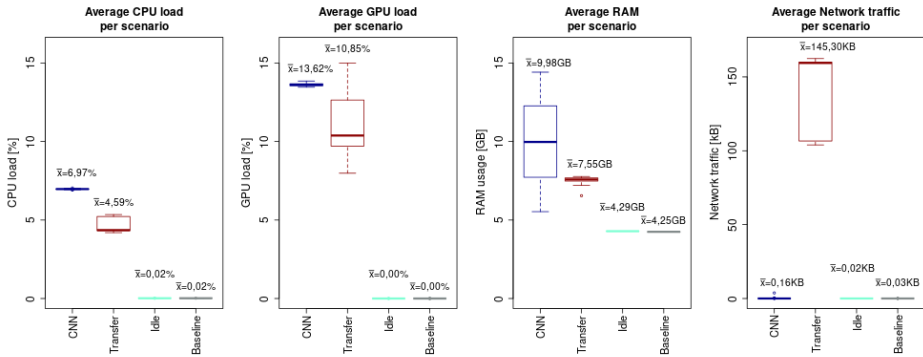


Fig. 7: Hardware usage overview.

For fig. 7, we chose to plot the CPU and GPU load, RAM usage, and network traffic as the main hardware usage indicators. Here, we find that the average hardware usage is also, for the most part, lower for the transfer scenario than for the CNN scenario. However, because the pre-trained model needs to be downloaded for the transfer scenario, this generates some network traffic and disk activity.

Another point we wanted to investigate was if the loss or accuracy would correlate with the energy consumption. This was, however, not the case in our measurements. What we did find was that the energy consumption per epoch correlates with the GPU temperature ($r = 0.98$), as can be seen in fig. 8. Likely, the GPU draws more power when it heats up.

From the collected data, it is also possible to calculate an *efficiency factor*, based upon the metrics proposed by Johann et al. [Jo12] as

$$\text{Energy efficiency} = \frac{\text{Useful work done}}{\text{Used energy}}.$$

For this paper, we used the validation accuracy as the “useful work done” in the training process. This results in the unit *validation accuracy per watt-hour*. It is depicted in fig. 8. Because the accuracy is not a sufficient measure for the “work” of the neural net, in the future more indicators should be taken into account (see section 5).

In summary, it seems advisable to include at least the following indicators for further analyses: Total energy consumption, GPU and CPU energy consumption (if available), CPU and GPU load, RAM and GRAM usage, network traffic and disk activity, as well as loss and

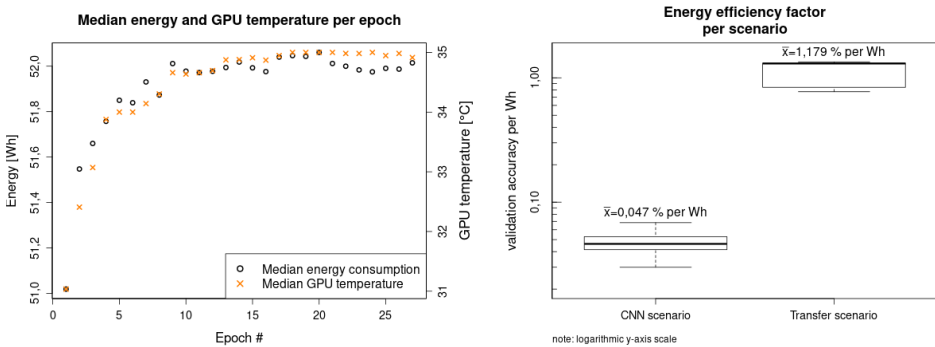


Fig. 8: Energy consumption per CNN scenario epoch (left) and efficiency factor comparison (right).

accuracy. The baseline measurement also proves useful to assess the energy consumption and hardware usage that is induced by the algorithms.

5 Discussion and outlook

In this section, we discuss the goals set in section 1: Applying our approach to the field of AI/ML, comparing our method to the existing approaches, and providing an outlook on our planned work. The devised method seems promising for the systematic and in-depth assessment of the sustainability of ML algorithms. Where the trackers, especially CodeCarbon, provide a good tool for familiarizing oneself with the energy demands of the implemented algorithms, using a PM makes it possible not only to estimate the energy consumption, but to measure it. Combined with hardware usage logs, the calculation of efficiency factors, and baseline measurements to assess the consumption induced by the ML algorithms, the method allows comparing scenarios, down to source code level. This allows further experiments to assess how changes to the algorithms effect the resource efficiency.

Since this is the first evaluation of our method in the area of ML, there are some threats to validity that we now discuss. We used only one ML library (TensorFlow), one SUT, and one programming language (Python). To mitigate this, our scenarios were based on standard use cases and, as future work, we plan to extensively evaluate the approach on a large set of ML applications, ranging different languages and technologies. We took precautions to minimize side effects in the measurements, performed baseline measurements of the SUT energy consumption, and executed each test case multiple times in order to ensure the statistical relevance of the data. We assessed the precision of the measurement setup by calculating the standard deviation of the power measurements and execution times and checked the plausibility of the gpu measurements. There is still potential for measurement- and coding errors and conclusions may be prone to subjective interpretations. In order to

mitigate these threats, we jointly inspected and discussed the results in order to identify potential occurrences of divergent interpretations and provide a replication package.

For the comparison with “CodeCarbon” (CC) [Sc21] and “experiment-impact-tracker” (EIT) [He20], we wrapped the test runs with the EmissionsTracker from CC and the ImpactTracker from EIT, respectively and then executed them as before, while recording the same three scenarios again (see fig. 9). Baseline measurements cannot be recorded with the trackers, because they require python to be running (= idle scenario). Of course, since our SUT contains an AMD CPU, RAPL could not be used and only the graphics card was available to the trackers. The CC tracker worked as expected and reported approx. 78 % of the PM measurements, which is close to the nvidia-smi readings.

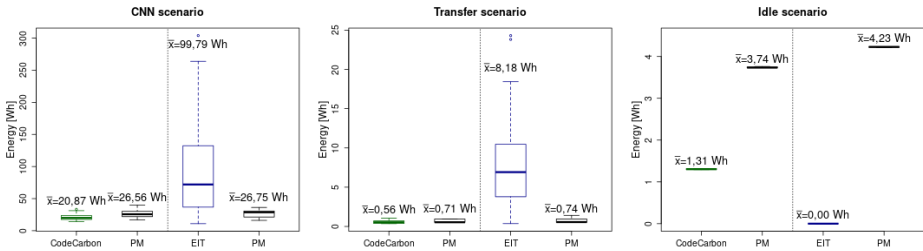


Fig. 9: Comparison of measured energy consumption of CC, EIT and our PM.

Although we used the EIT as described in the documentation with standard settings, the results cannot be correct. It is not the focus of this work to go into detail on why these errors occurred. From what we could see in a brief code review of EIT, the error is likely due to the fact that the power and energy averages are not calculated over time, but the first result from nvidia-smi over a measurement interval is used. This can produce large measurement errors because power spikes or drops can be missed. Furthermore, the tracker did not record any consumption for the idle scenario. This is the case, because the tracker considers the gpu-processes and bases the estimates on those. Since there are no processes running in the idle scenario, the tracker records no consumption. Compared to CC, EIT itself induces approx. 0.49 Wh additional energy consumption in the idle scenario, likely because it produces frequent console outputs and large amounts of log files with the standard settings. This results in additional power draw.

Since Strubell et al. [SGM19] do not provide or use a tool, but rather directly log the data with nvidia-smi and RAPL. Thus, the method seems directly comparable and should lead to similar estimation results if the scenarios we defined here were used. The focus here was on bringing the issue to the attention and quantifying the approximate financial and environmental costs. With our approach, this can be taken one step further, showing how adjustments to the net relate to energy and hardware consumption.

In the future, we plan to extend the method in two directions: influence factors along the AI life cycle and further comparison approaches. So far, only the training process of ML

algorithms was considered. We propose to stretch these considerations to a holistic approach along the complete life cycle of AI systems. This starts with data collection, taking sensors, networks, data reduction, etc. into account, which may influence the algorithms and their resource usage. The way the data is pre-processed and managed also needs to be addressed. This includes questions of data quality, storage, the use of synthetic data, data augmentation, the proportion of training, validation, and test data, etc. In the usage and optimization phase, influences from different classification, detection, and synthesis applications should be addressed, as well as possible saving that can be achieved with edge-AI. Furthermore, we plan to investigate the balance between the energy consumption during the training and usage phase. E.g. if a model is used only infrequently or over a short time, maybe it does not need high accuracy levels and thus shorter training, etc.

When modelling the algorithms, we plan to compare optimization algorithms and ML-libraries (TensorFlow, Theano, PyTorch, Keras, OpenCV, Caffe2, scikit, MXNet, etc.) with each other and with minimal implementations. Different ML-algorithms that can perform the same or similar tasks should be compared, as well as, implementations in different programming languages. In the area of deep learning, we plan to evaluate impacts from factors like network topology (e.g. layer structure, layer sizes, dropout layers), activation functions, gradient descent algorithms, performance measures, epoch and batch sizes, etc. For CNNs especially, the influence of factors like kernels, strides, pool sizes, pooling methods, number of pooling layers, number of filters, etc. should be assessed through comparison experiments.

Finally, the relation between energy consumption and ML success indicators needs to be evaluated, depending on the “useful work” of the ML-model for the energy efficiency factor. Here, measures like *precision*, *recall*, or *F-measure* during tests should be taken into account. This would result in e.g. the unit *correct classifications per Wh*.

References

- [Ca21] Canilang, H. M.; Caliwag, A.; Kwon, J.; Lim, W.: DNN Power and Energy Consumption Analysis of Edge AI Devices. In: KICS Winter Conference. 2021.
- [Di06] Dirlwanger, W.: Measurement and Rating of Computer Systems Performance and of Software Efficiency. Kassel Univ. Press, Kassel, 2006.
- [He20] Henderson, P.; Hu, J.; Romoff, J.; Brunskill, E.; Jurafsky, D.; Pineau, J.: Towards the Systematic Reporting of the Energy and Carbon Footprints of Machine Learning. *Journal of Machine Learning Research*, 2020.
- [Jo12] Johann, T.; Dick, M.; Kern, E.; Naumann, S.: How to Measure Energy-Efficiency of Software: Metrics and Measurement Results. In: 1. Intl. Workshop on Green and Sustainable Software (GREENS), Zurich, Switzerland. Pp. 51–54, 2012.
- [JOJ] Jin, D.; Ocone, R.; Jiao, K.: Energy and AI, *Journal*, URL: <https://www.journals.elsevier.com/energy-and-ai/>.

- [KB15] Kingma, D. P.; Ba, J.: Adam: A Method for Stochastic Optimization, 2015.
- [Ke18] Kern, E.; Hilty, L. M.; Guldner, A.; Maksimov, Y. V.; Filler, A.; Gröger, J.; Naumann, S.: Sustainable software products - Towards assessment criteria for resource and energy efficiency. *FGCS 86/*, pp. 199–210, 2018.
- [Kh18] Khakurel, J.; Penzenstadler, B.; Porras, J.; Knutas, A.; Zhang, W.: The Rise of Artificial Intelligence under the Lens of Sustainability. *Technologies 6/4*, 2018.
- [Na11] Naumann, S.; Dick, M.; Kern, E.; Johann, T.: The greensoft model: A reference model for green and sustainable software and its engineering. *Sustainable Computing: Informatics and Systems 1/4*, pp. 294–304, 2011.
- [NGK21] Naumann, S.; Guldner, A.; Kern, E.: The Eco-label Blue Angel for Software - Development and Components. In: *Advances and New Trends in Environmental Informatics. Progress in IS*, Springer, pp. 79–89, 2021.
- [Sc21] Schmidt, V.; Goyal, K.; Joshi, A.; Feld, B.; Conell, L.; Laskaris, N.; Blank, D.; Wilson, J.; Friedler, S.; Luccioni, S.: CodeCarbon, <https://github.com/mlco2/codecarbon>, 2021.
- [SGM19] Strubell, E.; Ganesh, A.; McCallum, A.: Energy and Policy Considerations for Deep Learning in NLP. In: *57th Annual Meeting of the Association for Computational Linguistics*. Pp. 3645–3650, 2019.
- [Vi20] Vinuesa, R.; Azizpour, H.; Leite, I.; Balaam, M.; Dignum, V.; Domisch, S.; Felländer, A.; Langhans, S. D.; Tegmark, M.; Nerini, F. F.: The role of artificial intelligence in achieving the Sustainable Development Goals. *Nature Communications 11/1*, 2020.

Integration von KI-Algorithmen in Umweltinformationssysteme mittels SensorThings API

Philipp Hertweck¹, Hylke van der Schaaf¹, Desiree Hilbring¹, Jonas Weis², Tanja Liesch²,
Matthias Budde³

Abstract: Künstliche Intelligenz oder Maschinelles Lernen wird zunehmend in der Umwelt domäne eingesetzt. Eine offene Frage dabei ist, wie entwickelte Algorithmen flexibel in Umweltinformationssysteme integriert werden können. Dieser Artikel befasst sich mit dieser Frage und untersucht den Einsatz des offenen Standards SensorThings API des Open Geospatial Consortiums (OGC) für die Integration von KI-Algorithmen in eine Geodaten-Infrastruktur. Die entwickelte Methode trägt mittels Container-Technologie dem Einsatz unterschiedlicher Technologien und unabhängigen Entwicklergruppen in verteilten Systemen Rechnung. Entwickelt und Erprobt wird diese Methode im Projekt NiMo 4.0. Hierfür dient als Beispiel ein Prognose-Algorithmus für die räumliche Vorhersage von Nitrat-Daten.

Keywords: Nitratmonitoring; Umweltinformatik; SensorThings API; Künstliche Intelligenz; Algorithmenintegration; Container; Docker

1 Einleitung

In den letzten Jahrzehnten sind in den Bereichen Verwaltung, Forschung, Wirtschaft und für die Öffentlichkeit Umweltinformationssysteme (UIS) mit vielfältigen Arten von Daten für unterschiedliche Zwecke entstanden. Weiterhin ergeben sich mit den Entwicklungen im Forschungsgebiet der Künstlichen Intelligenz (KI) neue Analysemöglichkeiten für die Daten dieser UIS. Eine Integration dieser KI-Algorithmen in etablierte und bestehende UIS bringen Herausforderungen mit sich. Bestehende Umweltinformationssysteme auf Basis von Umweltdatenbanken sind häufig mit proprietären Schnittstellen für spezifische Zwecke ausgerüstet. Eine region- und domänenübergreifende Analyse ist üblicherweise nur unter Beteiligung mehrerer Behörden oder Versorgungsbetriebe möglich. Zusätzliche erschwert wird dies durch die im föderalen System verteilten Zuständigkeiten. Im behördlichen Umfeld stammen teilweise bereits verfügbare offene Schnittstellen meist aus dem Umfeld der Geodateninfrastrukturen und tragen den INSPIRE Richtlinien mit der Umsetzung von Standards des Open Geospatial Consortiums Rechnung.

¹ Fraunhofer IOSB, Fraunhoferstr. 1, 76131 Karlsruhe, Deutschland philipp.hertweck@iosb.fraunhofer.de

² Karlsruher Institut für Technologie (KIT), Institut für Angewandte Geowissenschaften (AGW), Kaiserstr. 12, 76131 Karlsruhe, Deutschland

³ Disy Informationssysteme GmbH, Ludwig-Erhard-Allee 6, 76131 Karlsruhe, Deutschland

Ziel ist daher eine leicht zu implementierende Schnittstelle für den Zugriff auf die notwendigen Quelldaten zu schaffen, um einerseits durch Wiederverwendbarkeit Synergien zu nutzen und andererseits um durch domänen- und zuständigkeitsübergreifende Analysen neue Erkenntnisse zu gewinnen. Darauf aufbauend können KI-Algorithmen in die vorhandenen UIS integriert werden. Sowohl bei der Datenbereitstellung als auch bei der Integration der Analysen soll auf offene Standards zurückgegriffen werden, um die Interoperabilität zu erhöhen.

2 NiMo 4.0

Im Rahmen des Forschungsprojektes Nitrat-Monitoring 4.0 – Intelligente Systeme zur nachhaltigen Reduzierung von Nitrat im Grundwasser (NiMo 4.0) [Li21] wird der Transfer von innovativen Lösungsansätzen von KI-Anwendungen im Grundwasser-Bereich von der universitären und institutionellen Forschung in die industrielle, anwendungsnahe Forschung und Praxis vorangetrieben. Übergeordnetes Ziel ist eine verbesserte Vorhersage von Nitrat im Grundwasser. Um die zu entwickelnden Vorhersagemechanismen und die dafür benötigten verschiedenen Datenquellen effektiv nutzen zu können, werden im Rahmen des Forschungsprojektes Fragestellungen aus IT-Sicht behandelt: z.B. die Bereitstellung und Integration verschiedener Datenquellen. Hierzu zählen im Kontext des NiMo-Projektes seit Jahren aufgebaute Fachinformationssysteme und Umweltdatenbanken für die Gewässerüberwachung der Umweltverwaltungen und Wasserversorger, die aufgebaut wurden um die Wasserrahmenrichtlinie umzusetzen. Eine länderübergreifende Datenauswertung ist bisher nicht möglich. Ein Ziel in NiMo als Leuchtturmprojekt ist es, durch die Nutzung von offenen Standards, die Übertragbarkeit der Lösung von den Pilotregionen des Projektes auf weitere Regionen, sicherzustellen.

Anhand eines Anwendungsbeispiels aus dem NiMo 4.0-Projekt zeigt dieser Beitrag, wie mit Hilfe des leichtgewichtigen OGC Geodatenstandards SensorThings API, Daten aus einem bestehenden UIS bereitgestellt und durch einen KI-Algorithmus analysiert werden können. Dazu wird anhand einer KI-Nitratprognose beispielhaft diskutiert, wie Daten über geeignete Schnittstellen aus bestehenden Gewässergütedatenbanken zugänglich gemacht und Messwerte verschiedener Quellsystemen kombiniert werden können, um damit ein Prognosemodell zu trainieren. Es wird gezeigt, dass auf Basis des zweiten Teils der SensorThings API (Tasking Core), eine flexible Integration von KI-Algorithmen möglich ist. Die Prognose soll im Rahmen des NiMo-Projektes in ein Decision Support System für Entscheidungsträger in Politik und Industrie einfließen.

Um den Anforderungen der verteilten Zuständigkeiten Rechnung zu tragen wird dabei ein Vorgehensmodell für die Integration von KI-Algorithmen angewendet.

3 Herausforderungen bei der Integration von KI-Algorithmen in Umweltinformationssysteme

Bei der Integration von KI-Algorithmen in Umweltinformationssysteme treten einige Herausforderungen auf, die im Folgenden zusammengefasst werden sollen um den Kontext und die Zielsetzung dieser Beiträge zu motivieren.

Die erste Schwierigkeit liegt darin, dass sich der Lebenszyklus von etablierten Umweltinformationssystemen und KI-Anwendungen unterscheidet. Erstere sind üblicherweise über Jahre, bzw. Jahrzehnte umgesetzt und erweitert worden. Weiterentwicklungen folgen klassischen Softwareentwicklungsprozessen. Das Einbringen und Evaluieren neuer Komponenten und Algorithmen ist daher mit einem hohen Aufwand verbunden. Dieses Vorgehen steht im Gegensatz zur agil geprägten Entwicklung von KI-Anwendungen. Aufgrund der Unsicherheit bezüglich Datenlage und Eignung verschiedener Lösungsansätze kann eine KI-Anwendung, insbesondere die eingesetzte KI-Methode nicht vollständig im Vorfeld geplant und umgesetzt werden. Diese kann nur in einem iterativen Proof-of-Concept Prozess, bestehen aus prototypischer Umsetzung und Evaluierung anhand vorhandener Daten (beispielsweise beschreiben in [JP20]) entwickelt werden. Um den produktiven Einsatz eines bestehenden Umweltinformationssystems nicht zu beeinträchtigen, ist eine passende Test- und Evaluationsumgebung zu schaffen.

Für die Entwicklung und Evaluierung von KI-Methoden müssen geeignete Daten zur Verfügung stehen. Hierfür können synthetische Werte generiert werden, besser geeignet sind jedoch reale Daten. Diese stehen in bestehenden Umweltinformationssystemen über teils proprietäre Schnittstellen zur Verfügung. Aufgrund der Unsicherheit ist im Vorfeld häufig nicht bekannt welche Datenquellen benötigt werden. Daher lassen sich verschiedene Datenquellen mit unterschiedlichen Zugriffsprotokollen nur mit erheblichem Aufwand anbinden. Um flexibel den Einfluss verschiedener Quellen evaluieren zu können, müssen diese einfach zugreifbar zur Verfügung stehen. Neben einer einheitlichen Zugriffsschnittstelle wird ein gemeinsames Datenmodell mit definierter Semantik benötigt.

Zusätzlich zu den Herausforderungen bei der Bereitstellung der Daten, muss die Integration und der Betrieb von KI-Algorithmen betrachtet werden. Das beschriebene agile Vorgehen bei der Entwicklung führt dazu, dass diese häufig in Form von Jupiter Notebooks⁴ oder Python-Skripten bereitgestellt werden. Diese, aus Softwaretechnik-Sicht *ad hoc*-Entwicklung führt zu fehlender Wiederverwendbarkeit und Nachvollziehbarkeit. Hat sich eine Proof-of-Concept Implementierung bewährt, wird ein transparentes Vorgehen zur Überführung in den produktiven Betrieb benötigt. Erschwerend kommt hinzu, dass im Bereich der Datenauswertung und bei der Entwicklung von KI-Algorithmen eine Vielzahl unterschiedlicher Programmiersprachen, Frameworks und Bibliotheken eingesetzt werden.

Im Rahmen des NiMo-Projektes werden daher folgende Ziele zur Integration von KI-Algorithmen in Umweltinformationssysteme verfolgt: Erstens, eine einheitliche Bereitstel-

⁴ <https://jupyter.org/>

lung unterschiedlicher Datenquellen. Zweitens eine Plattform zur einfachen Integration von KI-Algorithmen. Diese soll es erlauben bestehende Entwicklungen ohne größere Anpassung integrieren zu können. Ein Anwender soll außerdem den KI-Algorithmus ohne tiefergehendes System- und Administrationswissen zur Ausführung bringen können. Hierzu soll sowohl über die genaue Funktion und technische Umsetzung der KI-Methode, als auch von der Infrastruktur zu Ausführung abstrahiert werden. Die Berechnung soll durch eine einfache Auswahl des KI-Algorithmus und der notwendigen Konfigurationsparameter angestoßen werden können.

4 Related Work

Das Umweltbundesamt hat mit der Studie „Künstliche Intelligenz im Umweltbereich“ Zukunftsperspektiven für den Einsatz von KI in der Umweltdomäne untersucht [TJ19]. Haupttreiber der Entwicklung von KI-Awendungen sind wirtschaftliche Akteure. Es wird herausgestellt, dass intelligente Systeme komplexe Zusammenhänge deuten und Muster erkennen können und damit als flexible Warnsysteme für den Umweltschutz dienen können. Ein Anwendungsbeispiel für die Herausforderung des Einhaltens von Grenzwerten ist die im Anwendungsbeispiel gewählte Nitratprognose. Die Ergebnisse können nicht nur Einflüsse auf die Prozesse der Wasserversorger, sondern auch Einfluss auf politische Entscheidungen haben. Der Artikel „Übertragung des Vorgehensmodells KI auf Umweltinformationssysteme“ beleuchtet ein mögliches Vorgehensmodell um diesen Prozess im behördlichen Umfeld zu unterstützen [HP21]. Die technische Umsetzung bleibt aber offen. Ein open source Ansatz für die Modellierung von Schadstoff wird mit MoRE (Modelling of Regionalized Emissions) verfolgt [Fu17]. MoRE koppelt einen Rechenkern mit einer Datenbank und stellt die Ergebnisse über einen Webserver tabellarisch und in Form von Diagrammen zur Verfügung. Eine offene, standardisierte Schnittstelle zur Datenabfrage stellt das System nicht zur Verfügung. Ein gängiger Standard zur Bereitstellung solcher Daten ist der OGC Sensor Observation Service (SOS) [BSE12]. Im Gegensatz zur SensorThings API [LHK16] setzt der SOS auf Prinzipien einer Service-oriented architecture [SW04]. Dadurch steigt die Komplexität und der Aufwand SOS einzusetzen. Eine offener Standard zu Integration von Auswertungen ist der OGC Web Processing Service (WPS) [Sc07]. Beispielsweise wird dieser von Wiemann et al. zur Modellintegration [Wi18] aufgegriffen. Ähnlich wie der Sensor Observation Service ist auch WPS komplex in der Verwendung und hat daher in diesem Kontext keine breitere Verwendung gefunden.

Für das Einbringen und Ausführen von Verarbeitungsschritten zur Auswertung von heterogenen Datenquellen gibt es insbesondere aus dem Bereich der BigData-Verarbeitung unterschiedliche existierende Plattformen [Ma17]. Hierzu zählt beispielsweise Apache Hadoop⁵, das auf MapReduce [DG08] als Programmiermodell setzt. Als Alternative existiert mit Spark ein weiteres Werkzeug für die Verarbeitung großer Datenmengen [Za16]. Im Gegensatz zu Hadoop, das auf MapReduce ausgerichtet ist, unterstützt Spark neben der

⁵ <https://hadoop.apache.org/>

Stream- und Graphverarbeitung auch SQL-Abfragen sowie Machine Learning Modelle. Für die Umsetzung von Verarbeitungsschritten stehen sowohl bei Hadoop als auch bei Spark geeignete Schnittstellen zur Integration bereit. Dadurch sind sowohl das Programmiermodell als auch die unterstützten Programmiersprachen vorgegeben. Eine freie Technologiewahl ist nicht möglich. Daher lassen sich bestehender Implementierungen häufig nur mit großem Aufwand integrieren.

Speziell für den Einsatz von KI-Methoden steht mit Kubeflow [Bi19] eine Plattform für das Entwickeln, Trainieren und Ausführen von Machine Learning Anwendungen zur Verfügung. Kubeflow unterstützt KI-Algorithmen von der Entwicklung und dem Trainieren bis hin zur skalierbaren Ausführung. Dadurch ist Kubeflow ein mächtiges Werkzeug für den kompletten Lebenszyklus einer KI-Anwendung. Dies setzt jedoch eine entsprechende Einarbeitung und Aufwand bei der Integration bestehender Entwicklungen voraus.

5 Integration über SensorThings API

Grundlage für die Daten- und KI-Integration in diesem Beitrag bildet der offene und leichtgewichtige Geodatenstandard *OGC SensorThings API*. Dabei wird für die Bereitstellung der messreihenbasierten Daten auf den ersten Teil (*Part 1: Sensing*) des Standards gesetzt. Zur Einbindung der KI-Algorithmen, deren Konfiguration und Ausführung wird auf Teil zwei (*Part 2: Tasking Core*) zurückgegriffen.

5.1 SensorThings API

Die OGC SensorThings API ist ein Standard zur Verwaltung von Sensormess und -metadaten. Im Gegensatz zu dem bereits genannten Sensor Observation Service baut die SensorThings API auf moderne Konzepte und Technologien auf. Hierzu zählt eine JSON-basierte Datenrepräsentation, eine REST-Schnittstelle [Fi00] zur Datenabfrage und Manipulation sowie eine auf OData [PHZ14] aufbauende Abfrage- und Filtersprache. Dies ermöglicht eine einfache und technologieunabhängige Bereitstellung von Sensordaten. Dadurch hat die Verbreitung der SensorThings API für den Zugriff auf Sensordaten in den letzten Jahren zunehmen an Bedeutung gewonnen.

Die OGC SensorThing API besteht aus den folgenden Komponenten: 1. einem Datenmodell das vorhandene Entitäten und deren Beziehungen definiert 2. einer HTTP basierten Schnittstelle, die das Abfragen, Bearbeiten und Löschen von Daten erlaubt 3. mehreren Erweiterungen, wie beispielsweise eine MQTT-Schnittstelle die die Umsetzung ereignisbasierter Anwendungsfälle erlaubt.

Neben dem *OGC SensorThings API Part 1: Sensing*, existiert ein weiterer Standard, die *SensorThings API Part 2: Tasking Core* [LK19]. Dieser Teil des Standards verfügt über die interessante Möglichkeit mittels der leichtgewichtigen Schnittstelle nicht nur Daten

abzufragen, sondern auch Berechnungsprozesse anzustoßen. Das Datenmodell des ersten Teils wird um weitere Entitäten (Abbildung 1) ergänzt. Dadurch können mit Hilfe der SensorThings API nicht nur Sensor Mess- und Metadaten, sondern auch Aktuatoren modelliert werden. Dies bietet die Grundlage für die in diesem Beitrag betrachtete KI-Integration.

Ein *Actuator* ist häufig ein physikalisches Gerät, das gesteuert werden kann. Die Fähigkeiten des Aktors werden als *TaskingCapability* dargestellt. Diese wiederum sind einem *Thing* zugeordnet, was den Übergang in das Datenmodell des ersten Teils darstellt. Während eine *TaskingCapability* die Fähigkeiten beschreibt, werden Aktionen des Aktors mittels eines *Tasks* umgesetzt. Akzeptierte Konfigurationsparameter einer *TaskingCapability* werden über das Attribut *taskingParameters* definiert. Die konkreten Werte zur Ausführung werden im *Task* über das gleichnamige Attribut abgelegt. Betrachtet man einen Fensteröffner als *Thing*, so besitzt dieser einen Motor (*Actuator*), der das Fenster zu einem gewissen Prozentsatz öffnen kann (*TaskingCapability*). Soll das Fenster nun zu 50% geöffnet werden, wird eine *Task*-Instanz erzeugt, die als *taskingParameter* einen Wert *öffnung* = 50% enthält. An dieser Stelle sei angemerkt, dass der Standard keine Aussage über die Darstellung des Ausführungszustands macht.

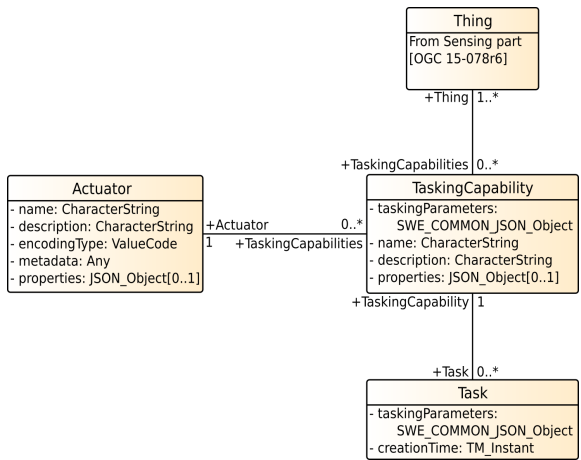


Abb. 1: Datenmodell der OGC SensorThings API Part 2: Tasking Core

Ein Aktor muss nicht zwingend ein physikalisches Objekt repräsentieren. Ebenfalls können reine Softwarekomponenten (im Kontext dieses Artikels *virtueller Aktor* genannt) betrachtet werden. Aufrufe an die Softwarekomponente können somit als *Task* und Aufrufparameter als *taskingParameter* dargestellt werden. Betrachtet man mit diesem Hintergrund einen KI-Algorithmus, so kann dieser als *virtueller Aktor* aufgefasst und das Anstoßen einer Berechnung als *Task* umgesetzt werden.

5.2 Bereitstellung der Daten über die SensorThings API

Mit Hilfe der SensorThings API Part 1 wird im Projekt NiMo die Bereitstellung von dynamischen Zeitreihendaten realisiert. Für die Umsetzung des Anwendungsbeispiels der Nitratprognose sind Gütemesswerte von Gewässern notwendig. Diese stehen in einer internen Umweltdatenbank zur Verfügung. In einem ersten Schritt wurde das vorhandene Datenmodell auf die SensorThings API abgebildet. Ein *Thing* repräsentiert eine Messstelle aus einer Datenbank mit Grundwasserqualitätsdaten. Die Koordinaten des Messortes wurden auf die *Location* abgebildet. Die Kombination des gemessenen Parameters (z.B. Nitrat) mit dem Medium in dem dieser gemessen wurde (z.B. im Anwendungsfall Wasser) entspricht der *ObservedProperty*. Auf den *Datastream* wurden die allgemeinen Attribute einer Zeitreihe eines Messortes abgebildet. Für die Ermittlung der Messwerte werden an klassischen Messstellen Wasserproben genommen, welche im Labor analysiert werden. Die Informationen über die Probenahme wurden im *FeatureOfInterest* abgebildet. Die Messwerte werden durch die *Observations* repräsentiert.

Da eine Anpassung am Quellsystem nicht möglich ist, um dies beispielsweise um eine SensorThings API Schnittstelle zu erweitern, wurden die vorhandenen Daten in eine Instanz von FROST^{®6} kopiert. Dieser ist eine Open Source Referenzimplementierung der SensorThings API und bildet im Rahmen von NiMo die Grundlage zur Daten und Algorithmenintegration.

5.3 PERMA: Integration von Verarbeitungsschritten

Im Abschnitt 1 wurde aufgezeigt, dass sich die SensorThings API Part 2: Tasking Core neben der Modellierung von physikalischen Aktuatoren auch dazu eignet Metadaten von Verarbeitungsschritten, bzw. KI-Algorithmen zu modellieren, deren Ausführung durch *Tasks* darzustellen und die Ausführungsparameter über *TaskingParameters*

In der praktischen Umsetzung ergeben sich an dieser Stelle, wie bereits in Abschnitt 3 motiviert, weitere Fragestellungen. *Virtuelle Aktoren*, bzw. KI-Algorithmen benötigen ein System oder eine Laufzeitumgebung, auf der sie ausgeführt werden können. Da KI-Algorithmen häufig von Datenwissenschaftlern oder Fachexperten umgesetzt werden, besitzen diese in der Regel nicht das nötige Wissen um die Implementierung (bspw. per Secure Shell (ssh)) auf einem Server zur Ausführung zu bringen.

Um dieser Herausforderung zu begegnen, wird im Rahmen von NiMo eine Software-Komponente namens PERMA entwickelt und eingesetzt. Mit Hilfe von PERMA können KI-Algorithmen ohne tieferes Wissen über die eigentliche Infrastruktur, bzw. Laufzeitumgebung zur Ausführung gebracht werden. Im Rahmen des Projektes werden hierfür zwei Technologien unterstützt. Erstens können KI-Algorithmen in der Programmiersprache Java

⁶ <https://github.com/FraunhoferIOSB/FROST-Server>

umgesetzt werden. Dazu muss lediglich ein einfaches Interface mit zwei Methoden (Abfrage der bereitgestellten *TaskingCapability* und Ausführen der Berechnung) implementiert werden. Anschließend kann dies als JAR-Artefakt gepackt, in PERMA hochgeladen und zur Ausführung gebracht werden.

Je nach Anwendungsfall werden KI-Algorithmen mit unterschiedlichen Technologien umgesetzt, wofür eine geeignete Abstraktionsebene notwendig ist. Hierfür setzt PERMA auf Container-Technologien. Dies erlauben es die KI-Algorithmus zusammen mit abhängigen Bibliotheken, Frameworks und Tools in ein austauschbares und ausführbares Artefakt zu packen. Neben einer Java-Laufzeitumgebung unterstützt daher PERMA als zweite Technologie den Einsatz von Containern. Dies ermöglicht es KI-Algorithmus in nahezu jeder Programmiersprache zu implementieren und zu integrieren.

Klassische Container-Plattformen wie Kubernetes oder openShift⁷ sind aufgrund ihrer vielfältigen Funktionen mächtig, setzen jedoch einen großen Einarbeitungsaufwand voraus. Ähnliches gilt für Machine Learning Frameworks wie Kubeflow. PERMA bietet hier eine zusätzliche Abstraktionsebene um diese Komplexität zu verstecken. Entwickler der KI-Algorithmen müssen lediglich ein Container Image erstellen, das Eingabeparameter als *TaskingParameter* definiert und die Werte aus den Umgebungsvariablen auslesen kann. Für das Ausführen der Berechnung kann, ohne weitere Kenntnisse, über die PERMA Weboberfläche angestoßen werden. Durch das Anlegen eines *Tasks* kann dies ebenfalls durch ein externes System erfolgen. Zur Ausführung der Container setzt PERMA im Hintergrund auf Kubernetes. Eine Erweiterung, dass lediglich Docker⁸ zur Verfügung steht, ist in Planung.

6 Anwendungsbeispiel Nitratmonitoring

Im Rahmen des Forschungsprojektes NiMo wird die Integration eines KI-Algorithmus mit Hilfe der SensorThings API anhand einer flächenhaften Prognose von Nitratdaten demonstriert. Hierzu wird das *Vorgehensmodell Machine Learning for Production (MLAP)* [JP20] aus dem Kontext der industriellen Produktion auf ein Umweltinformationssystem übertragen [HP21]. Die einzelnen Phasen des methodischen Vorgehens sind: 1. *Zieldefinition und Lösungsansatz*, 2. *Proof of Concept*, 3. *Systemspezifikation*, 4. *Umsetzung und Inbetriebnahme*, 5. *Übergabe* und 6. *Betrieb*.

Im ersten Schritt (*Zieldefinition und Lösungsansatz*) wurde im Rahmen des Projektes ein Anwendungsfall definiert, mit dem die Integration eines KI-Algorithmus in ein Umweltinformationssystem demonstriert werden kann. Hierfür wurde die flächenhafte Prognose von Nitrat im Grundwasser mit Hilfe von an Messstellen gemessenen Nitratwerten ausgewählt.

⁷ <https://www.openshift.com/>

⁸ <https://www.docker.com/>

Im Rahmen der *Proof of Concept* Phase wurde ein Lösungsansatz prototypisch umgesetzt und die Machbarkeit überprüft. Die in die Berechnung einfließenden Datenquellen standen zu diesem Zeitpunkt noch nicht endgültig fest. Daher wurden potentiell relevante Datenquellen bereitgestellt. Dies sind einerseits räumliche Daten, bzw. Rasterdaten, welche statisch sind und sich nicht regelmäßig ändern, beispielsweise Informationen über Flächennutzung, Stickstoffüberschuss und -hintergrunddeposition, Bodenbeschaffenheit, Sickerwasserrate und Wasserschutzgebiete. Andererseits wurden dynamische Zeitreihendaten beispielsweise die der Nitratmessstellen aus öffentlich verfügbaren Grundwasserqualitätsdaten betrachtet. Eine direkte Integration dieser Daten für das Vorhersagemodell ist aufgrund der uneinheitlichen und nicht standardisierten Schnittstelle nicht praktikabel. Aus diesem Grund wurden, wie in Abschnitt 5.2 beschrieben, Gewässerqualitätsdaten aus vorhandenen Quellsystemen auf das Datenmodell der SensorThings API abgebildet und in eine FROST-Instanz importiert. Damit ergeben sich einerseits Synergieeffekte für den Datenquellbezug bei der Entwicklung weitere KI-Algorithmen mit ähnlichen Daten und die Möglichkeit die verfügbaren Daten länderübergreifend in einem System zu nutzen.

Bei der prototypischen KI-Methode handelt es sich um einen Algorithmus zur räumlichen Vorhersage der Nitratkonzentration im Grundwasser. Die Grundlage stellt ein Random Forest dar, der in der Programmiersprache Python geschrieben wurde und ein Regressionsproblem löst. Für dieses Problem stehen dem Algorithmus die Nitratwerte der Grundwassermessstellen als Zieldaten sowie die genannten Rasterdaten als Eingangsdaten zur Verfügung. In einem ersten Schritt werden aus dem Gesamtdatensatz Trainings-Daten erzeugt. Dies geschieht durch die räumliche Abfrage der Rasterwerte an jeder der Nitratmessstellen. So entsteht eine 2D-Matrix, in der jede Zeile eine Nitratkonzentration sowie die zugehörigen Eingangsdaten enthält. Zusätzlich zu dieser Trainings-Matrix wird noch eine 3D-Matrix der Größe [Länge*Breite*Anzahl Eingangsdaten] erstellt, die als Grundlage der räumlichen Vorhersage dient. Anschließend wird der Random Forest mit vorher optimierten Hyperparametern erstellt und mit den Trainings-Daten trainiert, um die Zusammenhänge zwischen Eingangsdaten und Nitratkonzentrationen zu lernen. Nachdem das Modell trainiert wurde, erfolgt die Vorhersage der Nitratkonzentration über die Fläche. Hierfür iteriert der Algorithmus über die 3D-Matrix und berechnet an jeder Stelle einen Wert, sodass als Ergebnis eine Matrix mit räumlich vorhergesagten Nitratkonzentrationen entsteht. Fehlen an einer Stelle Informationen zu Eingangsdaten, wird an dieser Stelle kein Wert vorhergesagt, sondern diese Zelle übersprungen. Aus der berechneten Matrix mit den räumlich vorhergesagten Nitratkonzentrationen wird dann wieder eine Raster-Datei mit räumlichem Bezug erzeugt.

Die Entwicklung des KI-Algorithmus wurde in der Programmiersprache Python umgesetzt und vorhandene Entwicklungswerkzeuge eingesetzt. Die Funktionsfähigkeit des *Proof of Concepts* konnte somit ohne zusätzlichen Einarbeitungsaufwand gezeigt werden.

Im Rahmen der *Systemspezifikation* wurden die genannten Herausforderungen bei der Überführung der prototypischen Implementierung vom Entwicklersystem in eine produktive Laufzeitumgebung betrachtet, damit die Prognose weiteren Anwendern zur Verfügung gestellt werden kann. Die Grundstruktur der Systemspezifikation ist durch den Einsatz von

PERMA, bzw. der SensorThings API bereits vorgegeben. Sodass in diesem Schritt lediglich der Prognose-Algorithmus und dessen Parameter [SL18] durch eine *TaskingCapability* beschrieben werden mussten.

Zur *Umsetzung und Inbetriebnahme* wurde für die vorhandene Implementierung, zusammen mit der *TaskingCapability* aus der vorangegangenen Phase ein Container-Image erzeugt. Dies ist beispielsweise mit Hilfe von Docker⁹ ohne großen Aufwand möglich. Damit sind die Voraussetzungen erfüllt, um den KI-Algorithmus mit Hilfe der in Abschnitt 5.3 beschriebenen Komponente PERMA zur Ausführung zu bringen. In einem ersten Schritt muss dazu die Nitratprognose, unter Angabe des Containerimages, in PERMA registriert werden. Danach kann eine Ausführung, ohne größere Kenntnisse und technisches Wissen, über die PERMA-Weboberfläche angestoßen werden. Die abschließenden Phasen *Übergabe* und *Betrieb* wurden im Projekt noch nicht durchlaufen. PERMA selbst setzt auf Kubernetes¹⁰ als Containerlaufzeitumgebung. Sodass in diesen Phasen eine entsprechende Umgebung in der Zielinfrastruktur bereitgestellt werden muss.

Basis für die Datenbereitstellung und KI-Integration ist die SensorThings API. Aufgrund dieser standardisierten und offenen Schnittstelle kann dies problemlos in einen größeren Systemkontext integriert werden. Im Rahmen des NiMo-Projektes soll die Vorhersage der Nitratbelastung beispielsweise in ein Entscheidungsunterstützungssystem einfließen.

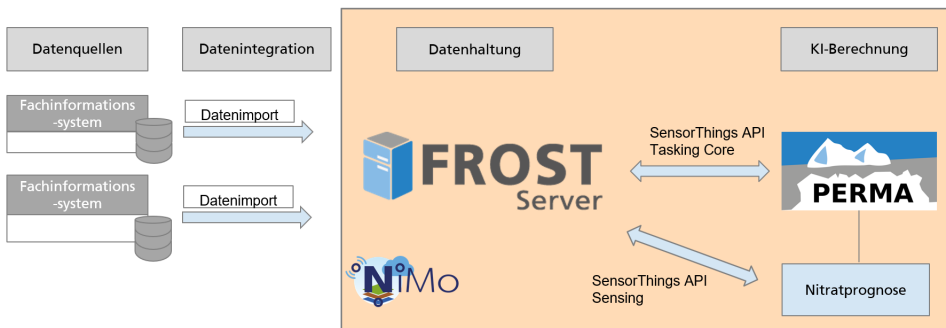


Abb. 2: NiMo Architektur

Die aus dem ersten Anwendungsfall abgeleitete NiMo-Architektur ist in Abbildung 2 verfügbar. Es hat sich gezeigt, dass alle bisher betrachteten Datenquellen mit Zeitreihendaten problemlos und ohne größeren Aufwand in das SensorThings API Datenmodell überführt werden können. Aufgrund der unterschiedlichen Eigenschaften und Datenrepräsentationen ist für jedes Quellsystem ein individueller Datenimport notwendig. Aktuell laufen weitere Arbeiten um diesen Schritt durch geeignete Import-Werkzeuge zu vereinfachen. Durch die standardisierte Bereitstellung der Daten konnte die Entwicklung der Nitratprognose deutlich vereinfacht werden. Zusätzlich konnte in der *Proof of Concept*-Phase mit geringem Aufwand der Einfluss unterschiedlicher Eingabedaten auf das

⁹ <https://www.docker.com>

¹⁰ <https://kubernetes.io/>

Vorhersagemodell evaluiert werden. Im Vergleich zur Datenbereitstellung über das lokale Dateisystem führt der Datenzugriff über die SensorThings API, bzw. das HTTP-Protokoll zu Performanz-Einschränkungen. In der Praxis wurden jedoch keine erheblichen Einbußen beobachtet, die dieses Zugriffsmuster unpraktikabel oder ein Zwischenspeichern der Daten erforderlich machen. Vielmehr wurde dadurch die Flexibilität bei der Entwicklung der Prognose erhöht. Durch den Einsatz von PERMA ist es möglich die Nitratprognose ohne großen Aufwand vom Entwicklungssystem in eine produktive Umgebung zu überführen und somit anderen Beteiligten schnell zur Verfügung zu stellen.

7 Zusammenfassung und Ausblick

In diesem Artikel wurde diskutiert, wie KI-Algorithmen über standardisierte offene Schnittstellen in bestehende Systeme integriert werden können. Beschrieben wurde ein Ansatz der dies mittels der SensorThings API umsetzt. Anhand des praktischen Anwendungsfalls zeigt sich, dass die SensorThings API mit ihrem zeitreihenbasierten Datenmodell für Integration der Gewässerqualitätsdaten gut geeignet ist. Für die Realisierung des endgültigen Systems werden außerdem Rasterdaten benötigt. Wie diese über weitere offene Schnittstellen behandelt werden können wird im weiteren Verlauf des Projektes NiMo untersucht. Durch den Einsatz von PERMA und der SensorThings API: Part 2 konnte die implementierte Nitratprognose erfolgreich und ohne großen Mehraufwand in einer Zielumgebung zur Ausführung gebracht werden. Dadurch kann eine vom Umweltinformationssystem unabhängige Entwicklung der KI-Methoden gewährleistet und eine Erprobung dieser ermöglicht werden. Im weiteren Verlauf des Projektes wird dieses Vorgehen auf weitere Anwendungsbereiche angewendet und anhand von Performance-Tests der Einfluss des Zugriffs auf die SensorThings API weiter untersucht.

8 Acknowledgement

Diese Arbeit wurde im Rahmen des vom BMU geförderten Projekts „NiMo 4.0“ [Li21] (Förderkennzeichen 67KI2048) erstellt.

Literaturverzeichnis

- [Bi19] Bisong, Ekaba: Kubeflow and kubeflow pipelines. In: Building Machine Learning and Deep Learning Models on Google Cloud Platform, S. 671–685. Springer, 2019.
- [BSE12] Bröring, Arne; Stasch, Christoph; Echterhoff, Johannes: OGC Sensor Observation Service Interface Standard, Version 2.0. 2012.
- [DG08] Dean, Jeffrey; Ghemawat, Sanjay: MapReduce: simplified data processing on large clusters. Communications of the ACM, 51(1):107–113, 2008.

- [Fi00] Fielding, Roy Thomas: Architectural styles and the design of network-based software architectures. University of California, Irvine, 2000.
- [Fu17] Fuchs, Stephan; Weber, Tatyana; Wander, Ramona; Toshovski, Snezhina; Kittlaus, Steffen; Reid, Lucas; Bach, M; Klement, L; Hillenbrand, T; Tettenborn, F: Effizienz von Maßnahmen zur Reduktion von Stoffeinträgen. Endbericht. UBA-Texte, 5:2017, 2017.
- [HP21] Hilbring, Désirée; Pfrommer, Julius: Übertragung eines Vorgehensmodells zur KI-Integration von der Industrie auf Umweltinformationssysteme. INFORMATIK 2020, 2021.
- [JP20] Julius Pfrommer, Jürgen Beyerer, Marlene Eisenträger Christian Frey Andreas Herzog Ali Moghiseh Lukas Morand Henrike Stephani Anke Stoll Lars Wessels: ML4P – Vorgehensmodell Machine Learning for Production. White paper, Fraunhofer, 10 2020.
- [LHK16] Liang, Steve; Huang, Chih-Yuan; Khalafbeigi, Tania: OGC SensorThings API Part 1: Sensing, Version 1.0. 2016.
- [Li21] Liesch, Tanja; Bruns, Julian; Abecker, Andreas; Hilbring, Désirée; Karimanzira, Divas; Martin, Tobias; Wagner, Martin; Wunsch, Andreas; Fischer, Thilo: Nitrat-Monitoring 4.0 – Intelligente Systeme zur nachhaltigen Reduzierung von Nitrat im Grundwasser. In (Reussner, Ralf H.; Koziolk, Anne; Heinrich, Robert, Hrsng.): INFORMATIK 2020. Gesellschaft für Informatik, Bonn, S. 1069–1079, 2021.
- [LK19] Liang, Steve; Khalafbeigi, Tania: OGC SensorThings API Part 2–Tasking Core, Version 1.0. 2019.
- [Ma17] Marjani, Mohsen; Nasaruddin, Fariza; Gani, Abdullah; Karim, Ahmad; Hashem, Ibrahim Abaker Targio; Siddiqa, Aisha; Yaqoob, Ibrar: Big IoT data analytics: architecture, opportunities, and open research challenges. *ieee access*, 5:5247–5261, 2017.
- [PHZ14] Pizzo, Michael; Handl, Ralf; Zurmuehl, Martin: OData Version 4.0 Part 1: Protocol Plus Errata 02. 2014.
- [Sc07] Schut, Peter: OpenGIS® Web Processing Service. 2007.
- [SL18] Steve Liang, Tania Khalafbeigi, Kan Luo: OGC SensorThings API Tasking CoreDiscussion Paper. Discussion paper, OGC, 2018.
- [SW04] Sprott, David; Wilkes, Lawrence: Understanding service-oriented architecture. *The Architecture Journal*, 1(1):10–17, 2004.
- [TJ19] Tobias Jetzke, Stephan Richter, Jan-Peter Ferdinand und Samer Schaat: , Künstliche Intelligenz im Umweltbereich. https://www.umweltbundesamt.de/sites/default/files/medien/1410/publikationen/2019-06-04_texte_56-2019_uba_ki_fin.pdf, 2019.
- [Wi18] Wiemann, Stefan; Al Janabi, Firas; Eltner, Anette; Krüger, Robert; Luong, T; Sardemann, Hannes; Singer, Thomas; Spieler, Diana; Kronenberg, Rico: Entwicklung eines Informationssystems zur Analyse und Vorhersage hydro-meteorologischer Extremereignisse in mittleren und kleinen Einzugsgebieten. In: *Forum Hydrol. Wasserbewirtsch.* Jgg. 39, S. 357–367, 2018.
- [Za16] Zaharia, Matei; Xin, Reynold S; Wendell, Patrick; Das, Tathagata; Armbrust, Michael; Dave, Ankur; Meng, Xiangrui; Rosen, Josh; Venkataraman, Shivaram; Franklin, Michael J et al.: Apache spark: a unified engine for big data processing. *Communications of the ACM*, 59(11):56–65, 2016.

Paving the road to a circular textile economy with AI

Katharina Rudisch,¹ Sebastian Jüngling,² Ricardo Carrillo Mendoza,³ Ulrike K. Woggon,¹
Ina Budde,² Mario Malzacher,² Karsten Pufahl¹

Abstract: The textile industry has a detrimental impact on the environment and faces challenges to meet climate and environmental goals. The core problem is its linear system where only a small fraction of used garments is reused or recycled. To facilitate recycling of textiles, a sorting method is proposed that provides the necessary information about a garment to sort it into the appropriate recycling channel. Current methods have difficulties identifying garments polluted by chemicals used during dyeing and finishing which hinder recycling into new high-quality fibers. In this concept paper, the Circular Textile Intelligence (CRTX) project is outlined which addresses the central problem of linear structures in the textile industry. A spectroscopy and AI-based method for textile sorting will enable more informed decisions about optimal reuse and recycling channels, thus closing the loop. It will build upon a material database containing information about materials, textiles, dyes, finishings, and unwanted hazardous substances.

Keywords: Circular Economy; Textile Recycling; Chemometrics; Spectroscopy; Artificial Intelligence; Neural Networks

1 Introduction and background

The textile industry has a major negative impact on the environment, due its incapacity of reusing raw materials, as well as its extensive use of pollutants. The total greenhouse gas emissions from textile production currently exceed those of all international flights and shipping combined, at 1.2 billion tons of CO₂ eq. per year.[14, 5]

The textile industry's central problem is its primarily linear production pattern. Currently, over 100 billion garments are produced annually worldwide, predominantly from virgin raw materials (> 97%). At the end of their use cycle, these are generally disposed of in landfills or incinerated (87%) or mechanically recycled to a lower value (downcycling) (12%).[2] Less than 1% is recycled using innovative recycling processes so that it can be used again as a raw material for new garments.[2] In this context, fiber-to-fiber recycling would save the extraction and production of virgin fibers and prevent the waste of resources

¹ Technical University of Berlin, Institute for Optics and Atomic Physics, Straße des 17. Juni 135, 10623 Berlin, Germany

² circular.fashion UG, Skalitzer Str. 97, 10997 Berlin, Germany

³ Dahlem Center for Machine Learning and Robotics – Freie Universität Berlin, Arnimallee 7, 14195 Berlin, Germany

through landfill and incineration. Fiber regeneration would have been technically possible for several years through chemical recycling processes[13, 3]. However, it requires very precise information about the composition of the garments to be recycled, which cannot be determined by current sorting methods. Fiber-to-fiber recyclers have very specific ban lists on certain chemicals, fibers and material blends that must not be contained in the feedstock.[1]

Contamination and pollutants pose a serious problem for existing fiber-to-fiber recycling processes. For example, coatings based on fluorocarbons, such as those found in impregnated outdoor materials, pose a major problem for chemical recycling processes – even in very low concentration.[1] Also, a separation or exclusion of certain dyes is required by some recyclers. This limits the range of suitable feedstock to either well characterized production waste (pre-consumer) or easily identifiable garments with known characteristics (e.g. indigo dyed blue jeans).

According to the German Federal Environment Agency (Umweltbundesamt), up to one kilogram of harmful chemicals is used per kilogram of processed textiles during textile finishing.[16] A large proportion of the textile auxiliaries subsequently end up in wastewater[4], having detrimental effects on the environment. As stated by Greenpeace e.V., the textile industry is responsible for approximately 20% of the discharge of organic pollutants into water bodies. Some of these substances are difficult to degrade and can only be reduced to a limited extent in biological wastewater treatment plants.

Conventional sorting plants focus on the extraction of reusable garments to be sold to the second-hand market. A small portion of the remaining textiles undergoes material specific sorting for textile recycling, which is predominantly conducted through haptic recognition. Better results could be enabled by optical detection with near infrared (NIR) spectroscopy, which is currently tested in pilot sorting plants.[18, 3] These sorting plants focus on the detection of cotton, polyester and wool and do not identify possible contaminants. Here, the vision for a project is presented, which aims to reliably detect the material composition and contamination of used textiles to enable fiber-to-fiber recycling. A multispectral approach will be combined with Artificial Intelligence (AI) by using convolutional neural networks (CNN) for the material classification. Used clothing contaminated with harmful substances could be sorted out and sent for hazardous waste recycling, thus preventing contamination in future garments. Furthermore, a contaminant analysis in the recycling process provides a concrete and reliable source on consumer exposure to contaminants. Guidelines and decision-making principles for environmental protection could be derived from the analyses.

2 Goal of the Project

The overarching goal is to close the missing gap that exists in textile recycling, thereby enabling a more sustainable textile industry. In order to capture all facets of this diverse challenge, the project focuses on three parts: (i) optimizing the reuse of garments (small

cycle) by classifying the type of second hand garments via computer vision, (ii) enabling textile fiber recycling (large cycle) and (iii) identifying hazardous substances in textiles. Achieving these three aspects corresponds to several sustainability dimensions and would have significant impact at the social and economic level. The content of the remaining paper focuses on parts (ii) and (iii), because an extensive description and discussion of part (i) would exceed the scope of this report.

The aim of part (ii) of the project is to demonstrate how a spectroscopy-based automated detection that is backed by an AI-classifier can identify raw materials and pollutants to enable a sustainable circular economy for textiles. In order for sorting to be optimized, relevant information from the garments will be identified through the development of a combined approach merging spectroscopic analysis and an appropriate AI-based classification.

The project continues to address the issue of pollutants in textiles (subproject (iii)), both in terms of recycling requirements and environmental targets. On one hand, a mapping and market overview of commonly used chemicals and their influence on recyclability will be developed within the scope of this subproject. On the other hand, an algorithm will be developed that compares the chemical list of textiles with recycling specifications and environmental guidelines. The results of this working part will then be made accessible to the spectroscopy AI as new features. Considering current environmental targets and including possible alternatives for harmful substances, project results are envisioned to lead to changes in the product design as actions for manufacturers. In the long term, this approach should be able to identify and decrease the use of pollutants from the textile industry's material flows.

3 Current advances in AI and spectroscopy for textile analysis

Researchers can choose from a wide variety of spectroscopic methods to investigate the characteristics of a textile sample. Near infrared (NIR) spectroscopy is able to distinguish different fiber materials of textiles and is indeed suitable for analysis of a large sample set. Automated sorting of used clothing based on material composition has been tested using visible and near infrared spectroscopy (VIS-NIR spectroscopy) and has succeeded in separating textiles by color and material.[17, 18, 3] However, there are certain requirements of high-quality recycling processes that this technology cannot meet.[20] One shortcoming of the mentioned sorting technologies is their inability to analyze other characteristics, especially finishings and contaminants.

Other methods, such as nuclear magnetic resonance spectroscopy (NMR), chromatography and mass spectrometry, can deliver exhaustive and reliable information about textiles samples but are not suitable for use on a sorting line due to very high efforts in terms of sample preparation and experimental setup.[12]

Another method that is used to analyze textile samples in criminal forensics on a laboratory scale is Raman spectroscopy in the VIS and NIR range.[15, 7] Raman spectra yield

information about the fiber type as well as dyes and other finishings. Publications show that the spectra provide sufficient information for optimized recycling.[12, 6] However, the common interference of fluorescence with the Raman signal is an obstacle that needs to be tackled in the future.[12] Raman spectroscopy is advantageous as its signals have a high degree of information, which allows very accurate differentiation of materials. Compared to NIR spectroscopy, materials have a significantly lower cross-entropy due to the distinctive and discrete spectral peaks in the signal.

Machine learning methods are becoming more and more popular for the analysis of Raman spectra. Liu et al. [8] highlighted the use of convolutional neural networks (CNN) for the purpose of material identification in Raman spectra. The approach shows that CNNs, unlike traditional chemometric methods (e. g. principal component analysis (PCA)), can effectively identify substances without the need for complex preprocessing, even with fluorescence background from the spectra. Several recent studies report on the advantages of using a CNN for analyzing Raman spectra compared to other methods such as partial least square regression [21] and conventional Raman processing methods [19].

In relation to the project, the work of J. Luo et al.[9] should be highlighted, which aimed to analyze cotton purity combining NIR Raman spectroscopy and AI analysis. By using a simple random forest model, a decision tree was generated that can detect the contamination of cotton with polyester and thus verify the manufacturer's claims of the fiber composition. The focus of the study is on quality assurance of cotton textiles, some of which are stretched with synthetic fibers due to cost pressure.

4 A spectroscopy- and AI-based sorting method for fiber-to-fiber recycling

By combining multi-spectral analysis tools, a reliable optical detection method for identifying the material composition and contamination of used textiles is to be created. The combination of several spectral ranges will enable sorting in compliance with the requirements of fiber-to-fiber recycling. NIR spectroscopy provides broad and continuous spectra that allow for coarse presorting based on the fiber material. On the other hand, Raman spectra in the visible and UV range consist of distinct and discrete lines that are more difficult to interpret but allow for much finer material specific classifications. Raman spectroscopy can therefore produce a molecular fingerprint of the textile and can provide information about the occurrence of substances that have been explicitly excluded from the list of substances permitted for the recycling process.

Within the scope of this project, a neural network based classifier will be developed to classify a large amount of different materials. This approach is believed to be superior to classical chemometric analysis methods such as principal component analysis (PCA). In particular, the vast amount of possible substances can increase the dimensionality of PCA. This leads to a large number of configurations where those with low eigenvalues can be

missed even if they are essential to represent the original data. Instead, neural networks have already been successfully applied outside of the textile sector for the interpretation of Raman spectra and have proven their worth as an analysis tool for chemical fingerprinting. In particular, the use of different data sources is important to gain maximum reliability. In addition to the multi-spectral data, the data set will be enriched by garment type information from subproject (i) determined via computer vision. Based on this extensive and diversified data set, the AI will classify the material composition and estimate the pollutant load in order to make decisions regarding the appropriate recycling channel.

In contrast to conventional spectroscopic analysis routines, it is not the quality of the fit that is crucial here, but the certainty of unique classified features. For example, recyclers have different requirements for the accuracy of component determination. Some recyclers allow mass fractions of elastane in the single-digit percent range in their processes, but must be able to exclude other contaminants in the range of a few ppm. A conventional method requires very high accuracy over the entire spectrum to solve the task and achieve sufficient convergence of the fit to make a confident conclusion. A deep neural network can be used to consider the existing requirements of the recycler. Hence, the approach considers the expectations of the results already for the analysis. It integrates the classically separated processes of data analysis and result interpretation into a holistic spectroscopy. In addition, mutually exclusive components such as material-specific staining can be taught. Thus, the strong correlations and anti-correlations of features are taken into account. Furthermore, exclusionary substances such as fiber-specific dyes and finishings serve to resolve degeneracies of the spectral components and make the method more robust.

The certainty of the estimation takes an integral part and decides whether a textile has been classified with sufficient accuracy. Here, the limits defined by the recycling methods do not have to be met per individual garment, but only per sorting unit (e.g. container or bale). In this way, a stochastic approach can be used to achieve a maximum recycling rate with optimized utilization of the limit values. In the case of an uncertain estimate, the system can decide to measure again or to adjust the integration time of the measurement independently in order to improve the signal-to-noise ratio. The necessary integration time can be used as feedback to estimate the optimal integration time depending on previous knowledge (e.g. absorption coefficients from UV-VIS-NIR pre-characterization). This results in a time-optimized evaluation. If the signal-to-noise ratio is already maximized without sufficient classification, this may indicate a still unknown substance. For resolution, the textile can be directed to further laboratory analysis (e.g. chromatography or NMR). The results of this analysis are used as a feedback for the evaluation. In this way, the system continuously learns previously unknown Raman spectra and improves its recognition performance over the initially learned spectra from the development phase.

The data on contaminants derived from sorting will also be used for monitoring compliance of the textile industry with the REACH regulation[10] or stricter standards such as ZDHC[22]. This way, real pollutant levels in the textile cycle can be identified and reduced, independently of manufacturer data, in order to improve the safety of textiles in the long term. Textiles that

are not REACH compliant, e. g. those that originate from non-EU markets or were produced before the REACH implementation, are sorted out. Similarly, textiles contaminated by use (e. g. contaminated work clothing) can be recycled as hazardous waste, so that harmful substances are successively removed from the textile material cycle.

5 Outlook

The shift to a circular economy is a significant building block for achieving the UN Sustainable Development Goals (SDGs) and will contribute to solving current ecological challenges. The results of this project aim to create the necessary tools to foster a systematic change from linear to circular practices in the fashion industry. The textile industry has the potential to act as a role model in this regard, as 60 million people worldwide are professionally involved in textile production[11]. Since every person is wearing clothes, everyone can relate to the challenges of this industry and might even be affected personally. This makes the topic particularly tangible and shows the possibility of direct influence on the consumer side, thus creating the potential to spread the concept of circularity to other sectors.

Acknowledgement

This work has been funded within the framework "KI-Leuchttürme für Umwelt, Klima, Natur und Ressourcen"(AI lighthouses for environment, climate, nature and resources) by the German Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (BMU) under grant number 67KI2016.

References

- [1] Circle Economy. *Recycled post-consumer textiles - an industry perspective*. Interreg NWE, Fibersort project, 2020.
- [2] Ellen MacArthur Foundation. *A new Textiles Economy*. 2017.
- [3] Fibersort - Textiles Programme - Circle Economy. URL: <https://www.circle-economy.com/programmes/textiles/fibersort> (visited on 05/08/2021).
- [4] Greenpeace e.V. *Schmutzige Wäsche: Die Belastung chinesischer Flüsse durch Chemikalien der Textilindustrie*. 2011.
- [5] International Energy Agency IEA. *Energy, Climate Change and Environment 2016 Insights – Analysis and key findings*. 2016.
- [6] G. Jochem and R. J. Lehnert. "On the potential of Raman microscopy for the forensic analysis of coloured textile fibres". In: *Science & Justice* 42.4 (Oct. 1, 2002), pp. 215–221. DOI: 10.1016/S1355-0306(02)71831-5.

- [7] L. Lepot et al. “Application of Raman spectroscopy to forensic fibre cases”. In: *Science & Justice* 48.3 (Sept. 1, 2008), pp. 109–117. DOI: 10.1016/j.scijus.2007.09.013.
- [8] Jinchao Liu et al. “Deep convolutional neural networks for Raman spectrum recognition: a unified solution”. In: *Analyst* 142.21 (Oct. 23, 2017). Publisher: The Royal Society of Chemistry, pp. 4067–4074. DOI: 10.1039/C7AN01371J.
- [9] Jun Luo, Jian Fei Xie, and Wei Fan. “Study on Identification of 100% Cotton Fabric by Raman Spectroscopy and Random Forest”. In: *Advanced Materials Research* 1033-1034 (2014). Conference Name: Advances in Chemical Engineering and Advanced Materials IV ISBN: 9783038352693 Publisher: Trans Tech Publications Ltd, pp. 439–443. DOI: 10.4028/www.scientific.net/AMR.1033-1034.439.
- [10] *Official Journal of the European Union, L 136*. Volume 50. ISSN 1725-2555. European Union, 2007.
- [11] Kiran Pandey. *Fashion industry may use quarter of world’s carbon budget by 2050*. 2018. URL: <https://www.downtoearth.org.in/news/environment/fashion-industry-may-use-quarter-of-world-s-carbon-budget-by-2050-61183> (visited on 05/11/2021).
- [12] Anna Peterson. “Towards Recycling of Textile Fibers - Separation and Characterization of Textile Fibers and Blends”. Master’s thesis in Materials Chemistry and Nanotechnology. Göteborg: Chalmers University, 2015.
- [13] Ida Marie Sandvik and Wendy Stubbs. “Circular fashion supply chain through textile-to-textile recycling”. In: *Journal of Fashion Marketing and Management: An International Journal* 23.3 (July 8, 2019), pp. 366–381. DOI: 10.1108/JFMM-04-2018-0058.
- [14] “The price of fast fashion”. In: *Nature Climate Change* 8.1 (Jan. 2018), pp. 1–1. DOI: 10.1038/s41558-017-0058-9.
- [15] J. Thomas et al. “Raman spectroscopy and the forensic analysis of black/grey and blue cotton fibres”. In: *Forensic Science International* 152.2 (Sept. 2005), pp. 189–197. DOI: 10.1016/j.forsciint.2004.08.009.
- [16] Umweltbundesamt. *Schwerpunkte 2016, Jahrespublikation des Umweltbundesamtes*. 2016.
- [17] Vinnova. *SIPTex Swedish Innovation Platform for Textile sorting*. 2016. URL: <https://www.vinnova.se/en/p/siptex--swedish-innovation-platform-for-textile-sorting/> (visited on 05/08/2021).
- [18] Vinnova. *Swedish innovation platform for textile sorting (SIPTex)*. 2019. URL: <https://www.vinnova.se/en/p/swedish-innovation-platform-for-textile-sorting-siptex/> (visited on 05/08/2021).
- [19] Joel Wahl, Mikael Sjö Dahl, and Kerstin Ramser. “Single-Step Preprocessing of Raman Spectra Using Convolutional Neural Networks”. In: *Applied Spectroscopy* 74.4 (Apr. 1, 2020). Publisher: SAGE Publications Ltd STM, pp. 427–438. DOI: 10.1177/0003702819888949.

- [20] Helena Wedin et al. *Trash 2 Cash, Can automated NIR technology be a way to improve the sorting quality of textile waste?* EU Project Report. 2017.
- [21] Xu Yan et al. "Combining convolutional neural networks and on-line Raman spectroscopy for monitoring the Cornu Caprae Hircus hydrolysis process". In: *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* 226 (Feb. 5, 2020), p. 117589. DOI: 10.1016/j.saa.2019.117589.
- [22] *ZDHC MRSL V2.0*. URL: <https://mrs1.roadmaptozero.com/> (visited on 06/23/2021).

Adaptives luftqualitätsgewichtetes Fahrradrouting mittels Land-use Regression auf Basis offener Daten

Julian Janßen,¹ Paul Tremper,¹ Till Riedel¹

Abstract: Luftschadstoffen ausgesetzt zu sein hat langfristige negative gesundheitliche Folgen, denen besonders Fahrradfahrer im urbanen Raum ausgesetzt sind. Dabei gibt es wahrscheinlich keine unschädliche Dosis: weniger ist immer besser. Diese Arbeit zeigt, dass luftqualitätsgewichtete Fahrradrouten die persönliche Exposition gemäß dem Regressionsmodell deutlich reduzieren können, wobei die errechneten Umwege zumeist nur minimal sind. Auf Basis offener Daten wird ein neuronales Netzwerk zur Schätzung der Luftqualität trainiert. Dabei werden PM_{10} -Daten aus mobilen Messungen als Indikator der Luftqualität verwendet. Das entstehende Land-Use-Regression-Modell bezieht dabei sowohl zeitliche als auch räumliche Features mit ein. Anschließend wird dieses Modell verwendet, um luftqualitätsgewichtete Routen zu berechnen. Dabei wird gezeigt, wie ein solches feingranulare Modell im Routing verwendet werden kann. Anhand von zufällig gewählten Start/Ziel Paaren werden die luftqualitätsgewichteten Routen mit der jeweils kürzesten Strecke verglichen.

Keywords: Neuronales Netzwerk; Land Use Regression; Luftqualität; Routing

1 Warum brauchen Radfahrer künstliche Intelligenz

Luftqualität im städtischen Raum ist ein aktuelles und komplexes Thema. Dauerhafte, erhöhte Exposition an Luftschadstoffen kann zu gesundheitlichen Folgen führen, weswegen in Deutschland die innerstädtische Luftqualität staatlich überwacht wird und bei Grenzwertüberschreitungen Fahrverbote oder andere Maßnahmen verhängt werden können.

Die Schadstoffbelastung in der Stadt variiert aufgrund der Volatilität der Schadstoffe stark. In einer belgischen Studie [Pe13] wurde gezeigt, dass die PM_{10} (Feinstaub kleiner gleich 10 Mikrometer Durchmesser) Konzentration auf der Straße stark von der Verkehrsdichte abhängt. Dies führt zu der Frage, ob man durch die Wahl der Route eine deutliche Reduktion der Schadstoffexposition erreichen kann ohne dabei unzumutbare Umwege in Kauf nehmen zu müssen. Wir fokussieren uns in unserer Arbeit auf eine Routinganwendung für Radfahrer, da diese sowohl der Schadstoffexposition des Verkehrs direkt ausgesetzt sind, als auch sehr flexibel in der Wahl ihrer Route sind. Wir untersuchen dabei, ob durch die eine KI-gestützte Auswahl der Route die individuelle Schadstoffbelastung effektiv reduziert werden kann.

¹ Karlsruhe Institute of Technology (KIT), TECO, Vincenz-Prießnitz-Str. 1, 76131 Karlsruhe lastname@teco.edu

2 Hintergrund und Verwandte Arbeiten

2.1 Datensatz

Im Projekt *SmartAQnet* [Sm21] wurden über einen Zeitraum von drei Jahren Luftqualitätsdaten durch ein heterogenes Sensornetzwerk erhoben². Das Ziel war es, durch die Heterogenität des Netzwerks sowohl eine zeitlich wie räumlich hohe Auflösung zu erreichen, als auch durch qualitativ hochwertige Sensoren die Messungen der in der Breite ausgebrachten Low Cost Sensoren zu kalibrieren und zu validieren. Insgesamt enthält der Datensatz über 250 Millionen Datenpunkte. Der von uns für das maschinelle Lernen verwendete Teildatensatz enthält 16377 PM_{10} -Messwerte mit den dazugehörigen GPS-Positionen. Für kartographische Daten nutzen wir Daten des OpenStreetMap (OSM) Projekts.

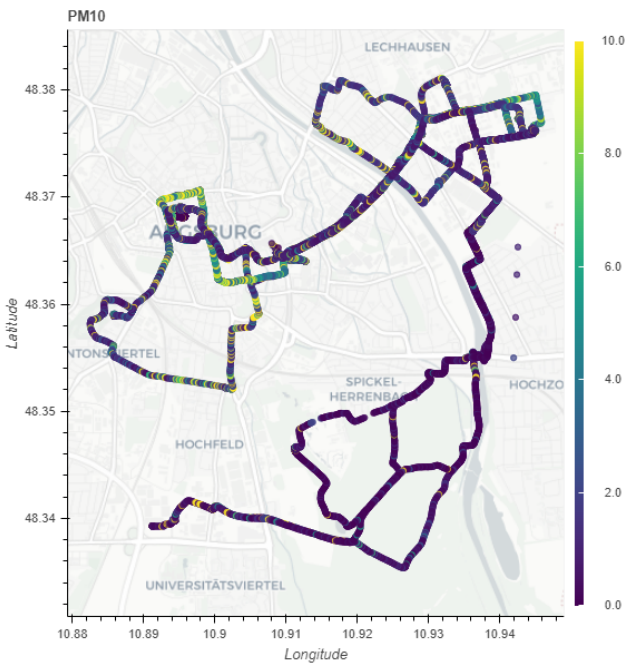


Abb. 1: Beispielhafte auf Fahrrädern gemessene PM_{10} -Belastungen in Augsburg

2.2 Land-Use-Regression

Das Land-Use-Regression-Modell (kurz LUR-Modell) ist eine häufig verwendete Methode, basierend auf maschinellem Lernen, um z.B. Luftqualitätsvorhersagen zu generieren.

² Alle Daten sind offen mittels einer OGC SensorThings-konformen API unter <https://api.smartaq.net/v1.0> abrufbar.

Innerhalb eines Regressionsmodells wird anhand von korrelierten Prädikatoren die räumliche Schadstoffverteilung auf Basis begrenzter Messkampagnen statisch hergeleitet.[Ho14]

Das LUR-Modell basiert auf zwei grundlegenden Annahmen:

1. Es genügt bereits eine kleine Anzahl an gemessenen Prädiktorvariablen, um die für die Zielgröße relevanten Umweltbedingungen abzuleiten.
2. Innerhalb eines begrenzten Trainingsgebiets kann auf zuverlässiger Weise die Beziehung zwischen der Zielgröße und den Prädikatoren ermittelt werden.[Br00]

Eine Studie in Auckland zeigt die Anwendbarkeit von LUR-Modellen im Mikromaßstab um Stickstoffdioxide vorherzusagen. Hierzu wurden für den ländlichen Raum die Kriterien Straßenbreite, Entfernung zur Hauptstraße, das Vorhandensein einer Markise und die Anzahl der Bushaltestellen (in der besagten Region ist der Einfluss der Dieselemission im Busverkehr nicht zu unterschätzen) definiert. Die Studie zeigt jedoch auch, dass die Parameter sich nicht auf den städtischen Raum übertragen lassen.[We18]

2.3 Exposition und Routing von Radfahrern

Luo et. al [LBB20] beschreiben eine neue Methode, die verkehrsbedingte Luftverschmutzung in die Routenplanung für Fahrradfahrer miteinbezieht. Die Studie evaluiert Fahrradrouten durch die Stadt Riverside, Kalifornien, anhand von Luftqualitätsdaten, welche mit einem Emissionsmodell auf geschätzter Verkehrsdichte simuliert wurde. Die Studie kommt zu dem Ergebnis, dass die Einbeziehung der Luftqualitätsdaten zu anderen Fahrradrouten führt.

Die Studie betrachtet nur von Autos befahrbare Straßen. In dieser Arbeit betrachten wir jedoch auch mögliche Routen durch Parks etc., was flächendeckende Luftqualitätsdaten notwendig macht.

Auch an anderer Stelle wird thematisiert, dass die Luftqualität zu wenig bei der Streckenwahl der Fahrradstraßen berücksichtigt wird [Tr20]. Bei der Planung der Radwege wird der negative Einfluss der Luftverschmutzung auf die Fahrradfahrer oft außer Acht gelassen. Luftqualität wird dabei mithilfe von Open-Source-Daten, Land-Use-Regression, einem Neuronalem Netz und einer räumlichen Analyse geschätzt. Eine Studie [LBB18] zeigt, dass die Luftverschmutzung bei der längeren Route um 44 - 48 % kleiner ist, abhängig von der Tageszeit. Dadurch zeigt sich, dass durch die Wahl der Route die Konzentration der Luftverschmutzung der man ausgesetzt wird, verringert werden kann. Die Studie weist noch darauf hin, dass durch die Integration weiterer Faktoren wie beispielsweise die Verkehrs-, Wetter- und bessere Luftqualitätsdaten in Echtzeit, die Wahl der Route noch verbesserte werden könnte. [Ma19] behandelt ebenfalls die Möglichkeit durch ein optimiertes Routing, die Exposition der Luftschadstoffe zu verringern. Die Studie entwickelte den so genannten CAR-Algorithmus (Clean Air Routing), der dafür verwendet wird, ein

gesundheitsoptimierendes Routing-Empfehlung zu ermitteln. Dafür werden Open-Source-Daten des Feinstaubs $PM_{2,5}$ herangezogen. Mithilfe des Straßennetzdiagramms von Taiwan und den gemessenen Feinstaubdaten wird nun eine räumlich-zeitliche Interpolation gemacht, um die $PM_{2,5}$ -Werte an den Straßennetzkreuzungen zu ermitteln.

3 Routing mittels Neuronalem Netzwerk als Land-Use-Regressor zur Streckengewichtung

Das Vorgehen bei der Erstellung des Systems ist schematisch in Abbildung 2 dargestellt. Beginnend auf der linken Seite werden die Messdaten als Basis für das Modell der Luftqualität verwendet. Zusammen bilden diese beiden Komponenten die Thematik der Land-Use-Regression ab. Das erstellte Vorhersagemodell wird anschließend im Routing verwendet um die eine luftqualitätsgewichtete Route zu bestimmen.



Abb. 2: Schematischer Systemaufbau

Um das Fahrradrouting mit Luftqualitätsdaten gewichten zu können, benötigen wir eine flächige Regression. Diese erhalten wir durch eine Land Use Regression mithilfe eines neuronalen Netzes.

Die geographische Umgebung stellt dabei eine Repräsentation der örtlichen Gegebenheiten an einem Punkt da. Dabei werden die geographischen Gegebenheiten in Kategorien eingeteilt. Für jeden Punkt und jede Kategorie wird anschließend ein numerischer Wert berechnet. Abhängig von den vorliegenden Daten zu einer Kategorie, handelt es sich hierbei entweder um die Summe der Längen bzw. der Flächen innerhalb der zu bestimmenden Umgebung.

Wir beschreiben die geographische Umgebung durch einen Kreis mit Radius r um den Messpunkt. Da im Vorhinein unklar ist welcher Umgebungsradius r für die Modellierung am geeignetes ist, wird dieser explorativ bestimmt. Dies erfolgt in Abschnitt 3. So wird die geographische Umgebung für alle Messpunkte mit verschiedenen Radien berechnet.

Die Oberkategorien „Land Use“ und „Water“, des OSM-Datensatzes, enthalten die tatsächliche geographische Ausprägung in Form von Polygonen. Die Summe der im Umgebungskreis enthaltenen Flächen für eine bestimmte Kategorie bildet hier den Wert der jeweiligen Kategorie. Elemente der Oberkategorien „Railways“ und „Road“ werden als Linien modelliert.

Hier wird die Summe der im Umgebungskreis enthaltenen Entfernungen als Wert der Kategorie verwendet.

Die Daten werden 80/20 in Trainings- und Validierungsset aufgeteilt. Hierbei werden die Trainingsdaten für die Optimierung der Modelle verwendet und zu diesem Zweck noch einmal in Trainings- und Testset aufgeteilt. Die Zuordnung zu dem jeweiligen Set, wird dabei randomisiert. Die Daten aus dem Validierungsset werden zum Schluss für die Bewertung des Modells verwendet.

Als Features für das Neuronale Netz werden die kategorisierten Anteile der geographischen Umgebung, um einen Messpunkt gewählt. Diese werden linear unter Verwendung des Minimums sowie des Maximums in den Bereich $[0,1]$ skaliert. Zusätzlich wird Wochentag und Stunde des Tages als Feature hinzugefügt. Diese wird dabei mit dem One-Hot Verfahren encodiert. Die PM_{10} Messungen werden dabei als Ausgangsgröße für das Lernen verwendet und ebenfalls Minimum/Maximum skaliert.

Auf dieser Grundlage werden verschiedene Modellierungsparameter variiert. Eines der untersuchten Parameter ist der Abstand, welcher für die Umgebungsbeschreibung verwendet wird. So werden verschiedene Radien des Umgebungskreises untersucht. Ebenso werden zwei Loss-Funktionen betrachtet. Als Aktivierungsfunktionen werden ReLU, Sigmoid sowie Tanh verwendet. Auch verschiedene Netzstrukturen werden untersucht. So wird auch eine eher Breite (Anzahl Neuronen), sowie eine etwas tiefere Struktur (Anzahl Schichten) erstellt. So werden insgesamt 90 verschiedene Modelle trainiert. Tabelle 1 gibt einen Überblick über die untersuchten Parameter.

Tab. 1: Variierte Hyperparameter

Radius der Umgebung	3, 5, 10, 15, 20
Loss-Funktion	Mittlerer absoluter Fehler (MAE), Mittlere quadratische Abweichung (MSE)
Aktivierungsfunktion	ReLU, Sigmoid, Tanh
Struktur	[128, 64, 32, 1] [128, 64, 1] [64, 32, 16, 8, 1]

Anschließend werden die trainierten Modelle miteinander verglichen. Den geringsten MAE weist das Neuronale Netz mit einer vollverbundenen Struktur von [128, 64, 1] auf. Als Loss-Funktion wird der MAE, als Aktivierungsfunktion wird Sigmoid verwendet. Dabei wird ein Umgebungsradius von 15 Metern verwendet (MAE: 0.04241, MSE: 0.01259). Die als Baseline dienenden Modelle weisen dabei einen schlechteren MAE auf. Der Random Forest weist einen MAE von 0.05863, die SVM sogar einen von 0.09297 auf.

3.1 Routing

Als Grundlage für den Routing-Algorithmus berechnen wir aus den vorliegenden Basisdaten einen gerichteten und gewichteten Graphen der Strecken. Alle für die Routing-Entscheidung nicht benötigten Details werden dabei verworfen. Die Schritte zur Bestimmung einer Route sind schematisch in Abbildung 3 dargestellt. Die Luftqualitätsdaten aus der Land Use Regression fließen dabei wie dargestellt in die Bestimmung der Gewichte ein.

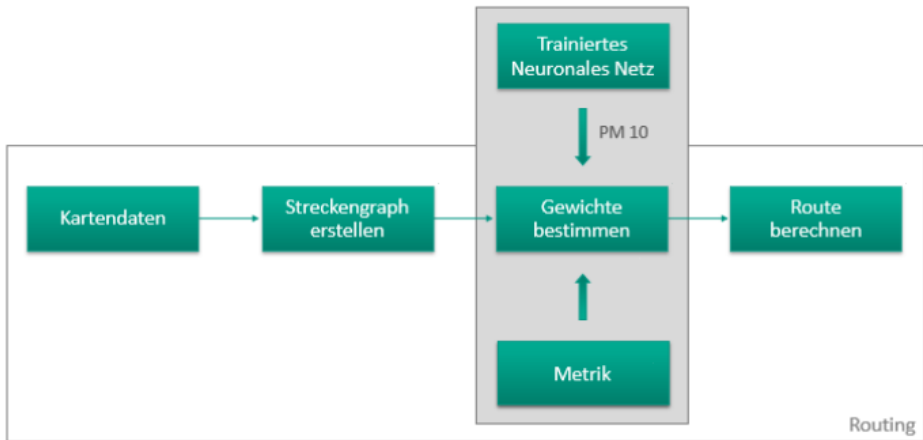


Abb. 3: Routingpipeline

Wir laden alle mit dem Fahrrad befahrbare Strecken aus dem OSM-Datensatz herunter und erstellen daraus den gerichteten Streckengraphen, bestehend aus Knoten und Kanten. Anschließend werden die Kanten mit Luftqualitätsdaten wie folgt gewichtet:

Im ersten Schritt wird ein Streckenabschnitt in vorhersagbare Punkte aufgeteilt. Hierbei wird zunächst die Anzahl der Punkte bestimmt, die auf dem Abschnitt liegen können ohne dass der Umgebungskreis sich überschneidet. Dazu wird die Länge des Abschnitts durch den Durchmesser des Umgebungskreises geteilt. Der Radius des Umgebungskreises entspricht dabei dem des Radius der für das Training des Modells verwendet wurde. In diesem Fall entspricht der Radius 15m. Anschließend werden die Lage der Punkte auf dem Teilstück so bestimmt dass diese den gleichen Abstand zueinander haben. Für jede Lokation wird nun die geographische Umgebung berechnet. Im zweiten Schritt wird diese verwendet um die Luftqualität vorherzusagen. Der Durchschnitt der Vorhersagen ergibt nun den PM_{10} Wert einer Kante.

Aufgrund der Performance werden zuerst alle vorhersage Punkte bestimmt. Anschließend werden die dazugehörigen geographischen Umgebungen berechnet. Nun werden für alle Umgebungen die Vorhersagen für die Luftqualität gemacht, die dann in das Kantengewicht einfließen. Die in dieser Arbeit verwendete Gewichtungsfunktion soll den Trade-Off

zwischen der Fahrzeit auf der einen, sowie der Exposition von Luftschadstoffen auf der anderen Seite abbilden. Dabei wurde folgende Formel für die Implementierung verwendet:

$$\text{Gewicht} = \text{Dauer Exposition} \times (1 + (\text{Präferenzfaktor} \times \text{Prädiktion})) \quad (1)$$

Wir vereinfachen die Formel durch die Annahme, dass der Fahrradfahrer mit konstanter Geschwindigkeit fährt und erhalten:

$$\text{Gewicht} = \text{Länge Streckenabschnitt} \times (1 + (\text{Präferenzfaktor} \times \text{Prädiktion})) \quad (2)$$

Wobei Prädiktion der entsprechende berechnete Luftqualitätswert für den Streckenabschnitt ist. Der eingebaute Präferenzfaktor ermöglicht hierbei eine individualisierung des Trade-Offs.

Für die Implementierung des Routing Abschnitts wurde ebenfalls die Programmiersprache Python verwendet. Das Paket OSMnx wurde für die Berechnung der Routen verwendet. Dieses basiert auf dem Paket NetworkX welches die Modellierung von Netzwerken in Form von Graphen ermöglicht. Als Algorithmus für die Bestimmung der optimalen Route wird der Dijkstra-Algorithmus verwendet.

4 Evaluierung

In dieser Arbeit untersuchen wir die Frage, ob ein luftqualitätsgewichtetes Fahrradrouting für gegebene Start- und Zielpunkte Routen mit deutlich reduzierter Exposition bei geringer zusätzlicher Weglänge produzieren kann. Die vagen Bezeichnungen "reduzierte Exposition" und "geringe zusätzliche Weglänge" beziehen sich dabei auf einen subjektiven Trade-Off, der durch den Präferenzfaktor frei wählbar ist.

Zur Überprüfung dieser Hypothese werden 100 verschiedene Routen berechnet. Dabei wird sowohl der Start- als auch der Zielknoten zufällig ausgewählt. Für jedes dieser Paare wird sowohl die kürzeste Strecke als Baseline sowie die luftqualitätsgewichtete Strecke bei fixem Trade-Off Faktor berechnet. Anschließend wird für beide Routen jeweils sowohl die Länge als auch die PM_{10} -Exposition berechnet. Es wird die Prozentuale Veränderung der Exposition wie auch die Prozentuale Veränderung der Länge berechnet.

In Abbildung 4 ist ein Beispiel einer Route in blau zu sehen, die nach der Länge der Strecke optimiert ist.

Die blaue (grüne) Strecke in Abbildung 4 ist die kürzeste (luftqualitätsgewichtete) Strecke für ein beispielhafte Wahl der Start- und Zielpunkte. Im Fall der luftqualitätsgewichteten Strecke macht die Routenführung nun einen Bogen um die Innenstadt. Im vorliegenden Fall weist die kürzeste Route eine um 27.8 % höhere Exposition auf als die luftqualitätsgewichtete Route. Die kürzeste Strecke von 5.096 km wird dabei lediglich um 1.1 % verlängert. Der Präferenzfaktor wurde auf 0.8 festgesetzt.



Abb. 4: Für ein Beispiel von Start- und Zielpunkten: Kürzeste Route in blau, luftqualitätsgewichteten Route in grün.

Abbildung 5 zeigt 100 zufällig gewählte Start- und Zielpaare. Die blauen Punkte repräsentieren dabei ein Start- und Zielpaar, für das die kürzeste sowie die luftqualitätsgewichtete Route bestimmt wurde. Die rote Gerade ist eine Gerade mit einer Steigung von 1. Sie lässt die unabhängige Skalierung der Achsen erkennen. Die grüne Gerade stellt eine einfache lineare Regression dar.

Dabei ist eine Häufung der Ergebnisse im Bereich nahe $[0, 0]$ zu erkennen. Dies bedeutet das viele der Luftqualitätsgewichteten Routen entweder genau der kürzesten Route entsprechen oder nur eine minimale Verbesserung erzielt werden konnten. Zum einen hängt dies natürlich mit kurzen Routen zusammen, für die wenn überhaupt eine Verbesserung nur durch große Umwege erreicht werden könnte (weil z.B. Start oder Ziel die Hauptursache für eine Exposition sind). Auch ist eine große Streuung der Ergebnisse zu beobachten, die natürlich erscheint aufgrund der nichtlinearen Zusammenhänge auf Basis der Randbedingungen in einem Straßennetz. Dennoch werden vorwiegend Routen gefunden die die Exposition deutlich reduzieren, dafür aber nur einige wenige Prozentpunkte längerer Umweg in Kauf genommen werden muss. Wichtig dabei ist zu verstehen, dass jede zusätzliche Exposition mit Feinstaub sich potentiell negativ auf die Gesundheit des Fahrradfahrers auswirken kann, da es keine unschädliche Dosis gibt ("wahrscheinlich" lineare Dosis-Wirkung-Beziehung, no adverse effect level [Or13]). Daher ist jede Einsparung ein wichtiger Beitrag zur Gesundheit (umgekehrt macht es keinen Sinn auf die Person bezogen lediglich einen Grenzwert zu unterschreiten).

Die kürzesten Strecken haben eine durchschnittliche Streckenlänge von 6.25 km, wobei die zugehörigen luftqualitätsgewichteten Strecken durchschnittlich nur um 0.028 km (0.44 %) länger sind. Von den betrachteten Start- Zielpunktpaaren waren die luftqualitätsgewichteten Strecken nie mehr als 3 % länger als die jeweilige kürzeste Strecke. Gleichzeitig weisen die kürzesten Strecken eine im Durchschnitt um 36,24 % höhere Exposition auf.

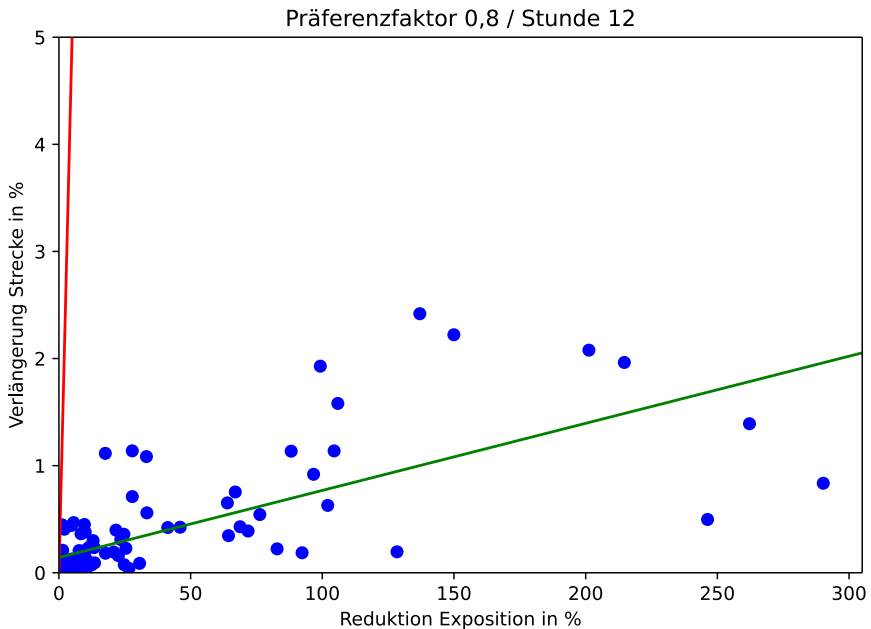


Abb. 5: Ergebnis Plot zufälliger Routen. Die Reduktion in % bezieht sich auf die erhöhte Exposition auf der kürzesten Route relativ zur luftqualitätsgewichteten Route.

5 Diskussion und Ausblick

Im Rahmen dieser Arbeit haben wir auf Basis eines Luftqualitätsdatensatzes für die Stadt Augsburg untersucht, ob durch die intelligente Auswahl einer Fahrradrouten die persönliche Feinstaubexposition deutlich reduziert werden kann ohne unpraktikable Umwege in Kauf nehmen zu müssen. Wir finden, dass hierbei zumeist zwar nur eine kleine Reduktion der Exposition erreicht werden kann, teilweise jedoch bis zu 75%, wobei die in Kauf zu nehmenden Umwege nie mehr als +3% der kürzesten Strecke betragen.

Bei der Bewertung der Ergebnisse ist zu beachten, dass diese stark von den vorhandenen Luftqualitätsdaten abhängig sind. Zudem wurde der Modellfehler des grundlegenden Neuronalen Netzes in der Arbeit nur anhand von mobilen Messungen bewertet. Um das luftqualitätsgewichtete Verkehrsrouting weiter zu entwickeln sind weitere Untersuchungen notwendig. Wird der Fehler besser bewertbar, macht es auch Sinn andere Netzarchitekturen als Regressoren zu testen. So enthält der Datensatz bereits ganze Fahrradrouten, die zur Evaluation herangezogen werden können. Hier ist jedoch noch eine Kalibrierung notwendig, da nur die fixen Sensoren im Datensatz bis jetzt kalibriert wurden. Eine weitere Frage ist, inwieweit ein Vorhersagehorizont (Now-Casting) zusätzlich zur Land-use Regression die gewählte Route verbessert, da diese gerade zu Hochverkehrszeiten stark

schwanken kann. Eine Alternative, die wir momentan betrachten, ist die Nutzung von auf Simulationen gelernten Now-Casting-Modelle als schnell ausführbare Surrogate für die Expositionsrechnung. Grundlage ist jedoch, dass die Emission wirklich vorhersagbar ist, was beispielsweise bei starken Belastungen wie Baustellenaktivitäten oft nur mit einem geringen Horizont der Fall ist. Kritisch zu hinterfragen ist weiterhin, inwieweit überhaupt eine dynamische Vorhersage das Routing verändert, oder ob es nicht ausreicht ein typisches Modell für Wochentage und Uhrzeit anzunehmen. Hierzu planen wir weitere Untersuchungen. Der Vorteil von Land.Use-Regression-Modellen und anderen Lernmodellen zum Scoring von Routenteilstücken ist prinzipiell die partielle, bedarfsgerechte Auswertung innerhalb des Routingalgorithmus.

Literaturverzeichnis

- [Br00] Briggs, David J.; de Hoogh, Cornelis; Gulliver, John; Wills, John; Elliott, Paul; Kingham, Simon; Smallbone, Kirsty: A regression-based method for mapping traffic-related air pollution: application and testing in four contrasting urban environments. *Science of The Total Environment*, 253(1):151–167, Mai 2000.
- [Ho14] de Hoogh, Kees et al.: Comparing land use regression and dispersion modelling to assess residential exposure to ambient air pollution for epidemiological studies. *Environment International*, 73:382–392, Dezember 2014.
- [LBB18] Luo, Ji; Boriboonsomsin, Kanok; Barth, Matthew: Reducing pedestrians’ inhalation of traffic-related air pollution through route choices: Case study in California suburb. *Journal of Transport & Health*, 10:111–123, September 2018. WOS:000449750400013.
- [LBB20] Luo, Ji; Boriboonsomsin, Kanok; Barth, Matthew: Consideration of exposure to traffic-related air pollution in bicycle route planning. *Journal of Transport & Health*, 16:100792, März 2020. WOS:000539174500012.
- [Ma19] Mahajan, Sachit; Tang, Yu-Siou; Wu, Dong-Yi; Tsai, Tzu-Chieh; Chen, Ling-Jyh: CAR: The Clean Air Routing Algorithm for Path Navigation With Minimal PM2.5 Exposure on the Move. *IEEE Access*, 7:147373–147382, 2019.
- [Or13] Organization, World Health et al.: Health effects of particulate matter. Policy implications for countries in Eastern Europe, Caucasus and central Asia. World Health Organization Regional Office for Europe, Copenhagen, 2013.
- [Pe13] Peters, Jan; Theunis, Jan; Van Poppel, Martine; Berghmans, Patrick: Monitoring PM10 and Ultrafine Particles in Urban Environments Using Mobile Measurements. *Aerosol and Air Quality Research*, 13(2):509–522, April 2013. WOS:000315327400008.
- [Sm21] SmartAQnet: , Website SmartAQnet. <https://www.smartaq.net/de>, Februar 2021.
- [Tr20] Tran, Phuong T. M.; Zhao, Mushu; Yamamoto, Kohei; Minet, Laura; Nguyen, Teron; Balasubramanian, Rajasekhar: Cyclists’ personal exposure to traffic-related air pollution and its influence on bikeability. *Transportation Research Part D-Transport and Environment*, 88:102563, November 2020. WOS:000591708400002.

- [We18] Weissert, L. F.; Salmond, J. A.; Miskell, G.; Alavi-Shoshtari, M.; Williams, D. E.: Development of a microscale land use regression model for predicting NO₂ concentrations at a heavy trafficked suburban area in Auckland, NZ. *The Science of the Total Environment*, 619-620:112–119, April 2018.

A meta analysis of the status of AI in environmental computer science

Lukas Sinnwell¹ Julian Bruns² Matthias Budde³ Andreas Abecker⁴

Abstract: Artificial Intelligence is hyped as one of the key enablers of the future which will help to solve key challenges of humanity. The research field which already addresses many of those challenges is Environmental (Computer) Science. Therefore, it seems like a natural fit to combine both fields. AI can provide algorithms and methods to ease the computations and extend the existing (simulation) models of the environmental sciences. The environmental sciences already have in-depth knowledge of the problems at hand, can evaluate and interpret both the relative importance of input data as well as the results. However, to achieve this synergy, a strong foundation and knowledge of previous work is needed. This work aims to contribute to this foundation by providing a data-driven overview of AI-based research activities in the field of environmental computer science.

Keywords: AI; Computational Sustainability, Environmental Computer Science, Meta Analysis

1 Introduction

Most of today's medium- to long-term problems are environmental problems. 8 out of the 17 United Nation's Sustainable Development Goals (SDGs) [SD18] are directly connected to environmental issues. Environmental Computer Science (ECS) plays part in solving these problems [Fi17], as it deals with modelling, understanding, and ultimately predicting environmental processes by using methods of information technology [Hi95]. However, given the complex, and often non-linear causalities in natural processes, building these models and generating usable results, i.e. predictions, is a difficult task. This is where Artificial Intelligence (AI) could make a difference, as it offers a different approach to data modeling than conventional methods do. But while AI postulates itself as a universal field"[RN10], it is not yet fully part of the tool set of the ECS field[Go19]. And while there are many publications, projects and visions, such as [Li21] in the proceedings of this conference last year, as well as even broader calls for actions such as the idea of computational sustainability [Go09] - an enhanced integration of Computer Science (CS), Mathematics and Statistics in a quest for a more effective use of natural resources - there does not yet exist an overview of the state of AI in ECS. To cite the German Ministry of Environment: There is "[. . .] no publication that allows a comprehensive interpretation

¹ Disy Informationssysteme GmbH, Ludwig-Erhard-Allee 6, 76131 Karlsruhe lukas.sinnwell@disy.net

² Disy Informationssysteme GmbH, Ludwig-Erhard-Allee 6, 76131 Karlsruhe julian.bruns@disy.net

³ Disy Informationssysteme GmbH, Ludwig-Erhard-Allee 6, 76131 Karlsruhe matthias.budde@disy.net

⁴ Disy Informationssysteme GmbH, Ludwig-Erhard-Allee 6, 76131 Karlsruhe andreas.abecker@disy.net

of the research activities in the context of environment and sustainability - neither on a European nor on an international level (status March 2019)” [Je19]. Hence, the aim of this work is to provide a first data-driven overview of the current state of the art of AI-related work in ECS. While there are many surveys, theses and textbooks regarding AI in general [RN10][Ch20], its status[Mo20] as well as future [Ka20][Sc19]. And while there are specific surveys for the use of AI in specific sub fields of the ECS, such as [MD00] or [Sh18], their approach is guided by the applicability to their field and not about overall status of AI in their sub field.

We aim to achieve this by (1) Gather a corpus of publications with regard to ECS and AI, (2) Mapping the publications to their respective research fields and (3) evaluate the methodological approaches of the 5 most cited publications in each field in regard to 12 criteria derived from the AI literature.

2 Publication corpus

To get an impression of active fields in AI, and AI and ECS in particular, we used the records from Web of Science (WoS) to gather the different publications. WoS groups each publication into specific categories and rank these by the number of records per category. These are not assigned exclusively. A first generic query for just AI and its native fields returns about 200.000 results. After investigating the top 82 results, we refined our query to focus on publications with an environmental science background, see fig. 1. This query returns all records in which one of the words marked in purple above is present in either title, abstract or keywords. These are basic AI terms, in order to include as many records as possible. We chose those four terms as they encompass the most undisputed AI terms in regard to methods. The second part of the query (after AND) specifies the desired WoS categories. Every record on WoS is assigned at least one of more than 250 categories. We then selected all categories from the field of environmental science which belongs to the top 100 AI categories. Eight categories re-main using this approach, namely Environmental Sciences, Water Resources, Geosciences Multidisciplinary, Remote Sensing, Green Sustainable Science Technology, Geography Physical, Agriculture Multidisciplinary and Environmental Studies. Overall, the query returned 18.254 records, as of November 2020. This is slightly less than 10% of the overall AI literature found in WoS. The addition of further terms would deviate from the data driven approach used in this paper.

3 Clustering of the field

Based on this corpus of publications, we are able to build a network of the publications based on their title, keywords, categories as well as citation links. This allows us to cluster the overall publications and then map each publication to a cluster. We use the tool VOSviewer [VEW13] to perform these tasks. We use the association strength as a measure for the

```

(TS = ("Artificial Intelligence") OR TS = ("Machine Learning") OR
TS = ("Deep Learning") OR TS = ("Artificial Neural Network*"))
AND
(WC = ("ENVIRONMENTAL SCIENCES") OR WC = ("WATER RESOURCES") OR WC = ("GEO-
SCIENCES MULTIDISCIPLINARY") OR WC = ("REMOTE SENSING") OR
WC = ("GREEN SUSTAINABLE SCIENCE TECHNOLOGY") OR
WC = ("GEOGRAPHY PHYSICAL") OR WC = ("AGRICULTURE MULTIDISCIPLINARY") OR
WC = ("ENVIRONMENTAL STUDIES"))

```

Fig. 1: Query for Web of Science to gather all publications for analysis

mapping and similarity of the publications to each other. We then applied the Smart Local Moving Algorithm (SLM) from VOSviewer to create the different clusters. However, there is no inherent selection criteria for the optimal number of clusters in the algorithm or software. We used a variant of the so-called elbow criteria often used in the k-means clustering - to decide the different parameter for the clustering. This results in seven different classes of cluster and it can be seen in fig. 2.

Having established suitable parameters for mapping and clustering, the problem of labelling the clusters emerges. VOSviewer does not label the clusters natively. However, VOSviewer allows to save a created mapping into a map-file and a net-file. The map-file contains all items, their coordinates, information (title, authors, source), and most important for this, the cluster they belong to. The different terms used for the mapping of the items for each cluster are extracted from the map-file into single, tab delimited text files. These files can in turn be analysed in VOSviewer to extract frequent term via NLP. After extracting all the titles into a separate file, VOSviewer detects just over 10.000 different terms. We select only 25 most frequent domain terms for each cluster. Generic Terms like 'model', 'prediction', etc. are not meaningful in this context and therefore are removed beforehand. By performing this analysis we are able to name the cluster based on their content and most often used terms. Fig. 3 shows the top 5 terms per cluster and the cluster names derived. We chose those terms based on existing knowledge about the fields and our pre-existing domain knowledge as well as discussions with environmental scientists.

4 Methodological Evaluation

Based on the seven different clusters and their mapped publications we are now able to investigate more in-depth the approach to AI in the different sub fields. We investigate the five most cited publications in each field. Reviews and surveys were removed in order to analyse publications carrying out a specific study. Experience has shown that AI often only appears marginally in reviews and is one of many explored methods. This way, the use of AI and the methods are defined more clearly and are more accessible. In this context, it is

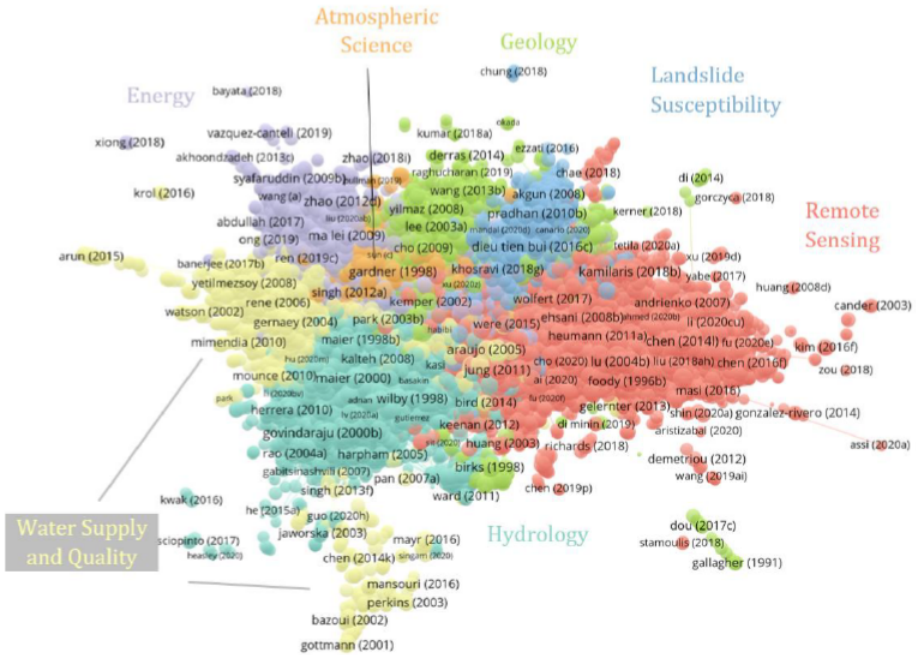


Fig. 2: Map of the WoS dataset created in VOSviewer using 0.30 resolution and a minimum cluster size of 300. The biggest cluster (red) contains 4839 paper.

cluster 1	cluster 2	cluster 3	cluster 4	cluster 5	cluster 6	cluster 7
sentinel	river	solar radiation	adsorption	soil	emission	landslide susceptibility mapping
remote sensing	river basin	building	wastewater	rock	city	gis
forest	basin	energy	aqueous solution	blast	air quality	landslide
hyperspectral image classification	reservoir	energy consumption	treatment	earthquake	ozone concentration	landslide susceptibility assessment
remote sensing image	temperature	photovoltaic system	water	mine	ozone	landslide susceptibility
<i>Remote Sensing</i>	<i>Hydrology</i>	<i>Energy</i>	<i>Water Supply & Quality</i>	<i>Geology</i>	<i>Atmospheric Science</i>	<i>Landslide Susceptibility</i>

Fig. 3: Top 5 domain terms per cluster and cluster names derived from this.

also interesting how the studies were carried out and what steps and information were given. We use the most cited publications as our selection criteria as the number of citations can be used as measure of their importance for the field. Either by their innovative character, foundation of the use of AI in the field or by relevance of the topic for the overall field. It also enhances the chance that it is indicative for the field. However, this approach has a trade-off. We implicitly provide more weight to older publications and thereby ignore more recent developments and improvements in the fields. We argue that this trade-off is acceptable for the aim of this paper to provide a first overview for our reasons stated before.

For the final, in-depth analysis we had to perform several small changes from our selection of publications for the clusters. For the "Water Supply and Quality" cluster the five papers with the most citations all revolve around the topic of climate change. However the majority of papers in this cluster, deal with topics evolving around water, waste water, and chemistry. While they are often relevant in the context of climate change, they address issues beyond that in most cases. Since the most cited papers would not be adequate representation of this cluster, only the most cited publication is chosen in this cluster with explicit connection to climate change. The selection is filled up with the next most cited publication, which are thematically address a broader spectrum and provide a more representative view of the cluster. Tab. 1 provides an overview of the selected publications for each cluster.

Cluster	Most cited paper
Landslide Susceptibility	[PL10], [Yi09], [Bu16], [PPG12], [Le04]
Remote Sensing	[HDT02], [Ro12], [Ch14], [Ch16], [CZH16]
Hydrology	[HGS95], [MH96], [VBB13], [Wa09], [Na04]
Geology	[KS02], [Zo08], [KS07], [SMJ02], [Br00]
Energy	[Br10], [Gu12], [Gh09], [MRH98], [Ka09]
Atmospheric Science	[AWBAA05], [So07], [GD99], [Fe15], [GC06]
Water Supply	[Ar05], [HKH04], [Ma16], [YD08], [NSD04]

Tab. 1: Overview of papers selected for analysis grouped by clusters.

To evaluate those 35 publications, we developed a set of methodological criteria to compare those publications objectively on their approach to AI. Those criteria are based on the common steps in an AI workflow commonly used in the literature or in practical AI projects, for example proposed in a similar form by [Za19] for human centred AI models. The criteria are extended by characteristics proposed by [YBL11], [Yu20] and [RS20]. 4 shows the used criteria in this publication to evaluate the publications. We divided the criteria in the general criteria for statistical and bibliographic aspects and in methodological criteria, which are the focus in this work. The 12 different criteria allow the reader to quickly identify the key information of the AI model building process. While additional criteria, such as the type of data, the preprocessing and so on could be added, we decided to focus on the most common and overarching criteria which can be investigated objectively. They allow to investigate key means of reproducibility of the studies as well.

The aggregated results of this comparison to our criteria is shown in fig. 5. Only two studies,

General criteria	Methodological criteria	
Research Area	type of data	<i>Kind of data is used in study</i>
Title	sample size	<i>size of dataset</i>
Authors	Input Selection	
Times Cited	Input Significance	<i>methods for deciding which inputs to use in model</i>
Year	Input Independence	<i>methods for checking correlation among inputs</i>
Countries	Data Division	
Type	Data Division	<i>method for dividing data into train / test set</i>
Journal	Training [%]	<i>Percentage of whole dataset used for training</i>
Conference	Test [%]	<i>Percentage of whole dataset used for testing</i>
Keywords	Validation	<i>validation technique or percentage of dataset used for validation</i>
AI methods in focus	Model Building	
Key notes	Hyperparameter tuning	<i>methods for setting HPs</i>
Outcome	NN architecture	<i>if available, methods of determining NN architecture</i>
	Evaluation & Reference	
	Evaluation metrics	<i>metrics used for model evaluation</i>
	Reference model	<i>reference for model developed in study</i>

Fig. 4: Criteria used for in-depth analysis.

namely [Na04] and [Zo08] met all criteria. A large portion of studies met all but one, two or three criteria. As the number of missing information grows, the number of affected papers declines. The focus of this in-depth analysis is on completeness, transparent documentation, reproducibility, and replicability. The criteria are chosen to reflect these characteristics. By not meeting all the criteria, not all of these characteristics can be achieved by the studies. However, it is important to say that when a study does not mention certain information about the ML process, it does not automatically imply that the respective step was neglected. It might as well be that the authors deemed this information as not important enough to appear in the publications. None of the papers we investigated provide access to materials i.e., neither the data which was used nor the code or fragments of it. This means an exact replication these studies is not possible. In the following subsection we provide a more in-depth view of each criteria.

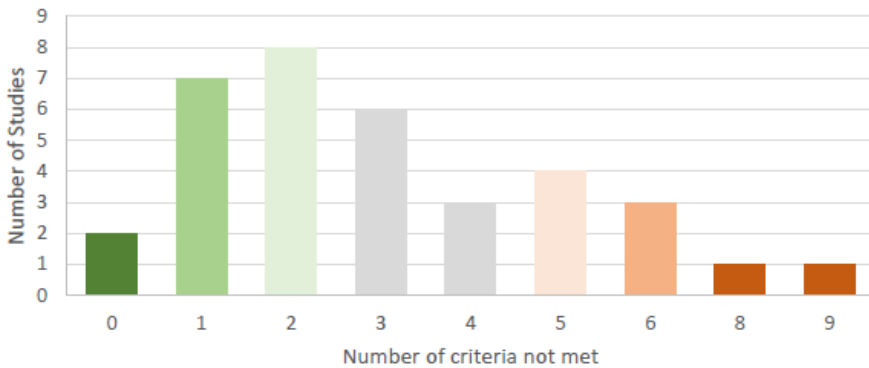


Fig. 5: Distribution of number of studies not meeting a number of criteria.

4.1 Sample size

All but three studies reported on their sample size. It ranges from 36 soil samples to 12.000 environmental drugs and chemicals.

4.2 Input significance

Concerning Input Significance, 12 studies made no clear statement, whether input significance measures were applied and if so, which ones. One has to consider that input significance can be a highly domain specific matter. For some studies, it may be the case that from a domain perspective the selection of inputs is generally agreed on. Studies that explicitly used domain knowledge for input selection were marked with 'ad hoc'. This relates to eight studies. This leaves 15 studies which used some numerical tool to determine significance of input features; four studies used two procedures. Four studies used training

models with different combinations of inputs i.e., trial and error, and three times each regression analysis, correlation analysis, and Principal Component Analysis were used.

4.3 Input independence

Input independence was neglected more often. But here one can also not rule out that some generally recognized domain knowledge was assumed to be known by the reader. Only five studies made explicit statements regarding input independence, of which three used correlation analysis, one trial and error approach and one variance inflation factor.

4.4 Data division

In terms of data division, a random division is most used among the sample studies (12 times). However, for time series forecasting, a random division of data is not advised. In these cases, the time series were divided arbitrarily at some point in time according to the chosen training and test sizes. All selected studies concerned with time series forecasting (11) used this technique. The remaining three cases used stratified random division, supervised division or a given division, respectively. Nine studies performed no data division. The percentage division ranges from a 50:50 split to 80:20 split (training/testing). Six publications provided no information about a data division and two studies provided the number of training samples but no further information.

4.5 Validation

Validation was mentioned significantly fewer times than training, testing or model building related facts. Nearly half of the studies (17) did not mention this topic. It was shown that some textbooks recommend a three-way split, where a partition of the data is re-served for validation purposes. However, in recent years, the cross-validation technique gained popularity. This technique does not require a static validation set. The studies are almost evenly distributed between these two approaches with eight using a validation set and seven using cross-validation. One study used the built-in Out-of-bag Error (OOB) of the Random Forest algorithm, one used “[. . .] random samples [. . .]” [Le04], and one switched test and training set for validation [KS02].

4.6 Hyperparameter tuning

In 24 studies remarks were made regarding the tuning of these parameters. Out of the 24 which made remarks, ten were ad hoc settings, just providing information on the values

that were chosen or referring to other works that used the same settings. This leaves 14 studies where a technique was used to determine the optimal hyper parameters. Here, Grid Search was the most frequent technique (five times), followed by trial and error (twice) and stepwise adjustment (twice).

4.7 NN architecture

The architecture refers to the structure of neural networks, more specifically to the determination of the number of hidden layers and hidden units used. Therefore, it applies only to studies which used NNs. 30 of the 35 publications used some kind of neural network. All gave information about the architecture itself i.e., the number of neurons and layers. Ten publications were marked as ‘ad hoc’, where no clear approach could be identified. A total number of 11 studies used a trial-and-error approach to determine network architecture, whereas three used a controlled, stepwise approach. Three studies referred to literature to justify their network architecture. Other approaches (Grid Search, predefined) were used once, respectively.

4.8 Evaluation metrics

This choice of the evaluation metric is dependent on the type of problem (classification, regression, clustering). Generally, 25 different metrics were used and only nine studies used just one metric, whereas 4 studies did not provide this information at all.

4.9 Reference model

The use of a reference is important in order to properly evaluate the result of a study. Usually, the performance of a trained model is compared to the performance of some pre-existing models, approaches, algorithms or simulation approaches. It has to be trained or computed with the same data basis and evaluated on the same metrics to allow a fair comparison. A major part of studies did so in some way, namely 27, leaving 8 studies without reference for the evaluation. Three publications developed multiple ML models of different kinds and compared them to each other. This can be seen as a form of reference, as these studies aim to find the best performing algorithm for their data and use case. Most of the other studies concentrated on a maximum of two ML methods and compared them to a simpler model or, in two cases, to common models and methods of their domain[KS07][SMJ02]. Most regression based studies used simple regression models as reference.

5 Conclusion

To survey a field as large as Artificial Intelligence in Environmental Computer Science a data driven approach is a suitable first step. We showed that there is indeed a widespread use of methods of Artificial Intelligence in Environmental (Computer) Science. There are 18.000 publications concerning these topics to be found in the databases Web of Science alone. According to our analysis the research activities of ECS in regard to AI are driven mainly by seven different fields. Based on the interconnectivity of the terms these clusters could be derived and their fields are based on the most common terms in each cluster. Finally, we analysed the five most cited publications of each field to investigate in detail their ML processes and documentation of these processes.

The way of publication suggests that the field of AIECS is predominantly populated by domain researchers who use AI as a tool. The way of publication suggests that the field of AIECS is predominantly populated by domain researchers who use AI as a tool, which is the most important finding of this work. AI in ECS is mostly seen as a set of methods, a helpful tool like GIS, Statistics or Data Visualization. It is a means to an end, a way to get the best out of one's own data. This bears a second important realization. The goal of most publications in our sample for in-depth analysis is to get a good prediction from a very specific dataset. During our literature research we often found that most researcher are mainly interested to find a good prediction for an environmental problem at a specific location. Seldom, the goal is to develop a general applicable model for a specific issue. This is entirely legitimate and is a question that naturally comes up again and again in location-related sciences. However, this signifies a divergence between research in pure Artificial Intelligence and Environmental Science. Research in pure AI aims to design models and develop algorithms and methods that are applicable as broadly as possible. On the other hand, Environmental Sciences use AI as one of many tools to achieve their actual goal – often times, the determination of variable Y in place Z. There are of course cases in which the application of AI in ECS is analogous to developments in AI.

However, this often specific application can also be seen more critical. Many if not most of the publications investigated can not be reproduced and therefore truly evaluated. The problems detected throughout this work resemble those that [FW19] describe as reproducibility crisis - studies not being reproducible and missing transparency in description of procedure. In particular, establishing guidelines and rules for data sharing and transparency could be ways to achieve this in the long term, see e.g. [FW19]. And while there are already steps taken by different agencies such as the explicit publication and citation of data sets, requirements by journals and conferences, we want to emphasize this problem.

This study is only a first step to provide understanding of the field and is based on a pure data based approach to identify the clusters of the field and pre-select the most cited studies in each field. While we argue that it provides a starting point in particular in regard to the difference in approaching AI for computer scientist and environmental scientists - and their differences in their goals - a more in depth survey paper which investigates this field of

AI in ECS is still missing. We only focussed on the most cited publications, which could have led to a bias to older publications. For future work, it would also be interesting to investigate the seven sub fields more in-depth as to why those fields seem to publish the majority of AI related work in the ECS field. While remote sensing can be easily deduced to be influenced by the increasing availability of remote sensing data and their huge amount of data as well as their similarity to the AI field of vision, the other fields are not so obvious for us. Other potential evaluation criteria, such as the use of the well-known KDD process or CRISP-DM, were considered, but proved to be difficult to asses in an automated manner. The aim was to first focus purely on the AI models and the core-application of those by the researchers. Further studies, which investigate the use of standardized approaches to the application of AI in their broader sense, would provide interesting insights. Finally, a more in-depth evaluation of different time periods would provide a benefit to our understanding of the evolution of the field. Almost a third of the examined publications, around 5600, were published in 2019 and 2020. This would allow to capture the rapid advances in the field. However, this was beyond the scope of this work. In our view this can only be created collaboratively by researchers from both fields to provide the overview and understanding needed to provide this overview. It could then be possible to truly leverage the benefits of AI in ECS.

Acknowledgement

This publication was done in the context of the project Nitrat-Monitoring 4.0 - Intelligente Systeme zur nachhaltigen Reduzierung von Nitrat im Grundwasser (NiMo 4.0), <https://nimo-projekt.de/>, funded by the BMU with Förderkennzeichen FKZ 67KI2048.

Bibliography

- [Ar05] Araújo, Miguel B; Pearson, Richard G; Thuiller, Wilfried; Erhard, Markus: Validation of species–climate impact models under climate change. *Global change biology*, 11(9):1504–1513, 2005.
- [AWBAA05] Abdul-Wahab, Sabah A; Bakheit, Charles S; Al-Alawi, Saleh M: Principal component and multiple regression analysis in modelling of ground-level ozone and factors affecting its concentrations. *Environmental Modelling & Software*, 20(10):1263–1271, 2005.
- [Br00] Brown, Warick M; Gedeon, TD; Groves, DI; Barnes, RG: Artificial neural networks: a new method for mineral prospectivity mapping. *Australian journal of earth sciences*, 47(4):757–770, 2000.
- [Br10] Brekken, Ted KA; Yokochi, Alex; Von Jouanne, Annette; Yen, Zuan Z; Hapke, Hannes Max; Halamay, Douglas A: Optimal energy storage sizing and control for wind power applications. *IEEE Transactions on Sustainable Energy*, 2(1):69–77, 2010.

- [Bu16] Bui, Dieu Tien; Tuan, Tran Anh; Klempe, Harald; Pradhan, Biswajeet; Revhaug, Inge: Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. *Landslides*, 13(2):361–378, 2016.
- [Ch14] Chen, Yushi; Lin, Zhouhan; Zhao, Xing; Wang, Gang; Gu, Yanfeng: Deep learning-based classification of hyperspectral data. *IEEE Journal of Selected topics in applied earth observations and remote sensing*, 7(6):2094–2107, 2014.
- [Ch16] Chen, Yushi; Jiang, Hanlu; Li, Chunyang; Jia, Xiuping; Ghamisi, Pedram: Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 54(10):6232–6251, 2016.
- [Ch20] Chowdhary, KR: *Fundamentals of Artificial Intelligence*. Springer, 2020.
- [CZH16] Cheng, Gong; Zhou, Peicheng; Han, Junwei: Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 54(12):7405–7415, 2016.
- [Fe15] Feng, Xiao; Li, Qi; Zhu, Yajie; Hou, Junxiong; Jin, Lingyan; Wang, Jingjie: Artificial neural networks forecasting of PM_{2.5} pollution using air mass trajectory based geographic model and wavelet transformation. *Atmospheric Environment*, 107:118–128, 2015.
- [Fi17] Fisher, Douglas: A selected summary of AI for computational sustainability. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. volume 31, 2017.
- [FW19] Fidler, Fiona; Wilcox, John: Reproducibility of Scientific Results. In (Zalta, Edward N., ed.): *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Winter 2019 edition, 2019.
- [GC06] Grivas, G; Chaloulakou, A: Artificial neural network models for prediction of PM₁₀ hourly concentrations, in the Greater Area of Athens, Greece. *Atmospheric environment*, 40(7):1216–1229, 2006.
- [GD99] Gardner, MW; Dorling, SR: Neural network modelling and prediction of hourly NO_x and NO₂ concentrations in urban air in London. *Atmospheric Environment*, 33(5):709–719, 1999.
- [Gh09] Ghobadian, Barat; Rahimi, Hadi; Nikbakht, AM; Najafi, Gholamhassan; Yusaf, TF: Diesel engine performance and exhaust emission analysis using waste cooking biodiesel fuel with an artificial neural network. *Renewable energy*, 34(4):976–982, 2009.
- [Go09] Gomes, Carla P: Computational sustainability: Computational methods for a sustainable environment, economy, and society. *The Bridge*, 39(4):5–13, 2009.
- [Go19] Gomes, Carla; Dietterich, Thomas; Barrett, Christopher; Conrad, Jon; Dilkina, Bistra; Ermon, Stefano; Fang, Fei; Farnsworth, Andrew; Fern, Alan; Fern, Xiaoli et al.: Computational sustainability: Computing for a better world and a sustainable future. *Communications of the ACM*, 62(9):56–65, 2019.
- [Gu12] Guo, Zhenhai; Zhao, Weigang; Lu, Haiyan; Wang, Jianzhou: Multi-step forecasting for wind speed using a modified EMD-based artificial neural network model. *Renewable Energy*, 37(1):241–249, 2012.

- [HDT02] Huang, Chengquan; Davis, LS; Townshend, JRG: An assessment of support vector machines for land cover classification. *International Journal of remote sensing*, 23(4):725–749, 2002.
- [HGS95] Hsu, Kuo-lin; Gupta, Hoshin Vijai; Sorooshian, Soroosh: Artificial neural network modeling of the rainfall-runoff process. *Water resources research*, 31(10):2517–2530, 1995.
- [Hi95] Hilty, LM; Page, Bernd; Radermacher, FJ; Riekert, W-F: Environmental informatics as a new discipline of applied computer science. In: *Environmental Informatics*, pp. 1–11. Springer, 1995.
- [HKH04] Hamed, Maged M; Khalafallah, Mona G; Hassanien, Ezzat A: Prediction of wastewater treatment plant performance using artificial neural networks. *Environmental Modelling & Software*, 19(10):919–928, 2004.
- [Je19] Jetzke, Tobias; Richter, Stephan; Ferdinand, Jan-Peter; Schaaf, Samer: *Künstliche Intelligenz im Umweltbereich*. Umweltbundesamt, 2019.
- [Ka09] Karatepe, E; Hiyama, T et al.: Artificial neural network-polar coordinated fuzzy controller based maximum power point tracking control under partially shaded conditions. *IET Renewable Power Generation*, 3(2):239–253, 2009.
- [Ka20] Kayid, Amr: *The role of Artificial Intelligence in future technology*. 2020.
- [KS02] Kemper, Thomas; Sommer, Stefan: Estimate of heavy metal contamination in soils after a mining accident using reflectance spectroscopy. *Environmental science & technology*, 36(12):2742–2747, 2002.
- [KS07] Khandelwal, Manoj; Singh, TN: Evaluation of blast-induced ground vibration predictors. *Soil Dynamics and Earthquake Engineering*, 27(2):116–125, 2007.
- [Le04] Lee, Saro; Ryu, Joo-Hyung; Won, Joong-Sun; Park, Hyuck-Jin: Determination and application of the weights for landslide susceptibility mapping using an artificial neural network. *Engineering Geology*, 71(3-4):289–302, 2004.
- [Li21] Liesch, Tanja; Bruns, Julian; Abecker, Andreas; Hilbring, Désirée; Karimanzira, Divas; Martin, Tobias; Wagner, Martin; Wunsch, Andreas; Fischer, Thilo: *Nitrat-Monitoring 4.0*. INFORMATIK 2020, 2021.
- [Ma16] Mayr, Andreas; Klambauer, Günter; Unterthiner, Thomas; Hochreiter, Sepp: DeepTox: toxicity prediction using deep learning. *Frontiers in Environmental Science*, 3:80, 2016.
- [MD00] Maier, Holger R; Dandy, Graeme C: Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Environmental modelling & software*, 15(1):101–124, 2000.
- [MH96] Minns, AW; Hall, MJ: Artificial neural networks as rainfall-runoff models. *Hydrological sciences journal*, 41(3):399–417, 1996.
- [Mo20] Mondal, Bhaskar: Artificial intelligence: state of the art. *Recent Trends and Advances in Artificial Intelligence and Internet of Things*, pp. 389–425, 2020.

- [MRH98] Mohandes, M; Rehman, S; Halawani, TO: Estimation of global solar radiation using artificial neural networks. *Renewable energy*, 14(1-4):179–184, 1998.
- [Na04] Nayak, Puma C; Sudheer, KP; Rangan, DM; Ramasastri, KS: A neuro-fuzzy computing technique for modeling hydrological time series. *Journal of Hydrology*, 291(1-2):52–66, 2004.
- [NSD04] Nadal, M; Schuhmacher, M; Domingo, JL: Metal pollution of soils and vegetation in an area with petrochemical industry. *Science of the total environment*, 321(1-3):59–69, 2004.
- [PL10] Pradhan, Biswajeet; Lee, Saro: Landslide susceptibility assessment and factor effect analysis: backpropagation artificial neural networks and their comparison with frequency ratio and bivariate logistic regression modelling. *Environmental Modelling & Software*, 25(6):747–759, 2010.
- [PPG12] Pourghasemi, Hamid Reza; Pradhan, Biswajeet; Gokceoglu, Candan: Application of fuzzy logic and analytical hierarchy process (AHP) to landslide susceptibility mapping at Haraz watershed, Iran. *Natural hazards*, 63(2):965–996, 2012.
- [RN10] Russell, Stuart J; Norvig, Peter: *Artificial Intelligence: A Modern. Approach*, 2nded., PrenticeHall, December2002, 2010.
- [Ro12] Rodriguez-Galiano, Victor Francisco; Ghimire, Bardan; Rogan, John; Chica-Olmo, Mario; Rigol-Sanchez, Juan Pedro: An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 67:93–104, 2012.
- [RS20] Raghu, Maithra; Schmidt, Eric: A survey of deep learning for scientific discovery. *arXiv preprint arXiv:2003.11755*, 2020.
- [Sc19] Schuchmann, Sebastian: *Analyzing the Prospect of an Approaching AI Winter*. PhD thesis, PhD thesis, 05 2019, 2019.
- [SD18] SDG, UN: , *Sustainable development goals*, 2018.
- [Sh18] Shen, Chaopeng: A transdisciplinary review of deep learning research and its relevance for water resources scientists. *Water Resources Research*, 54(11):8558–8593, 2018.
- [SMJ02] Shahin, Mohamed A; Maier, Holger R; Jaksa, Mark B: Predicting settlement of shallow foundations using neural networks. *Journal of Geotechnical and Geoenvironmental Engineering*, 128(9):785–793, 2002.
- [So07] Sousa, SIV; Martins, Fernando Gomes; Alvim-Ferraz, MCM; Pereira, Maria C: Multiple linear regression and artificial neural networks based on principal components to predict ozone concentrations. *Environmental Modelling & Software*, 22(1):97–103, 2007.
- [VBB13] Valipour, Mohammad; Banihabib, Mohammad Ebrahim; Behbahani, Seyyed Mahmood Reza: Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir. *Journal of hydrology*, 476:433–441, 2013.
- [VEW13] Van Eck, Nees Jan; Waltman, Ludo: *VOSviewer manual*. Leiden: Univeristeit Leiden, 1(1):1–53, 2013.

- [Wa09] Wang, Wen-Chuan; Chau, Kwok-Wing; Cheng, Chun-Tian; Qiu, Lin: A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. *Journal of hydrology*, 374(3-4):294–306, 2009.
- [YBL11] Yue, Tao; Briand, Lionel C; Labiche, Yvan: A systematic review of transformation approaches between user requirements and analysis models. *Requirements engineering*, 16(2):75–99, 2011.
- [YD08] Yetilmezsoy, Kaan; Demirel, Sevgi: Artificial neural network (ANN) approach for modeling of Pb (II) adsorption from aqueous solution by Antep pistachio (*Pistacia Vera L.*) shells. *Journal of hazardous materials*, 153(3):1288–1300, 2008.
- [Yi09] Yilmaz, Işık: Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: a case study from Kat landslides (Tokat—Turkey). *Computers & Geosciences*, 35(6):1125–1138, 2009.
- [Yu20] Yu, Manzhu; Bambacus, Myra; Cervone, Guido; Clarke, Keith; Duffy, Daniel; Huang, Qunying; Li, Jing; Li, Wenwen; Li, Zhenlong; Liu, Qian et al.: Spatiotemporal event detection: a review. *International Journal of Digital Earth*, 13(12):1339–1365, 2020.
- [Za19] Zaldivar, Andrew; Hutchinson, Ben; Spitzer, Elena; Raji, Inioluwa Deborah; Vasserman, Lucy; Mitchell, M.; Barnes, Parker; Wu, Simone Sanoian McCloskey; Gebru, Timnit, eds. *Model Cards for Model Reporting*, 2019.
- [Zo08] Zorlu, K; Gokceoglu, C; Ocakoglu, F; Nefeslioglu, HA; Acikalin, SJEG: Prediction of uniaxial compressive strength of sandstones using petrography-based models. *Engineering Geology*, 96(3-4):141–158, 2008.

9. Workshop Umweltinformatik zwischen Nachhaltigkeit und Wandel (UINW 2021)

Digitale Schlüsseltechnologien für eine Nachhaltige Digitalisierung

Digitale Leuchtturmprojekte der Umweltverwaltung Baden-Württemberg

Jan Tomaschek¹ und Ulrike Kugler²

Abstract: Die Digitalisierung verändert unser Leben mit großer Geschwindigkeit. Die neuen digitalen Möglichkeiten bieten vielfältige Chancen für Umwelt-, Klima- und Naturschutz. Gleichzeitig kann der digitale Wandel auch zu einem gesteigerten Ressourcenverbrauch führen. Daher ist es wichtig, dass wir die richtigen Rahmenbedingungen für eine nachhaltige Digitalisierung setzen. Den digitalen Wandel zu nutzen, um die Nachhaltigkeit zu fördern, ist ein zentrales Anliegen der Landesregierung Baden-Württembergs. Das Ministerium für Umwelt, Klima und Energiewirtschaft Baden-Württemberg nimmt sich dieser Leitlinie mit großem Einsatz an. Durch digitale Leuchtturmprojekte soll der Nutzen der Digitalisierung erlebbar gemacht und aufgezeigt werden, welche Chancen der digitale Wandel für den Schutz von Umwelt, Klima und Natur bietet. Digitale Schlüsseltechnologien spielen dabei eine wichtige Rolle.

Keywords: Digitalisierung, Nachhaltigkeit, Schlüsseltechnologien, Green IT, Umweltinformationen

Einleitung

Der digitale Wandel ändert die Art wie wir leben und arbeiten auf eine radikale Weise und mit rasender Geschwindigkeit. Die Corona-Pandemie zeigt uns unmittelbar, wie wichtig ein systematischer Digitalisierungsprozess für uns ist. Die Pandemie darf uns aber nicht davon ablenken, dass der voranschreitende Klimawandel und die Zerstörung intakter Ökosysteme unser Handeln dringend erfordern.

Informationen über den Zustand und die Veränderungen der Umwelt bereitzustellen, ist eine wesentliche Voraussetzung für einen effektiven Klima-, Natur- und Umweltschutz. Seit mehr als drei Jahrzehnten erfüllt das Umweltinformationssystem Baden-Württemberg (UIS BW) diesbezüglich vielfältige staatliche Umweltaufgaben. Es hat sich in den vergangenen Jahren zu einem strategischen Instrument der Landespolitik für einen vorsorgenden Umwelt- und Klimaschutz entwickelt. Die Fachverfahren werden

¹ Ministerium für Umwelt, Klima und Energiewirtschaft Baden-Württemberg, Referat 15 – Informationstechnik, IT-Leitstelle, UIS, Nachhaltige Digitalisierung, Kernerplatz 9, 70182 Stuttgart. E-Mail: jan.tomaschek@um.bwl.de

² Ministerium für Umwelt, Klima und Energiewirtschaft Baden-Württemberg, Referat 15 – Informationstechnik, IT-Leitstelle, UIS, Nachhaltige Digitalisierung, Kernerplatz 9, 70182 Stuttgart. E-Mail: ulrike.kugler@um.bwl.de

in der Regel zentral entwickelt und entsprechend ihrer Aufgaben von den Behörden (ca. 3500 Nutzer) der Umweltverwaltung genutzt. Für die Auswertung und den verwaltungsinternen Zugriff auf die Umweltinformationen steht mit dem Berichtssystem eine gemeinsame Anwendung zur Verfügung, die von ca. 4500 Anwenderinnen und Anwendern genutzt wird. Aggregierte und standardisierte Berichtsdaten sind Entscheidungsgrundlage für Führungskräfte in Politik und Verwaltung und dienen zur Erfüllung zahlreicher Berichtspflichten. Entsprechend dem Umweltverwaltungsgesetz des Landes eröffnet das UIS BW auch allen Bürgerinnen und Bürgern den freien Zugang zu aktuellen Informationen. Es nutzt eine serviceorientierte IT-Architektur aus vielen Einzelkomponenten und Fachanwendungen, die modular ineinandergreifen. Mit neuen digitalen Schlüsseltechnologien verbinden sich Herausforderungen und Chancen für eine UIS-Weiterentwicklung.

Digitale Schlüsseltechnologien, wie „Künstliche Intelligenz“, Blockchain, Sensorik und Robotik können den Umwelt-, Klima- und Naturschutz weiter voranbringen und unterstützen, etwa, indem sie die erneuerbaren Energien effizient in die Energiewirtschaft integrieren und Prozesse zum Beispiel in der industriellen Produktion optimieren. Eine jüngst veröffentlichte Studie kommt zu dem Ergebnis das digitale Technologien mehr als die Hälfte dazu beitragen, dass Deutschland seine Klimaziele bis zum Jahr 2030 erfüllt [KI21].

Dabei darf jedoch nicht außer Acht gelassen werden, dass der Digitalisierungsprozess selbst zu einem wachsenden Bedarf an Rechenkapazität führt, welcher wiederum mit einem stetig steigenden Energieverbrauch einhergeht. Berechnungen zeigen, dass nicht zuletzt durch die Corona-Pandemie der Stromverbrauch von Rechenzentren in Deutschland im Jahr 2020 - trotz Effizienzgewinnen - weiter angestiegen ist und mittlerweile bei 16 TWh liegt [Hi21]. Dies entspricht etwa 3 % der Bruttostromerzeugung Deutschlands [Br21].

Nicht vergessen darf man auch den großen Bedarf an Hochtechnologierohstoffen, welche in elektronischen Endgeräten verbaut sind. In einem Smartphone befinden sich beispielweise über 50 Arten von Metallen, darunter Gold, Palladium, Kupfer, Nickel, Silizium, Magnesium, Platin, Neodym, Aluminium, Tantal, Zinn und Eisen [Bo20]. Diese Rohstoffe werden immer noch viel zu oft mit großen Schäden für Natur und Umwelt und unter schlechten sozialen Bedingungen dem Boden entnommen. Am Ende des mitunter kurzen Lebenswegs eines Produktes landen große Rohstoffmengen im Abfall. Die unterschiedlichen Elemente zurückzugewinnen und wiederzuverwerten gelingt nach wie vor nur in sehr begrenztem Umfang [Sa19].

Um die Möglichkeiten des digitalen Wandels zu nutzen und gleichzeitig negative Auswirkung zu minimieren, bedarf es demnach einer konkreten Herangehensweise und guter Rahmenbedingungen.

Nachhaltige Digitalisierung für einen effizienten Umwelt-, Klima- und Naturschutz

Den digitalen Wandel zu nutzen, um die Nachhaltigkeit zu fördern, ist ein zentrales Anliegen der Landesregierung Baden-Württembergs. Bereits im Juli 2017 wurde deshalb die Digitalisierungsstrategie digital@bw beschlossen. Dabei wird die Digitalisierung als eine Chance für Innovation und Nachhaltigkeit in Baden-Württemberg gesehen, die mit und für die Menschen im Land genutzt werden soll. So soll Baden-Württemberg zu einer Leitregion bei intelligenten, ressourcensparenden und klimaschonenden Technologien werden, und es soll gelingen, wirtschaftliches Wachstum vom Ressourcenverbrauch zu entkoppeln.

Das Ministerium für Umwelt, Klima und Energiewirtschaft Baden-Württemberg (UM BW) nimmt sich dieser Leitlinie mit großem Einsatz an. Durch digitale Leuchtturmprojekte soll der Nutzen der Digitalisierung erlebbar gemacht und aufgezeigt werden, welche Chancen der digitale Wandel für den Schutz von Umwelt, Klima und Natur bietet. Digitale Schlüsseltechnologien spielen dabei eine wichtige Rolle. Einige dieser Förderprogramme und Maßnahmen seien im Folgenden exemplarisch dargestellt.

Beispielweise unterstützt das UM BW im Förderprogramm „Digitalisierung & Ultraeffizienz“ Projekte mit dem Ziel, Produktionsprozesse in Betrieben so zu optimieren, dass Wachstum und Verbrauch weitestgehend voneinander entkoppelt werden. Materialien und Energie müssen dabei so effektiv und effizient wie möglich eingesetzt werden. Durch Digitalisierung und neue Technologien soll eine leise, emissions- und abfallarme Produktion entstehen. Digitale Vernetzung kann beispielweise dabei helfen, Wertschöpfungsketten zu optimieren. So kann zeit- und ortsnahe eine Möglichkeit zur Weiterverwertung von Materialien gefunden werden, die bei der Produktion wortwörtlich „abfallen“ (Stichwort: Kreislaufwirtschaft 4.0).

In einem der vom UM BW geförderten Projekte sollen daher Daten für die Kreislaufwirtschaft mittels Distributed Ledger beziehungsweise Blockchain-Technologie abgebildet und so digitale Zwillinge kreiert werden. Hierdurch soll es gelingen, den kompletten Lebensweg eines Produktes von der Rohstoffgewinnung über die Produktion bis die Entsorgung lückenlos zu verfolgen. Dies ermöglicht es, die Einhaltung von Umwelt- und Sozialstandards zu überwachen, auch kleine Rohstoffströme sicher zu verfolgen und Rohstoffe besser wiederzuverwerten. In einem anderen Förderprojekt soll durch die Anwendung „Künstlicher Intelligenz“ in der industriellen Produktion die Ressourceneffizienz gesteigert werden, indem etwa optimierte Steuerungen für prozesstechnische Anlagen entwickelt werden [Ha21a].

Im Förderschwerpunkt „Smarte Umweltdaten“ sollen neue digitale Technologien zur Erfassung, Analyse und Bereitstellung von Umweltdaten erprobt werden. Ein Beispiel hierfür ist das Projekt „SAMOSEE-BW“, das das Umweltministerium Baden-Württemberg gemeinsam mit dem Institut für Seenforschung (ISF) der Landesanstalt für Umwelt Baden-Württemberg (LUBW) durchführt. Das Projekt zielt darauf ab,

Fernerkundungsdaten für ein landesweites Gewässermonitoring aufzubereiten. Satellitenbasierte Fernerkundungsdaten ermöglichen es beispielsweise, den Zustand von Gewässern landesweit und flächendeckend zu erfassen. Sie helfen dabei, die räumliche und zeitliche Entwicklung der Seen zu bestimmen und Karten zu erstellen. Die regelmäßige Beurteilung des Gewässerzustands ist wichtig, denn die Wasserqualität hat zum Beispiel direkte Auswirkungen auf die Trinkwasserentnahme, die Artenvielfalt und die Eignung eines Gewässers als Badesee. Mit herkömmlichen In-situ-Messverfahren ist ein Monitoring der Gewässer logistisch nicht zu realisieren. Daher sind Fernerkundungsdaten, die Satelliten von den Gewässern erstellen, für die Wasserwirtschaft und die Umweltverwaltung ein wichtiges Hilfsmittel. Aktuell können Wasserwirtschaft und Umweltverwaltung diese Daten jedoch ohne Spezialsoftware nur schwer in ihre tägliche Arbeitsroutine einbinden. Daher sollen im Projekt aus den Daten Informationsprodukte und -formate erstellt werden, die möglichst intuitiv und ohne besondere Software genutzt werden können. Mögliche Daten aus der Fernerkundung sind dabei zum Beispiel Sichttiefe, Trübung, Oberflächentemperatur, Algenwachstum und Eisbedeckung [Ha21b].

Um den Zustand von Natur und Umwelt regelmäßig bestimmen zu können, erhebt die Umweltverwaltung Baden-Württemberg große Mengen an Umweltdaten. Die Daten werden mit Hilfe verschiedener Softwareanwendungen im Umweltinformationssystem Baden-Württemberg (UIS BW) zentral gespeichert, analysiert und verfügbar gemacht. Mit Hilfe neuer digitaler Technologien kann das Umweltinformationssystem Baden-Württemberg zukünftig den Vollzug innerhalb der Verwaltung besser unterstützen, aber auch den Zugang der Öffentlichkeit zu Umweltinformationen erleichtern. Durch "Citizen Science" können Bürgerinnen und Bürger helfen, selbst Umweltdaten zu erheben. Der wachsende Markt für kostengünstige Sensoren mit Anbindung zum Internet (Internet of Things (IoT)) macht dies möglich. Im Projekt „CrowdSensing für Bodensee Online“ haben Bootsbesitzerinnen und Bootsbesitzer mit Hilfe mobiler Sensorik und Karten Temperaturmessungen im See vorgenommen. Diese Messungen werden schließlich mit offiziellen Messdaten abgeglichen. Sind die gesammelten Daten der Bürgerinnen und Bürger ausreichend valide, kann das Vorhaben auf weitere Gewässer und Messdaten ausgeweitet werden [Ha21c].

Im Zusammenspiel mit der Digitalisierungsstrategie digital@bw wirkt die Landesstrategie Green IT, die 2014 initialisiert wurde. Sie hat zum Ziel, die Informations- und Kommunikationstechnologie der Landesverwaltung ressourceneffizienter zu gestalten und nimmt dabei den gesamten Lebensweg (Produktion, Nutzung, Weiterverwertung) in den Blick [Di18]. Ein Thema, das mehr und mehr in den Fokus rückt, ist der softwareinduzierte Stromverbrauch beziehungsweise die ressourceneffiziente Programmierung. Während es im forschungsnahen Bereich des High Performance Computing dazu bereits Vorgehensmodelle gibt [En18], [Hp17], [En17], steht der verwaltungsnahen Bereich der Anwendungssoftware, z.B. des UIS BW, hier eher noch am Anfang. In einer Machbarkeitsstudie soll daher anhand eines in der Umweltverwaltung Baden-Württemberg verwendeten Softwareprogramms beleuchtet werden, wie sich der

Kriterienkatalog des „Blauen Engels für ressourcen- und energieeffiziente Softwareprodukte“ anwenden lässt und welche technischen und personellen Aufwände damit verbunden sind [Re20]. Daraus sollen Erkenntnisse abgeleitet werden, wie sich zukünftig entsprechende Anforderungen in Entwicklungsvorgaben zu verwaltungseigenen Programmen integrieren lassen.

Ein weiteres Themenfeld der Landesstrategie ist die Beschaffung. Hierfür konnten entsprechende Vorschriften und Richtlinien wie die Verwaltungsvorschrift (VwV) Beschaffung und die VwV IT Standards erweitert werden (Vergabe und IT-Richtlinien des Landes), um der Beschaffung nachhaltiger IT eine rechtliche Grundlage zu geben. Auf Basis dieser Vorschriften wurde im Jahr 2020 eine Computermaus als Teil der Standardausrüstung der IT-Arbeitsplätze des Landes beschafft, die unter transparenten und fairen Arbeitsbedingungen hergestellt wird [Fa20].

Fazit

Digitale Schlüsseltechnologien wie „Künstliche Intelligenz“, Blockchain, Sensorik und Robotik können eine Nachhaltige Entwicklung unterstützen, wenn die entsprechenden Rahmenbedingungen dafür beachtet werden. Wichtig ist dabei, dass es sich nicht um einen rein technokratischen Prozess handelt, sondern dass konkrete Mehrwerte erzielt und Technologien nicht um ihrer selbst willen eingesetzt werden. So darf nicht außer Acht gelassen werden, dass eine effiziente Verwaltung auf eine ebenso effiziente IT angewiesen ist. Die durch die Corona-Pandemie gewonnene Geschwindigkeit muss auch zukünftig für einen kontinuierlichen Digitalisierungsprozess genutzt werden.

Innovative digitale Methoden und Werkzeuge werden durch das Ministerium für Umwelt-, Klima- und Energiewirtschaft Baden-Württemberg in verschiedenen Leuchtturmprojekten erprobt und weiterentwickelt. Den Rahmen hierfür bildet die landesweite Digitalisierungsstrategie digital@bw. Die gewonnenen Erkenntnisse helfen dabei, das bestehende Umweltinformationssystem Baden-Württemberg (UIS BW) zu ertüchtigen und den Bedarf der Gesellschaft nach Umweltinformation besser zu decken. Mittlerweile nutzen etwa 1 Mio. Bürgerinnen und Bürger das öffentlich zugängliche System Umweltdaten- und Karten-Online (UDO) um sich über Umweltbelange zu informieren.

Gleichzeitig benötigt der Digitalisierungsprozess aber auch entsprechende Leitplanken, damit er nicht zu einem steigenden Verbrauch von Ressourcen führt. Mit der Landesstrategie Green IT sollen innerhalb der Verwaltung die negativen Digitalisierungseffekte hinsichtlich Energie- und Rohstoffverbrauch, die über den Lebensweg von Hard- und Software auftreten, möglichst minimiert werden. So konnten durch die Umsetzung von Effizienzmaßnahmen in einem der Landesrechenzentren 100 MWh Strom jährlich eingespart werden [Bi19].

Literaturverzeichnis

- [Bi19] BITBW spart Energie durch Kaltgangeinhausung. Hrsg. Ministerium für Umwelt, Klima und Energiewirtschaft Baden-Württemberg, 2019. <https://green-it.baden-wuerttemberg.de/kompetenzstelle-green-it/die-landesstrategie>, zuletzt abgerufen am 22.04.2021
- [Bo20] Bookhagen, B.; Bastian, D.: Metalle in Smartphones, in: Commodity TopNews (65), Hrsg: Bundesanstalt für Geowissenschaften und Rohstoffe, Hannover, Oktober 2020.
- [Br21] Bruttostromerzeugung in Deutschland. Hrsg. Statistisches Bundesamt. <https://www.destatis.de/DE/Themen/Branchen-Unternehmen/Energie/Erzeugung/Tabellen/bruttostromerzeugung.html>, zuletzt abgerufen am 22.04.2021.
- [Di18] Die Landesstrategie „Green IT 2020“. Hrsg. Ministerium für Umwelt, Klima und Energiewirtschaft Baden-Württemberg, 2018. <https://green-it.baden-wuerttemberg.de/kompetenzstelle-green-it/die-landesstrategie>, zuletzt abgerufen am 22.04.2021
- [En17] ENVELOPE - Effizienz und Zuverlässigkeit: Selbstorganisation in HPC-Systemen. BMBF-Projekt im Rahmen des Vorhabens "Grundlagenorientierte Forschung für HPC-Software im Hoch- und Höchstleistungsrechnen". Hrsg. Karlsruher Institut für Technologie, 2017. <http://envelope.itec.kit.edu/>, zuletzt abgerufen am 22.04.2021.
- [En18] Energy Efficiency - Enable energy efficient usage of high-performance computing systems through research on infrastructure requirements and numerical. Solutions. Hrsg. Höchstleistungsrechenzentrum Stuttgart, 2018. https://www.hlrs.de/fileadmin/sys/public/aboutus/research/brochures/Energy_Efficiency_at_HLRS.pdf, zuletzt abgerufen am 22.04.2021.
- [Fa20] Fairer Handel bei Computermäusen - 60.000 faire Mäuse. Hrsg. Taz, 18.3.2020. <https://taz.de/Fairer-Handel-bei-Computermäusen/!5671183/>, zuletzt abgerufen am 22.04.2021.
- [Ha21a] Handlungsfeld Ressourcen- und Energieeffizienz - Förderprogramm Digitalisierung & Ultraeffizienz. Hrsg. Ministerium für Umwelt, Klima und Energiewirtschaft Baden-Württemberg, 01.04.2021. <https://um.baden-wuerttemberg.de/de/umwelt-natur/nachhaltigkeit/nachhaltige-digitalisierung/projekte/digitalisierung-und-ultraeffizienz/>, zuletzt abgerufen am 22.04.2021
- [Ha21b] Handlungsfeld Smarte Umweltdaten – Projekt SAMOSEE-BW. Hrsg. Ministerium für Umwelt, Klima und Energiewirtschaft Baden-Württemberg, 19.02.2021. <https://um.baden-wuerttemberg.de/de/umwelt-natur/nachhaltigkeit/nachhaltige-digitalisierung/projekte/samosee-bw/>, zuletzt abgerufen am 22.04.2021
- [Ha21c] Handlungsfeld Smarte Umweltdaten – Projekt „CrowdSensing“ für Bodensee Online. Hrsg. Ministerium für Umwelt, Klima und Energiewirtschaft Baden-Württemberg, 01.04.2021. <https://um.baden-wuerttemberg.de/de/umwelt-natur/nachhaltigkeit/nachhaltige-digitalisierung/projekte/crowdsensing-bodensee/>, zuletzt abgerufen am 22.04.2021

- [Hi21] Hintemann, R.: Rechenzentren 2020 - Cloud Computing profitiert von der Krise. Energiebedarf der Rechenzentren steigt trotz Corona weiter an. Hrsg: Borderstep Institut für Innovation und Nachhaltigkeit gGmbH, Berlin, 2021.
- [Hp17] HPC2SE - Hardware- und Leistungsorientierte Codegenerierung für Informatik und Ingenieurwesen. Hrsg. Gauß-Allianz, 2017. <https://gauss-allianz.de/de/project/title/HPC2SE>, zuletzt abgerufen am 22.04.2021.
- [Kl21] Klimaeffekte der Digitalisierung Studie zur Abschätzung des Beitrags digitaler Technologien zum Klimaschutz. Hrsg. Bitkom, 2021. https://www.bitkom.org/sites/default/files/2021-03/bitkom_studie_klimaeffekte-der-digitalisierung_final_210318.pdf, zuletzt abgerufen am 22.04.2021.
- [Re20] Ressourcen- und energieeffiziente Softwareprodukte. Hrsg. RAL gGmbH, 2020. <https://www.blauer-engel.de/de/produktwelt/elektrogeraete/ressourcen-und-energieeffiziente-softwareprodukte/ressourcen-und-energieeffiziente-softwareprodukte>, zuletzt abgerufen am 22.04.2021.
- [Sa19] Sander K., et al.: Abfallwirtschaftliche Produktverantwortung unter Ressourcenschutzaspekten (RePro). Hrsg: Umweltbundesamt, Dessau-Roßlau, 2019.

Intelligente PV-Modul Datenbank-Applikation für Predictive Maintenance von PV-Anlagen

Felix Meyer ¹, Joachim Rüter ², Grit Behrens ³, Matthias Diehl ⁴

Abstract: Während der Laufzeit von mehreren Jahren können Photovoltaik-Anlagen (PV-Anlagen) diverse Fehler entwickeln. Um diese frühzeitig zu erkennen und maximale Erträge zu gewährleisten, sind regelmäßige Leistungsüberprüfungen erforderlich. Aktuelle Wartungsmethoden sind jedoch teuer und zeitaufwendig. Dieses Paper beschreibt die Entwicklung einer Datenbank-Webapplikation, die derzeitige wissenschaftliche Erkenntnisse der KI und Big Data in der PV-Praxis nutzbar machen soll. So können PV-Wartungs-ExpertInnen bei der Fehlersuche und der Entscheidungsfindung unterstützt werden. Die Applikation ermöglicht es, PV-Moduldaten abzuspeichern und abzurufen und mittels künstlicher Intelligenz Prognosen für die Leistung und die Kennlinie eines Moduls durchzuführen. Des Weiteren hilft sie bei der Vernetzung von PV-ExpertInnen untereinander sowie zur Vertrauensbildung der NutzerInnen in KI-Systeme.

Keywords: Photovoltaik; Dunkelkennlinie; IV-Kennlinien; Predictive Maintenance; Künstliche Intelligenz; Datenbank; Applikation; Web

1 Motivation

Der Zugang zu sauberer und bezahlbarer Energie für jeden Menschen auf unserer Erde ist eine der wichtigsten Aufgaben unserer Zeit und deshalb auch eines der erklärten Ziele der Vereinten Nationen (“Sustainable Development Goal” 7 der “Agenda 2030 für Nachhaltige Entwicklung” der Vereinten Nationen, [Un15]). Ein von vielen ExpertInnen als vielversprechend angesehener Ansatz ist das Erzeugen von Energie aus Sonneneinstrahlung mittels PV-Anlagen. In der Geschwindigkeit der heutigen Entwicklung könnte PV bis zur Mitte des Jahrhunderts 30-50% des weltweit erzeugten Stroms liefern (vgl. z.B. [Qu13], S. 159). Trotz der vielen Vorteile von PV, gibt es allerdings auch noch Problemfelder, die gelöst werden müssen, um Energie aus solchen Anlagen effektiv als Teil einer erneuerbaren und sauberen Stromversorgung nutzen zu können. Da die Herstellung von PV-Modulen viel Energie benötigt, müssen Solaranlagen eine gewisse Zeit Strom erzeugen, um über

¹ FH Bielefeld, Campus Minden, Solar Computing Lab, Artilleriestraße 9, 32427 Minden, Deutschland, felix.meyer@fh-bielefeld.de

² FH Bielefeld, Campus Minden, Solar Computing Lab, Artilleriestraße 9, 32427 Minden, Deutschland, joachim.rueter@fh-bielefeld.de

³ FH Bielefeld, Campus Minden, Solar Computing Lab, Artilleriestraße 9, 32427 Minden, Deutschland, grit.behrens@fh-bielefeld.de

⁴ Photovoltaikbüro Ternius und Diehl GbR, Schönauerhofstraße 27, 65428 Rüsselsheim, Deutschland, md@pvbuero.de

ihren Lebenszyklus hinweg Energie erzeugt zu haben. Jedoch können im Laufe der Nutzung Fehler und Beschädigungen auftreten, welche die Erträge mindern. Aus diesem Grund sind regelmäßige Leistungsüberprüfungen wichtig, um auftretende Fehler frühzeitig zu erkennen und dadurch die Leistungsverluste zu minimieren. Die momentanen Testmethoden sind allerdings teuer, zeitaufwendig oder erfordern günstige Wetterbedingungen, die nicht häufig gegeben sind.

An diesen Problemen wollen wir ansetzen. In dieser Arbeit beschreiben wir die Entwicklung einer intelligenten PV-Modul Datenbank-Webanwendung, kurz “iPVModule”, die die Fehlererkennung und Leistungsbestimmung von PV-Anlagen in der Praxis vereinfachen, beschleunigen und günstiger machen soll. Sie tut dies, indem sie aktuelle wissenschaftliche Erkenntnisse einfacher in der Praxis nutzbar macht und aktuelle Methoden der KI und Vorteile von Big Data unkompliziert zur Verfügung stellt. Auf der anderen Seite profitiert auch die Wissenschaft durch neue, aktuelle Daten und die Verbindung mit den praxiserfahrenen ExpertInnen.

2 Hintergründe und verwandte Arbeiten

Ein sehr genaues Bild über die Leistungsfähigkeit und den Zustand eines Moduls geben exakte Messungen unter Laborbedingungen. Bei diesen werden die Strom-Spannungskennlinie und weitere Kenngrößen des PV-Moduls bei definierten Umgebungsbedingungen ermittelt. Einige beispielhafte Kennlinien sind in Abbildung 1 (links) zu sehen.

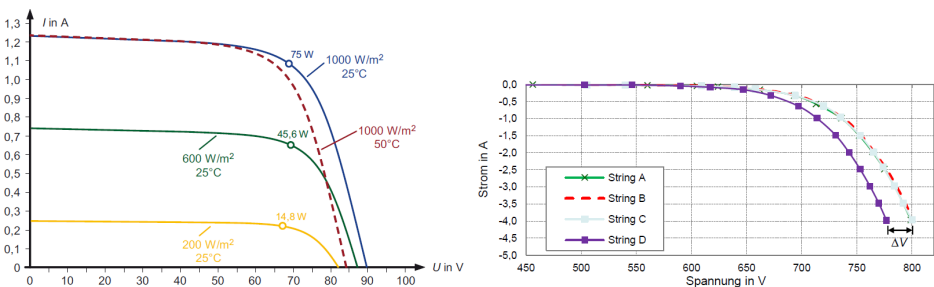


Abb. 1: Links: Hellkennlinien unter verschiedenen Umweltbedingungen (Quelle: [Me20]). Rechts: Dunkelkennlinien mit Abweichung (Quelle: [MAD18])

Auf Grund einer starken Preisreduktion von PV-Modulen über die letzten Jahre wird eine Untersuchung im Labor, nachdem die Module einmal im Feld installiert sind, jedoch wirtschaftlich immer unattraktiver. Aus diesem Grund wurden einige Vor-Ort Messverfahren entwickelt, um Fehler und Leistungsminderungen an PV-Modulen direkt erkennen zu können. Jedoch haben auch diese Verfahren erhebliche Nachteile. So benötigen einige davon stabile Wetterverhältnisse und hohe Einstrahlungen, die nicht immer vorhanden sind, andere sind sehr zeitaufwendig.

Um die Probleme der Wetterabhängigkeit und der Zeitdauer zu umgehen, wurde in [MAD17] bzw. [MAD18] ein neues Verfahren, die *String-Dunkelkennlinien-Messung*, vorgestellt. Bei diesem Verfahren wird die Dunkel-Strom-Spannungs-Kennlinie eines ganzen Strings bei Nacht aufgezeichnet. Dieses Verfahren erlaubt eine schnelle, wetterunabhängige und wiederholgenaue Analyse von PV-Modulen. Es zeigt außerdem vielversprechende Resultate bei der Fehlererkennung (vgl. [MAD17], [MAD18]). Dafür werden die Dunkelkennlinien aller Strings einer Anlage aufgezeichnet und anschließend die einzelnen Kurven miteinander verglichen. Abweichungen von der “normalen Dunkelkennlinie” können dann direkt auf Fehler hinweisen oder als Hinweis für die Notwendigkeit von tiefer gehenden Untersuchungen angesehen werden. In Abbildung 1 (rechts) ist diese Abweichung erkenntlich. Hier zeigt sich allerdings ein Problem dieses Verfahrens: Es ist stets ein Vergleich mit anderen Dunkelkennlinien erforderlich. Bei großen Anlagen mit vielen Strings können diese aufgezeichnet werden, aber vor allem bei kleinen Anlagen ist dies teilweise nicht möglich. Somit sind keine Vergleichskennlinien vorhanden.

Neben der Detektion von einigen Fehlern, wäre es außerdem wünschenswert, aus der Dunkelkennlinie auf die Hellkennlinie und damit auf die mögliche Leistung eines Strings schließen zu können (vgl. [Te19], [Me19], [Ba20a]). Damit würden - ggf. in Kombination mit anderen Daten der PV-Module, wie z.B. Monitoring-Daten - weitere Fehler erkannt werden können. Für solch ein Verfahren gibt es in der aktuellen Forschung verschiedene Ansätze (siehe [Me19], [Te19], [Ba20a], [Ba20b], [RMB21]). Einer der vielversprechendsten und am weitesten erforschten Ansätze ist der Einsatz von Neuronalen Netzen, die es ermöglichen, aus einer Strom-Spannungs-Dunkelkennlinie als Eingabe, auf eine Strom-Spannungs-Hellkennlinie bei vorgegebener Einstrahlung und Modultemperatur schließen zu können. Auf einem Testdatensatz sind damit bereits gute Prognose-Ergebnisse möglich.

Bisher wurden jedoch noch keine Verfahren für eine Nutzung in der Praxis umgesetzt, sondern finden bis jetzt nur Verwendung im wissenschaftlichen Kontext. So haben wir bereits in [Te19] ein User Interface zum Hochladen von Kennlinien-Daten entwickelt. Dieses zeigt auch erste Visualisierungen, ist aber nicht für die Nutzung in der Praxis ausgelegt und aufgebaut. Es wurde vor allem verwendet, um Daten zu sammeln, auf der eine KI trainiert werden konnte. In [Me19] wurde die Entwicklung eines Matlab-Programms beschrieben, das die Ersatzschaltbild-Parameter der Hellkennlinie aus der Dunkelkennlinie herleitet und damit eine erste Leistungsprognose ermöglicht. Des Weiteren gibt es einige lokale Modelle mit Neuronalen Netzen und andere Verfahren (siehe [Ba20a], [Ba20b], [RMB21]).

3 Komponenten der Applikation und Benefits

An den in Abschnitt 2 aufgezeigten Punkten wollen wir mit iPVMModule anknüpfen. Wir beschreiben die Entwicklung und die Möglichkeiten einer intelligenten PV-Moduldatenbank als Webapplikation, die einen Brückenschlag zwischen Wissenschaft und Praxis ziehen soll. So sollen zum einen wissenschaftliche Erkenntnisse für die Praxis nutz- und anwendbar gemacht und auf der anderen Seite ein kontinuierliches Feedback mit neuen Daten für die

Begleitforschung in der PV erreicht werden. Die Applikation umfasst dabei die folgenden Komponenten:

PV-Moduldatenbank Die PV-Moduldatenbank ermöglicht einen zentralisierten Zugriff auf verschiedene Daten von Modulen. Durch die anderen Projekte unserer Arbeitsgruppe konnte bereits eine große Datensammlung erfasst werden, welche der PV-Moduldatenbank zur Verfügung steht. Dazu zählen sowohl Daten und Spezifikationen zu Modulen als auch gemessene Hell- und Dunkelkennlinien. Monitoring-Daten von verschiedenen Anlagen können ebenfalls abgerufen werden.

Wie oben beschrieben, ermöglicht dies das Nutzen der String-Dunkelkennlinien-Messung auch für kleinere Anlagen in der Praxis. Auch für große Anlagen ergeben sich Vorteile, indem mehr Vergleichskurven zur Verfügung stehen. Bisher mussten sachkundige ExpertInnen ihre eigenen Daten erheben und intern speichern. Mit dieser gemeinsamen Datenbasis wird der Fehlerdiagnose-Prozess vereinfacht und effektiver gestaltet. Außerdem kann die Wissenschaft durch Zugang zu den Daten die KI-Modelle kontinuierlich verbessern und an aktuelle Daten und Gegebenheiten anpassen.

Des Weiteren wird dadurch ermöglicht, die Langzeitentwicklung von PV-Modulen im Feld zu untersuchen. Die Wissenschaft kann darauf KI-Modelle entwickeln und Data-Mining-Methoden anwenden. Es kann untersucht werden, welche Module bestimmte Fehler entwickeln und zu welchem Zeitpunkt. Außerdem können die Einflüsse von externen Parametern untersucht werden. Abbildung 2 zeigt den aktuellen Stand der Modul-Detailansicht im Entwicklungsprototypen. Hier ist es bereits möglich, Vergleiche zwischen Referenzkennlinien und selbstständig gemessenen Kennlinien zu ziehen.

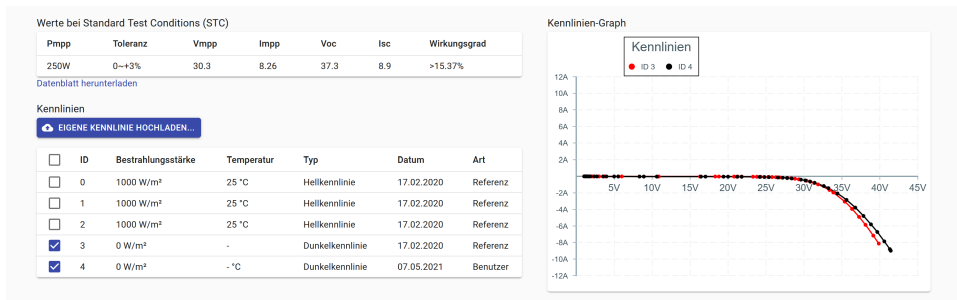


Abb. 2: Vergleich einer gemessenen Kennlinie mit einer Referenzkennlinie

PV-ExpertInnen-Forum Eine weitere Komponente der Applikation ist ein Forum für PV-ExpertInnen. Dies soll den ExpertInnen die Möglichkeit bieten, sich untereinander über Fehler auszutauschen und gemeinsam, auf Grundlage der Kennlinien und ggf. weiterer Daten, die Ursachen für Fehler oder anomales Verhalten zu finden.

Darüber hinaus bietet dies für die KI-Anwendungsentwicklung die Möglichkeit, an gelabelte Daten von Fehlerkurven zu gelangen, die essenziell für das Trainieren von Klassifikationsmethoden der KI zum Erkennen von Fehlern sind.

Schnittstellen und Vertrauensbildung zu KI-Modellen Die Applikation stellt außerdem eine Schnittstelle bereit, um die in Abschnitt 2 beschriebenen Neuronale Netze zum Schließen auf den Verlauf der Hellkennlinie durch die Dunkelkennlinie für die Praxis nutzbar zu machen.

Damit die NutzerInnen Vertrauen in die KI gewinnen und einen Eindruck für dessen Genauigkeit bekommen können, wird außerdem die Möglichkeit geboten, Prognose-Ergebnisse zu evaluieren. Dafür nimmt die Schnittstelle eine gemessene Dunkel- und Hellkennlinie der NutzerInnen entgegen. Die KI schließt aus der Dunkelkennlinie auf die Hellkennlinie und vergleicht diese anschließend mit der echten, gemessenen Hellkennlinie. Die Kennlinien und etwaige Abweichungen werden dann grafisch dargestellt. Dies ist ein wichtiger Schritt für das Heranführen der NutzerInnen an die KI und soll ihnen die Ungewissheit hinter der “Black-Box”-KI nehmen. Abbildung 3 zeigt die dazu entwickelte Ansicht.

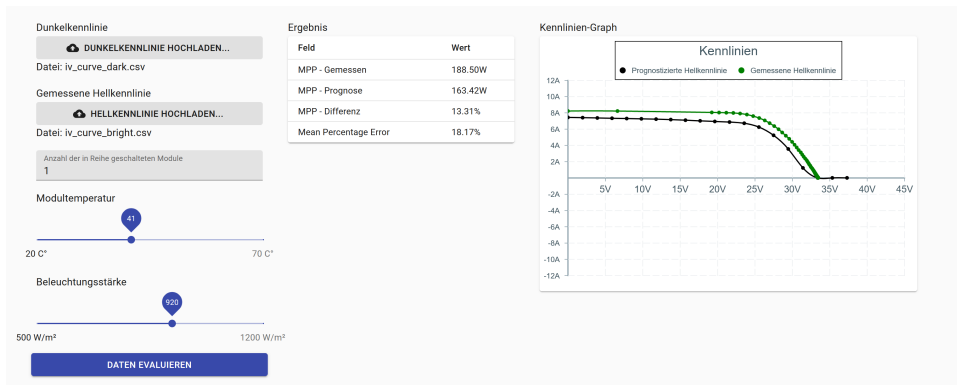


Abb. 3: Evaluation einer prognostizierten Kennlinie

4 Ergebnis und Diskussion

Die in dieser Arbeit vorgestellte Webapplikation ist aktuell noch in einem frühen Entwicklungsstadium und wird nur innerhalb unserer Arbeitsgruppe eingesetzt. Doch bereits jetzt verfügt die Anwendung über viele Funktionalitäten, die PV-ExpertInnen in Zukunft ein hohes Maß an Unterstützung bieten können.

Der zentralisierte Zugriff auf PV-Moduldaten sowie die Möglichkeit, in Echtzeit unterschiedliche Kennlinien mit Referenzwerten zu vergleichen, stellen einen großen Fortschritt gegenüber persönlich erfassten und nur lokal erreichbaren Datensilos dar. Das Forum

kann den Austausch der ExpertInnen untereinander verbessern und hilft parallel, unseren Datenbestand zu verbessern. Das Berechnen von Prognosen und die Evaluation dieser kann nicht mehr nur unter Laborbedingungen ausgeführt werden, stattdessen können diese Ergebnisse nun unterwegs bei der Arbeit der PV-ExpertInnen im PV-Feld abgefragt und ansprechend visualisiert werden.

Bevor iPVMModule auch öffentlich eingesetzt werden kann, sind noch weitere Arbeiten vorzunehmen. Neben der Implementierung weiterer Features ist auch das kontinuierliche Verbessern der User Experience eine wichtige Aufgabe. Denn nur, wenn die Webapplikation eine hohe Akzeptanz durch die PV-ExpertInnen erfährt, entfaltet sich ihr volles Potential.

Literaturverzeichnis

- [Ba20a] Basoglu, C.; Behrens, G.; Mertens, K.; Diehl, M.: "Anlagenevaluation mittels Hell- und Dunkelkennlinien und Methoden des Maschinellen Lernens", 35. PV-Symposium, Bad Staffelstein. 2020.
- [Ba20b] Basoglu, C.; Behrens, G.; Mertens, K.; Diehl, M.: Comparison of Knowledge-based Feature Vector Extraction and Geometrical Parameters of Photovoltaic I-V Curves. In: Proceedings of the Conference "Lernen, Wissen, Daten, Analysen". Jgg. 2738 in CEUR Workshop Proceedings, S. 70–75, 2020.
- [MAD17] Mertens, K.; Arnds, A.; Diehl, M.: "Quick and Effective Plant Evaluation Using Dark-IV String Curves". In: Proceedings of 33rd European Photovoltaic Solar Energy Conference, Amsterdam, 2017. S. 2346 – 2348, 2017.
- [MAD18] Mertens, K.; Arnds, A.; Diehl, M.: "String-Dunkelkennlinien: Eine neue effiziente Methode zur Anlagenevaluation", 33. Symposium Photovoltaische Solarenergie, Staffelstein. 2018.
- [Me19] Mertens, K.; Gercken, L.; Diehl, M.; Behrens, G.: "Fortschritte bei der Anlagenevaluation mittels String-Dunkelkennlinientechnik", 34. PV-Symposium, Bad Staffelstein. 2019.
- [Me20] Mertens, K.: Photovoltaik - Lehrbuch zu Grundlagen, Technologie und Praxis. Carl Hanser Verlag München, 5. Auflage, 2020.
- [Qu13] Quaschnig, V.: Erneuerbare Energien und Klimaschutz. Carl Hanser Verlag München, 3. Auflage, 2013.
- [RMB21] Rüter, J.; Meyer, F.; Behrens, G.: Weather-independent MPP prediction of PV plants using on-site measured dark IV string curves and deep learning. In: 2021 48th IEEE Photovoltaic Specialists Conference (PVSC). 2021.
- [Te19] Teuber, M.; Schlender, K.; Fehring, F.; Behrens, G.; Mertens, K.; Diehl, M.: "Anlagenevaluation mittels Data-Mining für Hell- und Dunkelkennlinien", 34. PV-Symposium, Bad Staffelstein. 2019.
- [Un15] United Nations: THE 17 GOALS - Sustainable Development. <https://sdgs.un.org/goals>, 2015. Accessed: 2021-04-15.

Nachhaltigkeit in der Softwareentwicklung

Ein Vorschlag zur alltagstauglichen Integration des ökologischen Fußabdrucks

Malte Räuchle, Jochen Wittmann¹

Abstract: Unter dem Aspekt der nachhaltigen Entwicklung von Software stellt das vorliegende Paper einen Widerspruch zwischen den üblicherweise nur im Nachhinein gemessenen Nachhaltigkeitskriterien für ein Software-Projekt und dem prozessbegleitenden Feedback- und Optimierungsgedanken des Agilen Softwareparadigmas fest. Durch die Heranziehung von sehr einfachen und pauschalisierenden Kriterien zur Abschätzung des ökologischen Fußabdrucks eines Software-Projektes gelingt es, den Projektteilnehmern bereits während der Projektlaufphase die Möglichkeit zu geben, die Nachhaltigkeitsmasse dynamisch zu beeinflussen. Damit wird der Gedanke des kontinuierlichen Feedbacks auch auf das Thema Nachhaltigkeit übertragen und bietet einen Ansatzpunkt für weitere unternehmenspolitische Maßnahmen zur Förderung einer nachhaltigen Software-Entwicklung.

Keywords: Nachhaltigkeit, Metriken, Softwareentwicklung; Agiler Prozess

1 Nachhaltigkeit und Software-Entwicklungsprozess

Der technische Fortschritt als Abbild der Globalisierung und der ständig wachsenden und international verflochtenen Gesellschaft unserer Zeit hält fortschreitend in verschiedensten gesellschaftlichen Teilaspekten Einzug [BU10]. Diese voranschreitende weltumspannende Vernetzung verändert nicht nur das alltägliche Leben, sondern beeinflusst darüber hinaus die Umwelt. Falls die Menschheit den Herausforderungen der Zukunft und den enormen ressourcentechnischen Ansprüchen gewachsen sein will, muss der Aspekt Nachhaltigkeit, der Umgang mit natürlichen Ressourcen und der Schutz der Umwelt in den Vordergrund gestellt werden [BU17]. Dabei ist es essentiell, nicht nur die klassischen Bereiche der Umweltpolitik, wie die Begrenzung, Verringerung und Vermeidung von Emissionen, zu betrachten, sondern auch den Blick auf den Technologiesektor zu richten. Der Grundgedanke beinhaltet die Vorstellung, dass sich innovative Gedankengänge auf technischer Ebene mit der Umweltpolitik vereinen, um andersartige Märkte für Umwelttechnologien und Nachhaltigkeit erschließbar zu machen [BU08].

Ziel dieses Beitrags ist es, diesen Nachhaltigkeitsgedanken auf den Softwareentwicklungsprozess zu übertragen. Für eine Definition des Nachhaltigkeits-Begriffes soll an dieser Stelle auf die Arbeit von Zschach verwiesen werden, wonach sich

¹ Hochschule für Technik und Wirtschaft, Fachbereich 2, Wilhelminenhofstraße 75A, 12459 Berlin,
[Malte.Raeuchle | Jochen.Wittmann@htw-berlin.de]

die Nachhaltigkeit im Wesentlichen auf die Nutzung von Ressourcen bezieht und als ökologische, wirtschaftliche und soziale Handlungsweise beschrieben werden kann [ZS20]. Typisch für die agile Softwareentwicklung ist es, dass iterativ, also in Zyklen entwickelt wird und die Entwicklung in kleinen Schritten und unter kontinuierlicher Einbeziehung von Feedback sämtlicher Akteure erfolgt [WR15].

Die Nachhaltigkeit von Software und Softwareentwicklungsprojekten ist Thema unterschiedlicher Arbeiten (z.B. [AL10], [AR15], [NA15]). Dort wird untersucht, inwieweit die Softwareentwicklung Einfluss auf ökologische, ökonomische und soziale Aspekte, also auf Nachhaltigkeit, hat und wie diese Einflüsse messbar gemacht werden können. Die konventionelle Bestimmung und Beurteilung von Nachhaltigkeit erfolgt dabei in der Regel im „Nachhinein“, was dem Gedanken des direkten Feedbacks innerhalb des agilen Software-Entwicklungsprozesses widerspricht und daher auch keine Möglichkeit besteht, das Nachhaltigkeitsmaß noch während des laufenden Prozesses zu beeinflussen. Dementsprechend ist das Ziel dieser Ausarbeitung die Bestimmung eines sehr einfach messbaren Kriteriums für eine transparente und datenbasierte Auswertung und Bewertung der Nachhaltigkeit innerhalb eines agilen Software-Entwicklungsprozesses.

2 Eine vereinfachte Abschätzung des ökologischen Fußabdrucks

Das Kriterium des “Ökologischen Fußabdrucks” unterstellt den Verbrauch natürlicher Ressourcen und die Betrachtung genereller Umweltauswirkungen einer Software oder eines Softwareproduktes innerhalb seines agilen Entwicklungsprozesses [AL10]. Es wird die Fragestellung behandelt, wie die Gesellschaft, speziell Unternehmen, mit der wachsenden Nachfrage nach Rohstoffen und deren Endlichkeit umgeht. Themen wie Konsum und Nachhaltigkeit, Energie- und Ressourcenverbrauch, persönliche Verantwortung und die Globalisierung stehen im Fokus. Zudem kann durch die Betrachtung und Bewertung des Fußabdrucks der Kraftstoffverbrauch und die damit verursachten Emissionen reduziert, die Büroflächenausnutzung optimiert und die Nutzung gemeinsamer Ressourcen maximiert werden [AL10]. Mithilfe des Kriteriums ist es interessierten Benutzergruppen möglich, die „Effektivität möglicher Maßnahmen auf eine nachhaltige Entwicklung abzuschätzen“ und aufgrund der Messbarkeit datengestützte, effektive Handlungsmaßnahmen zu erarbeiten, die direkt innerhalb der Softwareentwicklung angewandt werden können [SBS10]. Der ökologische Fußabdruck wird aktuell unter anderem durch Metriken, wie *Work-From-Home Days* und *Long-Haul Roundtrips*, definiert und dargestellt [AL10]. Mithilfe dieser Metriken ist die Nachhaltigkeit lediglich in einem geringen Maß transparent skalierbar und im Alltag nicht praktikabel, da sie nicht auf Abruf messbar sind. Daher beschreiben die folgenden Ausführungen hinsichtlich des ökologischen Fußabdrucks progressive Metriken, die die Nachhaltigkeit erfassbar, skalierbar und im Alltag anwendbar machen. Dahingehend beschreibt der

ökologische Fußabdruck eines vollständigen, agilen Softwareentwicklungsprozesses ein exzellentes Instrument zur Vermittlung einer nachhaltigen Entwicklungsideologie.

Im Folgenden wird ein Ansatz vorgestellt, der es erlaubt, das Kriterium des ökologischen Fußabdrucks unkompliziert zu messen bzw. zu steuern. Die Idee ist dem agilen SW-Entwicklungsparadigma nachempfunden: Die Messverfahren sollen so gut wie möglich in den bestehenden Softwareentwicklungsprozess integrierbar sein und der Aufwand möglichst klein gehalten werden. Die Wahl soll in einem Ausmaß erfolgen, dass die Verfahren trotz ihrer Einfachheit eine hohe Aussagekraft haben und einen Vergleich während der Projektlaufzeit, sowie auch einen Vergleich mit anderen Projekten erlauben. Außerdem soll aus dem Verfahren heraus ersichtlich sein, wo das Team ansetzen kann, um die Nachhaltigkeitsleistung zu erhöhen. Übergeordneter Ansatz dieser Metrik ist zum einen die Betrachtung der Messung vor Ort am Arbeitsplatz im Unternehmen und zum anderen die Messung in der eigenen Wohnung im Zuge der Ausübung mobiler Arbeit. Aufgrund der aktuellen Situation rund um die SARS-CoV-2 Pandemie und deren tiefgreifende Auswirkung auf alle Bereiche der Gesellschaft ist das Modell des mobilen Arbeitens alltäglicher Bestandteil unserer Arbeitskultur geworden. Unternehmen aus den unterschiedlichsten Wirtschaftszweigen bieten ihren Mitarbeitern daher flexible Arbeitszeitmodelle an, die eine Unabhängigkeit vom eigentlichen Standort des Unternehmens gewährleisten [DO18]. Vor diesem Hintergrund konzentrieren sich die folgenden Ausführungen allein auf die Erfassung der Art und Weise der An- und Abfahrt der Arbeitnehmer*innen zum Arbeitsplatz und die Schätzung des mittleren Stromverbrauches am Arbeitsplatz. Die beiden Verfahren zeichnen sich im Gegensatz zu bestehenden Vorgehensweisen, wie der Erfassung von *Work-From-Home Days* und *Long-Haul Roundtrips*, durch eine einfache und im Arbeitsalltag schnell durchführbare Anwendbarkeit und Messbarkeit aus.

2.1 An- und Abfahrt der Arbeitnehmer*innen zum Arbeitsplatz

Bei diesem Verfahren wird die Messung während der mobilen Arbeitsphase nicht betrachtet, da keine Fahrt zum Arbeitsplatz erfolgt.

Der tägliche Weg zur Arbeit stellt eine bedeutsame Quelle für die Produktion von Emissionen dar und verursacht zusammenhängend einen wesentlichen Betrag an auszugleichenden Umweltkosten. Aufgrund dessen ist es essentiell, eine tägliche Kontrolle der verursachten Emissionen zu gewährleisten, wenn die Nachhaltigkeit innerhalb eines Projektes einen zentralen Gegenstand verkörpert. Dadurch ist das Unternehmen in der Position, kurzfristig auf hohe Emissionswerte reagieren zu können. Die Erfassung des durch Mitarbeiter*innen genutzten Verkehrsmittels erfolgt direkt beim Check-In zu Arbeitsbeginn oder im Zuge der täglichen Zusammenkunft beim gemeinschaftlichen und täglich abgehaltenen Projektmeeting. Es empfiehlt sich die Nutzung des eigenen Intranets zur Eintragung der Daten und zur allgemeinen Veröffentlichung der Ergebnisse auf Basis der verursachten Treibhausgasemissionen. Dabei sind pro Person nur wenige Angaben notwendig, nämlich die für den Arbeitsweg

bevorzugten Verkehrsmittel und die mit dem jeweiligen Verkehrsmittel zurückgelegte Strecke. Gegebenenfalls kann die Liste der Alternativen noch um eine gemischte Verkehrsmittelnutzung im Sinne eines Modal Split erweitert werden. Zudem muss die zurückgelegte Distanz nicht für jeden Arbeitstag gesondert eingetragen werden, da das übliche Streckenprofil nur einmal direkt beim Projektstart hinterlegt wird. Diesbezüglich ist nur noch die Angabe der tagesaktuellen Wahl der Alternative sowie die Angabe der tagesaktuellen Auslastung als zusätzlicher Parameter zu tätigen. Die Einbeziehung der Auslastung als zusätzlichen Parameter erscheint auf den ersten Blick ungewöhnlich, ist aber sinnvoll, da dieser Wert mit großem Gewicht in die Berechnung der Kennzahlen einfließt. Das Resultat der Erfassung impliziert eine umgehende Rückmeldung bezüglich verursachter Treibhausgasemissionen und dem daraus abzuleitenden Beitrag zur Nachhaltigkeitsstrategie des Unternehmens. Die Einbeziehung der Auslastung ermöglicht so zwar eine realitätsnahe Berechnung der CO₂-Last, andererseits stellt sie jedoch einen Faktor dar, der durch das Unternehmen nur in einem geringen Maß beeinflusst werden kann. Gemäß dieser Berechnungsvorschrift generiert eine Fahrt zu Stoßzeiten eine bessere Emissionsbilanz. Dabei bleiben allerdings die Kosten und Ineffizienzen an anderer Stelle im Gesamtsystem, beispielsweise bei der Fahrtenplanung des ÖPNV, unberücksichtigt. Andererseits bleiben die Einflussmöglichkeiten des Managements auf der Ebene des Projektes darauf beschränkt, entweder durch eine Anpassung der Arbeitszeiten oder durch die Ermöglichung einer flexiblen Zeiteinteilung die Bilanz der Mitarbeiter*innen verbessern. Eine Übersicht über die wichtigsten Verkehrsmittel zusammen mit den entsprechenden Emissionswerten zeigt die folgende Tabelle (vgl. Abb. 1).

Vergleich der durchschnittlichen Emissionen einzelner Verkehrsmittel im Personenverkehr in Deutschland – Bezugsjahr 2019

Verkehrsmittel		Treibhausgase ¹	Kohlenmonoxid	Flüchtige Kohlenwasserstoffe ⁴	Stickoxide	Partikel ⁵	Auslastung
Pkw		143	0,95	0,14	0,39	0,006	1,5 Pers./Pkw
Flugzeug, Inland		214 ²	0,30	0,10	0,98	0,011	70%
Eisenbahn, Fernverkehr	g / Pkm	29 ³	0,02	0,00	0,03	0,001	56%
Linienbus, Fernverkehr		29	0,01	0,01	0,05	0,001	55%
sonstige Reisebusse ⁶		32	0,04	0,01	0,12	0,002	64%
Eisenbahn, Nahverkehr		55	0,04	0,01	0,19	0,004	28%
Linienbus, Nahverkehr		80	0,06	0,03	0,29	0,005	19%
Straßen-, Stadt- und U-Bahn		55	0,03	0,00	0,05	0,002	19%

g/Pkm = Gramm pro Personenkilometer, inkl. der Emissionen aus der Bereitstellung und Umwandlung der Energieträger in Strom, Benzin, Diesel und Kerosin

Quelle: TREMOD 6.14
Umweltbundesamt, 11/2020

Abb. 1: Abfrage nach Art der Auslastung des Verkehrsmittels [UM20]

Das Resultat der Dokumentation der Verkehrsmittelwahl begründet die Erstellung eines unternehmensinternen Punktesystems, welches die Umweltwirkung des verwendeten Transportmittels transparent für alle Mitarbeiter*innen des Unternehmens darlegt und somit den direkten Einfluss auf die Umwelt sichtbar darstellt. Ein im Intranet hinterlegter Emissionsrechner wandelt die eingetragenen Daten in Punkte oder eine Emissions-Ampel-Darstellung um, wodurch die Angestellten über den Zeitraum eines Projektes, eines

Monats oder eines Jahres ihren Umwelteinfluss stetig im Blick haben können. In Verbindung mit dem individuellen Punktstand könnte das Management Boni gewähren, die ein umweltfreundliches Verhalten belohnen (z.B. zusätzliche Urlaubstage, Vorteilskarten für den ÖPNV, Gutscheine für Kinderbetreuung, ...).

2.2 Schätzung des Stromverbrauches durch Arbeitsgeräte am Arbeitsplatz

In diesem Absatz wird die Messung des Stromverbrauches am Arbeitsplatz, abgebildet und erläutert, inwiefern dieses Verfahren den ökologischen Fußabdruck eines Projektes messbar darstellen kann. Zusätzlich wird der Heizverbrauch betrachtet, der als zweiter Aspekt seinen Teil zum Energieverbrauch beiträgt.

Im Gegensatz zum ersten Verfahren erfolgt nicht nur die Betrachtung des Arbeitsplatzes im Unternehmen, sondern auch die Messung des Stromverbrauches in der Wohnung während der Ausübung mobiler Arbeit. Die Messung und Quantifizierung der Energiemenge dient als direktes Verfahren zur Bestimmbarkeit der Nachhaltigkeitsleistung eines Unternehmens, da die Darstellung in Zahlen unmittelbare Schwächen bezüglich eines umweltschädlichen Bedarfes aufdecken kann. Die Messgrößen wurden bewusst einfach ausgewählt, um eine unkomplizierte, aber trotzdem aussagekräftige Aussage für die Einschätzung der Nachhaltigkeit treffen zu können.

Der nachstehende Abschnitt versucht nun relevante Faktoren für den Stromverbrauch am Arbeitsplatz zu identifizieren. Dabei muss zwischen Home-Office und Büro unterschieden werden. Innerhalb der mobilen Art der Arbeitsausübung werden die in Tabelle 3 aufgeführten Geräte und Nutzungscharakteristika betrachtet. Der Stromverbrauch des Laptops und des Monitors verweist auf die Nutzung an 220 Tagen im Jahr und wird mit einer aktiven Nutzungszeit von sechseinhalb Stunden (Laptop) und sechs Stunden (Monitor) am Tag bemessen [ST15]. Für eine bessere Vergleichbarkeit werden für die Verbräuche von Laptop und Monitor die standardisierten Referenzwerte des Büroarbeitsplatzes verwendet. Für den Verbrauch an Strom des Scanners und des Smartphones bezieht man sich auf eine Nutzung an 365 Tagen im Jahr, wobei für das Smartphone ein Ladezyklus pro Tag veranschlagt wird [ST15]. Die Beleuchtung im Home Office wird durch drei LED-Lampen repräsentiert, die an 1000 Stunden im Jahr genutzt werden [VE15]. Geräte wie Drucker, Scanner, Kopierer und Faxgeräte weisen eine so geringe Größenordnung des Strombedarfs auf, dass eine differenzierte Betrachtung im Rahmen der Untersuchung nicht erforderlich ist.

Faktor	Nutzungs- muster in Tagen	Nutzungs- muster in Stunden/Tag	Strom- verbrauch in kWh / Tag	Strom- verbrauch in kWh / Monat	Strom verbrauch in kWh / Jahr
Laptop	220	6,5	0,086	2,58	31
Monitor	220	6	0,111	3,33	40

Smartphone	365	1* Laden/Tag	0,008	0,24	2,9
Beleuchtung	41,66	4,5	0,125	3,75	45

Tab. 1: Stromverbrauch im Home Office [ST15]

Der Stromverbrauch, basierend auf den in der Tabelle 1 dargestellten Werten, summiert sich auf 118,9 Kilowattstunden/Jahr. Dabei wird vorausgesetzt, dass Arbeitnehmer*innen nicht im Besitz eines Desktop-PCs sind, sondern ihre arbeitsbezogenen Aufgaben ausschließlich über einen Laptop mit angeschlossenem zweitem Monitor ausführen. Zudem wird berücksichtigt, dass Arbeitnehmer*innen hauptsächlich ein Smartphone für berufliche Telefonate verwenden und nicht auf ein Festnetztelefon zurückgreifen.

Dieser Berechnung wird nun ein Büroarbeitsplatz gegenübergestellt. Tabelle 2 zeigt die berücksichtigten Faktoren. Der Stromverbrauch des Desktop-PCs und des Laptops verweist auf die Nutzung an 220 Tagen im Jahr und wird mit einer aktiven Nutzungszeit von sechseinhalb Stunden am Tag bemessen [ST15]. Der Einsatz des Monitors ist ebenfalls auf 220 Tage ausgelegt, allerdings mit einer Nutzung von sechs Stunden am Tag [ST15]. Im Gegensatz dazu erfolgt die Betrachtung des Multifunktionsgerätes, des Faxgerätes und des Telefons hinsichtlich des Stromverbrauches an 365 Tagen im Jahr [ST15]. Abschließend stellt die Beleuchtung des Arbeitsplatzes durchschnittlich die Hälfte des gesamten Stromverbrauches dar.

Faktor	Nutzungsmuster in Tagen	Nutzungsmuster in Stunden/Tag	Stromverbrauch in kWh /Tag	Stromverbrauch in kWh /Monat	Stromverbrauch in kWh /Jahr
Desktop-PC	220	6,5	0,18	5,42	65
Laptop	220	6,5	0,086	2,58	31
Monitor	220	6	0,111	3,33	40
Multifunktionsgerät	365	/	0,138	4,16	50
Faxgerät	365	0,5	0,188	5,66	68
Telefon	365	24	0,111	3,33	40
Beleuchtung	/	/	0,408	12,25	147

Tab. 2: Stromverbrauch im Büro [ST15]

Der Stromverbrauch am Arbeitsplatz im Unternehmen beläuft sich in der Summe nach den Angaben der Tabelle auf 294 Kilowattstunden/Jahr. Die Arbeit am Arbeitsplatz im Büro ist deutlich umfassender als die Arbeit im Home Office, da unterschiedliche Geräte genutzt werden können. Die Inanspruchnahme eines Desktop PCs, eines Faxgerätes oder eines Multifunktionsgerätes resultiert demzufolge in einem deutlich höheren Stromverbrauch als im Home Office.

2.3 Schätzung des Energieverbrauches am Arbeitsplatz durch Heizungsnutzung

Abgesehen von der Betrachtung des Stromverbrauches ist der Energieverbrauch durch das Beheizen der Arbeitsfläche ebenfalls von Bedeutung und wird im Folgenden erläutert. Dieser Verbrauch stellt ein weiteres Element dar, wird im Rahmen der Untersuchung zur Bewahrung der Verständlichkeit komprimiert erwähnt, jedoch nicht detailliert betrachtet. Im konkreten Fall einer näheren Untersuchung kann eine Anpassung des Detaillierungsgrades erfolgen. Die Tabelle 3 zeigt eine übersichtliche Darstellung des Energieverbrauches, angegeben in Kilowattstunden, durch die Heiznutzung in einer Wohnung und in einem Büro.

Faktor	Energieverbrauch in kWh je qm/Tag	Energieverbrauch in kWh je qm/Monat	Energieverbrauch in kWh je qm/Jahr
Wohnen	0,33	10,16	122
Büro	0,19	5,76	69,2

Tab. 3: Heizverbrauch Wohnung und Büro [UR17]

Aus den Werten lässt sich entnehmen, dass der Energieverbrauch pro Quadratmeter innerhalb eines Jahres in einer privaten Wohnung deutlich höher ausfällt als in einem Büro. Dies kann durch die größere Fläche eines Büros erklärt werden, wodurch der Verbrauch von Energie auf eine weiter gefasste Fläche verteilt wird. Allerdings bleibt hier für eine erste Überschlagsrechnung nicht berücksichtigt, dass auch die nicht genutzte heimische Wohnung auf eine Mindesttemperatur beheizt werden muss und hier eigentlich nur der Grenz-Aufwand für das Aufheizen von dieser Grundtemperatur auf eine Wohntemperatur zu berücksichtigen wäre.

3 Vergleichszahlen für ein fiktives Softwareentwicklungsprojekt

Für die Untersuchung erfolgt die Annahme, dass die Arbeitsleistung inklusive des Strom- und Heizenergieverbrauches und des Arbeitsweges am Büroarbeitsplatz, bezogen auf den Stand der Nachhaltigkeit und den damit einhergehenden Umweltauswirkungen, als konventioneller Ansatz angesehen wird und entsprechend den aktuellen Status Quo repräsentiert. Das mobile Arbeiten ist daher als nachhaltige Alternative zur konventionellen Ausübung der Arbeit definiert. Exemplarisch erfolgt abschließend eine Gegenüberstellung der erwähnten Ansätze für ein agiles Softwareentwicklungsprojekt mit der Dauer von drei Monaten, im Büro und Home Office. Der Arbeitsweg wird in CO₂-Äquivalenten, der Strom- und Heizverbrauch in Kilowattstunden angegeben.

Arbeitsweg [CO ₂ - Äquiv.]	Strom- verbrauch [kWh]	Heiz- verbrauch [kWh]
188,76	36,73	17,28

Tab. 4: Ansatz 1 – Büro

Arbeitsweg [CO ₂ - Äquiv.]	Strom- verbrauch [kWh]	Heiz- verbrauch [kWh]
/	29,7	30,48

Tab. 5: Ansatz 2 – Home Office

Aufgrund der dargestellten Messergebnisse erfolgt während einer Projektdauer von drei Monaten durch die Verlagerung der Arbeit in die eigene Wohnung eine Reduzierung des Stromverbrauches um 7,03 Kilowattstunden und ein Wegfall der Emissionsbelastungen durch die Vermeidung des Fahrtweges zum Unternehmen. Betrachtet man die gegenteilige Perspektive, kann lediglich ein verringerter Heizverbrauch um 13,20 Kilowattstunden innerhalb des definierten Zeitraums festgestellt werden. Insgesamt lässt sich dementsprechend hinsichtlich der Nachhaltigkeitsperformance sagen, dass durch die Umstrukturierung der Arbeit ins Home Office eine Verbesserung herbeigeführt werden kann und für eine Verbesserung der Nachhaltigkeitsleistung eine gezielte Abwägung von Arbeitstagen im Büro und Arbeitstagen im Home Office erfolgen muss.

4 Zusammenfassung und Ausblick

Ziel dieses Papers war es, ein Kriterium darzustellen, das die Nachhaltigkeit im Softwareentwicklungsprozess beschreibt und entsprechende Verfahren zur Messung dieses Kriteriums auszuarbeiten. Dabei lag der Fokus darauf, die Verfahren und Messungen möglichst einfach und praktikabel zu halten, sowie eine hohe Objektivität zu erreichen, sodass zu jedem Zeitpunkt eine mit Daten fundierte Aussage über die Nachhaltigkeitsleistung im Entwicklungsprojekt getätigt und Vergleiche mit anderen Projekten durchgeführt werden können. Im Rahmen der Arbeit wurde das bereits in der

Literatur bekannte Nachhaltigkeitskriterium des ökologischen Fußabdrucks ausgewählt. Die Ausarbeitung hat gezeigt, dass die beschriebenen Messverfahren in Bezug auf Praktikabilität, Objektivität und Aussagekraft verbessert werden konnten. Im Ergebnis liegen nun ein Vorschlag vor, mit dem eine Umsetzung der Messung dieser Kriterien in der täglichen Praxis möglich ist. Im nächsten Schritt wäre es nun sinnvoll, dies im Rahmen eines Modellversuches zu validieren. Dabei wäre es auch von Interesse zu prüfen, ob das Entwicklungsteam aus den Messungen tatsächlich Erkenntnisse gewinnen und die Nachhaltigkeitsleistung verbessern kann. Schließlich könnten Messverfahren und Metriken in ähnlicher Form für weitere Nachhaltigkeitskriterien ausgearbeitet werden. Albertao et al. (2010) und Naumann et al. (2015) schlagen das Kriterium *Portabilität* vor, das Aufschluss darüber gibt, auf wie vielen Hardware-Systemen eine Software läuft. Dies ist ökologisch relevant, weil eine starke Einschränkung auf gewisse Hardware zu einer verkürzten Lebensdauer von anderen Hardware-Systemen führen kann. Interessant wäre auch ein einfaches und praktisches Messverfahren für den Energieverbrauch der Software selbst zu finden. Ansätze dazu gibt es zum Beispiel bei Ardito et al. (2015). Dies ist allerdings nur eine Auswahl an weiteren Nachhaltigkeitskriterien, die im Zuge dieser Arbeit in der engeren Auswahl standen. Inwieweit diese eine Aussage über die Nachhaltigkeitsleistung geben oder überhaupt messbar sind, muss also noch in zukünftigen Arbeiten geprüft werden.

Literaturverzeichnis

- [AL10] Albertao, F., Xiao, J., Tian, C., Lu, Y., Zhang, K. Q., & Liu, C.: Measuring the sustainability performance of software projects. In (2010 IEEE 7th International Conference on E-Business Engineering), S. 369-373, 2010.
- [AR15] Ardito, L.; Procaccianti, G.; Torchiano, M.; Vetro, A.: Understanding green software development: A conceptual framework. In (IT Professional), S. 44-50, 2015.
- [BU17] Bundesministerium für Umwelt, Naturschutz und nukleare Sicherheit, <https://www.bmu.de/themen/europa-internationales-nachhaltigkeit-digitalisierung/nachhaltige-entwicklung/strategie-und-umsetzung/nachhaltigkeit-als-handlungsauftrag/>, Stand: 25.03.2021.
- [BU08] Bundesministerium für Umwelt, Naturschutz und nukleare Sicherheit, <https://www.bmu.de/themen/wirtschaft-produkte-ressourcen-tourismus/wirtschaft-und-umwelt/umwelttechnologien/masterplan-umwelttechnologien/>, Stand: 24.03.2021.
- [BU10] Bundeszentrale für politische Bildung, <https://www.bpb.de/nachschlagen/zahlen-und-fakten/globalisierung/>, Stand: 12.02.2021.
- [DO18] Dobmeier, N.: Home-Office vs. Büro – Der Kampf um die Effektivität., <https://www.accenon.de/blog-home-office-vs-buero-der-kampf-um-die-effektivitaet/>, Stand: 14.01.2021.
- [HA20] Haase, A.: Corporate Sustainability: So wichtig ist Nachhaltigkeit für Unternehmen., <https://blog.hubspot.de/sales/sustainability> Stand: 01.02.2021.

- [NA15] Naumann, S.; Kern, E.; Dick, M.; Johann, T.: Sustainable software engineering: process and quality models, life cycle and social aspects. In (ICT Innovations for Sustainability), S. 191-205, 2015.
- [SBS10] Schemel, B.; Draeger, I.; Schmidhals, M.: Der ökologische Fußabdruck. Multivision, Hamburg, 2010.
- [SFD20] Schelewsky, M.; Follmer, R.; Dickmann, C.: CO2-Fußabdrücke im Alltagsverkehr. In (Umweltbundesamt): Datenauswertung auf Basis der Studie Mobilität in Deutschland. Umweltbundesamt, Dessau-Roßlau, S. 38, 2020.
- [ST15] Stobbe, L. et.al.: Entwicklung des IKT-bedingten Strombedarfs in Deutschland. Studie im Auftrag des Bundesministeriums für Wirtschaft und Energie Projekt-Nr. 29/14, S.116 – 177, 2015.
- [SW14] Schmidt, B.; Wytzisk, A.: Software Engineering und Integrative Nachhaltigkeit. In (Informatik 2014), S. 1935-1937, 2014.
- [UM15] Umweltbundesamt: Umwelttipps für den Alltag, <https://www.umweltbundesamt.de/umwelttipps-fuer-den-alltag/mobilitaet/bus-bahn-fahren#unsere-tipps>, Stand: 10.01.2021.
- [UM17] Umweltbundesamt, Umwelttipps für den Alltag, <https://www.umweltbundesamt.de/umwelttipps-fuer-den-alltag/mobilitaet/bus-bahn-fahren#unsere-tipps>, Stand: 05.05.2021.
- [UM18] Umweltbundesamt: Umweltverträglichkeit von Software ist jetzt meßbar, <https://www.umweltbundesamt.de/presse/pressemitteilungen/umweltvertraeglichkeit-von-software-ist-jetzt>, Stand: 15.02.2021.
- [UR17] Urbansky, F.: Energiebedarf von Bürogebäuden ermitteln., <https://www.enwipo.de/2017/03/07/energiebedarf-von-buerogebaeuden-ermitteln/>, Stand: 25.02.2021.
- [VE15] Verbraucherzentrale Rheinland-Pfalz e.V.: Energieverbrauch von Lampen., https://www.verbraucherzentrale-rlp.de/sites/default/files/migration_files/media234424A.pdf, Stand: 18.02.2021.
- [WR15] Wolf, H.; Roock, S.: Agile Softwareentwicklung. Ein Überblick. dpunkt-Verlag, Heidelberg, S. 3, 2015.
- [ZS20] Zschach, M.: Nachhaltigkeit. In (Zeit im Lebensverlauf), transcript-Verlag, S. 227-228, 2020.

Maschinelles Lernen von CNN-Modellen zur Segmentierung von Störobjekten auf Gebädefassaden auf Infrarot- und Farbbildern

Klaus Schlender¹, Grit Behrens² und Reiner Creutzburg³

Abstract: Die Arbeit zeigt, dass es mit Hilfe von Technologien der künstlichen Intelligenz möglich ist, unerwünschte Objekte auf Infrarot- und Normalbildern zu segmentieren, um diese anschließend erfolgreich aus dem Bild entfernen zu können. Die in dieser Arbeit entwickelte Methode kann eingesetzt werden, um weitere Segmentierungsaufgaben zu übernehmen, die bei der Berechnung von thermischen Energieverlusten oder anderen Anwendungen hilfreich sind. Die vorgestellten Ergebnisse können im ENVIRON-Projekt "Environ - Entwicklung und Evaluation einer Intervention zur Vermeidung von durch energetische Sanierung ausgelösten Rebound-Effekten" anwendungsbezogen zu einer schnelleren Bewertung potentieller Energieverluste in sanierungsbedürftigen Gebäuden beitragen und entsprechend helfen, den Entscheidungsprozess für oder gegen eine Modernisierungsmaßnahme zu bewerten. In diesem Zusammenhang wird auch dazu beigetragen, unnötige CO₂-Emissionen von Gebäuden zu reduzieren und damit die Umwelt zu entlasten.

Keywords: Data Science, Deep Neural Network, Object Segmentation, Region Proposal, Convolutional Neural Network, thermal images.

1 Einleitung

Neubauten in Deutschland haben bereits einen hohen Energiestandard nach EnEV 2016 [En16] und werden nach der neuen EU-Gebäuderichtlinie [ETW16] gebaut. Häuser verlieren durch ihre Außenwände Energie in Form von Wärme. Die Höhe der Wärmeverluste ist abhängig vom Aufbau der Wand und viele alte Häuser mit geringer Wärmedämmung verlieren entsprechend mehr Energie als Sanierte oder neu gebaute Gebäude. Der Gebäudesektor verursacht etwa 30 % aller CO₂-Emissionen und die Wärmedämmung ist bei der Betrachtung deutscher Klimaschutzziele von großer Bedeutung [Um20]. Die Modernisierung von älteren Häusern ist teuer und sollte gut überlegt sein. Im ENVIRON-Projekt (BMBF5-Verbundforschungsprojekt, Förderkennzeichen: 01UT1703A) werden in Zusammenarbeit mit dem "Reallabor Bielefeld Sennestadt" Gebäude vor und nach einer Sanierung betrachtet [Se20]. Die

¹ FH-Bielefeld University of Applied Sciences, Artilleriestraße 9, 32427 Minden, klaus.schlender@fh-bielefeld.de

² FH-Bielefeld University of Applied Sciences, Artilleriestraße 9, 32427 Minden, grit.behrens@fh-bielefeld.de

³ Technische Hochschule Brandenburg University of Applied Sciences, Magdeburger Straße 50, 14770

Brandenburg an der Havel, reiner.creutzburg@th-brandenburg.de.

SRH Berlin University of Applied Sciences, Berlin School of Technology, Ernst-Reuter-Platz 10, D-10587 Berlin, Germany, Reiner.Creutzburg@srh.de

Wohnungen sind parallel zu gewonnenen Infrarotbildern mit einem SmartMonitoring-System versehen und liefern Daten. Es wurden Daten aus den Räumen der Gebäude vor und nach der Sanierung gesammelt [Be17, BSF18] und zusätzlich wurden Wärmebilder der Außenwand aufgenommen, um deren Wärmestrahlung zu erfassen. Die Thermografie gewinnt im Sanierungsprozess immer mehr an Bedeutung [PTT12]. So können z. B. Wärmebrücken aufgespürt, Wärmeübergänge beobachtet oder Wärmeverluste mit physikalischen Formeln berechnet werden. Wärmebilder werden zur Bewertung und Erkennung von Wärmebrücken in Hauswänden eingesetzt. Neben Wärmebrücken werden auch thermische Flächen und deren Wärmeabstrahlung analysiert, um den Sanierungsbedarf darstellen zu können. Neben wärmestrahlungsrelevanter Pixel der Außenwand befinden sich auf Thermografieaufnahmen ebenfalls verschiedene Störobjekte.

Aus diesem Grund liegt der Fokus dieser Arbeit darauf, eine Automatisierung der Wärmebildverarbeitung mittels CNN-Segmentierungen auf Infrarot-Input-Daten zu ermöglichen. Dies kann zum einen durch die Lokalisierung von Wärmebereichen und Wärmebrücken Expertinnen und Experten helfen, Gebäude vor und nach ihrer Sanierung zu analysieren bzw. miteinander zu vergleichen. Zum anderen wird das Segmentieren von Störobjekten wie Fenster, Pflanzen und Dachelemente, oder dem Himmel und das anschließende Entfernen dieser mittels Masken behandelt, um einen Mittelwert der abgestrahlten thermischen Energie über relevante Pixel zu berechnen.

2 Methoden

2.1 Vorverarbeitung und Normalisierung des Bildmaterials

Bevor die Thermobilder zum Trainieren eines Convolutional Neural Network (CNN) verwendet werden, müssen sie vorverarbeitet werden. Die Vorverarbeitung ermöglicht es, eine künstliche Intelligenz so gut wie möglich zu trainieren. In diesem Abschnitt werden verschiedene Methoden zur Vorverarbeitung von Bildern vorgestellt. Der erste Schritt besteht darin, die Farbräume von Wärmebildern zu standardisieren, da verschiedene Infrarotkameras ihre Daten in unterschiedlichen Farbräumen speichern. Der jeweilige Farbraum kann den Metadaten des Bildes entnommen werden. Die für das Projekt verwendete Infrarotkamera hat den Standard-RGB-Farbraum eingestellt und angewendet. Um die Bilder zu normalisieren, müssen die von der Infrarotkamera in den Bildern gespeicherten Farben als höchste und niedrigste Temperaturwerte dargestellt werden. Anschließend können die Farben in Temperaturen umgerechnet werden. Dadurch kann der Bildinhalt durch Umstellung von Gleichung (1)

$$I(x, y) = 255 * \frac{(T(x,y) - T_{min})}{T_{max}} \quad (1)$$

in absolute Temperaturen umgerechnet werden. Die Funktion I gibt das resultierende Graustufenbild mit $I(x, y)$ als Pixel an der Position (x, y) an. Analog dazu ist $T(x, y)$ die gemessene Temperatur an der Position. T_{\min} und T_{\max} sind die minimale und maximale gemessene Temperatur im Bild. Der resultierende Wert liegt zwischen 0 und 1. Da Bilder normalerweise eine Byte-Darstellung für ein Pixel verwenden, wird dieser Wert mit 255 multipliziert. Durch die Normalisierung wird der Kontrast zwischen den verschiedenen Temperaturen erhöht. Dadurch können kleinere Temperaturunterschiede erkannt werden, um optimiert labeln zu können (vgl. Abb. 1).

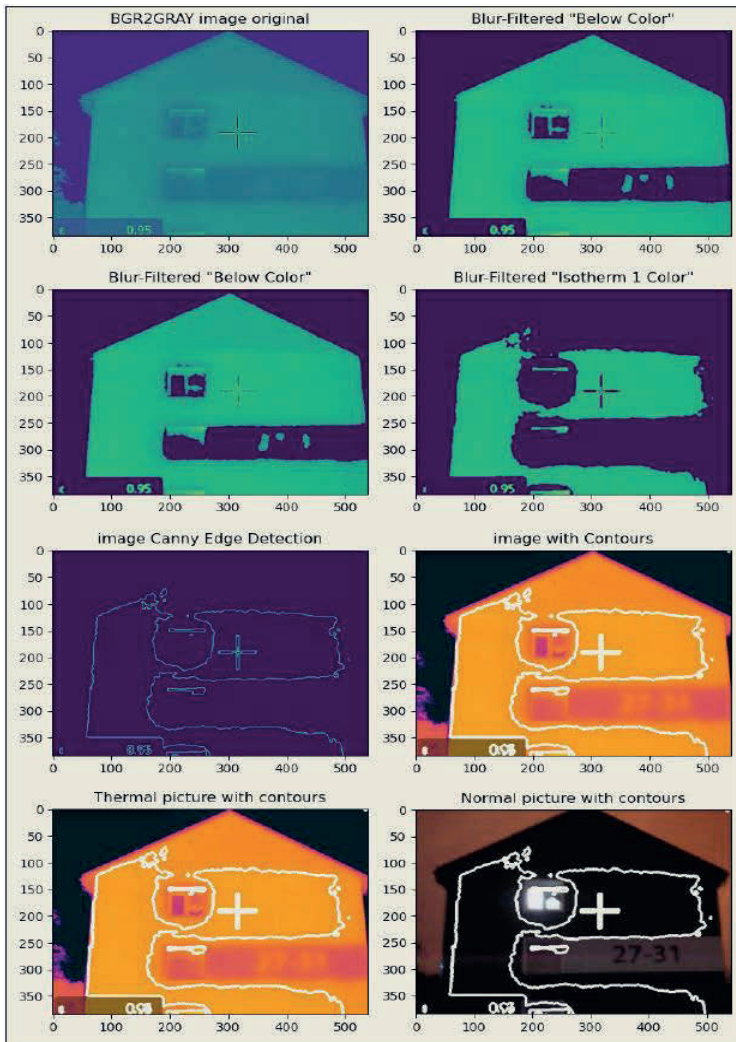


Abb. 1: Verfahren zum Labeln von Wärmebrücken

In den Bildern, die für das Projekt Environ aufgenommen wurden, zeigen Wärmebilder immer ein gewisses Grundwärmenniveau. Der Himmel hingegen ist immer viel kälter und meist in Schwarzönen dargestellt. Diese Information kann daher herausgefiltert werden, indem alle Bereiche eliminiert werden, deren Temperatur einen bestimmten Schwellenwert nicht erreicht. Die dazu verwendete Formel (2)

$$I(x, y) = \begin{cases} I(x, y), & I(x, y) > t \\ 0, & \text{sonst} \end{cases} \quad (2)$$

ermöglicht das automatisierte Entfernen von Himmelelementen bzw. ein einfaches und automatisiertes Labeln. Das Labeln von Wärmebereichen und -brücken erfolgt ebenfalls mit Schwellenwerten, die mit Hilfe von Metadaten aus den Thermografieaufnahmen entnommen sind.



Abb. 2: Input-Daten, mit Wärmebrücken-Labeln versehen

Zusätzlich zu den Wärmebildern stehen auch die zeitgleich erzeugten Farbbilder zur Verfügung, die an der gleichen Position aufgenommen wurden. Um sie später verwenden zu können, musste die Auflösung einheitlich skaliert werden. Für die Weiterverarbeitung im CNN-Training wurden alle Bilder aufbereitet und Bereiche automatisiert (Bildverarbeitung) und händisch (Online-Tool supervise.ly) gelabelt (vgl. Abb. 2), um das Trainieren der Objekterkennung bzw. -segmentierung zu ermöglichen und den

Trainingsprozess zu bewerten. Hierzu wurden je Anwendungsfall 407 Input-Daten gelabelt, um auf das Training zur CNN-Segmentierung von Pflanzen und Dachelementen, sowie Wärmebrücken und Wärmebereichen vorbereitet.

2.2 U-Net Architektur und Fenstererkennung

Der Segmentierungsaufgabe und dem Training dient als Basis eine U-Net-CNN-Architektur in Anlehnung an Keras U-Net starter - LB 0.277 [Am18], welche für das Training optimiert ist (vgl. Abb. 3). Auch wenn Transfer-Learning mit der Verwendung bestehender Modelle eine zunehmend beliebte Lösung für das Trainieren von neuen KI-Anwendungen darstellt [HBF19], wurden in dieser Ausarbeitung eigene Modelle (Kapitel 2.3 bis 2.5) für die jeweiligen Segmentierungsaufgaben unter Verwendung eines „Training-Endes“ trainiert, d.h. bei 100 Epochen ohne eine Verbesserung wird das Modell gespeichert und das Training beendet.

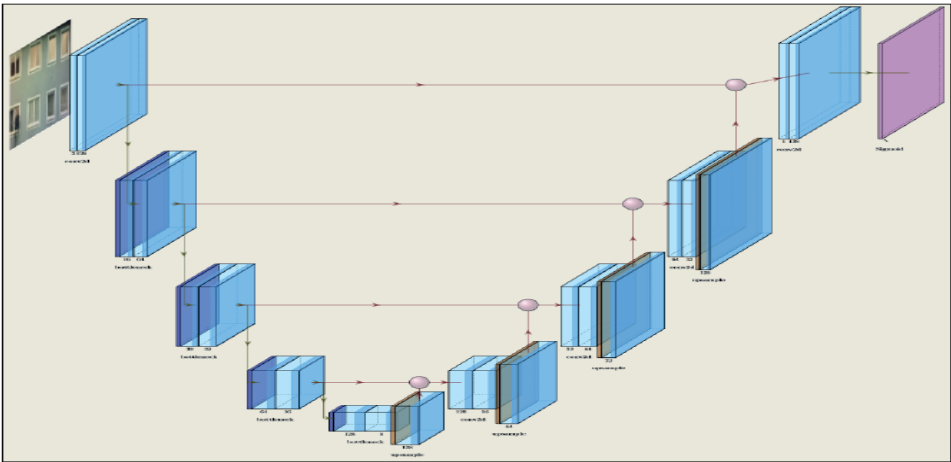


Abb. 3: Verwendete U-Net-CNN-Architektur für Erkennungs- bzw. Segmentierungsaufgaben

Hierzu bietet sich architektonisch ein U-Net Ansatz an, mit dem Unterschied, dass man nicht wie in Schlender et. al. [Sc21] ein vortrainiertes MobileNetV2-Modell auf einen neuen Anwendungsfall trainiert. Vergleicht man die Architektur vom MobileNetV2 [Sa19] und die verwendete U-Net Architektur in Abbildung 3, so erkennt man eine vereinfachte Version mit neun statt elf Convolutions. Auch beim „bottleneck“ (Flaschenhals-Filterung) und beim „up sampling“ skaliert die Architektur nicht so kleinteilig herunter und kommt gut mit den in dieser Arbeit behandelten Segmentierungsaufgaben zurecht.

2.3 Segmentierung von Fensterern

Beim Trainieren des CNN-Modells mit der Architektur wurden Normalbilder als Input-Daten und ein „Training-Ende“ verwendet, wenn das Modell bei der Minimierung des Validation-Loss über mehrere Epochen keine Fortschritte erzielt. Als Ergebnis wurde nach 136 Epochen ein Validation-Loss von $\approx 0,23$ (Abbildung 4) erzielt.

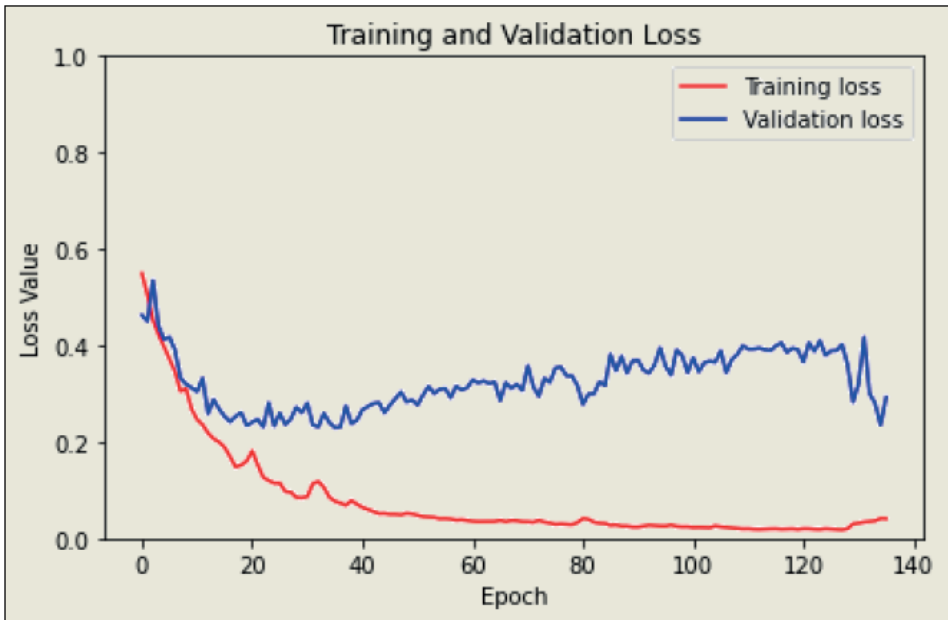


Abb. 4: Loss-Verlauf zur Fenstererkennung

Im Vergleich zu Experiment 2 von Schlender et. al. [Sc21] zeigt dieses Ergebnis mit einem Validation-Loss von 0,23, dass es möglich ist, mit einem eigenen Modell auf ein ähnlich gutes Ergebnis zu kommen wie beim Transfer-Learning (Validation-Loss von 0,22) eines vortrainierten MobileNetV2-Modells.

Die sich ergebenden Masken zeigen eine erfolgreiche Erkennung und Segmentierung (vgl. Abbildung 5) in bekannten Trainingsdaten (oben) und für das Modell mit unbekanntem Validierungsdaten (unten). Es ist jeweils ein Ergebnis mit dem Ursprungsbild (links), der gesuchten Maske (Mitte) und dem vom Modell erkannten Maskenbereich (rechts) zu sehen.



Abb. 5: Maskensegmentierung in den unbekanntem Validierungsdaten

Man kann auf den Bildern erkennen, dass sowohl auf bereits bekannten Bildern als auch auf dem Modell noch unbekanntem Bildern Fenster erfolgreich erkannt und mittels Masken segmentiert werden.

2.4 Segmentierung von Pflanzen und Dachelementen

Bei der Segmentierung weiterer Objekte wie Pflanzen oder einem Hausdach zeigt sich ebenfalls ein annehmbares Ergebnis. Es wurden 407 Bilder manuell mit Labels versehen und neue Input-Daten erzeugt. Nach der Vorverarbeitung sind entsprechende Input-Daten (vgl. Abbildung 6) zu sehen.



Abb. 6: Input-Daten-Bilder, manuell mit „Pflanze“-Label versehen

Der Umstand beim manuellen Labeln (Quadratische-Regionen) legt nahe, dass ein relativ gutes Ergebnis mit einer Validation-Loss von 0,37 (Abbildung 7) vom trainierten Modell mit Epoche 149 erreicht wird. Das Ergebnis bestätigt allerdings auch, dass es mit gleicher Architektur möglich ist, weitere Objekte wie Pflanzen oder ein Hausdach segmentieren zu können. Ein genaueres Labeln kann in diesem Fall zu einem genaueren Ergebnis führen.

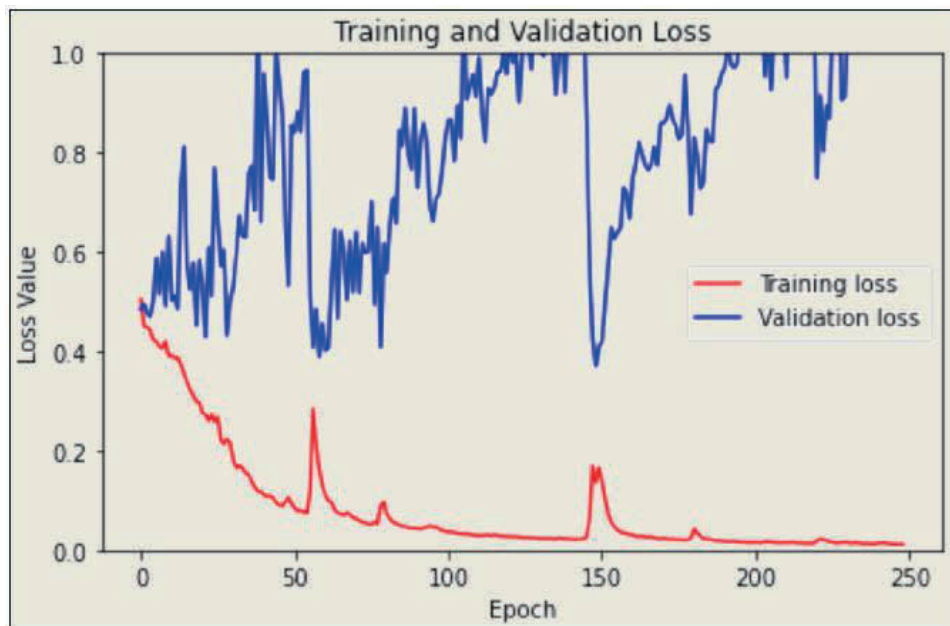


Abb. 7: Loss-Verlauf zur Pflanzen- und Hausdachsegmentierung

Anhand des Validation-Loss-Verlaufs in Abbildung 7 ist – wie schon im vorhergehenden Kapitel beim Modell-Training angewendet – erkennbar, dass der Validation-Loss nach Epoche 149 mit $\approx 0,37$ nicht besser wurde und dass das Training ein „Training-Ende“ verwendet, wenn das Modell bei der Minimierung des Validation-Loss über mehrere Epochen keine Fortschritte erzielt. Am Kurvenverlauf kann man erkennen, dass es vereinzelt Verbesserungen gegeben hat (wie bei Epoche 59 mit einem Validation-Loss von $\approx 0,39$ und anschließend ein Validation-Loss von $\approx 0,37$ bei Epoche 149); an dieser Stelle wurde das Modell gespeichert, weil keine Verbesserung mehr eintrat. Insgesamt kann man annehmen, dass es bei ca. 50 Epochen und bei ca. 150 Epochen gute Feature-Maps innehat und man das Training weiterführen könnte, weil ggf. bei Epoche 450 wieder eine Verbesserung zu erwarten ist.

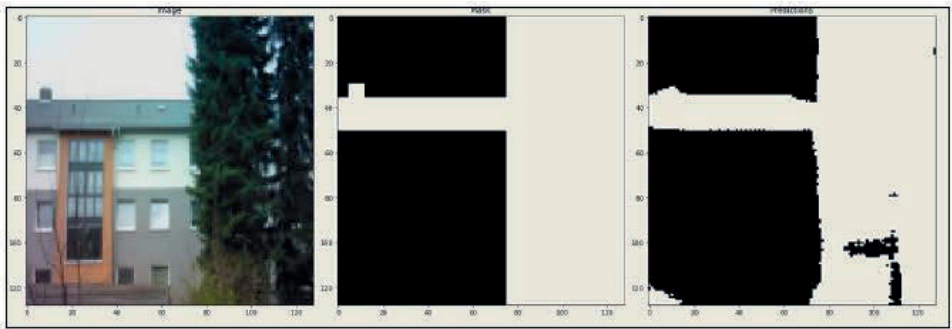


Abb. 8: Maskensegmentierung von Pflanzen und Hausdachelementen

Mit dem Anwendungsfall, Pflanzen und Hausdächer zu segmentieren, kommt das Modell mit der aktuellen Epoche gut zurecht (Abb. 8), sodass das trainierte Modell mit weiteren Hyperparameteroptimierungen [KSR20] und insbesondere besser gelabelten Input-Daten für seine Aufgabe weiter optimiert werden konnte. Man kann erkennen, dass das Label (Abb. 8, Mitte) weitestgehend gut segmentiert wird (Abb. 8, rechts).

2.5 Segmentierung von Wärmebereichen und Wärmebrücken

Um einen weiteren Anwendungsbereich von CNN-Technologien zu behandeln, lassen sich im Thermografie-Bereich auch Wärmebrücken in Infrarotbildern erkennen und segmentieren. Dieser Ansatz kann die Analyse von Thermographiebildern automatisieren. Anders als beim manuellen Labeln von Input-Daten der vorhergehenden Kapitel wurde in diesem Anwendungsbeispiel die Vorverarbeitung mittels klassischer Bildverarbeitungsalgorithmen unternommen und die Bereiche mit Wärmebrücken wurden manuell gelabelt, siehe Abbildung 2 aus Kapitel 2.1.

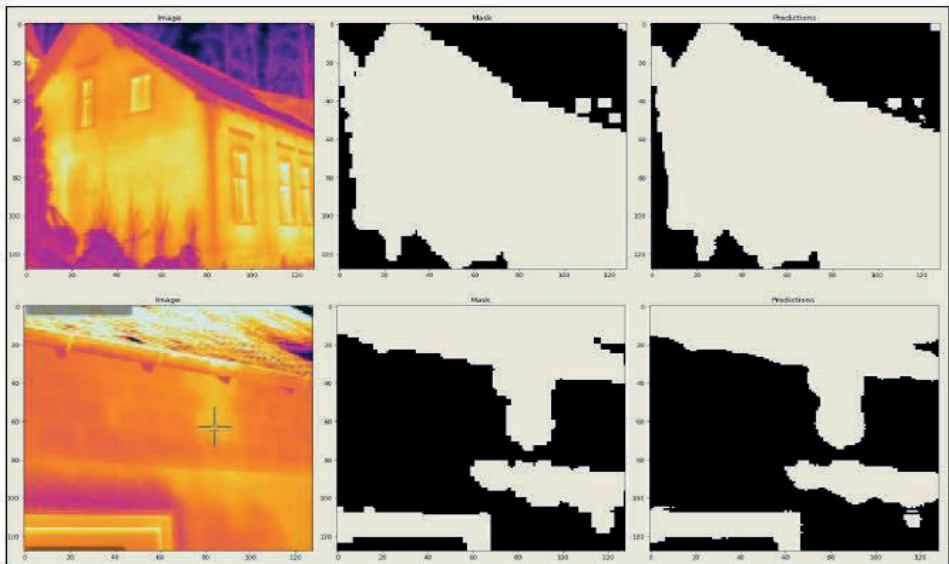


Abb. 9: Maskensegmentierung von Wärmebereichen und -brücken

Bei beiden trainierten CNN-Modellen sieht man ein relativ gutes Ergebnis. Beim Segmentieren von Wärmebereichen sind im Allgemeinen die wärmestrahrenden Flächen erkannt worden (Abbildung 9 erste Bildzeile rechts). Die Aufgabe, Wärmebrücken zu segmentieren, ist im Ansatz ebenfalls erfolgreich umgesetzt, potenzielle Schwachstellen in Fassaden konnten automatisiert lokalisiert werden.

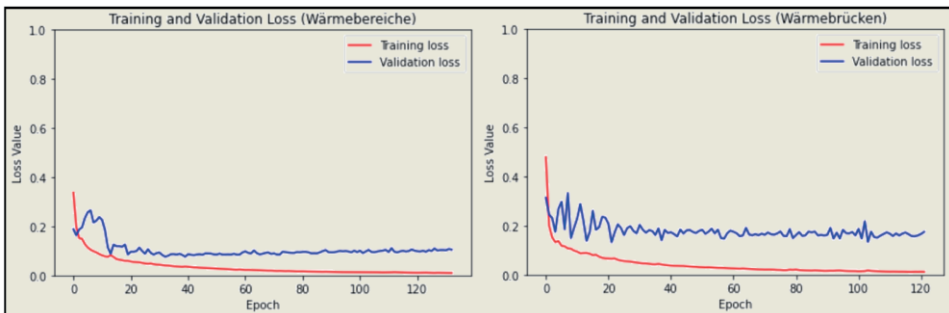


Abb. 10: Loss-Verlauf zur Wärmebereichs- und Wärmebrückensegmentierung

Die Validation-Loss beim Segmentieren von Wärmebereichen (Abbildung 10, links) liegt ab Epoche 33 bei $\approx 0,08$ und beim Segmentieren von Wärmebrücken ab Epoche 22 bei $\approx 0,13$. Dieses Ergebnis verdeutlicht, dass CNN bei der Segmentierung von Bereichen in Infrarotbildern sehr gut arbeiten können, denn jedem Pixel ist anhand eines Farbwertes ein Informationsgehalt zur Strahlungswärme zugeordnet. Ein noch genaueres Ergebnis lässt

sich erzielen, wenn Thermografie Expertinnen und Experten die Input-Daten manuell labeln. Mit der Erkennung von Wärmebrücken kommt die klassische Bildverarbeitung unter Verwendung einer Histogrammanalyse sehr gut zurecht [BAS17]. Die Tatsache, dass dieser mathematische Sachverhalt von einem CNN erlernt werden kann - um die Erkennung von Wärmebrücken zu erzielen - verdeutlicht, dass moderne CNN-Algorithmen eine gute Zukunft haben und ihnen gestellte Aufgaben übernehmen können.

3 Fazit und Diskussion

Basierend auf den erfassten, gemessenen und berechneten Daten kann eine automatisierte und einfache Bildverarbeitung durchgeführt werden. Die Daten können später von Expertinnen und Experten zur Abschätzung des Sanierungspotentials der Außenwände verwendet werden. Sie können dann besser entscheiden, ob eine Sanierung sinnvoll ist oder nicht. Es ist aber auch zu berücksichtigen, unter welchen Umgebungsbedingungen die Bilder aufgenommen wurden. Wurde das Bild beispielsweise bei wärmeren Außentemperaturen aufgenommen, kann der Sanierungsbedarf geringer ausfallen, als dies bei kälteren Temperaturen der Fall gewesen wäre.

Die Untersuchung zeigt, dass es mit Hilfe von künstlicher Intelligenz möglich ist, verschiedene Objekte in einem Thermographie-Bild zu segmentieren. Die Verwendung von Bildsegmentierung und damit die Entfernung von unerwünschten Objekten aus einem Thermographie-Bild mittels Objekterkennung führt zu einem gesäuberten Bild, in dem anschließend eine approximierte Durchschnittstemperatur berechnet werden kann.

Das bestehende System kann weiter verbessert und erweitert werden. Durch das Anlegen zusätzlicher und noch genauerer Label für verschiedene Objekte können auch weitere Segmentierungsaufgaben mit einem CNN-Modell erkannt werden. Im weiteren Environ-Projektverlauf soll eine Methode entwickelt werden, die eine schnellere und einfachere Berechnung der von Außenwänden abgestrahlten Wärmeenergie ermöglicht. Die Arbeit kann im Hinblick auf die Erkennung von Wärmebrücken ähnlich wie in Behrens et. al. [BAS17] zu einer schnelleren Bewertung möglicher Energieverluste in sanierungsbedürftigen Gebäuden beitragen und den Entscheidungsprozess für oder gegen eine Sanierungsmaßnahme in der Zukunft unterstützen. Dadurch wird auch ein Beitrag zur Reduzierung der gesamten CO₂-Emissionen geleistet und die Umwelt entlastet.

Literaturverzeichnis

- [Am18] Åmdal-Sævik, K. (2018). Keras U-Net starter - LB 0.277. Online verfügbar <https://www.kaggle.com/keegil/keras-u-net-starter-lb-0-277>; Datum letzter Zugriff 24.04.2021
- [BAS17] Behrens, G., Abele, O., Schlender, K. (2017): Visualisierung von 3D-Objekten anhand sensorfusionierter Thermografie- und Tageslichtaufnahmen mit automatischer

Lokalisierung von Wärmebrücken in der Gebäudeaußenhülle. Informatik 2017, WS30 UINW 2017 - 5. Workshop Umweltinformatik zwischen Nachhaltigkeit und Wandel, Chemnitz 2017 ISBN 978-3-88579-669-5

- [Be17] Behrens, G., Hamelmann, F., Thiel, C., Försterling, T., Weicht, J., Fehring, F., Schlender, K., Dreimann, R. (2017). Smart measuring system of air quality accompanying a renovation process of apartment buildings. *EnviroInfo* 2017, Luxemburg, Springer Nature 2018, imprint in "From Science to Society – new Trends in Environmental Informatics" ISBN 978-3-319-65686-1
- [BSF18] Behrens, G., Schlender, K., and Fehring, F. (2018). Data mining methods of healthy indoor climate coefficients for comfortable well-being. *Environmental Protection and Natural Resources; The Journal of Environmental Protection - National Research institute*. 29, 3, 7-12, Available From: De Gruyter - Sciendo. Online verfügbar <https://content.sciendo.com/view/journals/oszn/29/3/article-p7.xml>; Datum letzter Zugriff 19.02.2021
- [En16] EnEV: Neubau Wohngebäude ab 2016. https://enev-online.com/enev_praxishilfen/vergleich_enev_2016_enev_2014_neubau_wohnbau_15_04.06.html (2016); Datum letzter Zugriff 10.04.2021
- [ETW16] Engel K., Thomsen K., Wittchen K. (2015). Implementing the Energy Performance of Buildings Directive (EPBD). Lisbon, September 2015, ISBN 9789728646325. Online verfügbar <http://www.epbd-ca.eu/outcomes/2011-2015/CA3-BOOK-2016-A-web.pdf> (2016); Datum letzter Zugriff 10.04.2021
- [HBF19] Hussain M., Bird J.J., Faria D.R. (2019) A Study on CNN Transfer Learning for Image Classification. In: Lotfi A., Bouchachia H., Gegov A., Langensiepen C., McGinnity M. (eds) *Advances in Computational Intelligence Systems*. UKCI 2018. *Advances in Intelligent Systems and Computing*, vol 840. Springer, Cham. https://doi.org/10.1007/978-3-319-97982-3_16
- [KSR20] Krishnakumari, K., Sivasankar, E., Radhakrishnan, S. (2020). Hyperparameter tuning in convolutional neural networks for domain adaptation in sentiment classification (HTCNN-DASC). *Soft Comput* 24, 3511–3527. Online verfügbar <https://doi.org/10.1007/s00500-019-04117-w>
- [PTT12] Pleșu, R., Teodoriu, G., Taranu, G. (2012). INFRARED THERMOGRAPHY APPLICATIONS FOR BUILDING INVESTIGATION. *Bulletin of the Polytechnic Institute of Iasi - Construction & Architecture Section*. 62. S. 157-168. Online verfügbar https://www.researchgate.net/publication/236016131_INFRARED_THERMOGRAPHY_APPLICATIONS_FOR_BUILDING_INVESTIGATION
- [Sa19] Sandler, M., Howard, A., Zhu, M., Zhmoginov A., Chen L. (2019). MobileNetV2: Inverted Residuals and Linear Bottlenecks. Online verfügbar <https://arxiv.org/abs/1801.04381>; Datum letzter Zugriff 27.04.2021
- [Sc21] Schlender, K., Riechmann, M., Müller, J., Behrens, G. (2021). Machine learning methods for approximating the temperature of exterior walls using thermal images and colour images of building facades. *Environmental Informatics: New Perspectives in Environmental Information Systems: Transport, Sensors, Recycling*. Shaker Verlag. <https://doi.org/10.2370/9783844076288>.

- [Se20] Sennestadt GmbH (2020). Stadtlabor Sennestadt. Online verfügbar <https://www.sennestadt-gmbh.de/stadtlabor-sennestadt.html>; Datum letzter Zugriff 19.02.2021
- [Um20] Umwelt Bundesamt (2020). Treibhausgasminderungsziele Deutschlands. <https://www.umweltbundesamt.de/daten/klima/treibhausgasminderungsziele-deutschlands>; Datum letzter Zugriff 19.02.2021

Effizientere Flächennutzung in der Landwirtschaft mit vernetzten Drohnen

Matthias Nattke ¹

Abstract: Dieses Paper beschäftigt sich mit der Frage, wie im Schwarm vernetzte Drohnen zur nachhaltigen Nutzung von landwirtschaftlichen Flächen beitragen können. Dazu wird auf die Vorteile und höhere Krisenresistenz regionaler Lieferketten bei Grundnahrungsmitteln hingewiesen. Als Resultat wird aufgezeigt, wie vernetzte vernetzte Drohnen zur gezielten Beobachtung von vorhanden Flächen eingesetzt werden können. Sie lassen sich präziser und nachhaltiger bewirtschaften. Die Betrachtung umfasst technische Herausforderungen bei der Vernetzung und dem Schwarmflug. Weitere Verbesserungsmöglichkeiten durch längere Akkulaufzeiten und bessere Sensoren werden vorgeschlagen. Ein Anwendungsszenario in der Lausitz wird aufgezeigt. Im Ausblick wird die Integration in neuartige Bearbeitungsmethoden der Präzisionslandwirtschaft thematisiert.

Keywords: Drohnen; Landwirtschaft; Flächennutzung; Schwarmbildung; Lokalisierung; Vernetzung

1 Einführung

Die Hälfte der genutzten Fläche steht in Deutschland der Landwirtschaft zur Verfügung. Da die fruchtbare Ackerfläche in höherem Maße durch klimatische Einflüsse und zunehmende Bebauung in Ballungsgebieten beansprucht wird, sind vorhandene Flächen nachhaltig zu bewirtschaften. Dazu ist es essentiell Flächen und Pflanzen bodenkundlich zu beobachten. Dies trägt zur Erhaltung der Ertragskraft bei und sichert, in Kombination mit regionalen Lieferketten, die stets zuverlässige Versorgung mit landwirtschaftlichen Grundprodukten ab. Globale Lieferketten erhöhen durch gegenseitige Abhängigkeiten das Risiko, kurzfristig steigender Preise. Dies kann Exportbeschränkungen großer Exporteure und hoher Nachfrage durch Notkäufe zur Folge haben [OE20]. Insbesondere in ärmeren Ländern mit geringeren Möglichkeiten zum Anbau, führt dies zu Nahrungsknappheit [HG20].

In zahlreichen Anwendungsszenarien ermöglichen Drohnen (UAVs), durch Bodenkartierung aus der Luft, Erkenntnisse über die Eigenschaften und Qualität einer Fläche zu erhalten. Die Gebietsvermessung, Erkundung des Wasseranteils, die Erkennung des Bodentyps, die Erfassung der Vegetationsdichte und Zustandsbeschreibung von Pflanzen wird erleichtert [BSL11]. Große Flächen bedeuten komplexe, langandauernde Missionen. Die Aufgaben können durch mehrere zum Schwarm kombinierte Drohnen effizienter bewerkstelligt werden, da durch mehrere Flugbahnen geringerer Länge mehrere Messungen parallel durchgeführt werden können.

¹ BTU Cottbus - Senftenberg, Lehrstuhl Industrielle Informationstechnik, Universitätsstraße 22, 03048 Cottbus, Deutschland matthias.nattke@b-tu.de

2 Verwandte Arbeiten

Die bekanntesten Untersuchungen zu im Schwarm vernetzten Drohnen, finden sich gegenwärtig im Eventbereich zur Unterhaltung. Üblicherweise werden mehrere tausend Drohnen in aufwändigen Choreografien als Alternative zum Feuerwerk eingesetzt. Um dies zu erreichen werden die Flugbahnen im Vorfeld mit hoher Rechenleistung akribisch berechnet und die Missionen durch eine zentrale Steuereinheit zur Koordinierung der Drohnen ausgeführt [Co20]. Im Bereich der Landwirtschaft können diese Systeme nur schwer eingesetzt werden. Sie stehen allein durch die langwierige Berechnung der Flugpfade der Nachhaltigkeit entgegen und können schlecht auf spontane Gegebenheiten reagieren. Vielmehr sind Adhoc-Missionen mit dezentralen Steuersystemen und deutlich weniger Drohnen mit Lösungen für komplexe Aufgaben im Fokus der Betrachtung.

Zur Entwicklung von Algorithmen der Kollisionsvermeidung und als Basis der Beschreibung von Formationen, wurde in [Vá18] das Verhalten von Vogelschwärmen untersucht. Durch Simulation und die Integration in kollektive Formationsflüge mit 30 Drohnen innerhalb einer abgegrenzten Fläche, konnten durch veränderte Geschwindigkeiten Kollisionen dezentral vermieden werden. Nachteilig ist die fehlende Garantie am Ende der Mission jeden Bereich des begrenzten Gebietes tatsächlich abgeflogen zu haben.

3 Funktionsweise vernetzter Drohnen

Zum besseren Verständnis werden die Hauptherausforderungen bei der Steuerung von Drohnenschwärmen herausgearbeitet. Zudem wird die Funktionsweise der Teilkomponenten beschrieben. Sie umfassen automatisierte Drohnensteuerung, relative Lokalisierung, Kommunikation zwischen allen Schwarmbeteiligten, Flugstrategie und Kollisionsvermeidung. Durch die Zusammenwirkung aller Bestandteile und Prozesse, lassen sich die Drohnen über eine Basisstation beobachten, können automatisiert fliegen und werden zielgerichtet zum Schwarm gruppiert.

3.1 Aufbau und Schwarmkomponenten

Die Basis vernetzter Drohnen bildet die automatisierte Regelung des Flugverhaltens. Jeder Einzeldrohne muss es möglich sein, die Mission anhand der vorgegebenen Route autonom zu fliegen. Dies umfasst die Regelung der Flugrichtung, indem sie beibehalten oder durch Nick- und horizontale bzw. vertikale Drehbewegungen manipuliert wird. Dies erfolgt über eine an die aerodynamischen Eigenschaften des Flugobjektes angepasste Anordnung der Propeller und stetige Anpassung der Motordrehzahl. Es ist möglich die Geschwindigkeit zu erhalten oder bei Quadroptern adaptiv so zu verändern, dass sie auf der Stelle schweben [Ch14].

Zur gezielten Bewegung braucht es die relative Lokalisierung aller Drohnen im Schwarm. Das bedeutet das kollektive Wissen über den Standort und die Flugbewegung einer Drohne

mindestens in Abhängigkeit zu ihren Nachbarn. Im einzelnen Flugobjekt sind im Steuergerät zur genauen Orientierung und Lokalisierung mehrere Sensoren verbaut bzw. in dessen Nähe angebracht. Lokal gibt es Accelerometer, Kompass und Barometer, womit sich Beschleunigung, Richtung und Höhe der Drohne im Flug bestimmen lassen und die Orientierung der Drohne, in der Steuerungsroutine Berücksichtigung finden. Beispielsweise können Windturbulenzen in gewissen Grenzen ausgeglichen werden. Zur Positionierung im Außenbereich wird häufig eine Kombination aus GPS, Galileo und Glonass als GNSS eingesetzt. Die Position des Flugobjektes wird über einen Konsensusalgorithmus mit durchschnittlicher Genauigkeit von drei Metern bestimmt [MFZ19].

Bei Distanzen unter zehn Metern zwischen den Drohnen und in Bereichen mit schlechtem GNSS-Empfang, ist die Positionierung zusätzlich mit einem LiDAR-System oder über UWB-Sensoren denkbar. Denn es werden genauere Abstandsmessungen und Positionsdaten im Zentimeterbereich benötigt. Für die grobe Orientierung sind die globale Satellitenavigation zuständig, während im Nahbereich eine Art Karte aus Punktwolken exakte Distanzmessungen zu Objekten in der Nähe erlauben. Beim Lidar werden Laserimpulse ausgesendet, die an Objekten in deterministischer Weise reflektiert, gestreut oder absorbiert werden. Je nach Auflösung lassen sich detailreiche dreidimensionale Karten aus den Distanzmessungen generieren. Zugleich lassen sich Oberflächenreliefs und Anomalien, wie Äste von Bäumen erkennen [Zh21].

Im Unterschied dazu lässt sich mit Ultra-Breitband-Sensoren, über eine hochfrequente Laufzeitanalyse, die Distanz zwischen mehreren Flugobjekten bestimmen. Sie funkeln mindestens bei 500 MHz kurzen Distanzen [Ho16, HWO21]. Dieses Verfahren ist bei sehr kleinen Micro-Drohnen empfehlenswert, da hohe Energieeffizienz und die Übertragung geringer Datenmengen zur Identifizierung der Drohnen möglich sind. Über die zeitliche Betrachtung der einzelnen Flugbewegungen und der Positionsänderung, wird die relative Lokalisation von Drohnen im Schwarm über das Steuerungsgerät vorgenommen. Darauf aufbauend braucht es für die spätere Kollisionsvermeidung zuverlässige Kommunikation zwischen den Drohnen im Schwarm, sowie einer am Boden befindlichen Kontrollstation für den Drohnenpiloten. Mehrere Ansätze zur Kommunikation sind etabliert und miteinander kombinierbar. Zum Einen können die Drohnen des Schwarms über eine Art Relaydienst vollständig vom Boden kontrolliert werden und führen in diesem Fall exakte Kommandos aus. Hierbei ist die Bodenstation als Single Point of Failure betrachtet, dem Risiko des Signal- und Nachrichtenverlustes ausgesetzt. Je größer der Drohnenschwarm wird, desto mehr steigt der Aufwand zur Aufrechterhaltung der Kommunikation und Verarbeitung der Maneuver. Zum Anderen kann die Mission im Vorfeld geplant werden, sodass Pfade aus Koordinaten im dreidimensionalen Raum entstehen. Sie werden zur Ausführung vollständig an die Drohnen übergeben, sodass lediglich Korrekturkommandos vom Boden erwartet werden und die Beobachtung ermöglicht bleibt. Diese agieren folglich weitestgehend autonom [Le18, Co20]. Zur Absicherung der Kommunikation sollten wichtige Kommandos als Broadcast an alle Drohnen des Schwarm gesendet werden und kollektiv auf Plausibilität geprüft sein. Durch die Verwendung von Telemetrie-Signalen auf weiten Entfernungen und UWB-Kommunikation

im Nahbereich kann ein Mash-Netzwerk aufgespannt werden, welches den Schwarm und die Bodenstation umfasst. Die Bodenstation kann als Fernbedienung oder Kontrollsystem mit grafischen Bedienelementen ausgelegt sein. Sie dient der Planung der Mission, zur Generierung der geplanten Pflugpfade und ist ständige Beobachtungsinstanz der einzelnen Zustandsparameter innerhalb der Flugobjekte [Ko14]. Als Vermittlung zwischen Piloten und Drohnenschwarm ermöglicht sie die Beeinflussung des Flugs.

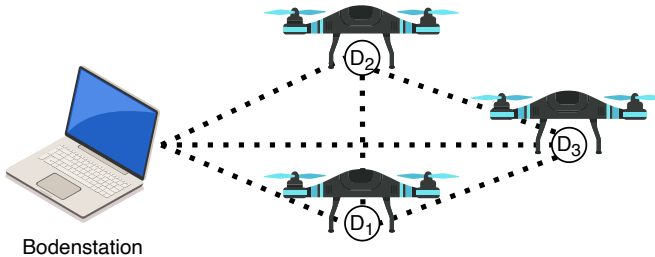


Abb. 1: Mash-Kommunikation mit Drohnenschwarm und Bodenstation [F117]

Sind alle Drohnen im Schwarm in der Lage in den Flugbetrieb überzugehen, werden Sensoren für spezifische Aufgaben verwendet. Besonders oft kommen optische Videoaufnahmen zum Einsatz. Infrarot-, Thermal-, 360-Grad-, UV-, Stereoskopie- oder Feuchtigkeitsaufnahmen sind möglich. Weiterhin sind bereits beschriebene Lidar-Systeme zur Oberflächenbestimmung und die Audioanalyse üblich bei Einzeldrohnen. Im Drohnenschwarm hat das Missionsziel Einfluss auf die Auswahl der an den Flugobjekten angebrachten Sensoren und deren Flugstrategie. Es lassen sich unterschiedliche Sensoren in definierten Abständen kombinieren und fliegende, drahtlose Sensornetze bilden [Ma18].

3.2 Missionsplanung

Die Flugstrategien umfassen die Pfadplanung und Zusammenstellung der Schwarmformation. In jedem Fall wird ein Gebiet über Geofencing durch seine Grenzen definiert und alle darin befindlichen Bereiche abgetastet [Sa20]. Trivial wird nach dem Zufallsprinzip durch Umsteuerung an den Grenzen oder Hindernissen überflogen. In diesem Fall wäre der Schwarm lediglich durch die kommunikative Zusammengehörigkeit definiert und alle beteiligten Drohnen würden selbstständig agieren. Die Bodenstation könnte global alle Pfade überwachen und bei Flächenabdeckung die Mission beenden. Vorteilhaft sind die einfache Steuerung, beliebig definierte Startpunkte und die vereinfachte Pfadplanung im Vorfeld. Nachteilig ist die lediglich grobe Schätzung der Missionsdauer, der benötigten Energie und die späte Verfügbarkeit des Gesamtergebnisses [Ta21]. Für das effiziente Monitoring von landwirtschaftlichen Flächen ist es von Vorteil, bereits während der Durchführung eine Auswertung der gewonnenen Daten vornehmen zu können. Dazu gilt innerhalb des beobachteten Gebietes die Strategie des „Auf und ab“-Fliegens. Die Pfade werden im Vorfeld auf die kürzeste Flugdauer optimiert, Teilpunkte definiert und mit

geringen Lokalisierungsabweichungen durch die Drohnen angefliegen [CBFJ19]. Von Vorteil ist die für den Piloten übersichtliche Erfassung des Missionsfortschrittes und Schätzung der Dauer. Nachteilig erscheint der höhere Koordinierungsaufwand bei größeren Flächen und mehr Flugobjekten.

Die Schwarmformation besteht aus den Grundmustern von in gleichmäßigen Abstand neben- oder nacheinander angeordneten Drohnen. Funktional werden gleichartige Aufgaben nebeneinander durch die Reihenformation parallelisiert, wodurch sie entsprechend der Drohnenzahl schneller abgearbeitet sind. Für verschiedene oder redundante, aufeinanderfolgende Aufgaben wird die Kettenformation eingesetzt. Sie kann auf die Einzeldrohne bezogen im Platoon energiesparender eingesetzt sein und exaktere Messungen ermöglichen [Ch15]. Aus den beschriebenen Grundmustern lassen sich die in Abbildung 2 dargestellten Mischformen zu Schwarmformationen zusammenfügen.

Die Vorteile der Formationen sind die Verknüpfung der Beobachtungen, Redundanz zur Fehlerkorrektur und höhere Ausfallsicherheit. Denn bei aktuellen Flugzeiten von 15 bis 20 Minuten müssen nach dieser Zeit die Akkus ausgetauscht und aufgeladen werden. Die Mission wird im Schwarm nahtlos weiter verfolgt. Beispielsweise können in der Kreuzformation die beiden vorderen Drohnen auf eine hochauflösende RGB-Aufnahme spezialisiert sein, die Mitteldrohne hochfrequente Schallmessungen vornehmen und die hinteren Drohnen jeweils einen Lidar-Scanner besitzen.

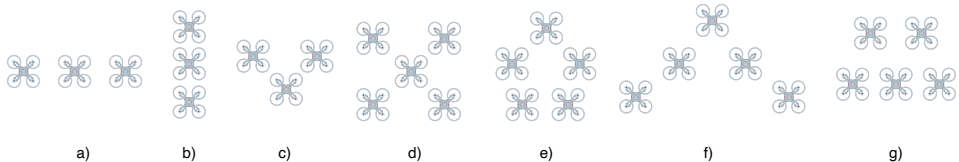


Abb. 2: a) Reihe, b) Kette, c) Dreieck, d) Kreuz, e) Kreis, f) Pyramide, g) Trapez

3.3 Kollisionsvermeidung

Wenn Drohnen in kürzeren Abständen von unter zehn bis zwanzig Metern zueinander fliegen und äußeren Hindernissen ausweichen müssen, braucht es zwingend Strategien zur Kollisionsvermeidung. Dabei gilt es Interferenzen zu vermeiden. Flüge in geringem Abstand zum Boden können die Steuerung erschweren, wenn die Drohne durch Luftströme Seiteneffekten ausgesetzt ist, abdriftet und am Boden aufprallt. Zwischen zwei Drohnen übereinander kann es für die untere Drohne zu Turbulenzen kommen, sodass die Steuerung erschwert wird und sie abdriftet. In der Landwirtschaft gibt es Einflüsse von Bäumen, Gebäuden, großen Windkraftanlagen oder schweren Landmaschinen. Dies muss allen Betrachtungen zwingend ausgeschlossen sein, da zerstörte Drohnen zur Gefahr für Unbeteiligte werden [Co20].

Die beschriebenen Effekte unkontrollierter Driftbewegungen zwischen Drohnen können vernachlässigt werden, wenn sie gleicher Höhe fliegen. Start- und Landemanöver erfolgen

zeitgleich bei gleichbleibender Position im Abstand von zehn Metern. Um die Kollisionsvermeidung zu realisieren wird aus der Abweichung der Eigengeschwindigkeit einer Drohne V_e und den Geschwindigkeiten der Nachbardrohnen V_n die relativen Geschwindigkeiten V_r , sowie der Abstand zum Nachbarn ermittelt. Ist die relative Geschwindigkeit bei einem Abstand von zehn Metern größer als ein Meter pro Sekunde, kann von einer Kollision ausgegangen werden [Vá18]. Zur Kollisionsvermeidung müssen die Geschwindigkeiten der Drohnen nahezu identisch sein, um im Schwarm zu agieren. Um Kollisionen mit nachfolgenden Drohnen zu vermeiden wird die Geschwindigkeit adaptiv reduziert. Es ist vorausschauender und durch geringere Aufwände zum Gegensteuern energiesparender. Bei Erkennung eines äußeren Hindernisses, findet ein horizontales Ausweichmanöver statt. Von Seiten des Hindernisses kann es die Pfadanpassung des Schwarms beeinflussen.

4 Anwendungsbeispiel

In der Lausitz sind die meisten Böden sehr leicht und sandig. Durch Witterungsbedingungen mit Niederschlägen mit etwa 500 l/m^2 Jahresniederschlag ist der Boden trocken und die Bewirtschaftung sehr großer Flächen üblich [LM03, Ge15]. Betriebe mit über tausend Hektar Anbaufläche und Einzelfelder über fünfzig Hektar sind keine Seltenheit, während überwiegend einheitlich Getreide, Öl- und Hackfrüchte wie Sonnenblumen oder Kartoffeln angebaut werden [St17]. Dies erfolgt mit landwirtschaftlicher, sehr schwerer Großtechnik. Der bisherige in der Forschung oder Industrie betriebene Einsatz von einzelnen Drohnen zur Beobachtung ist nicht praktikabel, da batteriebetriebene Geräte Flugzeiten von 15 bis 30 Minuten aufweisen und Beobachtungsaufgaben nicht parallel ablaufen [CBFJ19]. Dies führt bei großen Flächen zu Flugmissionen, die sich auf gezielte Einzelbeobachtungen beschränken oder mit geringerer Detailgenauigkeit aus großer Höhe erfolgen. Dies lässt den zeitlichen Aufwand der Beobachtung über den der Bewirtschaftung ansteigen.

4.1 Drohnen- und Systemkonfiguration im Schwarm

Es eignet sich die parallele Beobachtung der Fläche mit mehreren im Schwarm verknüpften Drohnen. Drei kleinere Drohnen mit einem Gewicht von zwei Kilogramm werden. Sie verfügen über einen PixHawk-Flugcontroller [Me12] und die zugehörige Telemetrie zur Kommunikation über MAVLink [Ko19] im Schwarm, sowie der Bodenstation als eine um Schwarmfunktionen erweiterte QGroundControl-Instanz. Die Drohnen fliegen in Dreiecksformation (Siehe Abbildung 2) mit gleichmäßigem Abstand von 25 Metern zueinander in 90 Metern Höhe. Die Kollisionserkennung kann sicher agieren. Zur Lokalisierung dient eine GNSS-System-Kombination. Die elektrischen Motoren sind aus 5 Metern 40 dB laut, Umwelteinflüsse auf Vögel sind gering. An den vorderen ist Drohnen je eine Multispektralkamera und an der hinteren eine RGB-Kamera angebracht.

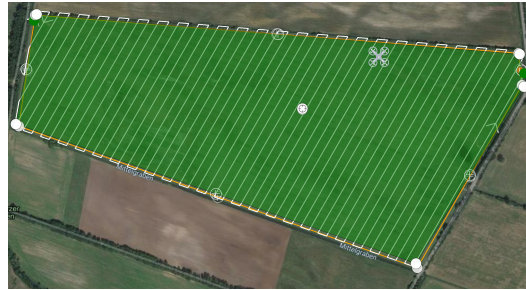
Die Pfadplanung zur Missionsdurchführung erfolgt auf Grundlage der Abgrenzung der betrachteten Fläche über Polygon-Geofences von der Umgebung. Die Zusammenfassung

des Schwarms als zusammenhängendes Flugobjekt aus mehreren in gleicher Höhe im sicheren Abstand vereinfacht die Betrachtung der Flugeigenschaften. Somit lassen sich ausgehend von der Schwarmmitte die Flugbahnen in wechselnden Korridoren festlegen. Die Drohnen berechnen aus ihrer Position die relativen Abstände zu ihren Nachbarn. Bei der jeweiligen Auflösung von 5472x3648 Pixeln erfassen die Sensoren bei einer Überlappung von vierzig Prozent, in Kombination 2,6 Zentimeter der Fläche pro Pixel. Einzeldrohnen arbeiten mit einer Überlappung von siebenzig Prozent, im Schwarm gibt es durch die Parallelaufnahmen bereits eine größere Überschneidung. Im Unterschied zu Einzeldrohnen sind die Korridorabstände deutlich weiter gefasst, wodurch die Anzahl der nacheinander geplanten Überflüge reduziert wird. Die Pfadplanung wird aus Sicht des Anwenders in ihrer Komplexität verringert, während die Steuerung der relativen Pfade von Drohnen über die Schwarmformation reguliert wird. Dies ist anspruchsvoller, da die Wendemanöver am Ende jedes Korridors die Neuordnung oder vollständige Drehung des Schwarms erfordert.

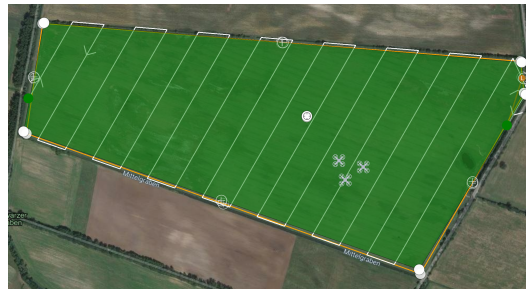
In Abbildung 3 sind die Pfade auf einer etwa 62 Hektar großen Fläche mit Winterweizen geplant. Je nach Pflanzenart würde die Flughöhe und die Sensorikkonfiguration variieren. Die Abbildung 3a geht von einer Einzeldrohne aus, welche am oberen Startpunkt mit dem Flug beginnt und die Mission bis zum Ende durchführt. Etwa alle zwanzig Minuten kehrt sie zurück zum Batterietausch. Anschließend wird die Mission am zuletzt abgetasteten oberen Flugkorridor fortgesetzt. Die vollständige Pfadlänge ohne Zwischenstopp beträgt 27 Kilometer und einschließlich der zusätzlichen Strecke für An- und Rückflug 32 Kilometer. Die Mission mit der Einzeldrohne benötigt ohne Zwischenstopps 1:55 Stunden und mit sechs Zwischenstopps 2:35 Stunden bis die Fläche vollständig durch die Beobachtung abgedeckt ist. Die Abbildung 3b zeigt die Mission in der zuvor beschriebenen Dreiecksformation. Die Pfadlänge beträgt ohne Zwischenstopps neun Kilometer und einschließlich zwei Vorgängen zum Akkutauch elf Kilometer. Die Missionsdauer liegt bei 0:57 Stunden bzw. 1:15 Stunden. Im Umkehrschluss müssen parallel drei Akkus getauscht werden.

Ziel der beschriebenen Mission ist die parallele Erfassung des Geländes als Multispektral- und RGB-Aufnahme, zur Datengewinnung mit möglichst vielen Informationen in hoher Detailgenauigkeit und kurzer Zeit. Das Monitoring der Fläche kann folglich regelmäßig, d.h. zweiwöchig oder monatlich, parallel zur Bodenbearbeitung stattfinden, damit stetige Messungen über die gesamte Vegetationsdauer Veränderungen adäquat abbilden. Es umfasst Bodenbeschaffenheit, Feuchtigkeit, Chlorophyllgehalt, Vegetationsintensität, Pflanzenhöhe, Pflanzengesundheit und Stickstoffaufnahme.

Dies hilft frühzeitig die Anbaufläche für die Ausbreitung des Saatgutes vorzubereiten und dieses zum richtigen Zeitpunkt mit präziser Saatstärke zu verteilen. Dies bedeutet die Mindestbodentemperatur von fünf Grad und die Bodenfeuchte von maximal 104 %nFK zur Keimung im September einzuhalten. Ansonsten brechen die Körner nicht auf oder sind von Wurzelfäule betroffen. Den Pflanzen wird bis zum Sommer des Folgejahres möglichst lange Zeit zum Wachstum zu gegeben. Dies erhöht die Robustheit der Pflanzen gegen leichte Wetterschwankungen, lässt Erträge und den Nährstoffgehalt der Körner steigen.



(a) Einzeldrohne



(b) Drohnenschwarm

Abb. 3: Flugpfade und Korridore im Vergleich

Der Boden kann behutsam bearbeitet und ausgeglichen gedüngt werden. Düngemittel und Pestizide lassen sich reduzieren. Es birgt trotz robuster Weizenarten gleichzeitig höhere Risiken für Pflanzenschäden durch extreme Witterung und Schädlingsbefall. Durch die frühzeitige Bestimmung des Reifegrades und Schätzung der zu erwartenden Kornmenge, kann der richtige Erntezeitpunkt festgelegt werden. Die Ähren können schonend, über längere Zeiträume verteilt, geerntet werden [BST61].

Mit den gewonnenen Daten lässt sich der Gesamtzustand der Anbaufläche und der Feldvegetation im Sinne der Präzisionslandwirtschaft nach verschiedenen Dimensionen optimieren. Die Ertragskraft, nachhaltige Bewirtschaftung, Qualität der Pflanzen und deren Früchte, können in ihrer Entwicklung gesteuert werden. Durch die hohe Auflösung der Aufnahmen, entsteht ein präzises Modell, welches nahezu jede Pflanze einbeziehen kann. Es lässt sich prüfen, welche Bereiche besonderen Einfluss auf den Ertrag haben oder wo weniger Pflanzen wachsen. Der Boden kann geschont oder gezielt aufbereitet werden. In besonders fruchtbaren Bereichen können die Messungen Aufschlüsse über die Gründe dafür geben oder Einfluss auf den Ernte- und Aussaatprozess haben. Erkenntnisse über Schädlinge können die frühzeitige Prüfung der Herkunft und Strategien zur Abwehr dieser zur Folge haben [Sc16].

Im Kontext der Gesamtflächenbewirtschaftung können darüberhinaus Simulationsmodelle

entwickelt werden, um Prognosen zur künftigen Entwicklung des Feldes zu ermöglichen. Szenarien für Sondersituationen, wie hohe Temperaturen und lange Trockenperioden, Hageleinfluss oder die vorhandene tierische Fauna lassen sich prüfen. Der Einfluss hoher Erträge auf den Nährstoffgehalt des Bodens kann untersucht werden. Dies erweitert die Möglichkeiten der Risikofolgenabschätzung im Vorfeld und ist der langfristigen Planung zuträglich. Entwicklungstendenzen werden verdeutlicht, Saatfolgen lassen sich bestimmen und Pflanzenpaarungen zur Vermeidung von dauerhaften Monokulturen aufstellen.

5 Verbesserungsmöglichkeiten und Erweiterung

Im vorliegenden Paper wurden Möglichkeiten zur Verknüpfung von Drohnen im Schwarm betrachtet. Um effizientes und umfassendes Monitoring von großen landwirtschaftlichen Flächen zu ermöglichen, muss die Thematik in vielen Aspekten tiefergehend untersucht werden und bietet Perspektiven zur vielschichtigen Konzepterweiterung.

Haupthindernis für den Einsatz elektrisch betriebener Drohnen ist die kurze Flugzeit. Künftig kann eine Zellchemie mit höherer Energiedichte in den Akkus deutlich längere Flugzeiten ermöglichen [Ka20]. Im Bereich der Mobilität werden jährliche Steigerungsraten von fünf bis acht Prozent bei einer Kostenhalbierung in drei Jahren erwartet [JZ16]. Der Akkutausch ließe sich automatisieren.

Um den Detailgrad der Aufnahmen zu steigern, könnten die Eigenschaften und der Aufbau von Sensoren verbessert werden. Ausgehend von Sie sollten für widrige Bedingungen an Drohnen angepasst werden. Es ließen sich mit höheren Auflösungen und Optimierungen in den Algorithmen zur Bildverarbeitung, bei gleicher Flughöhe mehr Informationen erschließen. Die spezifischen Sensoren der Multispektral-Linsen könnten mit speziellen CMOS-Chips zu einer einheitlichen Linse, mit hohem Aufnahmespektrum durch Einzelpixelaufnahmen kombiniert werden [Bi16]. Dies reduziert die Anzahl der mit einer Aufgabe betrauten Drohnen und den Datenverarbeitungsaufwand.

Durch höheren Detailgrad und besserer Reichweite könnten die Drohnen höher fliegen und mit steigenden Fluggeschwindigkeiten in kürzerer Zeit einen größeren Beobachtungsbereich abdecken. Ein ähnliches Ziel könnte durch geringere Überlappungen der Aufnahmen erreicht werden. Die Optimierung der Routen könnte statt des kürzesten Pfades anhand weiterer Kriterien vorgenommen werden. Die Bildauswertung und Datenanalyse während des Fluges ermöglicht es aus Sicht der Bodenbearbeitung zu planen.

Um dies zu unterstützen ist die direkte Interaktion des Drohnenschwarms mit den Landmaschinen nötig. Die Vernetzung der Drohnen und Landmaschinen kann Informationen vom Boden und der Luft kombinieren. Saatgut könnte entsprechend der Messungen nach einer Ertragsprognose vor Ort bedarfsgerecht ausgebracht werden und Ernteverluste ließen sich verringern. In einem 3D-Modell des gesamten Feldes könnte Digitaler Zwilling zur Nachverfolgung jeder Einzelpflanze generiert werden.

6 Zusammenfassung und Ausblick

Im vorliegenden Paper wurde der Einsatz von Schwarmdrohnen für das Monitoring von großflächigen landwirtschaftlichen Flächen, zur nachhaltigeren und umweltgerechteren Bewirtschaftung diskutiert. Dieses Vorgehen ermöglicht die ganzheitliche Betrachtung und Optimierung aller Aspekte der Feldbewirtschaftung über bisherige Herangehensweisen mit Satelliten, Wetter und Flugdaten hinaus. Denn durch den erhöhten Detailgrad und die vernetzten Drohnen können verschiedene Beobachtungen parallel erfolgen.

Im ersten Abschnitt wurde auf die Funktionsweise von Drohnen im Generellen eingegangen, um die Herausforderungen bei der Zusammensetzung im Schwarm zu analysieren. Besonderes Augenmerk wurde auf die Kollisionserkennung gelegt, da sie in Wechselwirkung mit inneren und äußeren Flugaspekten der Aerodynamik steht. Es wurde beschrieben wie auf Mirco-Drohnen verteilte Sensoren kombiniert werden können. Auf dieser Basis wurden zusammenhängende Flugformationen entwickelt und in Hinblick auf geeignete Einsatzszenarien von aufeinanderfolgenden oder parallelen Beobachtungsmissionen diskutiert. Im Anwendungsteil wurde das Monitoring eines 62 Hektar großen Feldes in der Lausitz beschrieben. Es wurde ein Drohnenschwarm mit einer Kombination aus Multispektral- und RGB-Kameras zusammengestellt, um ein Modell des Feldes, inklusive Bewuchs zu erzeugen. Die Pfadplanung hat den Vergleich zwischen den Möglichkeiten einer Einzeldrohne und dem des Drohnenschwarms ermöglicht. Die Mission konnte durch die geringere Wegstrecke und parallele Untersuchungen in der Hälfte der üblichen Zeit und mit größerer Informationsausbeute abgeschlossen werden. Abschließend wurden künftige Verbesserungsmöglichkeiten durch neuartige Akkus, kombinierte Bildsensoren und die Möglichkeiten zur Vernetzung mit Landmaschinen beschrieben.

Vernetzte Drohnenschwärme können als Teil der Präzisionslandwirtschaft zur Sicherung der nachhaltigen Pflanzenproduktion und damit zur krisensicheren Versorgung mit Lebensmitteln, bei regionalen Lieferketten beitragen. Sie können als Kundschafter eingesetzt, die zentimetergenaue Lokalisierung und Begutachtung von Pflanzen ermöglichen. Dies wird genauere Ernteprognosen und neue Anbaumethoden ermöglichen. Felder werden zur digitale Produktions- und Lagermöglichkeit für natürliche Lebensmittel transformiert. Die Beobachtung großer Flächen kann den Weg zum dynamischen, ökologischen Pflanzenanbau mit der standortbezogenen Aussaat verschiedener Arten ebnen und Monokulturen vermindern. Das Spotfarming [We19] wird künftig in Wechselwirkung mit neuartigen, dezentralen und schonenden Erntemaschinen unterstützt.

Das diesem Bericht zugrundeliegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung, und Forschung unter dem Förderkennzeichen 16ES1128K im Rahmen des iCampus Cottbus gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt beim Autor.

Literaturverzeichnis

- [Bi16] Bian, L.; Suo, J.; Situ, G.; Li, Z.; Fan, J.; Chen, F.; Dai, Q.: Multispectral imaging using a single bucket detector. *Scientific Reports*, 6(1):24752, 2016.
- [BSL11] Blume, H.-P.; Stahr, K.; Leinweber, P.: *Bodenkundliches Praktikum: Eine Einführung in pedologisches Arbeiten für Ökologen, Land- und Forstwirte, Geo- und Umweltwissenschaftler*. Springer-Verlag, 2011.
- [BST61] Brouwer, W.; Schrimpf, C.; Taheri, T.: Untersuchungen über die Entwicklung der Ähre und der Bekörnung bei Winterweizen sowie Beeinflussung der Ähre durch verschiedene zeitlich gestaffelte Düngergaben. *Z. f. Acker- und Pflanzenbau*, 113:21–40, 1961.
- [CBFJ19] Cabreira, T.; Brolora, L.; Ferreira Jr., P.: Survey on Coverage Path Planning with Unmanned Aerial Vehicles. *Drones*, 3(1), 2019.
- [Ch14] Chovancová, A.; Fico, T.; Chovanec, E.; Hubinsk, P.: Mathematical modelling and parameter identification of quadrotor (a survey). *Procedia Engineering*, 96:172–181, 2014.
- [Ch15] Chen, M.; Hu, Q.; Mackin, C.; Fisac, J.; Tomlin, C.: Safe platooning of unmanned aerial vehicles via reachability. *IEEE*, S. 4695–4701, 2015.
- [Co20] Coppola, M.; McGuire, K. N.; De Wagter, C.; de Croon, G.: A Survey on Swarming With Micro Air Vehicles: Fundamental Challenges and Constraints. *Frontiers in Robotics and AI*, 7:18, 2020.
- [Fl17] Flat drone collection. *Freepik*, 2017.
- [Ge15] Gerwin, W.; Raab, T.; Bauriegel, A.; Nicolay, A.: Junge Böden der Niederlausitz. *Brand. Geowiss. Beiträge*, 22:135–146, 2015.
- [HG20] Heigermoser, M.; Glauben, T.: COVID-19, der Ölpreisverfall und die Ernährungssicherheit einkommensschwacher Staaten. *IAMO Policy Brief 37, Halle (Saale)*, 2020. urn:nbn:de:gbv:3:2-120843.
- [Ho16] Hoang, G. M.; Denis, B.; Härrä, J.; Slock, D. T. M.: Cooperative localization in GNSS-aided VANETs with accurate IR-UWB range measurements. In: *2016 13th Workshop on Positioning, Navigation and Communications (WPNC)*. S. 1–6, Oct 2016.
- [HWO21] Hiraga, M.; Wei, Y.; Ohkura, K.: Evolving collective cognition for object identification in foraging robotic swarms. *Artificial Life and Robotics*, 26:21–28, 2021.
- [JZ16] Janek, J.; Zeier, W. G.: A solid future for battery development. *Nature Energy*, 1(9):1–4, 2016.
- [Ka20] Kanand, T.; Kemper, G.; König, R.; Kemper, H.: Wildfire Detection and Disaster Monitoring System Using Uas and Sensor Fusion Technologies. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 43:1671–1675, 2020.
- [Ko14] Kopardekar, P.: Unmanned aerial system (UAS) traffic management (UTM): Enabling low-altitude airspace and UAS operations. 2014.
- [Ko19] Koubaa, A.; Allouch, A.; Alajlan, M.; Javed, Y.; Belghith, A.; Khalgui, M.: Micro air vehicle link (mavlink) in a nutshell: A survey. *IEEE Access*, 7:87658–87680, 2019.

- [Le18] Lee, W.; Lee, J. Y.; Lee, J.; Kim, K.; Yoo, S.; Park, S.; Kim, H.: Ground control system based routing for reliable and efficient multi-drone control system. *Applied Sciences*, 8(11):2027, 2018.
- [LM03] Liedtke, H.; Marschner, B.: Bodengüte der landwirtschaftlichen Nutzflächen. *Nationalatlas Bundesrepublik Deutschland–Relief, Boden und Wasser*, 2:104–105, 2003.
- [Ma18] Majd, A.; Ashraf, A.; Troubitsyna, E.; Daneshtalab, M.: Using optimization, learning, and drone reflexes to maximize safety of swarms of drones. *IEEE*, S. 1–8, 2018.
- [Me12] Meier, L.; Tanskanen, P.; Heng, L.; Lee, G. H.; Fraundorfer, F.; Pollefeys, M.: PIXHAWK: A micro aerial vehicle design for autonomous flight using onboard computer vision. *Autonomous Robots*, 33(1-2):21–39, 2012.
- [MFZ19] Mancini, A.; Frontoni, E.; Zingaretti, P.: Satellite and UAV data for Precision Agriculture Applications. *IEEE*, S. 491–497, 2019.
- [OE20] OECD: COVID-19 and the food and agriculture sector: Issues and policy responses. 2020.
- [Sa20] Santos, L. C.; Santos, F. N.; Solteiro Pires, E. J.; Valente, A.; Costa, P.; Magalhães, S.: Path Planning for ground robots in agriculture: a short review. In: 2020 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC). S. 61–66, 2020.
- [Sc16] Schirrmann, M.; Giebel, A.; Gleiniger, F.; Pflanz, M.; Lentschke, J.; Dammer, K.-H.: Monitoring agronomic parameters of winter wheat crops with low-cost UAV imagery. *Remote Sensing*, 8(9):706, 2016.
- [St17] Statistischer Bericht C IV 7 - 3j / 16 Größenstruktur, sozialökonomische Betriebstypen sowie Rechtsformen der landwirtschaftlichen Betriebe im Land Brandenburg 2016, 2017.
- [Ta21] Tang, G.; Tang, C.; Zhou, H.; Claramunt, C.; Men, S.: R-DFS: A Coverage Path Planning Approach Based on Region Optimal Decomposition. *Remote Sensing*, 13(8), 2021.
- [Vá18] Vászárhelyi, G.; Virágh, C.; Somorjai, G.; Nepusz, T.; Eiben, A. E.; Vicsek, T.: Optimized flocking of autonomous drones in confined environments. *Science Robotics*, 3(20), 2018.
- [We19] Wegener, J.; Urso, L.-M.; Hoersten, D.; Hegewald, H.; Minßen, T.-F.; Schattenberg, J.; Gaus, C.-C.; de Witte, T.; Nieberg, H.; Frerichs, L.; Backhaus, G.: Spot farming - an alternative for future plant production Spot Farming - eine Alternative für die zukünftige Pflanzenproduktion. *Journal für Kulturpflanzen*, 71, 05 2019.
- [Zh21] Zhou, B.; He, Y.; Qian, K.; Ma, X.; Li, X.: S4-SLAM: A real-time 3D LIDAR SLAM system for ground/watersurface multi-scene outdoor applications. *Autonomous Robots*, 45(1):77–98, 2021.

Impulse für die Weiterentwicklung des Blauen Engels aus einer verhaltenswissenschaftlichen Perspektive

Ergebnisse einer qualitativen Untersuchung der Erwartungen von NutzerInnen an Software und Software-Updates

Andreas Winter¹

Abstract: In den Vergabekriterien des Umweltzeichens Blauer Engel für „Ressourcen- und energieeffiziente Softwareprodukte“ liegt der Fokus auf der Nutzungsphase von Software und deren NutzerInnen als zentrale Zielgruppe. Den NutzerInnen wird eine besondere Rolle bei der Reduktion des Energieverbrauchs der Informations- und Kommunikationstechnik (IKT) und der Steigerung der Ressourceneffizienz zugesprochen. In der Forschung zur Green IT findet diese Zielgruppe bislang jedoch wenig Beachtung. Dem Ruf nach einer transdisziplinären Betrachtung folgend, ergänzt dieser Beitrag die bestehende Forschung um eine verhaltenswissenschaftliche Perspektive. Mit Hilfe von leitfadengestützten Interviews wird die Frage untersucht, welches implizite Verständnis NutzerInnen von Software und Updates haben und welche Erwartungen sie an diese stellen. Die abgeleiteten User Insights tragen zur Schließung der bestehenden Forschungslücke bei und liefern Impulse für eine Weiterentwicklung der Kriterien des Blauen Engels für „Ressourcen- und energieeffiziente Softwareprodukte“.

Keywords: Blauer Engel, Verhaltenswissenschaften, Erwartungen, User Insights, Green IT, Nachhaltige Software, Software-Updates, Update-Paradoxon

1 Einleitung und Motivation

Die privat genutzte Informations- und Kommunikationstechnik (IKT) verzeichnete im Jahr 2020 mit 23,9% erneut ein starkes Umsatzwachstum [gG21]. Ein Rekordwert eines Konsumtrends, dessen Schattenseite seit Jahren in Politik und Wissenschaft an Bedeutung gewinnt: Denn in dem Maß, wie der Konsum neuer Produkte wächst, steigt auch die Menge an Elektronikschrott, auch als e-Waste oder Waste Electrical and Electronic Equipment (WEEE) bezeichnet [Hi05, SWS21]. Als Gründe für diese Entwicklung werden produktseitig zunehmend kürzere Innovationzyklen sowie das steigende Angebot an günstigen Elektronikprodukten gesehen. Innerhalb der IKT befasst sich deshalb die Forschung zur Green IT mit Maßnahmen, durch die die IKT selbst nachhaltiger werden kann und legt den Fokus dabei insbesondere auf die Hardware. Welche Rolle die Software für einen steigenden Ressourcenverbrauch spielt, blieb zunächst wenig beachtet [Sc13].

¹ Universität Bayreuth, Rechts- und Wirtschaftswissenschaftliche Fakultät, Lehrstuhl für Betriebswirtschaftslehre III – Marketing & Konsumentenverhalten, Universitätsstraße 30, 95447 Bayreuth, andreas.winter@uni-bayreuth.de

Erst sukzessive entwickelte sich ein Diskurs zu grüner und nachhaltiger Software [Ta11, Na11] und ebenso deren nachhaltiger Entwicklung [Na15], der sich schließlich auch in der Forschung etablierte [Ke18b]. Aus einem breiten Nachhaltigkeitsverständnis heraus entstand unter anderem die Definition von Naumann et al., die grüne bzw. nachhaltige Software als Software, „(...) whose direct and indirect negative impacts on economy, society, human beings, and environment that result from development, deployment, and usage of the software are minimal and/or which has a positive effect on sustainable development“ verstehen [Na11, S. 296]. Auf einem solchen, inhaltlich umfassenden Verständnis aufbauend, wurden Messmethoden [Jo12] und Kriterien entwickelt, um mittels eines Labels oder Siegels allen Anspruchsgruppen, und insbesondere den NutzerInnen, eine Beurteilungsgrundlage zur Nachhaltigkeit von Software zu ermöglichen [Ke15]. Im Forschungsprojekt UFOPLAN-SSD 2015 wurde eine konkrete Methodik zur Bewertung der Ressourceneffizienz von Softwareprodukten entwickelt [Gr18], die im Folgeprojekt Refoplan 2018 zu einer Vergabegrundlage [KN20] und schließlich zu den ersten Vergabekriterien des Blauen Engels für „*Ressourcen- und energieeffiziente Softwareprodukte (DE-UZ 215)*“² ausgebaut wurde. Neben der Reduktion des Energieverbrauchs der IKT und der Steigerung der Ressourceneffizienz, will der Blaue Engel Orientierung geben sowie „*Aufmerksamkeit für die Rolle der Software im Bereich IKT*“ schaffen. Bei seinen Kriterien konzentriert er sich auf die Nutzungsphase von Software und deren NutzerInnen [RA20]. Mit diesen wird (ergänzend zur Rolle der BeschafferInnen) eine Zielgruppe in den Fokus gerückt, der in der Forschung zur Green IT zwar Bedeutung zugesprochen wird [LFF13], deren Rolle bislang aber kaum untersucht wurde [Ke11, Ke15]. Eine Ausnahme stellt eine Onlineumfrage von 2016 dar, in der NutzerInnen³ zu den Kriterien für ein Umweltzeichen sowie möglichen Einflussfaktoren auf dessen Akzeptanz befragt wurden [Ke18b, NGK21]. Dabei fällt auf, dass die Befragten – zusätzlich zu der Beantwortung der zentralen geschlossenen Fragen – ihre Erwartungen in den optionalen Freitextfeldern zum Ausdruck brachten. Dort wurden unter anderem ergänzende Wünsche zu Kriterien an Software-Updates⁴ angebracht, wie z. B. zum Zeitraum der Bereitstellung oder der Anpassung von Software an Hardwareveränderungen [KGN18]. Dieses, für eine Onlineumfrage eher ungewöhnliche Engagement [Zh17], spiegelt die Bedeutung der in den Freitextfeldern angesprochenen Erwartungen sowie der Funktion von Updates für die nachhaltige Nutzung von Software aus Sicht der NutzerInnen wider. Neben der technisch-normativen Perspektive erscheint deshalb, auch im Sinne einer gewünschten transdisziplinären Betrachtung [Ke16], eine Untersuchung der NutzerInnen aus einer verhaltenswissenschaftlichen Perspektive notwendig: Dieser Beitrag soll dafür der grundlegenden Frage nachgehen, welches implizite Verständnis NutzerInnen von Software und Updates haben und zudem klären, welche Erwartungen sie an diese stellen.

² Im Folgenden als „Blauer Engel“ bezeichnet.

³ Im Folgenden synonym mit „(Software-)KonsumentInnen“.

⁴ Im Folgenden als „Updates“ bezeichnet.

Die abgeleiteten User Insights⁵ sollen zur Schließung der aufgezeigten Forschungslücke beitragen und Impulse für eine Weiterentwicklung der Kriterien des Blauen Engels liefern.

2 Methode und Forschungsdesign

Zur Gewinnung dieser User Insights wurden zwischen dem 16. September und 8. Oktober 2020 mit zufälligen NutzerInnen von Software leitfadengestützte Telefoninterviews geführt, die eine Untersuchung des Forschungsgegenstands über den eher deskriptiven Charakter einer Umfrage hinaus erlauben [RR12]. Die gestellten Fragen wurden offen formuliert, um eine Beeinflussung der Antworten zu vermeiden. Es wurde darauf geachtet, dass die Zusammenstellung der TeilnehmerInnen trotz Network- bzw. Snowball-Sampling [Du19] möglichst heterogen ist, da auftretende Effekte in solchen Fällen eine höhere Generalisierbarkeit besitzen [Ro14]. Insofern waren die TeilnehmerInnen hinsichtlich ihres Alters, Geschlechts und Bildungsstands gleichmäßig verteilt. Bezugnehmend auf das Ziel der Untersuchung, der Gewinnung von User Insights, wurde darauf geachtet die Befragten nur in Bezug auf ihrer individuellen Rolle als NutzerInnen von Software zu befragen. Die berufliche Tätigkeit war kein Kriterium für die Auswahl der TeilnehmerInnen. Es wurden so viele Interviews geführt, bis eine theoretische Sättigung erreicht war – zusätzliche Interviews also keine neuen Erkenntnisse lieferten. Dies führte zur Befragung von 19 Personen, davon zehn Frauen und neun Männer, die zwischen 21 und 62 Jahre alt waren. Die im Durchschnitt 39-minütigen Interviews wurden mit Hilfe der Software MAXQDA 2020 transkribiert und in Anlehnung an die strukturelle qualitative Inhaltsanalyse von Mayring in mehreren Runden codiert und ausgewertet [Ma15].

3 Ergebnisse und Diskussion

Im Folgenden werden nur die **User Insights** aus den Interviews dargestellt, die das Verständnis der Kriterien des Blauen Engels ergänzen oder diesen widersprechen. Die jeweils nach den horizontalen Linien abgeleiteten **Impulse** und die dort in Klammern gesetzten Kapitelverweise orientieren sich an der Ausgabe des Blauen Engels vom Januar 2020, Version 1 [RA20], die nicht mehr gesondert zitiert wird.

3.1 Erweiterung des Softwareverständnisses und des Kontexts

In den Interviews zeigt sich zunächst ein **undifferenziertes Begriffsverständnis von Software**: Darunter werden zwar überwiegend Anwendungssoftware (einschließlich mobiler Apps), Betriebssysteme oder Spiele, zum Teil aber auch die Schnittstelle

⁵ Hier vereinfacht: Im Sinne von fundierten und anwendbaren Erkenntnissen über die Bedürfnisse der NutzerInnen sowie ihre zugrundeliegende Motivation [WL16, S. 246].

zwischen Hardware und NutzerInnen oder eine Ware im ökonomischen Sinn verstanden. Die eindeutige Beschreibung fällt den Befragten schwer, wobei sie auch eine über die Zeit steigende Definitionskomplexität wahrnehmen. Dazu trägt bei, dass Software zwar weiterhin primär auf dem Personal Computer (PC) sowie auf mobilen Endgeräten verortet wird, die Übergänge für die Befragten aber als immer fließender wahrgenommen werden. Einigen Befragten ist eine Unterscheidung von Software sogar grundsätzlich nicht möglich und vor allem jüngere Befragte nennen nur Betriebssysteme als Beispiele für Software – Anwendungssoftware und mobile Apps werden von ihnen offensichtlich nicht als Software verstanden. Eine Erkenntnis, die eine der Befragten auf eine sinkende Affinität und eine geringere Motivation zur Auseinandersetzung mit Software in der Gesellschaft bei gleichzeitig steigender Nutzung zurückführt.

Es wird weiter deutlich, dass die meisten Befragten zugleich die Rolle der **NutzerInnen und der BeschafferInnen** innehaben, oder zumindest bereits Erfahrung mit der Beschaffung von Software besitzen. Der Kaufsoftware wird aber, zugunsten von Abo-Modellen sowie freier und Open-Source-Software, von über der Hälfte der Befragten eine geringe oder sogar abnehmende Bedeutung beigemessen.

Für das Verständnis von Software ist bei den meisten Befragten zudem der **berufliche Kontext** prägend. Software wird im Umfeld des Arbeitsplatzes aus verschiedenen Gründen bewusster als im Privaten wahrgenommen: Weil man es beispielsweise mit (mutmaßlich) veralteter oder impraktikabler Software zu tun hat, sie Bedeutung für fundamentale (Arbeits-) Prozesse besitzt oder der Beruf für die Befragten den Einstieg in die intensive und oftmals erstmalig bewusste Nutzung von Software darstellt. Gleichzeitig fühlen sich die Befragten im beruflichen Kontext häufiger von Dritten, meist IT-AdministratorInnen und/oder -SpezialistInnen, abhängig.

Die Fokussierung des Blauen Engels auf PC-Anwendungssoftware erscheint vor dem Hintergrund der Erkenntnisse aus den Interviews nachvollziehbar. Dennoch könnte dieses **Verständnis** bei zukünftigen Weiterentwicklungen vor allem auf mobile Anwendungen (wie bereits in Kapitel 1.2 beschrieben) und Betriebssysteme **erweitert** werden, da diese ebenfalls der direkten Nutzung und Beeinflussung der NutzerInnen unterstehen. Durch die Abdeckung eines breiteren Software-Verständnisses könnte (vor allem bei jüngeren NutzerInnen) eine höhere Akzeptanz des Umweltzeichens und zugleich das Bewusstsein für die Bedeutung bislang unbeachteter oder nicht als solche verstandener Software erreicht werden.

Die bei vielen der Befragten kombinierte Rolle von NutzerInnen und BeschafferInnen bestätigt das vorherrschende Verständnis [Ke18a] und die bisherige Ausrichtung des Blauen Engels. Aufgrund des Bedeutungsverlusts von Kaufsoftware vor allem im privaten Kontext und im Sinne einer möglichst breiten Abdeckung verschiedener Software, erscheint es als sinnvoll, bei einer Weiterentwicklung des Blauen Engels gezielter als bisher die **AnbieterInnen von freier und Open-Source-Software** anzusprechen. Ihnen könnten ergänzende Anreize zur Teilnahme, aber vor allem überhaupt die Möglichkeit

einer Zertifizierung entsprechend ihrer verfügbaren Mittel und Möglichkeiten, gegeben werden.

Um die prägende Wirkung des beruflichen Kontexts zu nutzen, wird dort die Einbindung ergänzender **Kommunikationsmaßnahmen**, z. B. in Form eines digitalen Umweltzeichens, empfohlen: Durch diese könnte die Wahrnehmung des Blauen Engels und seiner zugrundeliegenden Ziele erhöht, und weitere „*Aufmerksamkeit für die Rolle der Software im Bereich IKT*“ (Kapitel 1.2) erreicht werden. Die Abhängigkeit der NutzerInnen von **IT-AdministratorInnen und -SpezialistInnen** zeigt jedoch, dass der berufliche Kontext nicht nur eine prägende Wirkung, sondern auch andere Strukturen und Abhängigkeiten besitzt. Dem Verständnis folgend, dass für eine erfolgreiche Green IT auch Unternehmen betrachtet werden sollten [Ke11], adressiert der Blaue Engel den privaten als auch den gewerblichen Kontext bereits gleichermaßen. Die Erkenntnisse aus den Interviews legen aber eine differenzierte Betrachtung und vor allem eine **gesonderte Ansprache der gewerblichen Zielgruppen** (eben auch der IT-AdministratorInnen und -SpezialistInnen) nahe. Für diese sind bei der Nutzung und Beschaffung von Software Kriterien wichtig, die bislang noch keine Beachtung im Blauen Engel finden, und sich beispielsweise an **technischen, organisatorischen oder ökonomischen Aspekten** orientieren [GW17, Lo10]. Exemplarisch können hier die technische Kompatibilität mit bestehenden Softwaresystemen, die Usability für die EndnutzerInnen oder die Transparenz von Rahmenverträgen als Kriterien betrachtet werden – für eine weitere Konkretisierung bedarf es einer separaten Untersuchung der gewerblichen Zielgruppe. Die (zumindest teilweise) Integration derartiger Aspekte könnte die Attraktivität des Blauen Engels für selbige erhöhen und damit indirekt zur Einhaltung seiner zentralen Kriterien zur ökologischen Nachhaltigkeit beitragen.

3.2 Berücksichtigung der Bedeutung des Update-Paradoxons

Während Software für die befragten NutzerInnen bereits durch ihre (meist eindeutige) Funktion Relevanz besitzt, lässt sich bei Updates eine starke **Diskrepanz zwischen der wahrgenommenen funktionalen und der persönlichen Relevanz** erkennen: Ihnen wird von der Mehrheit der Befragten eine hohe und zukünftig noch steigende Relevanz für die (technische) Funktion einer Software zugesprochen. Eine persönliche Bedeutung (für die NutzerInnen selbst) besitzen die Updates hingegen kaum. Diese wird nur dann als hoch erachtet, wenn sie sich aus der funktionalen Relevanz ergibt: beispielsweise beim Schließen von Sicherheitslücken einer für die NutzerInnen persönlich bedeutsamen Software. Diese ambivalente Wahrnehmung stellt ein technologisches Paradoxon [MF98] dar: Updates werden zwar als funktional nützlich und notwendig eingestuft, andererseits besitzen sie aber keine persönliche Relevanz für die NutzerInnen.

Dieses Phänomen lässt sich durch eine ausgeprägte **negative Einstellung gegenüber Updates** erklären: Updates werden durchweg als störend, nervig sowie als notwendiges Übel betrachtet; und das nicht nur in den Fällen, in denen ihr Nutzen nicht erkannt wird. Bei einigen Befragten bestehen sogar grundsätzliche Gefühle des Zwangs und der

Fremdbestimmung ihres Handelns durch Updates. Die Bedeutung dieser Erkenntnis zeigt sich besonders darin, dass Updates lediglich in einem Interview per se als hilfreich und wichtig beschrieben werden. Als Ursache hierfür können unerfüllte **Erwartungen an den Prozess** des Updatens gesehen werden: Neben dem problemlosen Ablauf (sowohl des Update-Prozesses an sich wie auch seiner Auswirkungen auf Nutzungsabläufe) erwarten einige der Befragten eine stärkere Einbindung in den Updatevorgang; insbesondere in Form der Bereitstellung ergänzender Informationen über den Ablauf sowie der selbstbestimmten Ausführung sichtbarer Updates. Die Befragten wünschen sich, dass Updates „so selten wie möglich, so oft wie nötig“ durchgeführt werden – eine Erwartung, die sich an der persönlichen Relevanz orientiert und die kaum eindeutig quantifizierbar ist. Dahinter steht der Anspruch vieler Befragter, dass Software bereits bei der Veröffentlichung fehlerfrei und „zukunftsfähig“ oder zumindest so programmiert sein soll, dass sie ohne Updates dauerhaft lauffähig ist. In einzelnen Fällen führt eine als zu hoch wahrgenommene Updatefrequenz sogar zum Wechsel der entsprechenden Software.

Derartige Auswirkungen auf das tatsächliche Verhalten der NutzerInnen verdeutlichen die Notwendigkeit einer stärkeren Berücksichtigung dieser Erwartungen in den Kriterien des Blauen Engels (Kapitel 3.1.3.3). Dadurch kann die Akzeptanz des Umweltzeichens und die Kontinuität der Softwarenutzung erhöht werden. **Sicherheitsupdates**, deren Nutzen oft nicht direkt von den NutzerInnen wahrgenommen [F116] und die vielfach ignoriert oder verschoben werden [RAG20, Ma18], könnten möglichst **automatisiert und unsichtbar** erfolgen, um Störungen in den Nutzungsabläufen zu vermeiden. Sofern eine Automatisierung (z. B. aus technischen oder rechtlichen Gründen) nicht möglich ist, könnten eine **höhere Transparenz und mehr Informationen zum Updatevorgang** den wahrgenommenen Nutzen eines Updates erhöhen. Dies wiederum kann zu einer höheren Zufriedenheit und damit einer längeren Nutzung der Software führen [F116]. Ergänzt werden könnte eine solche Maßnahme durch die **Mitbestimmung des Updatezeitpunkts**. Die im Blauen Engel aufgeführte Wahlmöglichkeit zwischen Sicherheitsupdates und sonstigen Updates dürfte auf Basis der Erkenntnisse hingegen eher negative Auswirkungen haben: Sie räumt den NutzerInnen zwar mehr Mitbestimmung ein, führt aber im Gegenzug zu einer intensiveren Beschäftigung mit ungeliebten Sicherheitsupdates. Um dies zu verhindern, könnte die Wahlmöglichkeit auf den Installationszeitpunkt der Software beschränkt werden. Grundsätzlich sollte stets versucht werden die **Frequenz von sichtbaren Updates so gering wie möglich zu halten**. Zwar zeigen Untersuchungen, dass eine Erhöhung der Frequenz bei Funktionsupdates unter Umständen zu positiven Effekten führen kann [F116] – die grundsätzliche Abneigung gegenüber Updates in den Interviews sowie die fehlende Reflexion des Updatevorgangs durch die NutzerInnen sprechen dem aber entgegen.

3.3 Erweiterung des Updateverständnisses

Die fehlende Reflexion spiegelt sich auch bei **Updates** in Form eines **undifferenzierten Begriffsverständnisses** wider: Sie werden meist allgemein als „neue Version“ oder

„Aktualisierung“ beschrieben. Wenn eine Differenzierung stattfindet, dann nach der Art der Aufgabe (z. B. dem Erhalt der Sicherheit oder der Funktionserweiterung) oder nach der aktiven oder passiven Ausführung (z. B. als automatisches Update im Hintergrund). Das mangelnde Begriffsverständnis führt dazu, dass Updates vom Großteil der Befragten vor allem indirekt durch ihre individuellen **Nutzenerwartungen** umschrieben werden. Für diese lassen sich zwei übergreifende Bereiche identifizieren:

Als grundlegendes Bedürfnis steht für die NutzerInnen das **Erhalten und Bewahren** im Vordergrund. Darunter fällt, dass Fehler behoben werden und die Sicherheit der Software (und damit auch oftmals des gesamten Systems) sowie die Kompatibilität (insbesondere mit weiterer Software) garantiert werden. Die Basisfunktionen, die dem zentralen Nutzen der Software entsprechen, das grundlegende Bedienkonzept sowie die persönlichen Daten und Einstellungen sollten bei einem Update bestehen bleiben. Ergänzend zu diesem grundlegenden Bedürfnis äußerten nahezu alle Befragten auch die Erwartung des **Verbesserns** und des **Erweiterns** der Software durch Updates. Oft werden darunter Performancesteigerungen, eine vereinfachte Handhabung und Funktionserweiterungen subsumiert. Es zeigt sich jedoch, dass der bei vielen NutzerInnen existente Wunsch nach neuen Funktionen oft kein Wunsch nach neuen Funktionen an sich ist – sondern nach einer Anpassung an die Anforderungen einer sich verändernden Umwelt. Dies könnte auch eine Erklärung dafür sein, warum es in einigen Fällen zu einer Verwechslung von Updates mit funktionserweiternden, meist kostenpflichtigen Upgrades kommt und nur in einem Interview zu einer aktiven Abgrenzung von diesen.

Für den Blauen Engel ergibt sich aus dem heterogenen Begriffsverständnis sowie den individuellen Nutzenerwartungen, dass eine Erweiterung der Update-Definition über das reine Verständnis von Sicherheitsupdates hinaus sinnvoll erscheint. Um die im Sinne der ökologischen Nachhaltigkeit langfristige Nutzung einer Software, und damit oft verbunden auch der zugrundeliegenden Hardware, sicherzustellen, ist ergänzend zur Abwärtskompatibilität (Kapitel 3.1.2.1) das **Bewahren der bestehenden Funktionalität, Sicherheit und Kompatibilität** notwendig [F115]. Zudem könnten weitere **Ursachen der (z. B. funktionalen oder psychologischen) Obsoleszenz** [Ma16] von Software untersucht und entsprechende Gegenmaßnahmen in den Kriterien (z. B. in Kapitel 3.1.2 und Kapitel 3.2) berücksichtigt werden.⁶ Es empfiehlt sich eine **Anpassung der Funktionalität der Software relativ zum Status Quo**, die im Sinne des erweiternden Verbesserns bereits in der Update-Definition (Kapitel 1.4) verankert werden könnte.⁷ Zudem könnten **Funktionsupdates oder kostenlose Upgrades** in den Kriterien des Blauen Engels (insbesondere in Kapitel 3.1.3.3) ergänzt werden; auch, um dem übergreifenden Begriff der Nutzungsautonomie (Kapitel 3.1.3) besser gerecht zu werden.

⁶ Hinweise auf deren Einfluss finden sich bereits im Abschlussbericht des Forschungsprojekts Refoplan 2018 [Gr18, S. 18, S. 30 f, S. 34].

⁷ Aktuell nur: „ggf. verbesserte, Version“ [RA20, S. 9].

3.4 Bedeutung des Bereitstellungszeitraums für die nachhaltige Nutzung von Software

Eng mit Updates und deren funktionaler Relevanz ist die Kontinuität der Software verbunden, die sich aus technischer Sicht unter anderem im **Bereitstellungszeitraum von Updates**⁸ widerspiegelt. Dessen Wahrnehmung ist bei fast allen Befragten durch ihre eigenen, und dabei meist negativen, Erfahrungen geprägt und besitzt folglich eine hohe Relevanz – gleichzeitig ist er im Kaufprozess aktuell aber kaum entscheidend. Das liegt maßgeblich daran, dass es für die NutzerInnen bislang nicht möglich ist, den Bereitstellungszeitraum bereits beim Kauf einzuschätzen. Entsprechend hoch ist auch das Interesse an zusätzlichen Informationen und höherer Transparenz. Als **akzeptabel** erscheint den meisten Befragten im Durchschnitt ein Zeitraum von sechs Jahren bei PC-Software und dreieinhalb Jahren bei mobilen Apps. Ein **wünschenswerter** Bereitstellungszeitraum kann hingegen nicht eindeutig quantifiziert werden. Er liegt aber in allen Fällen über dem akzeptablen Wert und wird von den Befragten oft durch dessen Verdopplung abgeleitet. Es zeigt sich außerdem, dass die Erwartungen an den Bereitstellungszeitraum durch eine **Vielzahl von interdependenten Faktoren** beeinflusst werden, auf die wiederum die zuvor genannten Erfahrungen einwirken. Am häufigsten werden die erwartete Nutzungsdauer bzw. Gesamtnutzenüberlegungen (= bis eine Software ihren Nutzen erbracht hat; z. B. für die Erstellung einer Steuererklärung), die Kosten für die Anschaffung einer Software (= je teurer, desto länger) sowie die Gebundenheit an Hardware oder andere Software genannt. Der Wunsch nach **ökologischer Nachhaltigkeit** spielt eine untergeordnete Rolle und wird vor allem in Bezug auf die Hardwarelebensdauer geäußert, die wiederum die Dauer der Softwarenutzung beeinflusst. Hingegen zeigen nahezu alle Befragten sogar Verständnis für die (ökonomischen) Aufwendungen der Software-Hersteller und folglich einer Begrenzung des Bereitstellungszeitraums.

Durch die hohe Relevanz des Bereitstellungszeitraums für die NutzerInnen ergeben sich weitere Handlungsoptionen zur Sicherstellung der Kontinuität der Software (Kapitel 3.1.3.3). Aufgrund der Ausrichtung des Blauen Engels auf die NutzerInnen und BeschafferInnen erscheint eine Änderung des Bezugspunktes **vom Bereitstellungsende zum Datum des Erwerbs** einer Software als sinnvoll: Dieser Bezugspunkt ist für die NutzerInnen nachvollziehbarer, weil er beispielsweise durch Garantien und Gewährleistungen bereits in deren Verständnis und Verhalten etabliert ist [MD02]. Dem Wunsch nach mehr Transparenz könnte mit einer expliziten **Benennung und Deklaration des Bereitstellungszeitraums** durch die Software-Hersteller begegnet werden. Zwar entspricht der bislang im Blauen Engel mit fünf Jahren festgelegte Zeitraum in etwa dem Mittel der von den Befragten als akzeptabel genannten Zeiträume. Die Erkenntnisse zu einem wünschenswerten Zeitraum zeigen aber, dass die Festlegung einer **deutlich längeren Nutzungsdauer** als sinnvoll erscheint. Schließlich unterstreicht die in

⁸ Im Rahmen der Untersuchung wurde dieser als Zeitraum definiert, wie lange nach dem Erwerb einer Software entsprechende Updates vom Software-Hersteller bereitgestellt werden [Da19].

diesem Kontext fehlende Bedeutung der ökologischen Nachhaltigkeit für die NutzerInnen und die zugleich starke Abhängigkeit von weiteren, von ihr unabhängigen und sogar teilweise konträren Faktoren (wie z. B. dem vorhandenen Verständnis für das ökonomisch orientierte Handeln der Software-Hersteller) die Notwendigkeit einer stärkeren **Kommunikation der Rolle der Software für die Ressourcen- und Energieeffizienz** – beispielsweise im Zuge der offensichtlichen Benennung und Deklaration des Bereitstellungszeitraums.

4 Fazit und Ausblick

Dieser Beitrag gibt Antworten darauf, welches implizite Verständnis NutzerInnen von Software und Updates haben und klärt, welche Erwartungen sie an diese stellen. Die aus einer verhaltenswissenschaftlichen Perspektive abgeleiteten User Insights tragen somit zur Schließung einer zentralen Forschungslücke in der Green IT bei. Hierbei ist insbesondere der Einfluss indirekter Faktoren (wie des beruflichen Kontexts oder ergänzender Formen der Obsoleszenz) sowie die Bedeutung von Updates und deren Wahrnehmung (Stichwort: Update-Paradoxon) für die Erreichung der Ziele einer Green IT hervorzuheben. Darauf aufbauend ergeben sich Anknüpfungspunkte für die **inter- und transdisziplinäre Forschung**: Um eine Gewichtung der in den Interviews genannten Einflussfaktoren vornehmen zu können, sollte eine explorative Untersuchung zum weiteren Verständnis der Obsoleszenz bzw. des Wechselverhaltens von Soft- und Hardware beitragen. Ebenso sollte in diesem Rahmen der Einfluss der Usability von Software auf die Obsoleszenz bzw. das Wechselverhalten untersucht werden. Diese Erkenntnisse sollten gemeinsam mit denen der vorliegenden Untersuchung in einer repräsentativen, quantitativen Studie überprüft werden, um entsprechende Aussagen über die Grundgesamtheit treffen zu können. Ebenso sollte untersucht werden, welchen Einfluss eine so definierte grüne bzw. nachhaltige Software im Verständnis von Green by IT⁹ auf gesamtgesellschaftliche Verhaltensentwicklungen, wie beispielsweise die Wegwerfmentalität, besitzt und welche Wechselwirkungen es in diesem Zusammenhang zwischen First-, Second- und Third-Order-Effekten gibt [Fo02]. Der besonderen Rolle der gewerblichen Zielgruppe und hier insbesondere der IT-AdministratorInnen und -SpezialistInnen, sollte in einer der vorliegenden Untersuchung komplementären Forschungsarbeit Beachtung geschenkt werden.

Die aus den User Insights abgeleiteten Impulse können zwar nicht direkt als Vergabekriterien (z. B. nach ISO 14024:2018) für den Blauen Engel übernommen werden und stehen teilweise sogar mit den bestehenden Kriterien im Widerspruch. Sie geben jedoch konkrete und empirisch gestützte Hinweise darauf, an welchen Stellen und in welcher Form die Rolle der NutzerInnen und deren Erwartungen bei einer **Weiterentwicklung des Blauen Engels** berücksichtigt werden können. Beispielsweise sollte eine Ausweitung der zugrunde gelegten Nachhaltigkeits- und Ressourcenverständnisse erfolgen: Zum einen, um Interdependenzen mit weiteren

⁹ Im Sinne der Unterstützung einer nachhaltigen Entwicklung durch Software bzw. IKT [Na15].

Nachhaltigkeitsdimensionen (wie bei KGN19 dargestellt) zu beachten. Zum anderen, um Software im Sinne einer übergreifenden Lebenszyklus-Betrachtung auch im Blauen Engel als wichtige Ursache für die Erzeugung von eWaste zu berücksichtigen [Ke18a, Hi06, Ta11]. In einem nächsten Schritt sollten diese Impulse im Rahmen einer inter- und transdisziplinären Diskussion, gestützt durch eine vertiefende Analyse der vorliegenden Interviewdaten, hinsichtlich ihrer Bedeutung und Umsetzbarkeit (z. B. als Eigenerklärungen der Software-Hersteller) priorisiert und durch weitere gezielte Untersuchungen ergänzt werden. Auf diese Weise können schließlich aktualisierte Vergabekriterien formuliert werden, die die Rolle der NutzerInnen in der Green IT auch im Blauen Engel repräsentieren.

Danksagung

Die Interviews wurden im Rahmen eines einjährigen Projektseminars am Lehrstuhl für BWL III: Marketing & Konsumentenverhalten von Prof. Dr. Claas Christian Germelmann an der Universität Bayreuth erhoben – der Dank gilt insbesondere den TeilnehmerInnen des Projektseminars für ihre Unterstützung bei der Datenerhebung und Auswertung.

5 Literaturverzeichnis

- [Da19] Das Europäische Parlament und der Rat der Europäischen Union: Richtlinie (EU) 2019/770 des Europäischen Parlaments und des Rates vom 20. Mai 2019 über bestimmte vertragsrechtliche Aspekte der Bereitstellung digitaler Inhalte und digitaler Dienstleistungen, 2019.
- [Du19] Durdella, N.: *Qualitative Dissertation Methodology. A Guide for Research Design and Methods*, 2019.
- [F115] Fleischmann, M. et al.: *Gains and Losses in Functionality – An Experimental Investigation of the Effect of Software Updates on Users’ Continuance Intentions: Thirty Sixth International Conference on Information Systems*, Fort Worth, S. 1–21, 2015.
- [F116] Fleischmann, M. et al.: *The role of software updates in information systems continuance — An experimental study from a user perspective*. *Decision Support Systems* 83, S. 83–96, 2016.
- [Fo02] *Forum for the Future: The impact of ICT on sustainable development: EITO Yearbook 2002*, S. 250–283, 2002.
- [gG21] gfu Consumer & Home Electronics GmbH; GfK SE: *Home Electronics Markt Index Deutschland (HEMIX)*. Januar 2020 – Dezember 2020.

- [Gr18] Gröger, J. et al.: Entwicklung und Anwendung von Bewertungsgrundlagen für ressourceneffiziente Software unter Berücksichtigung bestehender Methodik - Abschlussbericht, 2018.
- [GW17] Groher, I.; Weinreich, R.: An Interview Study on Sustainability Concerns in Software Development Projects: 2017 43rd Euromicro Conference on Software Engineering and Advanced Applications (SEAA). IEEE, S. 350–358, 2017.
- [Hi05] Hilty, L. M.: Electronic waste—an emerging risk? *Environmental Impact Assessment Review*, 5/25, S. 431–435, 2005.
- [Hi06] Hilty, L. M. et al.: Rebound effects of progress in information technology. *Poiesis & Praxis*, 1/4, S. 19–38, 2006.
- [Jo12] Johann, T. et al.: How to measure energy-efficiency of software: Metrics and measurement results: 2012 First International Workshop on Green and Sustainable Software (GREENS). IEEE, S. 51–54, 2012.
- [Ke11] Kern, E. et al.: Green Software and Green IT: An End Users Perspective. In (Golinska, P.; Fertsch, M.; Marx-Gómez, J., Hrsg.): *Information Technologies in Environmental Engineering*. Springer, Berlin, Heidelberg, S. 199–211, 2011.
- [Ke15] Kern, E. et al.: Labelling sustainable software products and websites: Ideas, Approaches, and Challenges: Proceedings of EnviroInfo and ICT for Sustainability 2015. Atlantis Press, Paris, France, S. 82–91, 2015.
- [Ke16] Kern, E.: The development of an eco-label for software products - a transdisciplinary process? In (Mayr, H. C.; Pinzger, M., Hrsg.): *Informatik 2016*. Gesellschaft für Informatik e.V, Bonn, S. 1285–1296, 2016.
- [Ke18a] Kern, E. et al.: Sustainable software products—Towards assessment criteria for resource and energy efficiency. *Future Generation Computer Systems* 86, S. 199–210, 2018.
- [Ke18b] Kern, E.: Green Computing, Green Software, and Its Characteristics: Awareness, Rating, Challenges. In (Otjacques, B. et al., Hrsg.): *From Science to Society*. Springer International Publishing, Cham, S. 263–273, 2018.
- [KGN18] Kern, E.; Guldner, A.; Naumann, S.: Bewertung der Nachhaltigkeit von Software: Entwicklung einer Umweltkennzeichnung. In (Arndt, H.-K. et al., Hrsg.): *Nachhaltige Betriebliche Umweltinformationssysteme*. Springer Fachmedien, Wiesbaden, S. 9–19, 2018.
- [KGN19] Kern, E.; Guldner, A.; Naumann, S.: Including Software Aspects in Green IT: How to Create Awareness for Green Software Issues. In (Kharchenko, V.; Kondratenko, Y.; Kacprzyk, J., Hrsg.): *Green IT Engineering: Social,*

- Business and Industrial Applications. Springer International Publishing, Cham, S. 3–20, 2019.
- [KN20] Kern, E.; Naumann, S.: REFOPLAN - Ressourceneffiziente Software. „40 Jahre Blauer Engel - Weiterentwicklung seines Produktportfolios“, <https://www.umwelt-campus.de/forschung/projekte/green-software-engineering/projekte/refoplan-2018>, Stand: 15.05.2021.
- [LFF13] Lami, G.; Fabbrini, F.; Fusani, M.: A methodology to derive sustainability indicators for software development projects. In (Münch, J.; Lan, J. A.; Zhang, H., Hrsg.): Proceedings of the 2013 International Conference on Software and System Process - ICSSP 2013. ACM Press, New York, New York, USA, S. 70–77, 2013.
- [Lo10] Loebbecke, C. et al.: Drivers of B2B Software Purchase Decisions, S. 535–546, 2010.
- [Ma15] Mayring, P.: Qualitative Inhaltsanalyse. Grundlagen und Techniken. Beltz, Weinheim, 2015.
- [Ma16] Manhart, A. et al.: Resource Efficiency in the ICT Sector. Final Report, November 2016.
- [Ma18] Mathur, A. et al.: Quantifying Users’ Beliefs about Software Updates. In (Acar, Y.; Patil, S., Hrsg.): Proceedings 2018 Workshop on Usable Security. Internet Society, Reston, VA.
- [MD02] Murthy, D.; Djameludin, I.: New product warranty: A literature review. *International Journal of Production Economics*, 3/79, S. 231–260, 2002.
- [MF98] Mick, D. G.; Fournier, S.: Paradoxes of Technology: Consumer Cognizance, Emotions, and Coping Strategies. *Journal of Consumer Research*, 2/25, S. 123–143, 1998.
- [Na11] Naumann, S. et al.: The GREENSOFT Model: A reference model for green and sustainable software and its engineering. *Sustainable Computing: Informatics and Systems*, 4/1, S. 294–304, 2011.
- [Na15] Naumann, S. et al.: Sustainable Software Engineering: Process and Quality Models, Life Cycle, and Social Aspects. In (Hilty, L. M.; Aebischer, B., Hrsg.): *ICT Innovations for Sustainability*. Springer International Publishing, Cham, S. 191–205, 2015.
- [NGK21] Naumann, S.; Guldner, A.; Kern, E.: The Eco-label Blue Angel for Software—Development and Components. In (Kamilaris, A. et al., Hrsg.): *Advances and New Trends in Environmental Informatics*. Springer International Publishing, Cham, S. 79–89, 2021.

- [RA20] RAL gmbH: Ressourcen- und energieeffiziente Softwareprodukte. DE-UZ 215, <https://produktinfo.blauer-engel.de/uploads/criteriafile/de/DE-UZ%20215-202001-de%20Kriterien-2020-01-16.pdf>, Stand: 13.05.2021.
- [RAG20] Rajivan, P.; Aharonov-Majar, E.; Gonzalez, C.: Update now or later? Effects of experience, cost, and risk preference on update decisions. *Journal of Cybersecurity*, 1/6, S. 1–12, 2020.
- [Ro14] Robinson, O. C.: Sampling in Interview-Based Qualitative Research: A Theoretical and Practical Guide. *Qualitative Research in Psychology*, 1/11, S. 25–41, 2014.
- [RR12] Rubin, H. J.; Rubin, I. S.: *Qualitative interviewing. The art of hearing data*. SAGE, Thousand Oaks, Calif., 2012.
- [Sc13] Schluep, M. et al.: Insights from a decade of development cooperation in e-waste management. ETH E-Collection.
- [SWS21] Shittu, O. S.; Williams, I. D.; Shaw, P. J.: Global E-waste management: Can WEEE make a difference? A review of e-waste trends, legislation, contemporary issues and future challenges. *Waste management (New York, N.Y.)* 120, S. 549–563, 2021.
- [Ta11] Taina, J.: Good, bad, and beautiful software-In search of green software quality factors. *Cepis Upgrade*, 4/12, S. 22–27, 2011.
- [WL16] Wieneke, A.; Lehrer, C.: Generating and exploiting customer insights from social media data. *Electronic Markets*, 3/26, S. 245–268, 2016.
- [Zh17] Zhou, R. et al.: Who tends to answer open-ended questions in an e-service survey? The contribution of closed-ended answers. *Behaviour & Information Technology*, 12/36, S. 1274–1284, 2017.

Workshop: Computer Science for Biodiversity (CS4BIODiversity)

Carpe Diem: A Lifelong Learning Tool for Automated Wildlife Surveillance

Implementing Active and Incremental Learning for Object Detection

Clemens-Alexander Brust,¹ Björn Barz,¹ Joachim Denzler^{1,2}

Abstract: We introduce Carpe Diem, an interactive tool for object detection tasks such as automated wildlife surveillance. It reduces the annotation effort by automatically selecting informative images for annotation, facilitates the annotation process by proposing likely objects and labels, and accelerates the integration of new labels into the deep neural network model by avoiding re-training from scratch.

Carpe Diem implements active learning, which intelligently explores unlabeled data and only selects valuable examples to avoid redundant annotations. This strategy saves expensive human resources. Moreover, incremental learning enables a continually improving model. Whenever new annotations are available, the model can be updated efficiently and quickly, without re-training, and regardless of the amount of accumulated training data. Because there is no single large training step, the model can be used to make predictions at any time. We exploit this in our annotation process, where users only confirm or reject proposals instead of manually drawing bounding boxes.

Keywords: Lifelong Learning; Object Detection; Automated Monitoring

1 Introduction

Large-scale biodiversity studies are an important tool in coordinating a global response to the biodiversity crisis [Car+12; VHS09]. While camera traps are less invasive than conventional capture-mark-recapture approaches [AMM10], the resulting image data still requires analysis. When the scale of camera trap data makes manual analysis infeasible, there are several alternatives. Citizen scientists can be employed to distribute the analysis workload [Swa+15] if the task is reasonably attractive and not too difficult for laypeople. Automatic analysis using machine learning can help overcome the scale limitations. Such methods successfully generalize from smaller amounts of human-annotated data to large-scale processing [Bru+17; Nor+17; GSV17; DKB18; Gir+19].

However, it can still be cost-prohibitive to acquire sufficient labeled training data for machine learning analysis if (expensive) domain expertise is required and funding is inadequate. Active learning [Set09; BKD19] can substantially reduce the number of annotated images

¹ Friedrich Schiller University Jena, Computer Vision Group, Jena, Germany firstname.lastname@uni-jena.de

² DLR Institute of Data Science, Jena, Germany firstname.lastname@dlr.de

required to train a machine learning model with a given accuracy. It intelligently decides which unlabeled images should be annotated by a human, and which should not, eliminating redundancies.

Ideally, active learning is integrated in a feedback loop such that only a very small number of images is selected and annotated at a time. After each batch of annotations, the model is updated quickly with the new data using an incremental learning method [Käd+16a]. Then, the loop repeats. New unlabeled images are selected using active learning and presented to the human annotator. Because active learning considers the model’s current knowledge, it selects different images in each iteration. This setup is called lifelong learning [Käd+16b].

In this work, we present a Carpe Diem, GUI that implements the lifelong learning loop. It enables efficient annotation and is a machine learning tool supporting all object detection tasks intended for use by biodiversity researchers. When a model is trained to satisfactory performance, a batch prediction mode can be used to perform analysis on large amounts of images.

2 Active and Incremental Learning Methods

We implement a selection of active learning methods that can be grouped in two major categories [BKD20]. First, a method that is in principle compatible with any object detection system where predicted class probabilities are available for individual detected instances. We evaluate an uncertainty-based active learning metric for classification such as margin sampling or entropy sampling [Set09] on each instance. The resulting set of metrics is then aggregated into a whole-image score by computing the sum, maximum or average.

Second, two heuristics that are specific to the YOLO object detector [Red+16]. One compares the confidence scores predicted by the classification part and the “objectness” scores and considers mismatches valuable, as they can result from unknown classes or missed detections. The other applies margin sampling to each YOLO grid cell individually and computes a sum weighted by the respective objectness scores. The user can select from the aforementioned implementations which are shown to improve mAP on their own by up to 5.7% in a representative fast exploration scenario. They change the methods at any time.

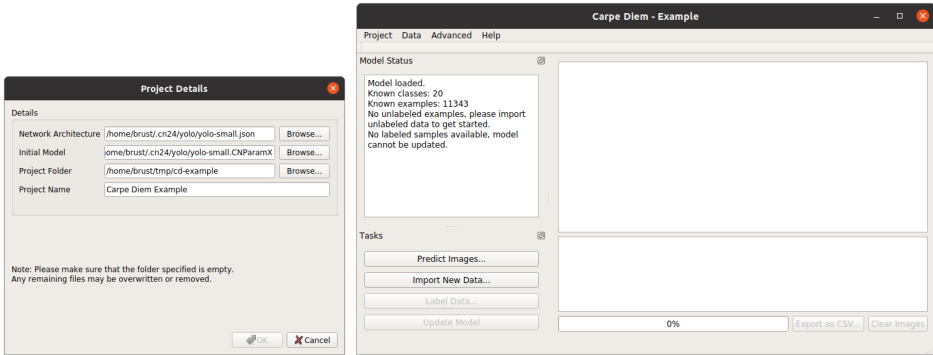
For incremental learning, we implement deep fine-tuning [Käd+16a], which is shown to alleviate catastrophic forgetting. It combines minibatches of previously seen and newly labeled examples in a fixed ratio λ , which defaults to 0.5. We allow the user to configure λ as well as the number of gradient steps per model update to trade off accuracy and runtime.

The combined system, including the weakly supervised annotation scheme proposed in [Pap+16] which further improves annotation time efficiency by a factor of five to nine, is validated in Brust, Käding, and Denzler [BKD20], including a case study in a biodiversity monitoring setting.

3 User Experience

Carpe Diem is built around a lifelong learning workflow. This workflow considers two major characteristics of large-scale automated monitoring projects. First, more data becomes available over time, as new images are transmitted from camera traps or observers. Second, the availability of annotators also changes over time, with funding inconsistencies and volunteers working in their spare time. Lifelong learning adapts to changing environments efficiently with incremental learning, and active learning strategies can intelligently select images with novel or interesting content to save annotation time.

Project Setup and Initial Model Before the lifelong learning process can begin, the user creates a new *project* (see Fig. 1a). A project captures the state of the machine learning component, i. e., the current model parameters, as well as the collected images and annotations. When creating a project, the user can choose from a selection of object detection models, e. g., YOLO [Red+16]. Because the lifelong learning process requires a working initial model, we include weights and the dataset that is used to train the respective model.



(a) Creating a new project.

(b) Initial state after creating a new project.

Fig. 1: Project setup and initial state.

Data Management The data stored in a project is divided into labeled and unlabeled data. Initially, only labeled data (from the initial model) is available. As indicated in a status message (see Fig. 1b), the next step is importing unlabeled images, e. g., from camera traps. However, it is also possible to import labeled images and fine-tune the model at any time. All data stored in the project can be reviewed and modified at all times (see Fig. 2), e. g., to remove faulty annotations. The user can rename, remove and add object classes known to the model.

Selection, Annotation and Model Update The lifelong learning cycle runs for as long as the project contains unlabeled images. Whenever the user is ready to annotate a small batch of images, they press the “Label Data...” button. Carpe Diem processes all unlabeled images and selects a user-configurable number of valuable examples. The selection is performed according to the active learning strategy chosen by the user. We implement several active learning methods as described in Sect. 2.

This batch is then presented to the user for annotation. There is no need for manual drawing of bounding boxes. Instead, the user only confirms, modifies, or rejects proposals from the object detection model (see Fig. 3a). This further increases annotation efficiency [Pap+16].

After each annotated batch, the model has to be updated to ensure a tight feedback loop. This results in fewer redundant annotations and a faster increase in accuracy. We use the incremental learning technique presented in Käding et al. [Käd+16a] to perform the model update quickly. Model snapshots can be stored to preserve prior states and revert to them in the event of errors, e. g., divergence during optimization.

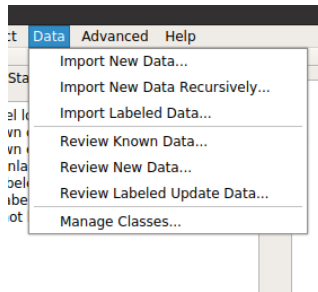
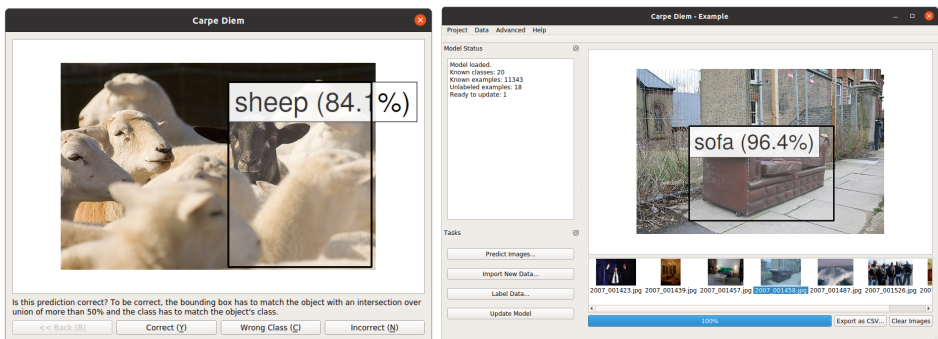


Fig. 2: Menu for data management.



(a) The annotation process with keyboard shortcuts. (b) Browsing predictions in the thumbnail gallery.

Fig. 3: Annotation and prediction.

Predictions Only a small fraction of the unlabeled images should ever be annotated. The main purpose of annotating data and training a model is to use its generalization capability for an analysis of the large-scale unlabeled data. After validating the model performance using a held-out hand-annotated validation set, the model can be used to make predictions. Carpe Diem can process very large batches of images for prediction and export the detected objects in comma-separated value (CSV) format. Before export, the user can browse the predictions in a gallery (see Fig. 3b).

4 Interoperability and System Requirements

Carpe Diem is built on the CN24 deep learning framework [Bru+15]. The framework is binary compatible with darknet³ models, e. g., Extraction and YOLO [Red+16]. Custom architectures can be specified in a human-readable JSON format. The same is true for image datasets. If needed, a user can easily construct and train an initial model on their own annotated data in CN24 and import it into Carpe Diem for lifelong learning.

Predictions are exported in CSV format for further analysis, with one line representing a single detected object. Each prediction includes the object class, a confidence score, and the coordinates of an axis-aligned bounding box around the object.

Carpe Diem is tested on Windows 10 and Linux (Ubuntu 18.04 LTS, 20.04 LTS) operating systems. It mainly depends on CN24 (which in turn requires OpenCL and cBLAS) and Qt 5. We recommend using a GPU that implements OpenCL 1.2 and is equipped with at least 4 GiB of VRAM. Carpe Diem supports both AMD and NVIDIA GPUs. Training the object detectors, even on the YOLO-Small variant, is not practical on CPUs only.

5 Conclusion

We present Carpe Diem, a fully integrated lifelong learning experience for object detection tasks. It is capable of adapting to changing environments and acquires annotations as efficiently as possible. The underlying methods are empirically validated individually [Käd+16a; BKD19; Pap+16] and as a complete system in a wildlife monitoring setting [BKD20]. The software will be made available publicly under a three-clause BSD license upon formal publication.

Future Work Carpe Diem is undergoing constant improvement. Current development efforts include a client-server separation layer to enable multi-user annotation scenarios. Further, we are enabling support for TensorFlow [Aba+15] models. Plans for the future include alternative annotation methods such as extreme clicking [Pap+17] and methods for integrating domain knowledge [BD19] to further increase label efficiency.


³ <https://pjreddie.com/darknet/>

References

- [Aba+15] Martin Abadi et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Software available from tensorflow.org. 2015. URL: <https://www.tensorflow.org/>.
- [AMM10] Steven C Amstrup, Trent L McDonald, and Bryan FJ Manly. *Handbook of Capture-recapture Analysis*. Princeton University Press, 2010. ISBN: 978-0-691-08967-6.
- [BD19] Clemens-Alexander Brust and Joachim Denzler. “Integrating Domain Knowledge: Using Hierarchies to Improve Deep Classifiers”. In: *Asian Conference on Pattern Recognition (ACPR)*. 2019. DOI: 10.1007/978-3-030-41404-7_1.
- [BKD19] Clemens-Alexander Brust, Christoph Käding, and Joachim Denzler. “Active Learning for Deep Object Detection”. In: *Computer Vision Theory and Applications (VISAPP)*. 2019. DOI: 10.5220/0007248601810190.
- [BKD20] Clemens-Alexander Brust, Christoph Käding, and Joachim Denzler. “Active and Incremental Learning with Weak Supervision”. In: *Künstliche Intelligenz* 34 (2 2020), pp. 165–180. DOI: 10.1007/s13218-020-00631-4.
- [Bru+15] Clemens-Alexander Brust et al. “Convolutional Patch Networks with Spatial Prior for Road Detection and Urban Scene Understanding”. In: *Computer Vision Theory and Applications (VISAPP)*. 2015. DOI: 10.5220/0005355105100517.
- [Bru+17] Clemens-Alexander Brust et al. “Towards Automated Visual Monitoring of Individual Gorillas in the Wild”. In: *International Conference on Computer Vision Workshops (ICCV-WS)*. 2017. DOI: 10.1109/ICCVW.2017.333.
- [Car+12] Bradley J Cardinale et al. “Biodiversity Loss and Its Impact on Humanity”. In: *Nature* 486.7401 (2012), pp. 59–67. DOI: 10.1038/nature11148.
- [DKB18] Joachim Denzler, Christoph Käding, and Clemens-Alexander Brust. “Keeping the Human in the Loop: Towards Automatic Visual Monitoring in Biodiversity Research”. In: *International Conference on Ecological Informatics (ICEI)*. 2018.
- [Gir+19] Jhony-Heriberto Giraldo-Zuluaga et al. “Camera-trap Images Segmentation Using Multi-layer Robust Principal Component Analysis”. In: *The Visual Computer* 35 (3 2019), pp. 335–347. DOI: 10.1007/s00371-017-1463-9.
- [GSV17] Alexander Gomez Villa, Augusto Salazar, and Francisco Vargas. “Towards Automatic Wild Animal Monitoring: Identification of Animal Species in Camera-trap Images Using Very Deep Convolutional Neural Networks”. In: *Ecological Informatics* 41 (2017), pp. 24–32. DOI: 10.1016/j.ecoinf.2017.07.004.

-
- [Käd+16a] Christoph Käding et al. “Fine-tuning Deep Neural Networks in Continuous Learning Scenarios”. In: *Asian Conference on Computer Vision Workshops (ACCV-WS)*. 2016. DOI: 10.1007/978-3-319-54526-4_43.
- [Käd+16b] Christoph Käding et al. “Watch, Ask, Learn, and Improve: a lifelong learning cycle for visual recognition.” In: *European Symposium on Artificial Neural Networks (ESANN)*. 2016.
- [Nor+17] Mohammed Sadegh Norouzzadeh et al. “Automatically Identifying, Counting, and Describing Wild Animals in Camera-trap Images with Deep Learning”. In: (2017). arXiv: 1703.05830v5 [cs.CV].
- [Pap+16] Dim P. Papadopoulos et al. “We Don’t Need No Bounding-boxes: Training Object Class Detectors Using Only Human Verification”. In: *Computer Vision and Pattern Recognition (CVPR)*. June 2016. DOI: 10.1109/CVPR.2016.99.
- [Pap+17] Dim P Papadopoulos et al. “Extreme clicking for efficient object annotation”. In: *International Conference on Computer Vision (ICCV)*. 2017.
- [Red+16] Joseph Redmon et al. “You Only Look Once: Unified, Real-time Object Detection”. In: *Computer Vision and Pattern Recognition (CVPR)*. 2016. DOI: 10.1109/CVPR.2016.91.
- [Set09] Burr Settles. *Active Learning Literature Survey*. Tech. rep. University of Wisconsin-Madison, 2009.
- [Swa+15] Alexandra Swanson et al. “Snapshot Serengeti, High-frequency Annotated Camera Trap Images of 40 Mammalian Species in an African Savanna”. In: *Scientific Data* 2 (1 2015). DOI: 10.1038/sdata.2015.26.
- [VHS09] Jean-Christophe Vié, Craig Hilton-Taylor, and Simon N. Stuart, eds. *Wildlife in a Changing World: An Analysis of the 2008 Iucn Red List of Threatened Species*. International Union for Conservation of Nature and Natural Resource, 2009. ISBN: 978-84-96553-63-7.

tRackIT OS: Open-source Software for Reliable VHF Wildlife Tracking

Jonas Höchst¹ , Jannis Gottwald² , Patrick Lampe¹ , Julian Zobel³ , Thomas Nauss² , Ralf Steinmetz³ , Bernd Freisleben¹ 

Abstract: We present *tRackIT OS*, open-source software for reliable VHF radio tracking of (small) animals in their wildlife habitat. *tRackIT OS* is an operating system distribution for *tRackIT stations* that receive signals emitted by VHF tags mounted on animals and are built from low-cost commodity-off-the-shelf hardware. *tRackIT OS* provides software components for VHF signal processing, system monitoring, configuration management, and user access. In particular, it records, stores, analyzes, and transmits detected VHF signals and their descriptive features, e.g., to calculate bearings of signals emitted by VHF radio tags mounted on animals or to perform animal activity classification. Furthermore, we provide results of an experimental evaluation carried out in the *Marburg Open Forest*, the research and teaching forest of the University of Marburg, Germany. All components of *tRackIT OS* are available under a GNU GPL 3.0 open source license at <https://github.com/nature40/tRackIT-OS>.

Keywords: Radio Tracking; Marburg Open Forest; Open Source; Movement Ecology

1 Introduction

In an increasingly densely populated and anthropogenically dominated environment, a scientific analysis of the consequences of human-wildlife interaction is essential for developing evidence-based guidelines for conservation [KA20]. Understanding the impact of altered habitats on the spatial distribution of species [Sa09], the effects of human infrastructures such as roads [Ho15, As19], and reasons for increased mortality of endangered species [Le19] is crucial for preserving biodiversity in a crowded world. Movement data of animals generated by recent technological advances support more detailed forms of analysis and insights into the behavior and ecology of threatened species than ever before [Wy18, Ca10, Wa18].

Wildlife observations can be realized with a variety of technologies. For example, GPS technology can be used to equip animals and record their movements independently of

¹ Department of Mathematics & Computer Science, University of Marburg, Germany, {hoechst,lampe,freisleb}@informatik.uni-marburg.de

² Department of Geography, University of Marburg, Germany, {gottwal5,nauss}@staff.uni-marburg.de

³ Multimedia Communication Lab (KOM), Technical University of Darmstadt, Germany {julian.zobel,ralf.steinmetz}@kom.tu-darmstadt.de

other communication infrastructures. However, size, weight, and battery life constraints prevent the use of GPS for most European songbirds and bats.

Manual radio telemetry is another option for observing small animals. However, it is extremely labor-intensive, limited to a small number of individuals that can be tracked simultaneously [Co99], and results in spatial and temporal data with poor resolution, which might not be sufficient for meaningful scientific analyses [Mo10].

Automated radio telemetry systems can minimize many of these disadvantages [Ka11, We16]. In previous work, some of us presented a system based on commodity-off-the-shelf (COTS) hardware for automatic radio tracking of small animals based on Very High Frequency (VHF) tags [Go19] as part of the open source project *radio-tracking.eu*⁴. However, three seasons of long-term stationary operating time of the system in the *Marburg Open Forest* (i.e., the teaching and research forest of the University of Marburg, Germany) revealed several deficits, such as the lack of failure handling, inadequate interfaces for data transmission and health-state monitoring, and problems with time synchronization of received signals between receivers of the same station and among different stations.

In this paper, we present *tRackIT OS*, an open-source operating system distribution for reliable VHF radio tracking of small animals. *tRackIT OS* runs on a *tRackIT station*; its basic hardware design is due to Gottwald et al. [Go19]. We developed *tRackIT OS* to provide new software functionality according to our experiences in studying the movement ecology of both diurnal and nocturnal wildlife with a network of 15 *tRackIT stations* in densely forested terrain. In particular, we present:

- a novel approach for automated signal detection of VHF radio tracking tags,
- means to provide reliable operation of *tRackIT stations* under harsh conditions,
- efficient live data transmission for monitoring data and detected signals,
- a novel web-based user interface for intuitive configuration of *tRackIT stations*,
- a comparative evaluation of *tRackIT OS* compared to the state-of-the-art.

The rest of the paper is organized as follows. Section 2 discusses related work. Section 3 discusses requirements, design decisions and implementation details, followed by experimental results in Section 4. Section 5 concludes the paper and outlines areas of future work.

2 Related Work

Ripperger et al. present a comprehensive overview of existing systems for localizing small animals using different technologies [Ri20]. The most recent projects on automated VHF

⁴ <https://radio-tracking.eu>

transmitter tracking are ARTS [Ka11], Atlas [We16], and Motus (also called SensorGnome) [Ta17].

ARTS consisted of towers with a height of 40 meters and top-mounted antenna arrays [Ka11], but the system was taken down in 2010 and replaced by camera traps and GPS transmitters. ARTS was able to determine the position of a tagged individual by triangulation with a spatial accuracy of 50 meters, but rotating through channels with different frequencies reduces the time span in which each individual can be observed. *tRackIT* supports more detailed observations of movements using a higher number of stations at lower cost and less effort in construction.

The Atlas project achieves great spatial accuracy by using the *time difference of arrival* (TOA) method for direction estimates as seen from the receiver, while costs for the developed tags are low [We16]. However, implementation of the receiving stations is quite expensive, a fact that probably explains why the system is only deployed in three areas in the Netherlands, England, and Northern Israel. *tRackIT* achieves comparable results with stations built from commodity off the shelf hardware at a lower price point.

Motus⁵ is a globally operating network of VHF receiver stations hosted by different collaborators and supporting researchers [Ta17]. Despite its open source character, an implementation of Motus at US\$ 3000 for a single SensorGnome⁶ receiver with three 9-element Yagi antennas, and US\$ 7500 for a Lotek SRX800 receiver station with four 9-element Yagi antennas is costly [LN18], leading to a trade-off between spatial resolution and coverage. By default, the implemented radio receiver listens at a single center frequency and can detect pulses from tags in a narrow band of ± 24 kHz around its center frequency. This limits the number of distinguishable frequencies, i.e., the number of detectable individuals, substantially. Motus has delivered great insights into the ecology of different species in more than 120 research projects [Ta17], but investigating fine-grained spatial movements by triangulation is not supported by the system. The wide frequency band that can be used by *tRackIT* supports both fine-grained temporal resolutions and observations of many individuals.

3 *tRackIT OS*

A *tRackIT* system consists of (a) VHF radio tags mounted on animals, (b) *tRackIT stations* for receiving signals emitted by VHF tags, (c) *tRackIT OS* running on *tRackIT stations* for detecting and matching signals received on multiple antennas, (d) *tRackIT servers* for collecting and presenting data transmitted from *tRackIT stations*, and (e) *tRackIT analytics modules* for deriving ecological knowledge from the collected data.

In this section, we present design and implementation issues of *tRackIT OS*, the operating system distribution for *tRackIT stations*.

⁵ Motus Wildlife Tracking System: <https://motus.org>

⁶ SensorGnome Project: <https://sensorgnome.org>

3.1 Requirements

Our experiences from three seasons of field work have indicated that automatic telemetry can only be a useful substitute of its manual counterpart if certain requirements are met:

1. *Low entry barrier.* To make automatic radio telemetry accessible to the widest possible user community, both hardware and software as well as data processing and analysis must be conveniently accessible, easy to use, and inexpensive.
2. *Reliability.* The used equipment must reliably record signals originating from VHF transmitters and minimize the amount of interference. Any component failures caused by adverse conditions, such as unstable power supplies, fluctuating temperatures, and hardware-based failures should be detected and handled automatically.
3. *Data availability.* In many application areas, like mortality studies [He20], fast data availability is highly important. Thus, direct data transmission from the field with the shortest possible delay between recording and transmission is desirable.

3.2 *tRackIT* Station

To deploy an operational installation in the field, a *tRackIT station* is equipped with directional antennas in the four cardinal directions, a solar panel, and a battery box. The basic hardware design is due to Gottwald et al. [Go19]. We have slightly adapted the hardware by including an active USB hub, a better LTE modem, and a LoRa (Long Range Wireless Radio Frequency Technology⁷) expansion board (LoRa HAT), as shown in Figure 1.

The 'brain' of a *tRackIT station* is a Raspberry Pi 3 Model B that consists of a quad-core 1.2 GHz ARM-Cortex-A53 and 1 GB of RAM. It offers various input/output options, including Wi-Fi and 4 USB ports. The system is powered through a 5V USB port and is capable of powering connected USB devices. The four directional antennas are connected to four software-defined radios (SDR) (Nooelec NESDR SMARt v4) for signal analysis. Since these SDRs require more power than provided by the Raspberry Pi, an active 4-port USB hub (Anker 4-port Ultra Slim USB 3.0 Data Hub, A7518) is used to connect the devices. An LTE modem (Huawei E3372H) and a local prepaid data plan is used to establish a mobile Internet connection. The battery box provides a 12 V source that is converted using a step down converter rated for 2×2.4 A at 5 V. For *tRackIT stations* relying on LoRa for data publishing, a Dragino SX127X GPS HAT⁸ is used. For receiving and forwarding *tRackIT stations*, the Dragino PG1301 LoRa Concentrator is used⁹. The basic hardware of a *tRackIT station* costs a total of about 200 €, consisting of 35 € for the Raspberry Pi 3B+,

⁷ Semtech: <https://www.semtech.com/lor/>

⁸ Dragino SX127X: <https://www.dragino.com/products/lor/item/106-lora-gps-hat.html>

⁹ Dragino PG1301: <https://www.dragino.com/products/lor/item/149-lora-gps-hat.html>

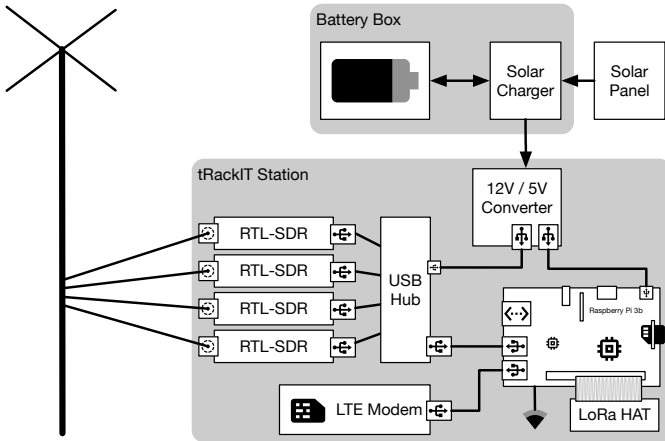


Fig. 1: The hardware components of a tRackIT station.

4 × 35 € for the Nooelec SDRs, 15 € for the active USB hub, and 10 € for the power supply unit. The optional communication modules cost 50 € in the case of the Huawei LTE modem and/or 35 € (LoRa HAT) / 110 € (LoRa Concentrator) for the LoRa publish / receive upgrade.

3.3 tRackIT OS Components

The operating system (OS) plays a crucial role in the reliable autonomous operation of the presented hardware. We developed a custom distribution of the Raspberry Pi OS, called *tRackIT OS*. The primary task of *tRackIT OS* is to execute a signal detection module, called *pyradiotracking*, in a reliable manner. The secondary task is to interface with users (a) interactively while setting up the station, and (b) continuously during autonomous operation for extended monitoring. *tRackIT OS* is built using PIMOD [Hö20b], which allows configuration of single-board computer system images in a reproducible manner. The resources required to build *tRackIT OS* as well as the OS image itself are released under a GPL 3.0 license¹⁰.

In Figure 2, the main software components of *tRackIT OS* are presented. Station-initiated communication is handled using the *Message Queuing Telemetry Transport* (MQTT) protocol, with *mosquitto* as an MQTT client and server implementation [Li17] for message distribution. It is configured such that incoming messages are forwarded to remote MQTT brokers for further processing. These brokers are also responsible for detecting and resolving connection failures.

¹⁰ tRackIT OS, available online <https://github.com/Nature40/tRackIT-OS>

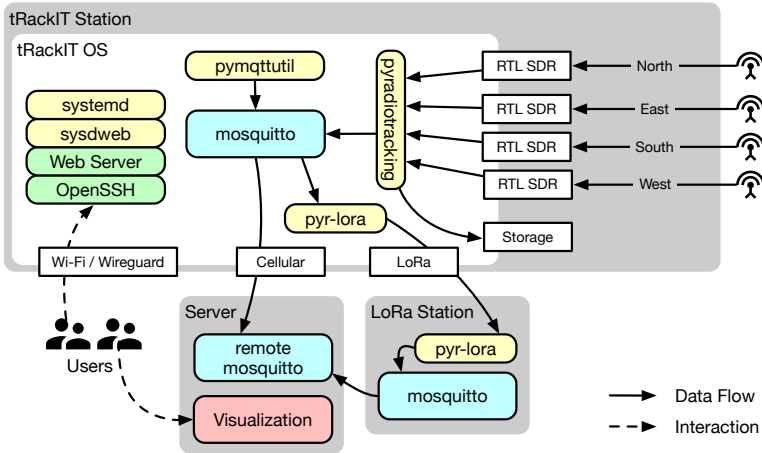


Fig. 2: Overview of the main software components of a tRackIT OS distribution.

The core software component for signal detection is called *pyradiotracking*. The component reads samples from all four SDRs, as well as detects, filters, and matches signals of VHF tags. Detected signals are saved to local storage, displayed via a custom web user interface, and published to a local message bus that is responsible for data distribution. Section 3.4 discusses the implementation details of *pyradiotracking*.

For system monitoring, we implemented a custom tool called *pymqtutil* in the Python programming language. It is released under a GPL 3.0 license¹¹. The tool executes configurable Python statements in a fixed schedule and publishes the corresponding results via MQTT. It is configured such that relevant system metrics are published in a 5 minute interval, i.e., temperature, system uptime, system load, memory usage, CPU frequency, network addresses, storage utilization, and cellular data usage.

All services are managed by *systemd*. The WebUI *sysdweb*¹² for *systemd* is configured to allow easy log access and service control for (mobile) users. The *Caddy* web server is used to provide convenient access to the local storage, *pyradiotracking* and *sysdweb*. Finally, OpenSSH provides direct system access for local and remote users. To allow secure remote access, *wireguard* is used as a virtual private network (VPN).

¹¹ *pymqtutil*, available online: <https://github.com/Nature40/pymqtutil>

¹² *sysdweb*, available online: <https://github.com/Nature40/sysdweb>

3.4 Signal Detection

The signal detection algorithm is implemented in the *pyradiotracking* Python package, which is released under a GPL 3.0 license¹³. In Figure 3, the stages of signal processing are presented in a block diagram. First, spectrograms of the incoming IQ samples are created, which are used to detect signals. The detected signals are then filtered for shadow signals of lower power in neighboring frequencies and sent to a central signal queue. The detected signals of multiple antennas are matched and written to a local file, published to the MQTT message bus, and visualized in the local dashboard.

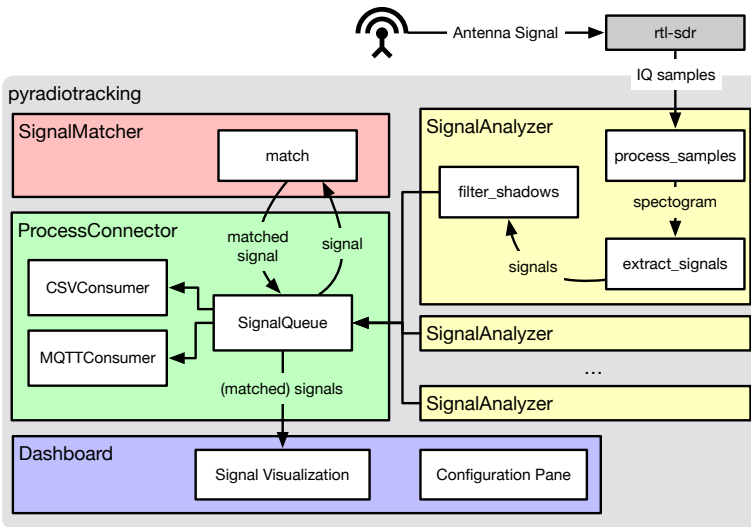


Fig. 3: Signal analysis stages implemented in *pyradiotracking*.

To illustrate how the different stages work, data of the length of one second is used as an example. An SDR is configured such that a center frequency of 150.150 MHz, a sample rate of 300 kHz, and a fixed gain of 49.6 dB are used. A test tag with the frequency of 150.172 MHz and a signal duration of 40 ms was placed near to the receiving antenna. In Figures 4, 5, and 6, three stages of signal processing are visualized.

Figure 4 shows the raw IQ samples received by the SDR. Following the example configuration described above, there are 300,000 samples, hence 600 kilobyte of data collected in one second. In the time interval of $t_0 = 0.45\text{ s}$ to $t_1 = 0.49\text{ s}$, the IQ samples contain high values, which appear as a rectangle in the visualization. This rather sharp rectangle indicates that the gain value is set too high and the signal is clipping. When setting up stations for regular operation, the gain value must be chosen such that a good compromise of gain and clipping is achieved.

¹³ *pyradiotracking*, available online: <https://github.com/Nature40/pyradiotracking>

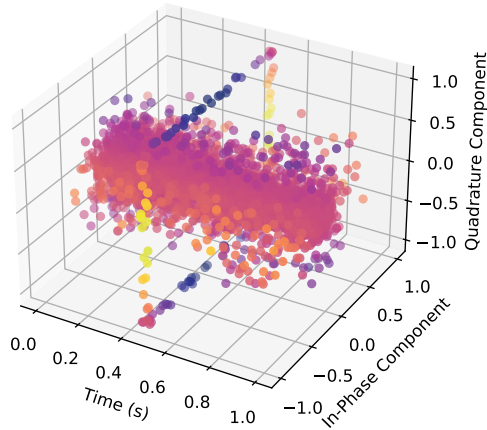


Fig. 4: IQ samples of one second, as received by RTL-SDR.

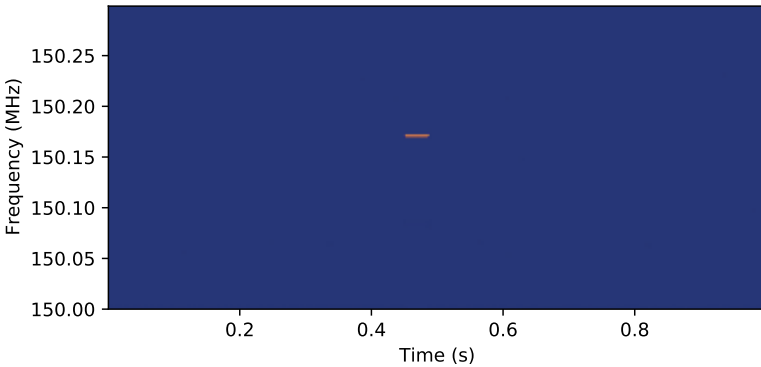


Fig. 5: Power spectral density (PSD) of samples computed via Short-time Fourier Transform (STFT).

To detect single signals from the received data, a spectrogram is computed and processed further. This is achieved by applying consecutive Short-time Fourier Transforms (STFT) [A177] to the data. Figure 5 shows the spectrogram computed from the previously presented example data. The STFTs are computed with 256 samples per Fast Fourier Transform (FFT) and no overlapping samples. The Hamming window function is applied to smoothen discontinuities at the start and the end of the processed FFT. In this configuration, the bandwidth of 300 kHz is divided into 256 bands and a frequency resolution of $1,171 \text{ kHz}$, to achieve a time resolution of $1.0 \text{ s}/1,171 = 0.853 \text{ ms}$.

In Figure 6, signal detection on individual frequencies is visualized. The signal power (dBW) in the logarithmic scale is plotted for four example frequencies near the test sender’s frequency, and the signal emitted by the test sender can be observed in three of those. The

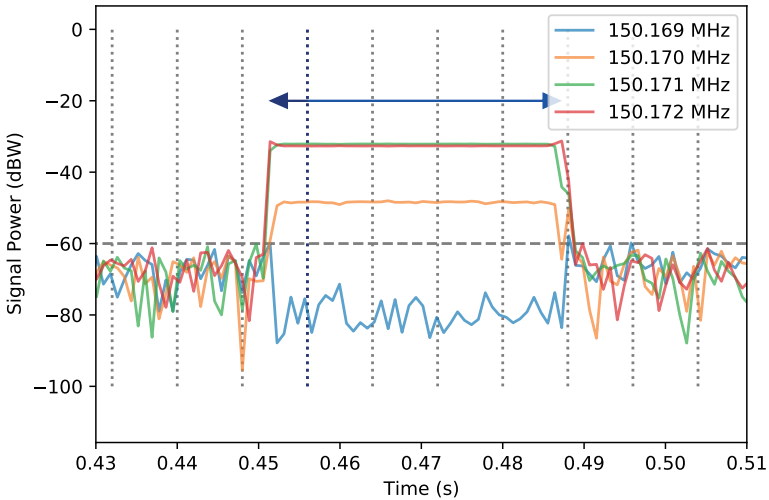


Fig. 6: Power spectral densities (PSDs) of selected frequencies, minimal signal power threshold, and signal power sampling points.

gray dashed horizontal line indicates the configured signal power threshold of -60 dBW. The gray dotted vertical lines show scan points used for initial signal detection. The blue arrow marks the total detected signal length. Signal detection is achieved by (a) iterating through all frequencies and (b) iterating through time using scan points placed according to the minimal detectable signal duration of 8 ms in our example. The signal-to-noise ratio (SNR) is calculated using the ratio of the current power and the average signal power of this frequency. If signal power and SNR at the scan points are above the configured thresholds, a potential signal is detected. The scan is then continued by evaluating the thresholds for all neighboring values until the thresholds are undershot, indicated by the blue arrows. If the duration of the detected signal is within the set limits, further complementary features are computed and added to a list for further processing.

After all signals of a spectrogram are extracted, shadow filtering is performed. We define a shadow signal as a signal that matches another signal in duration and time, but has a lower detected power. In the example of Figure 6, the signals detected at 150.170 MHz and 150.172 MHz would be shadow signals of the 150.171 MHz signal. The shadow signals are removed and the detected primary signals are added and written to disk, published via MQTT and sent to *pyradiotracking's* main process for signal matching and data presentation.

To improve reliability, a direct control component is introduced. The `librtlsdr` library used to retrieve data from an SDR works in such a way that as soon as requested data is available, a callback method is called. If the system load is too high and the callback method takes longer than the acquisition of the next samples, individual samples are omit-

ted. Hardware and library-specific errors may lead to no callbacks at all. The first problem is monitored by comparing the actual received samples with the expected number of samples using the system clock. In this way, dropped samples can be detected, even accumulated over longer periods of time. The second problem is solved by (re-)setting a periodic alarm, comparable to a dead man’s switch. If the callback method is not called in time, an alarm is triggered. This terminates the analysis process, which is then restarted by *pyradiotracking*’s main process.

3.5 Signal Matching

The detected and filtered signals of multiple antennas are consumed by the signal matcher, which works as follows. In a list, all currently active signal groups are held. When a new signal is detected, it is compared to each of the active signal groups in time, duration, and frequency. The SDR devices used in the project do not work synchronously and use individual quartz crystals as their clock sources, hence time and frequency mismatches are likely to happen. If all parameters of an active signal group are within the configured thresholds, the signal is added to the corresponding group. If no corresponding group is found, a new active signal group with this signal is created and added to the list. After a certain timeout, the active signal groups are removed from the list and the key features are written to disk and published.

3.6 Data Publishing

Detected signals are published directly to disk in CSV format and via MQTT in the CBOR format, which is a binary format and introduces smaller overheads compared to text-based formats. The MQTT broker running on a *tRackIT station* can be configured to forward published signals to other brokers, such as a central server via a cellular network or other IP-based networks.

Field	Accuracy	Size (bit)
Time	ms of current min	16
Frequency	offset to 150 MHz in kHz	9
Duration	ms	6
Signal Availability	flags	4
Signals	3-decimals	$[1 - 4] \times 17$

52 – 103

Tab. 1: *tRackIT station*’s LoRa matched signal payload: fields, accuracy, and sizes.

In addition to this IP-based data publishing, LoRa can be used to publish signals. LoRa is a physical layer protocol based on *chirp spread spectrum* (CSS) modulation, that is robust against channel noise, multi-path fading, and the Doppler effect. This allows transmission ranges of a few kilometers in urban environments and up to 15 kilometers in rural

areas, with minimal power requirements but also only low data rates between 300 bps and 50 kbps [Md17, Pe17, Hö20a].

The LoRa publishing service of *tRackIT OS* receives signals through the local MQTT broker, converts the data in a custom data-saving binary format, and sends it via LoRa. Table 1 shows the fields used for a matched signal’s payload, including accuracy and required size in bits. A matched signal contains a minimum of one and a maximum of four signals, depending on the number of antennas that received the signal, hence the final payload size is 52 up to 103 bits. Zeros are appended to the payload to reach the required byte boundaries, resulting in messages of 7, 9, 11, and 13 bytes, depending on the number of the contained matched signals. Compared to the already compact representation in CBOR of a 4-component matched signal (56 bytes + overhead), a reduction of up to 77% is achieved (13 bytes). In the most robust LoRa settings (SF:12, BW:250 kHz, CR:4/8), such a shortened message would require 594 ms Time-on-Air (ToA) (using Implicit Header mode with a 1-byte sender ID, the total length of the packet is 14 bytes). Following the duty cycle regulation of a maximum utilization of 1% (10%) per band, a message could be sent every 59 (5.9) seconds. While these settings do not allow continuous monitoring of individuals, sparse reporting of single observations are still of value, when trying to detect tags fallen off or with empty batteries. For stations in closer physical proximity to the receiving gateway, less robust settings may be chosen. Using a less robust LoRa setting (SF:8, BW:250 kHz, CR:4/8), the ToA drops down to 45.5 ms, allowing messages to be sent in an interval of 4.6 (0.46) seconds. From previous measurements in the Marburg Open Forest, signals could be reliably transmitted over 600 meters using this setting.

4 Experimental Evaluation

In this section, we evaluate *tRackIT OS* in benchmarking scenarios and in field experiments. The data of all experiments is publicly available at <https://github.com/Nature40/hoechst2021tRackIT-eval>.

4.1 Experimental Scenario

To evaluate *tRackIT OS* in a realistic manner, we use a system setup in the *Marburg Open Forest*, consisting of 15 *tRackIT* stations; 5 of them are used in our evaluation described below. The experiments are carried out twice: (a) with the most recent *tRackIT OS 0.7.0* and (b) using the most recent stable operating system version of the *radio-tracking.eu*¹⁴ project [Go19], called *paur 4.2*. We activated a test tag and carried it around in the area of the selected *tRackIT stations* together with a GPS receiver to receive ground truth data. The experiment took place over the course of 0:51:10 h with a VHF sender of 600 μ W power, 20

¹⁴ <https://radio-tracking.eu>

ms duration and an interval of one signal per second, which results in 3,193 sent signals. In Figure 7, the GPS trace of the conducted experiment is presented; stations are marked by the white circles, and the trace is colored to indicate the time component of the experiment.

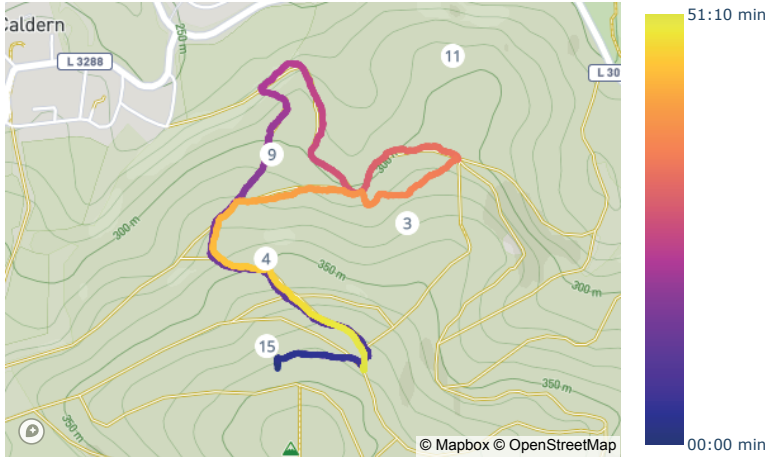


Fig. 7: GPS trace of the experimental evaluation track and the corresponding *tRackIT* stations.

Our observations using *paar* in two seasons of 2019 and 2020 indicated high numbers of falsely detected signals. We were not able to distinguish between true and false positives through the information available after signal detection. Thus, we used a power signal threshold. The *paar* experiments conducted in this paper showed the same low precision, hence all detected signals with a power lower than -78 dBW were removed for further processing. Using *tRackIT OS*, this threshold is not required, since we observed very low numbers of falsely detected signals.

4.2 Signal Delay

A second observation from our previous field seasons in 2019 and 2020 is a delay in signal detection using *paar* in the order of seconds to minutes. In Figure 8, an example of observed signal delay is visualized. The dots show the received signal strength measured on multiple antennas of the same *tRackIT station*. Every antenna received a series of signals with low variance in signal strength that appear to be a straight line, indicating that the tag is not moving. However, these straight lines on the individual receivers are offset in time from each other, which makes further processing of the data difficult and leads to worse to unusable bearing calculation. In the experiments of this paper, we measured a delay in signal detection of 8 seconds in *paar* and no recognizable delay in *tRackIT OS*.

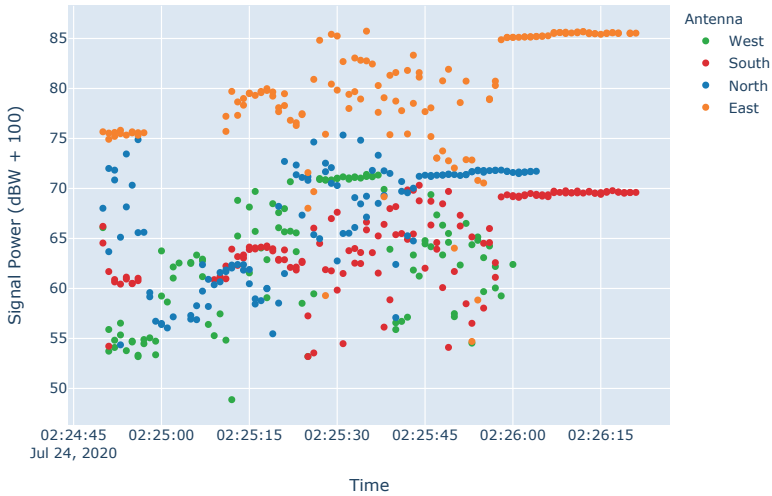


Fig. 8: Example of signal delay among different receivers observed in the 2020 field season using pair.

4.3 Signal Detection

In the *tRackIT OS* experiments, the selected five *tRackIT stations* detected 30,507 signals, each on potentially four antennas, resulting in an average of 1,525 signals (47.8%) detected per antenna. Signal detection depends on various factors, such as geographical and topographical conditions, the orientation of the antenna, the height of the transmitter above the ground, air humidity, and forest cover. Figure 9 shows the numbers of detected signals by station and antenna. Due to the positioning of the stations, the orientation of the antennas, and the selected test area, some of the antennas receive only a very small amount, others a large amount of the test signals. On the north antenna of station 11, only 95 (3.0%) signals could be detected, while 2,611 (81.8%) signals were detected on south antenna of station 9.

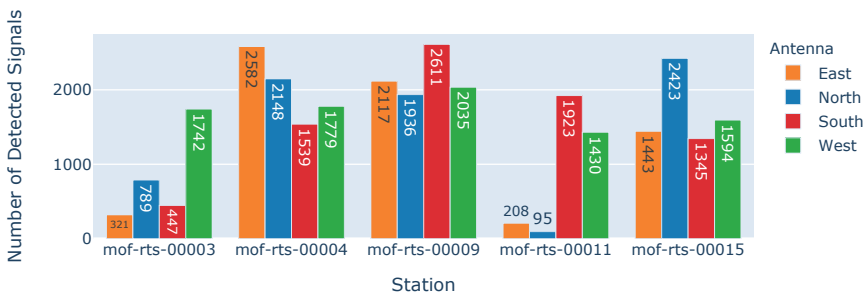


Fig. 9: Detected signals on tRackIT stations in the experimental scenario.

In addition to this quantitative analysis of signal detection, we evaluated the distances between the test tag and the *tRackIT* stations. Figure 10 shows the distance of the tag and stations measured via GPS and the power of the detected signal. While most stations can detect signals at distances of up to 400 meters, stations 4 and 11 detect signals up to 800 meters away. While the correlation of signal strength and measured distance is straightforward, a high variance can be observed from the data and signal strength alone, hence this is not a suitable estimator for distance in the presented experiment. Initially, the overall performance of the two systems appears comparable, especially for signals with high signal strength. While *paur* received 2,728 signals usable for bearing calculations, *tRackIT OS* received 4,438 such signals, an increase of 62.7%, when applying the same -78 dBW threshold. In addition, *tRackIT OS* received 1,108 signals of lower signal strength, which corresponds to an effective increase of 103.3% compared to *paur*.



Fig. 10: Signal power and distance to a receiving station.

4.4 Bearing Calculation

To reach the goal of signal triangulation, signals detected on multiple antennas of a station are used to calculate bearings. We use the method proposed by Gottwald et al. [Go19] to produce comparable results for our bearing calculation. First, the pair of neighboring antennas with the highest and second highest signal strength are selected (s_l, s_r) and the relative gain difference δg is computed using the maximum signal strength difference δm : $g = \frac{s_l - s_r}{\delta m}$. Second, the signal strength is used to calculate the bearing between the antennas following the formula derived from the cosine theorem $\omega = \frac{\pi}{90} \times \arccos(\delta g)$.

Figure 11 shows a histogram of bearing errors in *tRackIT OS* and *paur*. Due to an error on station 4 which was not resolved automatically, signal detection failed on this station in the *paur* experiment run, hence no data is presented in the histogram. While *tRackIT OS* has a mean bearing error of 23.7° and a standard deviation of 30.7° , *paur* not only has a

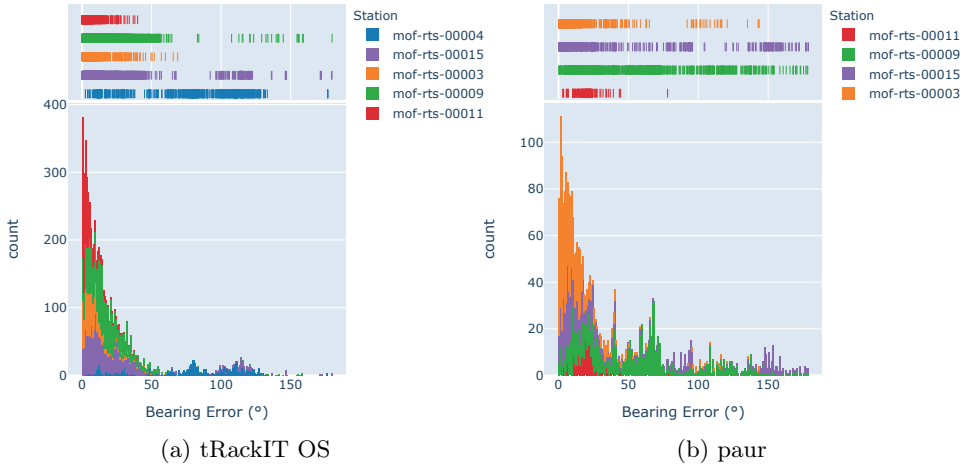


Fig. 11: Histogram of bearing errors.

lower total bearing count, but also results in 38.9° mean bearing error with 42.6° standard deviation. These results indicate that *tRackIT OS* is superior to *pair* that represents the current the state of the art in this field.

4.5 Power Requirements

To operate stations autonomously and to monitor and transmit data, a stable power supply is necessary. To get realistic values for the required power, we measured a *tRackIT* station at the 12 volts input using a Monsoon High Voltage Power Monitor¹⁵.

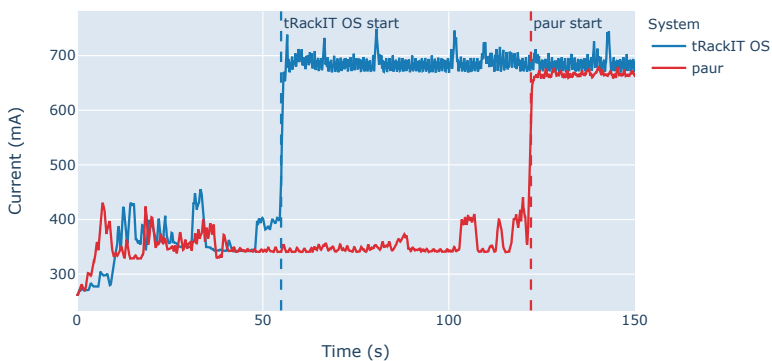


Fig. 12: Power measurements of tRackIT OS and pair in default settings.

¹⁵ Monsoon Solutions Inc. High Voltage Power Monitor: <https://www.msoon.com/online-store/High-Voltage-Power-Monitor-p90002590>

Figure 12 shows the power demands of *paur* and *tRackIT OS*. In contrast to *tRackIT OS*, *paur* does not start signal detection automatically. After all SDRs and signal analysis threads are running, *tRackIT OS* consumes an average of 8.23 W (684 mA), while *paur* consumes an average of 8.03 W (667 mA), which is an overhead of 2.55%. We also carried out experiments with varying sample rates (225 kHz – 300 kHz), but did not observe varying power demands. The systems used in the Marburg Open Forest use 12 V batteries with a capacity of 120 Ah (1440 Wh) of which only 80% should be used to limit wear, which allows a maximum theoretical runtime of 140 hours, or roughly 5.5 days. To allow a continuous operation, a 300 Watts peak solar panel is connected via a solar charger that even works during cloud cover. The presented results show only a slight increase in power consumption of *tRackIT OS* compared to *paur*, i.e., *tRackIT OS* meets the power requirements for continuous operation of the system.

5 Conclusion

We presented *tRackIT OS*, open-source software for reliable VHF radio tracking of small animals in their wildlife habitat. *tRackIT OS* is an operating system distribution for *tRackIT stations* that receive signals emitted by VHF tags mounted on animals. *tRackIT OS* encompasses components for VHF signal processing, system monitoring, configuration management, and user access.

We evaluated and compared *tRackIT OS* against a previous operating system distribution (called *paur*), in an experimental field evaluation carried out in the *Marburg Open Forest*. Our experimental results showed that compared to *paur*, *tRackIT OS* (a) enables reliable VHF signal detection for bearing calculation, (b) increases the number of usable signals by 103.3%, (c) improves the mean bearing calculation error from 38.9° to 23.7°, and (d) introduces only a slight overhead in power consumption of 2.55% or 0.2 W. *tRackIT* has the potential to substantially improve the quality of habitat usage studies and/or environmental assessments in the context of anthropogenic interventions in the environment, while massively reducing the time required for field work.

There are several areas for future work. For example, calculating exact bearings can be challenging, since signals are affected by multiple factors, such as vegetation, topology of the surrounding area, humidity, and rainfall. While bearings can be directly calculated based on a simple model, higher quality can be achieved by using data of multiple stations and further context information, such as a topology model and/or a calibration for the specific area of operation. Furthermore, it is quite challenging to transmit all detected signals under the given bandwidth limitations of the LoRa protocol. A coordinated selection and transmission approach for detected signals should be developed to increase the efficiency of stations connected via LoRa. Finally, the continuous preparation and further processing of the collected data is the next major task in creating a user-friendly and widely applicable animal tracking system for generating ecological knowledge.

6 Acknowledgements

This work is financially supported by the HMWK LOEWE program of the State of Hesse, Germany, within the LOEWE research cluster *Nature 4.0*.

References

- [Al77] Allen, Jonathan: Short term spectral analysis, synthesis, and modification by discrete Fourier transform. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 25(3):235–238, 1977.
- [As19] Ascensão, Fernando; Kindel, Andreas; Teixeira, Fernanda Zimmermann; Barrientos, Rafael; D’Amico, Marcello; Borda-de Água, Luís; Pereira, Henrique M: Beware that the lack of wildlife mortality records can mask a serious impact of linear infrastructures. *Global Ecology and Conservation*, 19:e00661, 2019.
- [Ca10] Cagnacci, Francesca; Boitani, Luigi; Powell, Roger A; Boyce, Mark S: Animal ecology meets GPS-based radiotelemetry: a perfect storm of opportunities and challenges. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1550):2157—2162, 2010.
- [Co99] Cohn, Jeffrey P: Tracking wildlife: high-tech devices help biologists trace the movements of animals through sky and sea. *BioScience*, 49(1):12–17, 1999.
- [Go19] Gottwald, Jannis; Zeidler, Ralf; Friess, Nicolas; Ludwig, Marvin; Reudenbach, Christoph; Nauss, Thomas: Introduction of an automatic and open-source radio-tracking system for small animals. *Methods in Ecology and Evolution*, 10(12):2163–2172, 2019.
- [He20] Hecht, Luke: , Methods for studying wild animals’ causes of death. *Wild Animal Initiative*, <https://www.wildanimalinitiative.org/blog/cause-of-death-2>, 11 2020. Accessed: 2020-11-02.
- [Ho15] Hothorn, Torsten; Müller, Jörg; Held, Leonhard; Möst, Lisa; Mysterud, Atle: Temporal patterns of deer–vehicle collisions consistent with deer activity pattern and density increase but not general accident risk. *Accident Analysis & Prevention*, 81:143–152, 2015.
- [Hö20a] Höchst, Jonas; Baumgärtner, Lars; Kuntke, Franz; Penning, Alvar; Sterz, Artur; Freisleben, Bernd: LoRa-based device-to-device smartphone communication for crisis scenarios. In: *17th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2020)*. Blacksburg, Virginia, USA, May 2020.
- [Hö20b] Höchst, Jonas; Penning, Alvar; Lampe, Patrick; Freisleben, Bernd: PIMOD: A tool for configuring single-board computer operating system images. In: *2020 IEEE Global Humanitarian Technology Conference (GHTC 2020)*. Seattle, USA, pp. 1–8, October 2020.
- [Ka11] Kays, Roland; Tilak, Sameer; Crofoot, Margaret; Fountain, Tony; Obando, Daniel; Ortega, Alejandro; Kuemmeth, Franz; Mandel, Jamie; Swenson, George; Lambert, Thomas et al.: Tracking animal location and activity with an automated radio telemetry system in a tropical rainforest. *The Computer Journal*, 54(12):1931–1948, 2011.

- [KA20] Katzner, Todd E; Arletaz, Raphaël: Evaluating contributions of recent tracking-based animal movement ecology to conservation management. *Frontiers in Ecology and Evolution*, 7:519, 2020.
- [Le19] Lees, Daniel; Schmidt, Tom; Sherman, Craig DH; Maguire, Grainne S; Dann, Peter; Ehmke, Glenn; Weston, Michael A: An assessment of radio telemetry for monitoring shorebird chick survival and causes of mortality. *Wildlife Research*, 46(7):622–627, 2019.
- [Li17] Light, Roger A: Mosquitto: server and client implementation of the MQTT protocol. *Journal of Open Source Software*, 2(13):265, 2017.
- [LN18] Lenske, Ariel K; Nocera, Joseph J: Field test of an automated radio-telemetry system: tracking local space use of aerial insectivores. *Journal of Field Ornithology*, 89(2):173–187, 2018.
- [Md17] Mdhaffar, Afef; Chaari, Tarak; Larbi, Kaouthar; Jmaiel, Mohamed; Freisleben, Bernd: IoT-based health monitoring via LoRaWAN. In (Karadzinov, Ljupco; Cvetkovski, Goga; Latkoski, Pero, eds): *IEEE EUROCON 2017 -17th International Conference on Smart Technologies*, Ohrid, Macedonia, July 6-8, 2017. IEEE, pp. 519–524, 2017.
- [Mo10] Montgomery, Robert A; Roloff, Gary J; Hoef, Jay M Ver; Millsbaugh, Joshua J: Can we accurately characterize wildlife resource use when telemetry data are imprecise? *The Journal of Wildlife Management*, 74(8):1917–1925, 2010.
- [Pe17] Petäjäjärvi, Juha; Mikhaylov, Konstantin; Pettissalo, Marko; Janhunen, Janne; Iinatti, Jari: Performance of a low-power wide-area network based on LoRa technology: Doppler robustness, scalability, and coverage. *International Journal of Distributed Sensor Networks*, 13(3):1550147717699412, 2017.
- [Ri20] Ripperger, Simon P; Carter, Gerald G; Page, Rachel A; Duda, Niklas; Koelpin, Alexander; Weigel, Robert; Hartmann, Markus; Nowak, Thorsten; Thielecke, Jörn; Schadhauer, Michael et al.: Thinking small: next-generation sensor networks close the size gap in vertebrate biologging. *PLoS Biology*, 18(4):e3000655, 2020.
- [Sa09] Sawyer, Hall; Kauffman, Matthew J; Nielson, Ryan M; Horne, Jon S: Identifying and prioritizing ungulate migration routes for landscape-level conservation. *Ecological Applications*, 19(8):2016–2025, 2009.
- [Ta17] Taylor, Philip; Crewe, Tara; Mackenzie, Stuart; Lepage, Denis; Aubry, Yves; Crysler, Zoe; Finney, George; Francis, Charles; Guglielmo, Christopher; Hamilton, Diana et al.: The Motus Wildlife Tracking System: a collaborative research network to enhance the understanding of wildlife movement. *Avian Conservation and Ecology*, 12(1), 2017.
- [Wa18] Walton, Zea; Samelius, Gustaf; Odden, Morten; Willebrand, Tomas: Long-distance dispersal in red foxes *Vulpes vulpes* revealed by GPS tracking. *European Journal of Wildlife Research*, 64(6):1–6, 2018.
- [We16] Weiser, Adi Weller; Orchan, Yotam; Nathan, Ran; Charter, Motti; Weiss, Anthony J; Toledo, Sivan: Characterizing the accuracy of a self-synchronized reverse-GPS wildlife localization system. In: 2016 15th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN). IEEE, pp. 1–12, 2016.
- [Wy18] Wyckoff, Teal B; Sawyer, Hall; Albeke, Shannon E; Garman, Steven L; Kauffman, Matthew J: Evaluating the influence of energy and residential development on the migratory behavior of mule deer. *Ecosphere*, 9(2):e02113, 2018.

Deep Learning Pipeline for Automated Visual Moth Monitoring: Insect Localization and Species Classification

Dimitri Korsch¹, Paul Bodesheim², Joachim Denzler^{3 4 5}

Abstract: Biodiversity monitoring is crucial for tracking and counteracting adverse trends in population fluctuations. However, automatic recognition systems are rarely applied so far, and experts evaluate the generated data masses manually. Especially the support of deep learning methods for visual monitoring is not yet established in biodiversity research, compared to other areas like advertising or entertainment. In this paper, we present a deep learning pipeline for analyzing images captured by a moth scanner, an automated visual monitoring system of moth species developed within the AMMOD project. We first localize individuals with a moth detector and afterward determine the species of detected insects with a classifier. Our detector achieves up to 99.01 % mean average precision and our classifier distinguishes 200 moth species with an accuracy of 93.13 % on image cutouts depicting single insects. Combining both in our pipeline improves the accuracy for species identification in images of the moth scanner from 79.62 % to 88.05 %.

Keywords: Biodiversity Monitoring; Deep Learning; Convolutional Neural Networks; Insect Detection; Species Classification; Unsupervised Part Estimation

1 Introduction

The discussions and concerns about environmental changes nowadays are both ubiquitous and necessary. We live in times in which ecosystems change drastically in a short time, and we, as humans, play a significant role in this development. One of these negative developments is the dramatic loss of insects [Ha17, Wa21]. One step towards a better understanding of insect die-off is monitoring species populations, which is time-consuming and often requires experts in that field. With about one million named species [St18], insects represent the vast majority of animals on our planet. Hence, it is also clear why manual counting of individuals for abundance estimations is not realistic. Though current developments in big data analysis and computer vision improved a lot, these technologies are not established in insect monitoring as they are, for example, in marketing or entertainment.

¹ Computer Vision Group, Friedrich Schiller University Jena, 07737 Jena, Germany; dimitri.korsch@uni-jena.de

² Computer Vision Group, Friedrich Schiller University Jena, 07737 Jena, Germany; paul.bodesheim@uni-jena.de

³ Computer Vision Group, Friedrich Schiller University Jena, 07737 Jena, Germany; joachim.denzler@uni-jena.de

⁴ German Aerospace Center (DLR), Institute for Data Science, Mälzerstraße 3, 07745 Jena, Germany

⁵ Michael Stifel Center Jena for Data-Driven and Simulation Science, Ernst-Abbe-Platz 2, 07743 Jena, Germany

Besides others, there are monitoring systems to observe insects [JFR14, Sv20, Bj21], great apes [Fr16, Br17, Kä18, YMB19, SB20], elephants [KBD18, KD19], or sharks [HB17]. Though such monitoring systems are already camera-assisted, a vast amount of data needs to be evaluated. Unfortunately, the daily work of many ecologists nowadays is still the manual inspection of hundreds or thousands of images, which is exhausting and time-consuming.

As a part of the AMMOD project⁶, we aim at automated monitoring of species in our immediate vicinity assisted by a computer vision system. In this paper, we cover the task of categorizing moth species (subset of *Lepidoptera*) by a non-invasive approach. Within the project, a so-called *moth scanner* is developed, which consists of an illuminated planar surface and an automated camera system. During the nighttime, special light sources illuminate the planar area to attract different moth species in the surrounding area. The automated camera system captures the attracted individuals that land on the illuminated surface. Finally, our task is to detect and classify the individuals in the taken images. With the detection and classification results, we assist the ecologists in analyzing the insect population trends.

For automatically analyzing the images, we propose a prototype for a deep learning pipeline consisting of two steps: (1) localization of individuals via moth detection and (2) species identification by classification. For the detection, we use a well-established detection model capable of identifying multiple objects in an image, namely the single-shot MultiBox detector (SSD) [Li16]. The mean average precision (mAP) of our moth detector is 88.88 % and 99.01 % for intersection over union (IoU) values above 0.75 and 0.5, respectively.

The subsequent classification of 200 common moth species in Central Europe is performed with the help of a convolutional neural network (CNN). In our experiments, we show the benefits of different design decisions for a classifier trained on cropped images, namely images depicting a single insect. The selection of a fine-tuning strategy, the pre-training dataset, and the extension of the classifier with an unsupervised part estimator improve the classification accuracy from 63.28 % to 93.13 %.

Finally, we show that the classification accuracy of our proposed pipeline improves with a preceding moth detector on uncropped images. These are the images captured by the moth scanner, where we cannot ensure that only a single insect has been photographed. Furthermore, in these images, the insects allocate only a small portion of the entire area. Hence, preceding a moth classifier with a moth detector, we can improve the classification performance from 79.62 % to 88.05 %.

⁶ AMMOD = Automated Multisensor Station for Monitoring of Biodiversity (<https://ammod.de/>)

2 Related Work

2.1 Insect Monitoring

In general, a commonly used method for monitoring insects is the usage of light traps. Jonason et al. [JFR14] presented a survey on the influence of weather, time of the year, and the type of the light source on the richness and abundance of species. While the authors identified the moth species manually, some of the first automated species identification systems were presented by Watson et al. [WOK04], Mayo and Watson [MW07], and Batista et al. [BCK10]. All of these works used the same dataset, namely 35 species with 20 individuals per species. Using support vector machines (SVMs) and nearest neighbor classifiers, they report an accuracy of up to 85 % [MW07] with leave-one-out cross-validation. Ding and Taylor [DT16] presented automated detection and classification of insect pests. They used a sliding window approach coupled with a CNN model. The CNN, followed by a non-maximum suppression as post-processing, performed a binary classification to identify a *codling moth* in the windows. They achieved an area under the precision-recall curve of 0.93. In contrast to these works, we perform the classification of much more classes, namely 200 moth species.

The works of Chang et al. [Ch17] and Xia et al. [Xi18] tackled more challenging classification tasks. Using images from the Internet, they classified 636 and 24 species, respectively. Chang et al. achieved for the 450 butterflies and 186 moths species an accuracy of 71.5 % with a ResNet-18 [He16] architecture. Xia et al. performed a joint detection and classification of individual insects and achieved with their variant of a VGG-19 CNN a mean average precision (mAP) of 89.22 %. First, we gather images in a more controlled environment. As a result, the background is more homogeneous, and the moths are photographed from above in a resting position. It is worth investigating how far the Internet images that do not represent our desired setup domain may enhance the classification performance. Anyway, this is out of the scope of this paper. Furthermore, unlike Xia et al., we aim to separate the detection and classification tasks since they will be performed on different physical devices in our setting.

Zhong et al. [Zh18] and Bjerger et al. [Bj21] presented detection and classification pipelines deployed on embedded systems, namely on Raspberry Pi variants. While Zhong et al. used the YOLO framework [Re16] for moth detection, Bjerger et al. presented a detection-by-thresholding approach. Zhong et al. achieved a classification accuracy of 90.2 % for six species with an SVM and shallow features (texture, shape, color, and HOG features). Bjerger et al. presented their own CNN architecture and report an F1-score of 93.00 % for the classification of nine classes. The authors perform additional counting and tracking of the insects, which is not part of this work. Furthermore, we outperform the classification results presented by Bjerger et al. in our experiments (Sect. 4.2).

2.2 Object Detection

Pre-CNN image-based object detection was dominated by Deformable Part Model (DPM) [FMR08] and Selective Search [Ui13]. The first approach uses a sliding window approach, whereas the latter uses region proposal selection as an object detection strategy. After the rise of CNNs, region proposal methods are dominating the object detection research field. One of the first was the R-CNN [Gi14] that combined selective search region proposals with a CNN-based classification of these regions. Many improvement and adaptations based on R-CNN were developed: SPPNet [He15], Fast R-CNN [Gi15], MultiBox [Er14], or Faster R-CNN [Re15]. Some of them improved the classification of the region proposals in quality and computation time [He15, Gi15]. Others improved the quality of the region proposals directly [Er14, Re15], especially with an integration of a region proposal CNN. Some of the methods skip the proposal step and predict bounding boxes directly with the confidences for multiple categories. Most popular examples are YOLO [Re16], OverFeat [Se13], and SSD [Li16]. While OverFeat implements a deep version of the sliding window approach, YOLO uses CNN features to predict bounding boxes and categories. SSD extracts features from multiple feature maps from multiple stages in the CNN and predicts bounding boxes based on a set of prior locations.

All of them have their advantages and disadvantages. Meanwhile, there are also dozens of adaptations and improvements to these methods. Nevertheless, in our work, we use the single-shot MultiBox detector (SSD) since it allows an exchange of the underlying backbone network and yields one of the best results on standard object detection benchmarks like Pascal VOC [Ev07, Ev12] and MS COCO [Li14].

2.3 Fine-grained Classification

Fine-grained classification is a special classification task, where the categories, which need to be classified, originate from the same object domain (e.g., bird species [Wa11], car models [Kr13], moth species [Ro15], or elephant individuals [KD19]). The challenge is now to distinguish closely related classes that differ only in subtle features. In the age of CNNs, it is common to use the data and let the network figure out what are relevant visual features that distinguish a class from the others. This kind of approach utilizes the input image as it is and performs either smart pre-training strategies [Cu18, Kr16] or advanced feature aggregation techniques [LRM15, Si18]. On the other hand, there are the part- or attention-based approaches [GLY19, HPZ19, Zh19b] that extract relevant regions already at the pixel level and use cropped image regions as additional features for the classification. Both classification strategies have their advantages and drawbacks. We use an unsupervised approach for part estimation proposed by Korsch et al. [KBD19] within our pipeline.

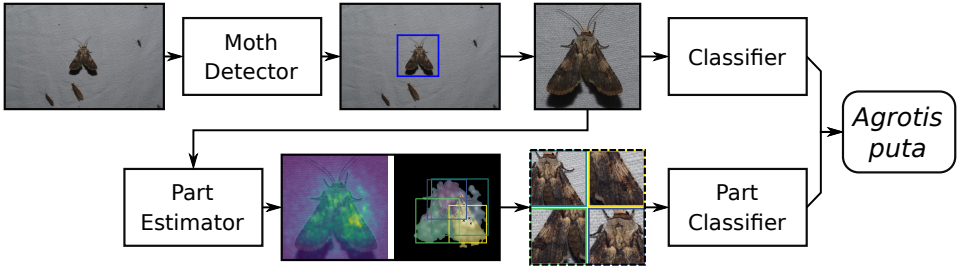


Fig. 1: Our deep learning pipeline for automatically analyzing images of the moth scanner: in the first step we use a moth detector to detect an individual in an image. On the one hand, we use this cropped image for the first species prediction. On the other hand, we estimate additional information in form of parts and perform a second part-based prediction. Finally, both predictions are combined to obtain the final species classification.

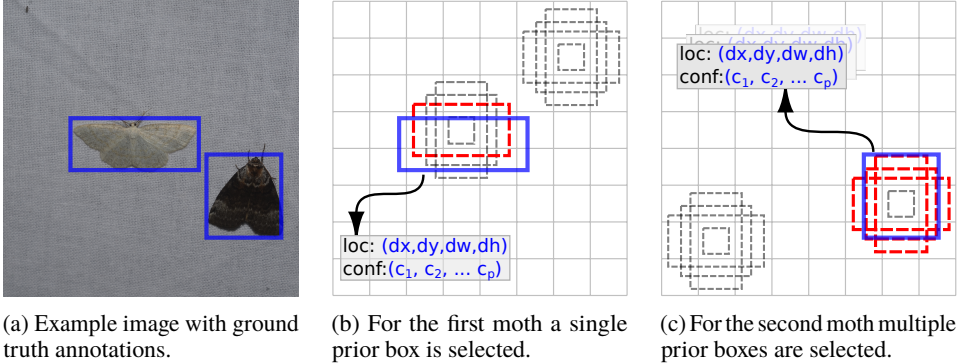
3 Methods

In this paper, we introduce an automated pipeline for moth species detection and classification. As visualized in Figure 1, the moth detector identifies bounding boxes around the insects given an input image. Afterward, the image patches identified by the detected bounding boxes are fed into a CNN classifier. The pipeline performs the classification either only on the input image or extracts informative regions, called parts, which it uses as additional information. In our experiments, we show that this additional information improves the classification performance (see Sect. 4.3).

In the following, we introduce the two stages of the pipeline: (1) moth detection based on the single-shot detector (Sect. 3.1), and (2) part-based classification with the help of classification-specific parts (Sect. 3.2).

3.1 Single-shot Detector

As already mentioned in Sect. 2.2, the main idea of the single-shot MultiBox detector (SSD) proposed by Liu et al. [Li16] is to utilize feature maps from multiple intermediate stages of the backbone CNN to predict location offsets and class confidences for a set of prior locations. More precisely, given a feature map with $F \in \mathbb{R}^{N \times M \times P}$ with P channels, K prior bounding boxes with different scales and aspect ratios are defined for each of the $N \cdot M$ locations. The feature map is transformed by a 3×3 convolution with $(C + 4) \cdot K$ output channels. This results for each of the K prior boxes in C per-class scores and four offset values $\Delta = \{dx, dy, dw, dh\}$. The offsets dx and dy describe the positional offset to the center of the prior box. The change of the width and the height of a prior box is modeled by dw and dh .



(a) Example image with ground truth annotations.

(b) For the first moth a single prior box is selected.

(c) For the second moth multiple prior boxes are selected.

Fig. 2: Example of SSD prior boxes (*dashed gray*) for an 8×8 feature map. The prior boxes have different scales and aspect ratios, which allows for detecting objects of various sizes and orientations. Different ground truth bounding boxes (*solid blue*) may be matched to a single (Figure 2b) or multiple (Figure 2c) prior boxes. For each selected prior box (*dashed red*) the offset values and the class scores are estimated. (Similar to Figure 1 in [Li16].)

During training, a prior box is selected as positive when there was a ground truth bounding box with an IoU score higher than 0.5. As visualized in Figure 2, prior boxes at different locations are assigned to different objects. Additionally, for a single object multiple prior boxes may be selected. The training objective for estimated location offsets Δ and class scores c given ground truth bounding boxes g is defined as a sum of the confidence loss and the localization loss:

$$L(\Delta, c, g) = \frac{1}{\mathcal{K}} (L_{conf}(c) + L_{loc}(\Delta, g)) \quad (1)$$

with \mathcal{K} being the number of matched prior boxes, and if no boxes are matched, the loss is set to 0. The class scores are unnormalized log-likelihoods of a class identified in a certain location. In our case, the detector distinguishes only the general *moth* class from the background. For more details about the loss functions, we refer to the original paper of Liu et al. [Li16].

3.2 Part-based Classification

Nowadays, neural networks like CNNs yield the best results in classification by extracting high-level features from the input image in the form of a high-dimensional feature vector (e.g., $D = 2048$ in case of InceptionV3). In the context of a fine-grained recognition task, the classifier has to focus on a specific feature dimension to distinguish a class from the others. Therefore, we first estimate the most informative features for the current classification task. It is realized by utilizing a linear classifier with a sparsity-inducing

L1-regularization. An optimization with L1-regularization forces the classifier’s decisions to perform the classification on a small subset of feature dimensions. This kind of implicit feature selection is classification-specific. Furthermore, it allows identifying for each class the feature dimensions that best distinguish this class from all other classes by selecting dimensions with classifier weights above a certain threshold.

Informative Image Regions: We utilize gradient maps [SVZ13] to estimate the most informative pixels in the image, identified by large gradients. As described previously, we restrict the computation of the gradients only to the feature dimensions used by the L1-regularized classifier. Thus, we incorporate the initial classification in the estimation of the part regions. Like Simonyan et al. [SVZ13] and Simon et al. [SR15], we use back-propagation through the CNN to identify the regions of interest for each selected feature dimension. We compute a saliency map $\vec{M}(\vec{I})$ for an image \vec{I} based on the feature dimension subset $\mathfrak{D} \subset \{1, \dots, D\}$ as follows:

$$M_{x,y}(\vec{I}) = \frac{1}{|\mathfrak{D}|} \sum_{d \in \mathfrak{D}} \left| \frac{\partial}{\partial I_{x,y}} f^{(d)}(\vec{I}) \right| . \quad (2)$$

Part Estimation: Next, we normalize the values of the saliency map to the range $[0 \dots 1]$, and discard regions of low saliency by setting values beneath the mean saliency value to 0. We use the resulting sparse saliency map to estimate the spatial extent of coherent regions. Like Zhang et al. [Zh19a], we achieve this by k -means clustering of pixel coordinates (x, y) and the saliencies $M_{x,y}$ (Eq. 2). Additionally, we also consider the RGB values at the corresponding positions in the input image. The clusters are initialized with k peaks computed by non-maximum suppression, identifying locations with the largest saliencies. Consequently, the number of peaks determines the number of parts to detect. Finally, it is straightforward to identify a bounding box around each estimated cluster, and the resulting bounding boxes serve as parts for the following classification.

Extraction and Aggregation of Part Features: In the final step, we extract image patches with the help of the estimated bounding boxes and treat them as regular images. The neural network should extract different features from these image patches than from the original image because the level of detail varies between these types of input. Therefore, we process the part images by the same CNN architecture as the original image but with a separate set of weights. Afterward, for every part image, the CNN extracts a feature vector, resulting in a set of part features for every single image. There are different ways to aggregate these features to a single feature vector and perform the classification. We have chosen to average over the part features, which results in a single feature vector with the same dimension as for the original image. This aggregation strategy yielded better results in our experiments than, for example, concatenation of part features. Finally, classification is performed based on the global feature and the aggregated part feature. For joint optimization of both CNNs, we average the cross-entropy losses of the global prediction \vec{p} and part prediction \vec{q} . It equals

to computing the geometrical mean of normalized class probabilities and enforces both classifiers to be certain about the correct class:

$$L_{final}(\{\vec{p}, \vec{q}\}, y) = \frac{1}{2} (L(\vec{p}, y) + L(\vec{q}, y)) \quad (3)$$

$$= -\frac{1}{2} \left(\sum_{i=1}^C y_i \cdot \log(p_i) + \sum_{i=1}^C y_i \cdot \log(q_i) \right) \quad (4)$$

$$= -\sum_{i=1}^C y_i \cdot \log(\sqrt{p_i \cdot q_i}) \quad (5)$$

4 Experiments

In the following, we evaluate the two parts of our moth scanner pipeline. We perform each experiment ten times and provide in Sections 4.2, and 4.3 the mean and the standard deviation of the evaluation metrics across the different runs. We fine-tune all CNNs for 60 epochs with the RMSProp [TH12] optimizer and L2-regularization with a weight decay of 5×10^{-4} . The starting learning rate of 1×10^{-4} is reduced by 0.1 after 20 and 40 epochs. Furthermore, we utilize standard image augmentation methods: random cropping, random horizontal and vertical flipping, and color jittering (contrast, brightness, and saturation). In the case of the classification, we further utilize label smoothing regularization [Sz16] with a smoothing factor of 0.1.

First, we evaluate the performance of the moth detector presented in Sect. 3.1. In Table 2, we report the mean average precision (mAP) as the evaluation metric for the detections. The precision is based on predictions of detected objects, where the intersection over union (IoU) of predicted and ground truth bounding boxes is above a certain threshold. The IoU describes how well two bounding boxes match by computing the ratio between the intersection and the union of the areas of the two bounding boxes. The two typical evaluation metrics used in one of the most common object detection benchmarks [Li14] are $mAP@0.5$ and $mAP@0.75$, with IoU thresholds of 0.5 and 0.75, respectively. We use a standard version of the detection network with an input size of 300×300 and the VGG16 [SZ14] architecture as a backbone that is pre-trained on ImageNet [Ru15]. All additional layers included for the detection task and not initially present in the VGG architecture are initialized randomly.

Second, we evaluate the classification performance. Since the datasets we use have a balanced sample distribution across the classes, we use accuracy as a primary evaluation metric. To be independent of the underlying detector, we use the ground truth bounding boxes and perform the classification on the cropped image patches. Further, we extract additional parts, as described in Sect. 3.2, and combine the predictions on these parts with the predictions on the entire image. For the classification, we use the InceptionV3 CNN



(a) The first two rows show two different moths species, whereas the third row shows images with more than one insect. These examples illustrate the versatility in the appearance of the moths. (b) Cropped images of the eight MCC classes. Bjerger et al. [Bj21] used an eight-megapixel web camera to capture the images. Hence, the details of the species are barely visible.

Fig. 3: Example images from the *EU-Moths* (a) and *MCC* (b) datasets.

architecture [Sz16]. Here, we also decided on the default input size of 299×299 for both the images and parts. Furthermore, we investigate the effect of two different pre-training methods. As the typical choice, we use ImageNet [Ru15] pre-training since most of the deep learning frameworks implementing different CNN architectures also provide weights for these architectures pre-trained on the ImageNet dataset. Additionally, we use pre-training on the iNaturalist [Va18] dataset provided by Cui et al. [Cu18]. Data used in this kind of pre-training is more related to the domain of animals, in our case to the domain of insects, which can also be seen in the improvement of the classification accuracy in Table 3.

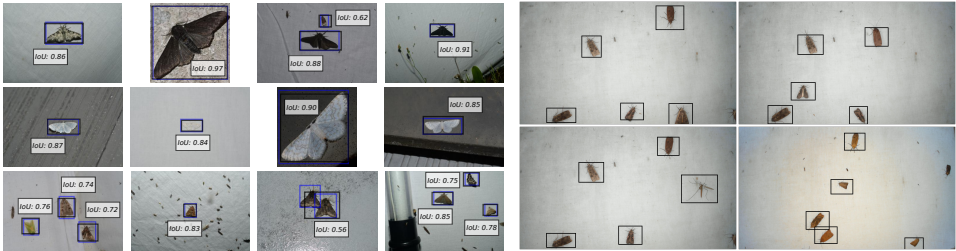
Finally, we evaluate the proposed pipeline as a whole: given an uncropped image, like in Figure 3a, we first apply the moth detector and then perform the classification on the resulting image patches. We compare this setup to a classifier that performs the classification on the initial uncropped images.

We further evaluate the classifier on the dataset provided by Bjerger et al. [Bj21]. Unfortunately, the authors report only F1-scores of their classification and do not provide their training-test split. Nevertheless, we performed five-fold cross-validation with the same train-test ratios as in the original paper. For each of the folds, we repeated the training ten times, like in the previous experiments.

4.1 Datasets

Moth Classification and Counting (MCC) Dataset⁷: Created by Bjerger et al. [Bj21], the subset for the classification consists of eight moth species with 250 images for each species, resulting in a dataset of 2000 images. The authors used an 80:20 training-test split of the data but did not provide their specific split. Additionally, there are two more classes: a background class and a class for a wasp species. We ignore these classes and perform

⁷ <https://github.com/kimbjerger/MCC-trap>



(a) Our manual ground truth annotations are visualized in *blue*. Text boxes contain the IoU of each detection with the ground truth. (b) The MCC dataset does not provide any ground truth annotations. Hence, only the detections are visualized.

Fig. 4: Visualized detection results (*black* bounding boxes) on the *EU-Moths* (a) and *MCC* (b) datasets.

the training and the evaluation only on the eight moths species. An individual from every class is shown in Figure 3b. Compared to the EU-Moths dataset, the images are of lower quality since the authors captured them with an eight-megapixel web camera connected to a Raspberry Pi 4.

European Moths (EU-Moths) Dataset⁸: From roughly 4700 moth species present in Central Europe⁹, this dataset consists of 200 species most common in this region. Each of the species is represented by approximately 11 images. We considered a random but balanced training-test split in eight training and three test images per species, resulting in roughly 1600 training and 600 test images in total. To evaluate the detector, we manually annotated the bounding boxes around the insects. Some examples of the images are shown in Figure 3a. The insects are photographed manually and mainly on a relatively homogeneous background. About 92 % of the images contain only a single individual like it is shown in the first two rows of Figure 3a. The last row of the same image depicts images with more than one insect of interest. This scenario may happen in the final moth scanner installation, and it is crucial to test how the detector performs in this case.

This dataset yields different challenges for the detector and the classifier. On the one hand, the detector should be able to detect insects of different sizes. Furthermore, we require a detector with MultiBox support. Both of these properties are fulfilled by the SSD. On the other hand, designing a classifier that can predict species from these raw images is difficult. As proposed in this paper, we decided to use a moth detector to locate single insects and classify these separately. Hence, the classifier is trained on images cropped to the bounding boxes identifying a single individual.

	F1-SCORE	ACCURACY	
		IMAGENET	INATURALIST
BJERGE ET AL. [Bj21]	93.00 %	–	–
OURS	99.69 % (± 0.34 %)	99.41 % (± 0.79 %)	99.55 % (± 0.40 %)

Tab. 1: Comparison of the classification results on cropped images (see Figure 3b) of the MCC dataset provided by Bjerger et al. [Bj21]. Besides the F1-Score that was reported by the authors in their work, we report additionally the accuracy for our trained classifiers.

4.2 Results on the MCC Dataset

Detection: Unfortunately, the authors do not provide any bounding box annotations. Hence, we were not able to evaluate the performance of our moth detector on this dataset. Nevertheless, we provide a qualitative evaluation on some of the images in Figure 4b.

Classification: In Table 1, we compare our classification method with the one proposed by Bjerger et al. [Bj21] on the MCC dataset. One can see that our classifier achieves near-perfect accuracies and F1-scores. We assume the reason for these results is in the composition of the dataset. Since Bjerger et al. do not provide any training-test split, we have used a random split with the same ratio (80:20) as the authors in their work. Nevertheless, one can see in Figure 4b some of the moth individuals do not move in different images captured in short intervals. As a result, extracting image crops from these images would result in near-identical images in different splits after the random splitting. Consequently, one would train and test the classifier on nearly the same data. Nevertheless, to be comparable to the results of Bjerger et al., we chose the same splitting strategy, even though it may not represent the correct evaluation of the model.

4.3 Results on the EU-Moths Dataset

Detection: We split this experiment into two parts: (1) we evaluated the detection performance on the entire dataset, and (2) we split the dataset into two subsets of distinct classes. The first part evaluates the standard performance of the detector. The second part of the experiment evaluates how good the detector performs on classes not seen during the training. This scenario is essential since not all species may be available at training time in the real-world setup. One could train the detector on a dataset captured at one location and deploy it at another one. Furthermore, a detector able to localize moth species not seen at

⁸ https://www.inf-cv.uni-jena.de/eu_moths_dataset

⁹ <http://lepiforum.org/> (accessed on 6th July, 2021)

	mAP@0.75	mAP@0.5
EVALUATED ON ENTIRE DATASET	TRAINED ON ENTIRE DATASET	
	88.88 % (± 0.77 %)	99.01 % (± 0.09 %)
EVALUATED ON SUBSET 1	TRAINED ON SUBSET 1	
	87.38 % (± 1.65 %)	98.53 % (± 0.13 %)
SUBSET 2	78.83 % (± 1.11 %)	99.19 % (± 0.28 %)
EVALUATED ON SUBSET 1	TRAINED ON SUBSET 2	
	85.88 % (± 1.04 %)	99.70 % (± 0.10 %)
SUBSET 2	82.45 % (± 1.23 %)	98.04 % (± 0.29 %)

Tab. 2: Detection results on the EU-Moths dataset. First row contains the evaluation of the detectors trained on the entire dataset (200 classes). The lower part of the table shows the capability of the detector to localize unseen species. For that, we split the dataset in two parts with distinct classes (SUBSET 1: classes 1 to 100, and SUBSET 2: classes 101 to 200) and perform cross-subset evaluations.

training time is beneficial for novelty detection, active learning, and incremental learning algorithms [Br17, BKD20].

Table 2 shows the detection results for both experiments. As previously mentioned, we report the mean average precision for IoU thresholds 0.75 and 0.5 (mAP@0.75 and mAP@0.5). The first row shows the results for the first experiment. The detector seems to be quite precise if we consider the challenges of the dataset. The lower part of Table 2 further shows the cross-subset results. Here we can see that the mAP@0.75 performance drops compared to the previous experiment, and the standard deviation increases. Both are explainable because, for the second experiment, we used only half of the classes for the training. Furthermore, mAP@0.5 performance remains comparable to the first experiment, which shows the moth detector’s reliability for unseen classes.

Additionally, Figure 4 depicts qualitative results of the detector. In Figure 4a, we estimated the bounding boxes (in *black*) for some of the images of the EU-Moths dataset. We visualized the ground truth bounding boxes (in *blue*) and the resulting IoU between the prediction and the ground truth. The detector’s most significant challenge seems to be insects located too close to each other (second last example in the final row).

Classification: Table 3 shows the results of the classification. We compare different training and pre-training methods and whether the additional information in the form of parts benefits the classification accuracy.

First, one can see that fine-tuning the entire CNN instead of using it only as feature extractor results in an improvement of the recognition rate by roughly 26 % and 4 % for CNNs pre-trained on ImageNet [Ru15] and iNaturalist 2017 [Va18] datasets, respectively.

	FINE-TUNING	ACCURACY	
		IMAGENET	INATURALIST
NO PARTS	<i>only FC layer</i>	63.28 % (± 0.45 %)	86.60 % (± 0.42 %)
	<i>entire CNN</i>	89.46 % (± 0.88 %)	90.54 % (± 1.10 %)
WITH PARTS	<i>only FC layer</i>	71.82 % (± 0.35 %)	87.96 % (± 0.38 %)
	<i>entire CNN</i>	91.50 % (± 0.61 %)	93.13 % (± 0.76 %)

Tab. 3: Classification results on cropped images of the EU-Moths dataset. The results show the effects of the different fine-tuning strategies, the two pre-training datasets, and the usage of additional information in the form of parts.

COMPOSITION	ACCURACY
CLASSIFIER ONLY	79.62 % (± 1.10 %)
DETECTOR + CLASSIFIER	88.05 % (± 0.58 %)

Tab. 4: Classification results on uncropped images as shown in Figure 3a. The effect of a preceding detector on the classification accuracy is clearly visible.

Though training the entire CNN results in longer training times and is computationally more expensive, the improvements are visible.

Second, the choice of the pre-training dataset is also crucial. Replacing the typical CNN weights provided by almost every deep learning framework pre-trained on the ImageNet dataset with the ones proposed by Cui et al. [Cu18] leads to a further improvement. The later pre-training increases the accuracy by approximately 1 % if the entire CNN is trained. It also yields remarkable benefits if choosing computationally cheaper training of only the final classification layer, namely an improvement of 20 %.

Finally, employing additional information in the form of parts improves the classification accuracies by approximately 2 to 2.6 % depending on the chosen pre-training. We achieved the best results with the part-based setup: 91.50 % and 93.13 % with ImageNet and iNaturalist pre-training, respectively.

Entire Pipeline: In this experiment, we evaluate the entire pipeline as presented in Sect. 3. For this purpose, we couple every trained detector with every trained classifier and observe the resulting classification performance. This way, we report the mean accuracy of 100 detector-classifier combinations in the last row of Table 4. As a baseline method, we trained ten CNN classifiers on the original uncropped images. One can see that the preceding detector improves the classification accuracy by approximately 9 %.

5 Conclusions

In this paper, we presented an automated detection and classification pipeline for 200 moth species. We plan to deploy this pipeline in the visual monitoring system of the AMMOD project, the so-called moth scanner, which will help ecologists observe the population trends of the insects. Since light sources easily attract moths, the moth scanner consists of an illuminated white surface and a camera that automatically captures images of insects resting on this surface. We first localized the moths with a single-shot MultiBox detector (SSD) in the recorded images and then classified the resulting detections using a CNN classifier. We also showed the effect of different training configurations on the final classification accuracy: the choice of fine-tuning strategy, the selection of the pre-training dataset, and the extension of the classification with an unsupervised part estimator. In our experiments, each part of the pipeline achieved promising results: a detection rate of up to 99.01 % (mAP@50) and classification accuracy on images depicting a single insect of up to 93.13 %. Finally, we evaluated both parts of the pipeline together and improved the classification accuracy on original images captured by the moth scanner from 79.62 % to 88.05 % compared to a setup without a preceding moth detector.

Acknowledgments

This work has been funded by the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung, BMBF, Deutschland) via the project “Development of an Automated Multisensor Station for Monitoring of Biodiversity (AMMOD) - Subproject 5: Automated Visual Monitoring and Analysis” (FKZ: 01LC1903E).

References

- [BCK10] Batista, Gustavo EAPA; Campana, Bilson; Keogh, Eamonn: Classification of live moths combining texture, color and shape primitives. In: 2010 Ninth International Conference on Machine Learning and Applications. IEEE, pp. 903–906, 2010.
- [Bj21] Bjerge, Kim; Nielsen, Jakob Bonde; Sepstrup, Martin Videbaek; Helsing-Nielsen, Fleming; Høye, Toke Thomas: An automated light trap to monitor moths (Lepidoptera) using computer vision-based tracking and deep learning. *Sensors*, 21(2):343, 2021.
- [BKD20] Brust, Clemens-Alexander; Käding, Christoph; Denzler, Joachim: Active and Incremental Learning with Weak Supervision. *Künstliche Intelligenz (KI)*, 2020.
- [Br17] Brust, Clemens-Alexander; Burghardt, Tilo; Groenenberg, Milou; Käding, Christoph; Kühl, Hjalmar; Manguette, Marie; Denzler, Joachim: Towards Automated Visual Monitoring of Individual Gorillas in the Wild. In: *ICCV Workshop on Visual Wildlife Monitoring (ICCV-WS)*. pp. 2820–2830, 2017.







- [Ch17] Chang, Qi; Qu, Hui; Wu, Pengxiang; Yi, Jingru: Fine-Grained Butterfly and Moth Classification Using Deep Convolutional Neural Networks. Machine Learning course project report, submitted to the Department of Computer Science, Rutgers University, 2017.
- [Cu18] Cui, Yin; Song, Yang; Sun, Chen; Howard, Andrew; Belongie, Serge: Large Scale Fine-Grained Categorization and Domain-Specific Transfer Learning. In: Proceedings of CVPR. 6 2018.
- [DT16] Ding, Weiguang; Taylor, Graham: Automatic moth detection from trap images for pest management. *Computers and Electronics in Agriculture*, 123:17–28, 2016.
- [Er14] Erhan, Dumitru; Szegedy, Christian; Toshev, Alexander; Anguelov, Dragomir: Scalable object detection using deep neural networks. In: Proceedings of CVPR. pp. 2147–2154, 2014.
- [Ev07] Everingham, Mark; Van Gool, Luc; Williams, Christopher K.I.; Winn, John; Zisserman, Andrew: The PASCAL visual object classes challenge 2007 (VOC2007) results. 2007.
- [Ev12] Everingham, Mark; Van Gool, Luc; Williams, Christopher K.I.; Winn, John; Zisserman, Andrew: The PASCAL visual object classes challenge 2012 (VOC2012) results. 2012.
- [FMR08] Felzenszwalb, Pedro; McAllester, David; Ramanan, Deva: A discriminatively trained, multiscale, deformable part model. In: Proceedings of CVPR. IEEE, pp. 1–8, 2008.
- [Fr16] Freytag, Alexander; Rodner, Erik; Simon, Marcel; Loos, Alexander; Köhl, Hjalmar; Denzler, Joachim: Chimpanzee Faces in the Wild: Log-Euclidean CNNs for Predicting Identities and Attributes of Primates. In: German Conference on Pattern Recognition (GCPR). pp. 51–63, 2016.
- [Gi14] Girshick, Ross; Donahue, Jeff; Darrell, Trevor; Malik, Jitendra: Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of CVPR. pp. 580–587, 2014.
- [Gi15] Girshick, Ross: Fast R-CNN. In: Proceedings of ICCV. pp. 1440–1448, 2015.
- [GLY19] Ge, Weifeng; Lin, Xiangru; Yu, Yizhou: Weakly supervised complementary parts models for fine-grained image classification from the bottom up. In: Proceedings of CVPR. pp. 3034–3043, 2019.
- [Ha17] Hallmann, Caspar A; Sorg, Martin; Jongejans, Eelke; Siepel, Henk; Hofland, Nick; Schwan, Heinz; Stenmans, Werner; Müller, Andreas; Sumser, Hubert; Hörrn, Thomas et al.: More than 75 percent decline over 27 years in total flying insect biomass in protected areas. *PloS one*, 12(10):e0185809, 2017.
- [HB17] Hughes, Benjamin; Burghardt, Tilo: Automated visual fin identification of individual great white sharks. *International Journal of Computer Vision*, 122(3):542–557, 2017.
- [He15] He, Kaiming; Zhang, Xiangyu; Ren, Shaoqing; Sun, Jian: Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE transactions on pattern analysis and machine intelligence*, 37(9):1904–1916, 2015.
- [He16] He, Kaiming; Zhang, Xiangyu; Ren, Shaoqing; Sun, Jian: Deep residual learning for image recognition. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 770–778, 2016.

- [HPZ19] He, Xiangteng; Peng, Yuxin; Zhao, Junjie: Which and How Many Regions to Gaze: Focus Discriminative Regions for Fine-Grained Visual Categorization. *IJCV*, pp. 1–21, 2019.
- [JFR14] Jonason, Dennis; Franzén, Markus; Ranius, Thomas: Surveying Moths Using Light Traps: Effects of Weather and Time of Year. *PLOS ONE*, 9(3):1–7, 03 2014.
- [Kä18] Käding, Christoph; Rodner, Erik; Freytag, Alexander; Mothes, Oliver; Barz, Björn; Denzler, Joachim: Active Learning for Regression Tasks with Expected Model Output Changes. In: *British Machine Vision Conference (BMVC)*. 2018.
- [KBD18] Körschens, Matthias; Barz, Björn; Denzler, Joachim: Towards Automatic Identification of Elephants in the Wild. In: *AI for Wildlife Conservation Workshop (AIWC)*. 2018.
- [KBD19] Korsch, Dimitri; Bodesheim, Paul; Denzler, Joachim: Classification-Specific Parts for Improving Fine-Grained Visual Categorization. In: *Proceedings of the German Conference on Pattern Recognition*. pp. 62–75, 2019.
- [KD19] Körschens, Matthias; Denzler, Joachim: ELPphants: A Fine-Grained Dataset for Elephant Re-Identification. In: *ICCV Workshop on Computer Vision for Wildlife Conservation (ICCV-WS)*. 2019.
- [Kr13] Krause, Jonathan; Stark, Michael; Deng, Jia; Fei-Fei, Li: 3D Object Representations for Fine-Grained Categorization. In: *4th International IEEE Workshop on 3D Representation and Recognition (3dRR-13)*. 2013.
- [Kr16] Krause, Jonathan; Sapp, Benjamin; Howard, Andrew; Zhou, Howard; Toshev, Alexander; Duerig, Tom; Philbin, James; Fei-Fei, Li: The unreasonable effectiveness of noisy data for fine-grained recognition. In: *ECCV*. Springer, pp. 301–320, 2016.
- [Li14] Lin, Tsung-Yi; Maire, Michael; Belongie, Serge; Hays, James; Perona, Pietro; Ramanan, Deva; Dollár, Piotr; Zitnick, C Lawrence: Microsoft COCO: Common objects in context. In: *ECCV*. Springer, pp. 740–755, 2014.
- [Li16] Liu, Wei; Anguelov, Dragomir; Erhan, Dumitru; Szegedy, Christian; Reed, Scott; Fu, Cheng-Yang; Berg, Alexander C: SSD: Single shot multibox detector. In: *ECCV*. Springer, pp. 21–37, 2016.
- [LRM15] Lin, Tsung-Yu; RoyChowdhury, Aruni; Maji, Subhransu: Bilinear cnn models for fine-grained visual recognition. In: *Proceedings of ICCV*. pp. 1449–1457, 2015.
- [MW07] Mayo, Michael; Watson, Anna T: Automatic species identification of live moths. *Knowledge-Based Systems*, 20(2):195–202, 2007.
- [Re15] Ren, Shaoqing; He, Kaiming; Girshick, Ross; Sun, Jian: Faster R-CNN: Towards real-time object detection with region proposal networks. In: *NIPS*. pp. 91–99, 2015.
- [Re16] Redmon, Joseph; Divvala, Santosh; Girshick, Ross; Farhadi, Ali: You only look once: Unified, real-time object detection. In: *Proceedings of CVPR*. pp. 779–788, 2016.
- [Ro15] Rodner, Erik; Simon, Marcel; Brehm, Gunnar; Pietsch, Stephanie; Wägele, J. Wolfgang; Denzler, Joachim: Fine-grained Recognition Datasets for Biodiversity Analysis. In: *CVPR Workshop on Fine-grained Visual Classification (CVPR-WS)*. 2015.

- [Ru15] Russakovsky, Olga; Deng, Jia; Su, Hao; Krause, Jonathan; Satheesh, Sanjeev; Ma, Sean; Huang, Zhiheng; Karpathy, Andrej; Khosla, Aditya; Bernstein, Michael et al.: Imagenet large scale visual recognition challenge. *International journal of computer vision*, 115(3):211–252, 2015.
- [SB20] Sakib, Faizaan; Burghardt, Tilo: Visual Recognition of Great Ape Behaviours in the Wild. *arXiv preprint arXiv:2011.10759*, 2020.
- [Se13] Sermanet, Pierre; Eigen, David; Zhang, Xiang; Mathieu, Michaël; Fergus, Rob; LeCun, Yann: Overfeat: Integrated recognition, localization and detection using convolutional networks. *arXiv preprint arXiv:1312.6229*, 2013.
- [Si18] Simon, Marcel; Rodner, Erik; Darell, Trevor; Denzler, Joachim: The whole is more than its parts? From explicit to implicit pose normalization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–13, 2018.
- [SR15] Simon, Marcel; Rodner, Erik: Neural Activation Constellations: Unsupervised Part Model Discovery With Convolutional Networks. In: *The IEEE International Conference on Computer Vision (ICCV)*. December 2015.
- [St18] Stork, Nigel E: How many species of insects and other terrestrial arthropods are there on Earth? *Annual review of entomology*, 63:31–45, 2018.
- [Sv20] Svenningsen, Cecilie S; Bowler, Diana E; Hecker, Susanne; Bladt, Jesper; Grescho, Volker; van Dam, Nicole M; Dauber, Jens; Eichenberg, David; Ejrnæs, Rasmus; Fløjgaard, Camilla et al.: Contrasting impacts of urban and farmland cover on flying insect biomass. *bioRxiv*, 2020.
- [SVZ13] Simonyan, Karen; Vedaldi, Andrea; Zisserman, Andrew: Deep inside convolutional networks: Visualising image classification models and saliency maps. *arXiv preprint arXiv:1312.6034*, 2013.
- [SZ14] Simonyan, Karen; Zisserman, Andrew: Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
- [Sz16] Szegedy, Christian; Vanhoucke, Vincent; Ioffe, Sergey; Shlens, Jon; Wojna, Zbigniew: Rethinking the Inception Architecture for Computer Vision. In: *Proceedings of CVPR*. June 2016.
- [TH12] Tieleman, Tijmen; Hinton, Geoffrey: Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude. *COURSERA: Neural networks for machine learning*, 4(2):26–31, 2012.
- [Ui13] Uijlings, Jasper RR; Van De Sande, Koen EA; Gevers, Theo; Smeulders, Arnold WM: Selective search for object recognition. *IJCV*, 104(2):154–171, 2013.
- [Va18] Van Horn, Grant; Mac Aodha, Oisín; Song, Yang; Cui, Yin; Sun, Chen; Shepard, Alex; Adam, Hartwig; Perona, Pietro; Belongie, Serge: The iNaturalist species classification and detection dataset. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 8769–8778, 2018.
- [Wa11] Wah, C.; Branson, S.; Welinder, P.; Perona, P.; Belongie, S.: The Caltech-UCSD Birds-200-2011 Dataset. *Technical Report CNS-TR-2011-001*, California Institute of Technology, 2011.

- [Wa21] Wagner, David L; Grames, Eliza M; Forister, Matthew L; Berenbaum, May R; Stopak, David: Insect decline in the Anthropocene: Death by a thousand cuts. *Proceedings of the National Academy of Sciences*, 118(2), 2021.
- [WOK04] Watson, Anna T; O'Neill, Mark A; Kitching, Ian J: Automated identification of live moths (Macrolepidoptera) using digital automated identification System (DAISY). *Systematics and Biodiversity*, 1(3):287–300, 2004.
- [Xi18] Xia, Denan; Chen, Peng; Wang, Bing; Zhang, Jun; Xie, Chengjun: Insect detection and classification based on an improved convolutional neural network. *Sensors*, 18(12):4169, 2018.
- [YMB19] Yang, Xinyu; Mirmehdi, Majid; Burghardt, Tilo: Great Ape Detection in Challenging Jungle Camera Trap Footage via Attention-Based Spatial and Temporal Feature Blending. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*. pp. 0–0, 2019.
- [Zh18] Zhong, Yuanhong; Gao, Junyuan; Lei, Qilun; Zhou, Yao: A vision-based counting and recognition system for flying insects in intelligent agriculture. *Sensors*, 18(5):1489, 2018.
- [Zh19a] Zhang, Jian; Zhang, Runsheng; Huang, Yaping; Zou, Qi: Unsupervised Part Mining for Fine-grained Image Classification. *arXiv preprint arXiv:1902.09941*, 2019.
- [Zh19b] Zhang, Lianbo; Huang, Shaoli; Liu, Wei; Tao, Dacheng: Learning a Mixture of Granularity-Specific Experts for Fine-Grained Categorization. In: *Proceedings of ICCV*. pp. 8331–8340, 2019.

Energy-efficient Mobile Sensor Data Offloading via WiFi using LoRa-based Connectivity Estimations

Julian Zobel¹  Paul Frommelt,¹ Patrick Lieser¹  Jonas Höchst² 
Patrick Lampe²  Bernd Freisleben²  Ralf Steinmetz¹ 

Abstract: Animal monitoring in natural habitats provides significant insights into the animals' behavior, interactions, health, or external influences. However, the sizes of monitoring devices attachable to animals strongly depends on the animals' sizes, and thus the range of possible sensors including batteries is severely limited. Gathered data can be offloaded from monitoring devices to data sinks in a wireless sensor network using available radio access technologies, but this process also needs to be as energy-efficient as possible. This paper presents an approach to combine the benefits of high-throughput WiFi and robust low-power LoRa communication for energy-efficient data offloading. WiFi is only used when connectivity between mobile devices and data sinks is available, which is determined by LoRa-based distance estimations without the need for additional GPS sensors. A prototypical implementation on low-end commodity-off-the-shelf hardware is used to evaluate the proposed approach in a German mixed forest using a simple path loss model for distance estimation. The system provides an offloading success rate of 87%, which is similar to that of a GPS-based approach, but with around 37% less power consumption.

Keywords: LoRa; Distance Estimation; Multi-RAT Wireless Sensor Networks; Data Offloading

1 Introduction

The monitoring of animals in their natural habitats provides valuable scientific insights for researchers. This includes animal health, territorial behavior, interactions with other animals or other species, and external influences like humans and human-made infrastructure [Xu16, Wy18, As19, KA20]. Monitoring can be performed using a range of sensors and technologies attached to animals, but these are typically strictly limited by weight and size, depending on the monitored animals. This is especially true for the battery size, which stands in contrast to a desired long-term performance of the monitoring sensors, and may also prohibit the use of GPS sensors [Go19].

Transmitting the monitored data from sensor nodes to data sinks is similarly demanding, and a considerable amount of the available battery power is usually consumed for the transmission [SIB12]. With limited storage size, data must be collected before it is overwrit-

¹ Multimedia Communication Lab (KOM), Technical University of Darmstadt, Germany {julian.zobel,patrick.lieser,ralf.steinmetz}@kom.tu-darmstadt.de,paul.frommelt@stud.tu-darmstadt.de

² Department of Mathematics & Computer Science, University of Marburg, Germany, {hoechst,lampep,freisleb}@informatik.uni-marburg.de

ten and lost. However, different *Radio Access Technologies* (RAT) or data collection approaches have various advantages and drawbacks. Cellular services can be used, but may not be available within the animal habitat and are usually too expensive in terms of hardware, provider costs, power consumption, and weight. Long-range low-power communications for *Low-Power Wide Area Networks* (LPWANs) like LoRa require significantly less power, weight, and have long transmission ranges, but can only send very small datasets due to their limited bandwidth.

Furthermore, data collection using Unmanned Aerial Vehicles (UAVs) is a recent and promising approach, but this would require the sensor node to be available at any time, which also increases power consumption [Xu16]. In addition, motorized UAVs yield severe audible disturbances in the animal habitat, negatively and artificially influencing animal behavior [Rü16]. A common approach is to integrate mobile nodes into a larger *Wireless Sensor Network* (WSN). If several communication technologies are available, they are also referred to as *multi-RAT WSNs*. Whenever a mobile node, e.g., attached to an animal, moves within the communication range of a static data sink, the mobile node attempts to offload its data. Such opportunistic offloading approaches are usually performed over high-throughput RATs like WiFi or Bluetooth. Nevertheless, regularly probing for the availability of a data sink increases power consumption, e.g., approximately 180 μ Ah for a single scan for available WiFi networks [Hu15].

In this paper, we overcome these challenges by combining the benefits of high-throughput WiFi and robust low-power LoRa communication in a multi-RAT WSN. Our approach predicts the WiFi connectivity between mobile sensor nodes and static data sinks based on LoRa distance estimations, without the need for additional GPS sensors on mobile nodes or time synchronization within the network. WiFi is only activated when a node determines a possible connection to a data sink, to further increase the node's energy efficiency. The approach is implemented in a prototype system based on widely available low-cost low-end hardware, and evaluated in a German mixed forest. We achieve an offloading success rate of 87%, which is similar to that of a GPS-based system, but with up to 37% lower power consumption for the communication and data offloading process.

To summarize, this paper makes the following contributions:

- We provide a design for a lightweight data offloading approach based on LoRa and WiFi communication that is independent of power-hungry GPS-based localization on mobile nodes and time synchronization in the WSN.
- We present the implementation of our prototype system using commodity-off-the-shelf low-end hardware and a simple path loss model for LoRa-based connectivity estimations.
- In a proof-of-concept evaluation, we highlight the applicability and benefits of our approach for data offloading in mobile multi-RAT WSNs.

The rest of this paper is structured as follows. Section 2 discusses related work for data offloading as well as localization and distance estimation in WSNs. Our protocol design is presented in Section 3, followed by the prototypical system implementation in Section 4. This prototype is then evaluated in a forest environment in Section 5 and compared to a GPS-based approach. Section 6 concludes the paper and gives an outlook on future work.

2 Related Work

Within multi-RAT applications, data offloading describes the general approach of redirecting traffic flows over a more suitable communication medium, i.e., a medium that provides higher throughput, increased reliability, reduced energy consumption, or better cost-efficiency. Typically, data offloading is studied in the field of cellular networks. Here, the common idea is to relieve the usually more expensive and more occupied cellular medium (e.g., LTE or 5G) by using other available networks such as an available *Wireless Local Area Network* (WLAN) provided by public WiFi hotspots [Hu15, Zh18], WiFi or Bluetooth using ad hoc communication for localized communication exchange [Ri14], or even access points of power line carriers [Wu17]. For WSN or IoT applications with small and infrequent data generation, offloading is also possible using long-range low-power communication technologies for LPWANs like SigFox and LoRa [Le19].

A typical research problem in the field of multi-RAT applications is the definition of conditions, which trigger the switch between communication technologies. This switch may be initiated by the user or the mobile device, which provides an individual choice of the used technology. However, this decision may be not optimal due to the lack of global knowledge on that device. In contrast, if the operator or a base station decides to switch based on a broader knowledge, the decision is often better for the overall network, but not necessarily for the individual user [Yu17]. Similar to handover procedures in cellular networks, the decision for switching between RATs can be based simply on the availability of another network, measured or estimated signal properties like *Received Signal Strength Indicator* (RSSI) and *Signal-to-Noise Ratio* (SNR), the throughput of the networks, or the distance between a mobile device and the access point [Wu17, Zh18, Le19].

In energy-constrained applications like mobile WSNs, constantly probing or listening for available network access points is not feasible due to the significant power consumption [Hu15, SIB12]. Mobile nodes can match their position to determine if they are in range of access points. Similarly, localization of mobile sensor nodes by *global navigation satellite systems* (GNSS) like GPS or Galileo may also be unavailable due to high energy consumption and comparably expensive hardware. Therefore, a considerable amount of research is conducted in GNSS-less localization of mobile nodes in WSNs. Triangulation is one of the most widespread approaches, where multiple base stations with known locations use incoming signals of a mobile node to estimate its position. By comparing differences in metrics like time of arrival, time of flight, RSSI, or SNR between base stations, this method achieves good results. However, triangulation requires a highly accurate

time synchronization within the WSN and several—i.e., three or better more—base stations to function properly and , which may be infeasible in certain scenarios [LCL18, DPT19]. Another well-known approach is fingerprinting the area of operation. Basically, a map of the area is created that contains important information like RSSI or SNR. Mobile nodes then compare their measurements of received signals with this map to obtain their location [AGM09, Ha19]. Obviously, this requires a detailed fingerprint map for accurate localization. The creation of such a map requires a tremendous effort — or even an impossible effort for large or inaccessible operation areas — before being able to deploy the WSN.

However, acquiring the exact position is often not necessary for data offloading, but estimating the distance to a base station is sufficient in many scenarios [Xu10]. Recent simulations have shown that localization over LoRa is possible using the time of flight or the RSSI values from received LoRa packets to estimate distances between base stations [LCL18, LSH19]. These distances themselves are obtained using a *path loss model* that matches RSSI values to a certain distance [J17, KT17, Wi17, Li18, Wu20]. A commonly used model is the *Log-Distance Model* that uses a logarithmic decrease of the RSSI with the distance. More sophisticated variants, like the *Log-Normal Shadowing Model*, include non-deterministic characteristics of real-world signal propagation like occasional obstruction, reflection, or interference [Xu10, Wu20]. The drawback of such models is that model parameters are highly scenario-specific and may change over the course of a year [Wu20]. Thus, they are usually derived from empirical data within the corresponding environment for a certain model, for example, as performed for the *Log-Normal Shadowing Model* in a representative Central European mixed forest area by Palaios *et al.* [PLM14]. In this work, we use these acquired parameters in the *Log-Normal Shadowing Model* for the distance estimation and demonstrate how this model can be used on low-end hardware with restricted performance for energy-efficient data offloading.

3 Design

In this section, we present our design of a lightweight, energy-efficient data offloading approach that is executable on low-end hardware. We assume a heterogeneous, multi-RAT WSN with small, battery-powered mobile sensor nodes, which are mounted, for example, on deer. More powerful, static sensor nodes are sparsely located throughout the animal habitat and act as distributed data sinks. Nodes and data sinks provide both LoRa and WiFi as communication technologies, but due to the size of the habitat, data sinks do not cover the whole area. As a restriction, we neglect the data acquisition within this work, because it is highly dependent on the specific application with its used sensors, measurement intervals, and more. We also assume that nodes and data sinks have no time synchronization.

Battery longevity is very important for animal monitoring, since it allows a long-term assessment without disturbing animals. However, it is also of high interest to acquire the collected monitoring data in a timely manner, which should be done energy-efficiently due to the first restriction. On the one hand, the obvious first choice of transmitting the data via

the long-range low-power technology LoRa is, however, infeasible. The limited bandwidth and the strict duty cycle restrictions in the used frequency bands do not allow transmitting large amounts of data. Additionally, best-effort transmissions of LoRa over long distances can possibly lead to lost messages and, thus, lost data. On the other hand, WiFi has very limited range and a large habitat cannot be fully covered by data sinks due to ecological and economical reasons. By utilizing the mobility of the observed animals, an opportunistic approach to upload data when the monitored animal moves in range of a data sink is possible. However, constantly using WiFi to probe for available data sinks would rapidly deplete the battery of a mobile node [Hu15]. In a similar fashion as presented in Section 2, using GPS sensors or RSSI fingerprinting is infeasible due to high power consumption and the vast size of animal habitats, respectively. Furthermore, the positions and the number of data sinks can change during long-term animal monitoring and, thus, rendering fingerprinting or location matching infeasible without replacing the information on the mobile nodes.

To overcome these problems, we propose an approach to combine the possibility for distance estimation over LoRa with the capability of WiFi for high-throughput data transmission on mobile nodes. The energy-efficient LoRa technology is used to opportunistically probe for available data sinks and estimate distances to them. Because LoRa is strictly limited by a duty cycle, due to the used frequency bands, and to minimize power consumption, LoRa messages should be as small as possible in general. If a monitored animal moves within the WiFi range of a data sink, a data offloading attempt is started. The activation of the energy-demanding WiFi connection is minimized and only used for a quick transmission. Our approach for combined distance estimation and data offloading is divided into three phases for mobile nodes. An overview of these phases is shown in Figure 1.

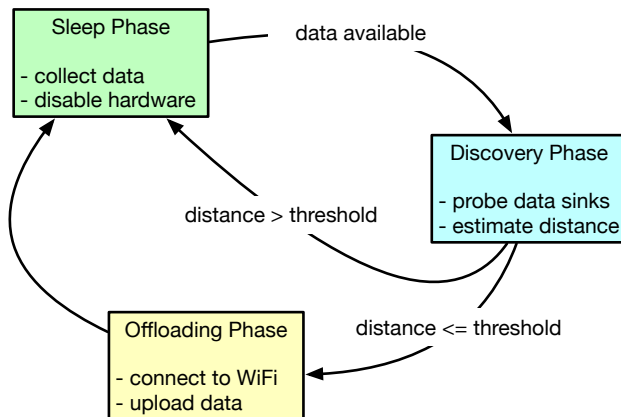


Fig. 1: Overview on the three-phased data offloading approach.

Sleep Phase The *sleep phase* is the basic condition of the mobile node, in which most components are turned off to save power. Within a predefined interval, the node comes out of hibernation to perform certain tasks, like collecting data or probing for data sinks. As stated beforehand, we do not incorporate a specified data collection phase into our communication protocol within the scope of this paper, since data acquisition is specific to the application and the used sensors. Similarly, the node should only proceed to the *discovery phase* if data is available for offloading.

Discovery Phase In the *discovery phase*, the mobile node opportunistically probes over LoRa for available data sinks using a beacon message. On reception at a data sink, the sink sends a response back that the mobile node will use to estimate the distance to the data sink with a path loss model and the RSSI of the response. The time of flight as additional metric to the RSSI cannot be used reliably without a highly accurate time synchronization between sender and receiver. Additionally, timestamps would significantly increase the size of LoRa packets. By using only the RSSI, our approach is independent of time synchronization. If the estimated distance based on the RSSI is less or equal to an offloading threshold, the mobile node switches to the next phase. However, if no response is received or the estimated distance exceeds the threshold, it goes back to the *sleep phase*.

Offloading Phase In the *offloading phase*, the node will activate its WiFi module and attempt to connect to the data sink's WiFi hotspot. When connected, available data is offloaded to the data sink. The node deactivates its WiFi directly afterwards to minimize WiFi usage. Then, the node switches back to the *sleep phase* to further save energy.

As an alternative, the distance estimation could also be performed on the data sink, for example if the mobile node hardware is not capable of calculating the path loss model. A response could then only be sent if the distance estimation on the sink is lower than the threshold. However, communication properties may not be bidirectionally balanced, e.g., due to more powerful hardware on the sink than on the node. It is, therefore, possible that the data sink will reach the node via WiFi, but not vice versa. A deeper analysis of this alternative approach is left open for future research. Using low-end commodity-off-the-shelf hardware, we realized this communication protocol in a prototypical implementation, as presented in the following section.

4 Prototypical Implementation

In this section, we present our prototype system for mobile sensor data offloading with LoRa-based connectivity estimations. First, we describe the realization of data sinks and mobile sensor nodes with a focus on the hardware components, followed by the communication protocol and interaction between them. The prototype is realized by using broadly available, inexpensive commodity-off-the-shelf hardware and open source libraries.

4.1 Data Sink

The foundation of the data sink is a single-board *Raspberry Pi 4B* computer. It provides a 64-bit quad-core 1.5 GHz ARM Cortex-A72 processor, 2 GB of RAM, and built-in WiFi capabilities. A standard installation of the Raspberry Pi OS³ as provided by Raspberry Pi Foundation⁴ is used. The prototype is powered over the Raspberry's USB-C port with a 99 Wh LiPo battery, but can also be powered for long-term use by a combination of a solar panel and a 12 V car battery as power source [Go19].

For the additional LoRa communication capabilities, we use a *Dragino PG1301 LoRa Concentrator*⁵, which is connected directly to the GPIOs of the Raspberry Pi. The PG1301 uses one SX1301 and two SX1257 LoRa modules to allow receiving on up to 10 parallel LoRa channels and sending on one LoRa channel, respectively, using the manufacturer's standard library⁶. It is used by our *LoRaPi* service that processes incoming messages and can also send out replies, as detailed in Sections 4.3 and 4.4. Furthermore, the *LoRaPi* service can also interact with the WiFi management service to activate or deactivate the sink's WiFi access point.

The WiFi management service *hostapd*⁷ runs in the background, managing the Raspberry's WiFi access point and IP addresses of mobile nodes. Each sink uses a standardized SSID which is extended by the specific sink's ID for a unique WiFi SSID. Additionally, a UDP service is running in parallel that monitors the connection of all mobile nodes and the progress of their data upload. The upload itself is performed over FTP, using the *proftpd*⁸ library that collects the data on an FTP server in the background.

4.2 Mobile Nodes

Our prototype of a mobile sensor node is realized on a *Heltec ESP32 LoRa WiFi v2*⁹ board. It has a 32-bit dual-core 240 MHz LX6 microcontroller and 529 kB internal SRAM with a weight of only 20 g. It provides Bluetooth, 2.4 GHz dual-mode WiFi, and LoRa using one SX1276 LoRa transceiver for communication. As a prototyping device, it also provides a 0.96 inch OLED display and can be powered either over a micro-USB port or a 1.25 mm JST LiPo battery interface.

³ <https://www.raspberrypi.org/software/>

⁴ <https://www.raspberrypi.org/>

⁵ <https://www.dragino.com/products/lora/item/149-lora-gps-hat.html>

⁶ https://github.com/dragino/pi_gateway_fwd

⁷ <https://w1.fi/hostapd/>

⁸ <http://www.proftpd.org/>

⁹ <https://heltec.org/project/wifi-lora-32/>

For data storage, we use a standard SD card breakout board¹⁰ connected via SPI. The SD card stores the used dummy sensor data and configuration files for the mobile node. The latter include a node's ID, LoRa settings, and configurations of the main program running on the node. Configuration files are stored in JSON format¹¹. For evaluation purposes, the prototype can be equipped with an optional GPS device¹² connected via SPI.

The main program is written in the C-like Arduino language and uses the standard library of the Heltec ESP32¹³ for LoRa communication. WiFi communication is provided by the standard Arduino library, the FTP transfer uses the ESP32 FTP Client¹⁴ library.

The software running on a mobile node includes the lightweight communication protocol over LoRa and WiFi, further discussed in detail in the following sections. The crucial part of the main program is the distance estimation. For evaluation purposes, this includes the possibility to use GPS localization and matching with a predefined location of the data sink.

However, for the main procedure, distance estimation is performed by a path loss model estimation. The code of the estimation can be easily exchanged with different models, but for our prototype, we used the already discussed *Log-Normal Shadowing Model*. In addition to the logarithmic reduction of a signal's RSSI over distance, this model also includes non-deterministic influences from the environment for the path loss PL :

$$PL(d) = PL(d_0) + 10n \log_{10}\left(\frac{d}{d_0}\right) + X_{\sigma} \quad (1)$$

where n is the path loss exponent, d is the distance between sender and receiver, X_{σ} is a zero-mean Gaussian-distributed random variable with standard deviation σ , and a reference measurement PL_{d_0} at a distance d_0 . The estimated distance d_E on a given RSSI is then calculated by

$$d_E(RSSI) = d_0 * 10^{(RSSI - PL(d_0) - X_{\sigma}) / (10n)}. \quad (2)$$

The fitted properties are $n = 4.96$, $\sigma = 6.13$ dB, and $PL(d_0) = 52.4$ dB with $d_0 = 10$ m as measured by Palaios *et al.* [PLM14] for a Central European mixed forest. Note that these measurements were conducted with a center frequency of 485 MHz, but our LoRa communication is performed in the 868 MHz EU SRD frequency band. However, we assume this to have only a minor influence, since possible external influences of different foliage or forestation in our work compared to the related work are expected to have a more significant influence.

¹⁰ Code based on: <https://iotdesignpro.com/projects/logging-temperature-and-humidity-data-on-sd-card-using-esp32>

¹¹ <https://rapidjson.org/>

¹² https://github.com/adafruit/Adafruit_GPS

¹³ https://github.com/HelTecAutomation/Heltec_ESP32

¹⁴ https://github.com/ladab/ESP32_FTIClient

4.3 LoRa Communication in the Discovery Phase

Within the discovery phase, all communication is performed over LoRa. We use a customized LoRa approach with very small messages, to minimize medium occupancy and power consumption. Each message is only 2 bytes in size, consisting of a 1-byte ID of the sending device and a 1-byte field for message flags that identify the type of message that is sent. Such a packet has an airtime of approximately 166 ms with a robust standard setting of spreading factor 10, coding rate 4/5, and bandwidth of 125 kHz of LoRa. This setting is also used in our evaluation setup. As a theoretical upper limit, each node and data sink could, therefore, send more than 200 of such packets each hour with a duty cycle of 1% in the 868 MHz EU SRD frequency band.

In our current implementation, four message types are available: *Beacon*, *Reply*, *Offload Initialization*, and *Offload Acknowledge*. *Beacons* are sent by mobile nodes containing their ID. Within a single discovery phase, only a single beacon is sent. When no *Reply* is received, the mobile node goes back to the sleep phase. Every *Beacon* that is received by a data sink is answered with a *Reply* message containing the ID of that sink. The mobile node listens to incoming replies for a certain reception window, as multiple data sinks may possibly overhear and answer its message. Only the reply with the largest RSSI is used, all others are discarded. This reply message is then used for the distance estimation, as presented above. If this estimated distance is below the given threshold, the mobile node sends an *Offload Initialization* message to the data sink with the sink's ID.

On reception of an *Offload Initialization* message at a data sink, the ID is checked first to assure that the message is intended for that device. Then, the sink starts up the WiFi access point and acknowledges the availability with an *Offload Acknowledge* message. If the mobile node does not receive the ACK within its reception window, it again goes back to the sleep phase. Otherwise, when the offload ACK is received, the mobile node also turns on its WiFi module and switches to the *offloading* phase.

4.4 WiFi Communication in the Offloading Phase

In the offloading phase, the mobile node tries to access the wireless network provided by the data sink. The WiFi SSID of each sink is a standard identifier followed by the sink's ID, e.g., "SSID-A1" for the data sink with the (hexadecimal) ID $0xA1$. The mobile node directly connects to that specific access point by using the ID information provided in the previously received *Reply* message. If the WiFi cannot be found, e.g., if the distance estimation was wrong or other circumstances prevent the mobile node from accessing WiFi, the mobile node will go back to the sleep phase.

If it is connected, however, a short acknowledgement is sent to the UDP service running on the data sink. This service blocks the sink's WiFi from deactivating until the mobile node has transferred its data to the background FTP service. After offloading the data, the

mobile node again sends a small UDP packet to confirm the transfer. The UDP service then unblocks the WiFi again if there are no other devices currently connected. Furthermore, it also manages timeouts, in case that the WiFi connection is lost and the second UDP message is received. The mobile node, either after confirming the transfer or after losing connection, turns off its WiFi module and goes to the sleep phase.

4.5 Limitations of the Prototype

In its current state, the prototype does not generate any meaningful sensor data and does not incorporate a distinct data acquisition phase. Since the used sensors and the type and frequency of measurements are specific for animals, scenarios, and applications, respectively, we deliberately omit a specification of the data acquisition, but focus on the communication. However, our implementation provides the possibility to easily integrate such a phase in the prototype.

The ESP32's LoRa transceiver theoretically provides the full range of spreading factors from 7 to 12. However, a significant amount of messages is not received for spreading factors 11 and 12. As stated in the transceiver's datasheet¹⁵, high spreading factors have a limited tolerance for frequency errors and require a reliable reference frequency, which we assume is not provided by the ESP32's XO. For higher spreading factors—or other LoRa settings with higher demands on the frequency stability, like low bandwidths—, an external TCXO is advised by the manufacturer. Within this work, we limit our prototype to a maximum spreading factor 10.

Another limitation of mobile nodes is the used FTP library for ESP32. It is only possible to reserve blocks of a maximum of 100 kB due to the construction of the ESP32's heap memory. Therefore, files that are transferred directly from the SD card to the FTP server can only have that same or a smaller size. An adaptation of the library to larger file sizes is out of the scope of this paper and, thus, left for future work.

5 Experimental Evaluation

The evaluation of the prototype was conducted as a field experiment in a forest near the city of Darmstadt, Germany¹⁶. As shown in Figure 2, the data sink was non-invasively attached to a tree at a height of approximately 1.80 m. The terrain around the chosen tree is slightly leveled without any artificial elements or forest aisles within 100 m, and mostly accessible with few bushes or shrubs within the direct vicinity. Since we measured in the winter months, there was no foliage.

¹⁵ The current datasheet for LoRa transceivers is provided by Semtech Corporation, www.semtech.com

¹⁶ Data Sink Coordinates: 49.857028 N 8.696744 E



Fig. 2: The data sink prototype was attached to a tree at around 1.80 m height. Distances were measured with a tape measure in a German mixed forest.

Parameter	Setting
TX Power	14 dB
Frequency	868 MHz
Spreading Factor	10
Bandwidth	125 kHz
Coding Rate	4/5
Preamble Length	8 symbols
CRC	disabled

Tab. 1: LoRa settings for the field experiments.

We used a mobile node including a GPS sensor for tracking the position of each measurement location in the experiment. LoRa communication was used on a robust standard setting, as summarized in Table 1. A total of 331 measurements were taken centered around the data sink tree over several days. Since the GPS localization has an inaccuracy of more than ± 10 meters in the forest, we used a tape measure to determine the distance between mobile node and data sink and took measurements every five meters. Shrubs, bushes, and other trees, however, did not allow to measure in all directions and usually only up to 40 to 60 meters, depending on the direction. Furthermore, we also took measurements at the approximate limits of the WiFi range (approx. 40–100 m) in a circle around the data sink

tree. As a result, 174 of the 331 measurements are taken with a tape measure, from which 144 data points are in a measured distance of 50 meters or less.

For each measurement, the mobile station is started in the *discovery phase*. As expected for LoRa, it is possible to send and receive LoRa messages at all measurement locations within that comparably short range. In addition, the evaluation system is adapted, such that it switched to the *offloading phase* regardless of the measured distance. With that, we determine if a WiFi connection to the data sink is possible at that location and if the decision of the mobile node to switch to the *offloading phase* would be correct. We use the acquired data points to first assess the quality of the distance estimation.

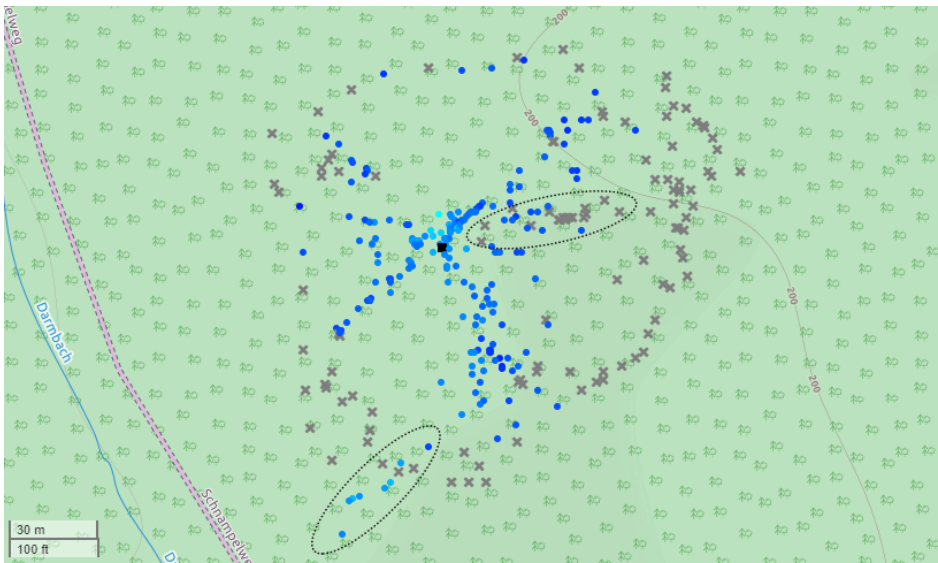


Fig. 3: Data points for 331 WiFi offloading attempts. The location of the data sink is denoted by the black square. Blue points denote positions with successful offloading, with lighter colors showing a higher LoRa RSSI, crosses where WiFi access is not possible. (Maps ©Thunderforest, Data ©OpenStreetMap contributors)

5.1 Distance Estimation and Data Offloading

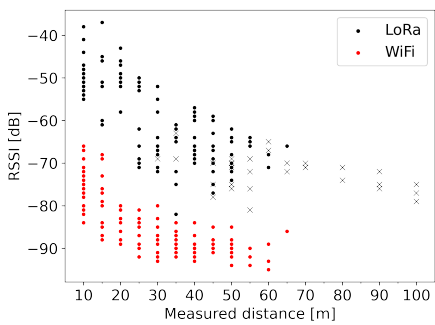
All data points are visualized in Figure 3, showing successful and unsuccessful WiFi offloading attempts, respectively. The location of the data sink is shown as a black square. Blue dots are locations where access to the sink's WiFi is possible, with lighter colors showing a stronger LoRa RSSI. Grey crosses denote where a WiFi connection is not possible. The measured area is approximately 200 m by 200 m in size.

As expected, a WiFi connection is typically possible within 50 meters, especially with line-of-sight to the data sink. To the east of the data sink, however, a large number of connection

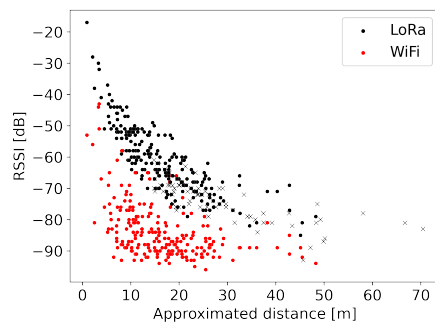
attempts fails even at close distance, most probably due to the the obstruction by several larger trees. The LoRa RSSI degrades similarly in this area. In contrast, a relatively large isle to the south-west allows line-of-sight to the data sink, with WiFi connections to around 100 m and reasonably good RSSI values for the LoRa communication. Nevertheless, nearby locations that are even closer to the data sink but without line-of-sight do not succeed in establishing a WiFi connection at all. It is therefore clear that the surrounding environment has a more significant impact on the connectivity than the pure distance to the data sink. Furthermore, this impact is similarly perceivable on the RSSI measurements for LoRa and WiFi, respectively.

Figure 4 highlights this coherence between LoRa RSSI and WiFi RSSI for the tape-measured distances and the approximated distances by the used path loss model, respectively. The general trend of decreasing RSSI values at longer distances is evident for both wireless technologies. The measured RSSI values in Figure 4a also show a significant variance, comparable to the results of Palaios *et al.* [PLM14]. With increasing distance, the number of failed WiFi connections increases. The longest exactly measured distance for a successful upload is 65 meters, although two out of three connections are unsuccessful at that same location. However, as already stated, longer distances are possible but highly dependent on the surrounding environment.

More interestingly, no WiFi connection is possible in any case where the approximated distance by the path loss model exceeds 50 m, as shown in Figure 4b. For less than 15 m approximated distance, every WiFi connection is successful. Approximations between 15 and 30 meters result in circa 30% failed connections, similar to the obstructed locations to the west of the data sink (cf. Figure 3), and the ratio of failed connections further increases with distance. Nevertheless, the number of failed connections between 15 and 30 meters cannot entirely be described by these measurements. We, therefore, also compare the ap-



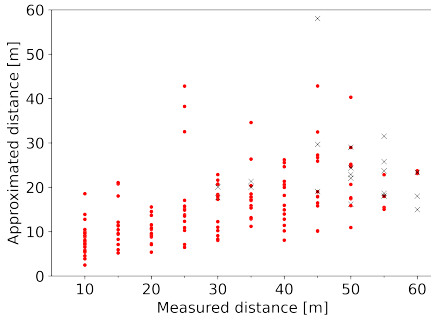
(a) RSSI for tape-measured distances.



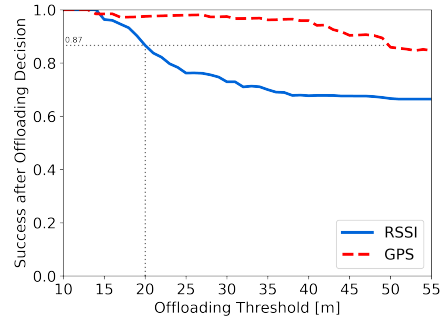
(b) RSSI for PLM distance approximation.

Fig. 4: LoRa RSSI and WiFi RSSI values for tape-measured and approximated distances, respectively. Unsuccessful WiFi connections are marked by a cross.

proximated distance with all measured distances in Figure 5a. Clearly, the used path loss model with the respective parameter settings (cf. [PLM14]) underestimates the real distance on the average. This might be a result of the used LoRa technology: it is more robust and uses a different frequency band compared to the related work. Other possible influences are differences in the terrain, the forestation, or foliage at the respective evaluation site. Clearly, the parameter settings or the path loss model must be adapted specifically for the used evaluation environment for a more realistic distance estimation, or may even require an adaptive adjustment to cope with varying influences of a diverse environment.



(a) Approximated distance vs. measured distance.



(b) Success after an offloading decision of the prototype w.r.t. the offloading threshold.

Fig. 5: The comparison of the approximated distance with the real distance shows a significant underestimation. Nevertheless, the results can be used to make a correct decision for offloading on the mobile node. The RSSI-based LoRa approach with an offloading threshold of 20 m provides a similar success rate to a GPS-based approach with an offloading threshold of 50 m.

5.2 Offloading Threshold

Despite the underestimation in the distances between data sink and mobile node, it is possible to use the approximated distance to decide whether the mobile node should start the offloading process or not. For that, we compare different offloading thresholds on the evaluation data, as depicted in Figure 5b. The ordinate shows the success rate of the offloading phase for both the LoRa RSSI-based approach and the GPS-based approach based on a fixed threshold. Success is defined as a completed WiFi connection and data offloading phase, happening after the respective approach calculates the distance to the data sink to be less or equal to that threshold. In case of the RSSI approach, this distance is approximated by the path loss model. In case of GPS, this is the difference between the measured GPS location and the location of the data sink. With a threshold of 50 m, the GPS approach has a success rate of 0.86, i.e., 86% of all offloading attempts were successful for that threshold. A similar success rate of 0.87 is achieved by the RSSI approach at 20 m. This resembles a similar offset than that between measured and approximated distance (cf. Figure 5a). Thus, due to the divergence of the RSSI approach, using an offloading threshold of 20 m for the

decision based on LoRa RSSI measurements allows to offload data within 50 m in reality. As a result, both approaches result in a similar success rate with appropriately defined thresholds.

The eventual decision on an offloading threshold must be based on the specific application scenario and must include factors like the monitored animals and their mobility, the area coverage by data sinks, as well as available power resources and the amount of generated data on mobile nodes. In general, a smaller threshold will yield a higher probability for successful WiFi connections, on the one hand, but will equally decrease the actual chance for data offloading in the context of animal monitoring significantly. On the other hand, a larger threshold increases the chances for offloading attempts, but similarly also the number of unsuccessful attempts, leading to a lower energy efficiency.

5.3 Power Consumption of the Prototype

Finally, we measure the power consumption of our prototype's mobile node using a USB multimeter in three scenarios for 60 seconds with and without GPS, respectively, resulting in a total of six settings. For all settings using GPS, a GPS satellite lock is established before starting the test. Acquiring a lock takes up to 5 minutes and, therefore, would have significantly increased the power consumption for each test. By acquiring the lock beforehand, our results only incorporate the difference of GPS usage compared to no GPS usage within the evaluation frame. This allows us to specifically evaluate and compare the communication and distance estimation part, without external influences on the GPS. Similarly, we do not include further sensors, measure any sensor's power requirements, or perform a long-term measurement on the prototype, due to the scope of this paper and the prototype which focus on the offloading procedure and the communication. The three scenarios, on which we measured the power consumption of the prototype, are (1) an idle mode of the ESP32 without the use of LoRa or WiFi, but activated built-in OLED display, (2) a full procedure of LoRa communication between mobile node and data sink but without activating WiFi, and (3) the same procedure with a connection and upload to the data sink over WiFi.

The results of the comparison of the scenarios is depicted in Figure 6. We find that the full bidirectional LoRa communication between mobile node and data sink consumes 109–120 μAh , which is less than for scanning for a WiFi access point once [Hu15]. Accessing the data sinks WiFi and uploading a 100 kByte dummy file to the FTP server consumes an additional 226–253 μAh . But as expected, the most significant impact on the overall power consumption is the running GPS. The idle mobile node without GPS consumes around 467 μAh , but with active GPS consumes an additional 431–469 μAh within the same measurement frame. As already stated, this does not include additional power required for acquiring a satellite lock and a position in the first place. Furthermore, note that also the data sink's WiFi has to be constantly activated for a GPS-based system due to the lack of a second communication channel, further increasing the overall power demand.

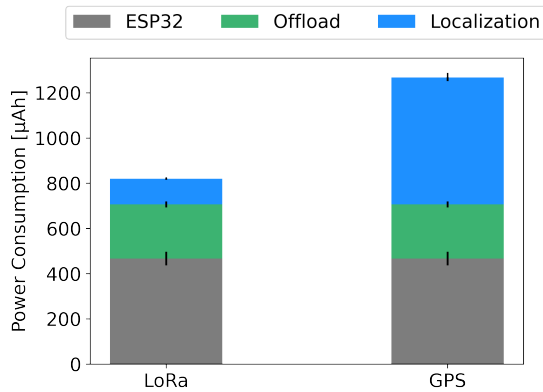


Fig. 6: Power consumption for the LoRa- and GPS-based localization, respectively, split for consumption of the ESP32 device, the WiFi offloading, and the localization approach. The LoRa-based approach requires only 63% of the power of the GPS-based approach.

It is evident that substituting the GPS-based approach with the RSSI-based approach has a significant impact on the overall power consumption, while maintaining a similar offloading success rate within 50 meters. Within the measurement setup, our approach only requires 63% of the power of the GPS approach. Even when omitting the power consumption of LoRa communication, which may not be required in a GPS approach, the power demand is still reduced to 70%. Additionally, our RSSI-based approach using LoRa communication is highly applicable for making the offloading decision with a sensibly chosen offloading threshold with a success rate of 87%.

6 Conclusion

In this paper, we presented an approach for energy-efficient data offloading via WiFi using LoRa-based connectivity estimations for mobile sensor nodes in multi-RAT WSNs. An experimental evaluation of our prototypical implementation demonstrated that our approach is capable of providing reasonable estimations whether data offloading via WiFi is possible or not, based on received LoRa messages and without further GPS capabilities. Our prototype uses available parameters for the log-normal shadowing path loss model [PLM14] for distance estimation, which has shown to underestimate real distances within our forestry evaluation setup. Nevertheless, it was possible to determine a reasonable offloading distance threshold in the evaluation results. In particular, we achieved an offloading success of 87%, which is similar to the performance of a GPS-based approach, while reducing the power consumption by up to 37%.

In our evaluation, we observed that the path loss model is highly dependent on its parametrization. Determining the environment-specific parameters is, therefore, of primary interest for future work to obtain realistic distance estimations. Furthermore, extended evaluations are required to determine if our rather static approach is also applicable in large-scale scenarios with very different terrain or types of forests in the same area, or for long-term monitoring under varying environmental conditions like rain and snow or the change of foliage over the year. Additionally, future work may encompass the influence of animal mobility, data sink distribution, area coverage, as well as adaptive offloading thresholds and more complex path loss models. Our next research goal is to compare sensor node-based, data sink-based, and cooperative distance estimation approaches and evaluate their influence on the offloading decision accuracy. Nevertheless, we have shown that wireless sensor networks with heterogeneous communication technologies are already capable of energy-efficient data offloading without the need for a GPS device and that a simple path loss model calculated on low-end commodity hardware yields valuable results.

Acknowledgment

This work has been funded by the LOEWE initiative (Hessen, Germany) within the LOEWE research cluster *Nature 4.0 – Sensing Biodiversity* and the *emergenCITY* center, and the German Research Foundation (DFG) within the Collaborative Research Center (CRC) 1053 MAKI.

Bibliography

- [AGM09] Arya, A.; Godlewski, P.; Melle, P.: Performance analysis of outdoor localization systems based on RSS fingerprinting. In: 2009 6th International Symposium on Wireless Communication Systems. 2009.
- [As19] Ascensão, Fernando; Kindel, Andreas; Teixeira, Fernanda Zimmermann; Barrientos, Rafael; D'Amico, Marcello; Borda-de Água, Luís; Pereira, Henrique M: Beware that the lack of wildlife mortality records can mask a serious impact of linear infrastructures. *Global Ecology and Conservation*, 19, 2019.
- [DPT19] Dieng, Ousmane; Pham, Congduc; Thiare, Ousmane: Outdoor Localization and Distance Estimation Based on Dynamic RSSI Measurements in LoRa Networks: Application to Cattle Rustling Prevention. In: 2019 International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob). IEEE, 2019.
- [Go19] Gottwald, Jannis; Zeidler, Ralf; Friess, Nicolas; Ludwig, Marvin; Reudenbach, Christoph; Nauss, Thomas: Introduction of an automatic and open-source radio-tracking system for small animals. *Methods in Ecology and Evolution*, 10(12):2163–2172, 2019.
- [Ha19] Ha, G. Y.; Seo, S. B.; Oh, H. S.; Jeon, W. S.: LoRa ToA-Based Localization Using Fingerprint Method. In: 2019 International Conference on Information and Communication Technology Convergence (ICTC). 2019.

- [Hu15] Hu, Xueheng; Song, Lixing; Van Bruggen, Dirk; Striegel, Aaron: Is there WiFi yet? How aggressive probe requests deteriorate energy and throughput. In: Proceedings of the 2015 Internet Measurement Conference. pp. 317–323, 2015.
- [J 17] Jörke, P.; Böcker, S.; Liedmann, F.; Wietfeld, C.: Urban channel models for smart city IoT-networks based on empirical measurements of LoRa-links at 433 and 868 MHz. In: 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC). 2017.
- [KA20] Katzner, Todd E; Arlettaz, Raphaël: Evaluating contributions of recent tracking-based animal movement ecology to conservation management. *Frontiers in Ecology and Evolution*, 7, 2020.
- [KT17] Kurt, S.; Tavli, B.: Path-Loss Modeling for Wireless Sensor Networks: A review of models and comparative evaluations. *IEEE Antennas and Propagation Magazine*, 2017.
- [LCL18] Lam, K.; Cheung, C.; Lee, W.: New RSSI-Based LoRa Localization Algorithms for Very Noisy Outdoor Environment. In: 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC). 2018.
- [Le19] Lemic, F.; Behboodi, A.; Famaey, J.; Mathar, R.: Location-Based Discovery and Vertical Handover in Heterogeneous Low-Power Wide-Area Networks. *IEEE Internet of Things Journal*, 2019.
- [Li18] Linka, Hendrik; Rademacher, Michael; Jonas, Karl; Aliu, Osianoh: Path Loss Models for Low-Power Wide-Area Networks: Experimental Results using LoRa. 2018.
- [LSH19] Lin, Yi-Cheng; Sun, Chi-Chia; Huang, Kuo-Ting: , RSSI Measurement with Channel Model Estimating for IoT Wide Range Localization using LoRa Communication, 2019.
- [PLM14] Palaios, Alexandros; Labou, Yann; Mahonen, Petri: A Study on the Forest Radio Propagation Characteristics in European Mixed Forest Environment. 2014 IEEE Military Communications Conference, 2014.
- [Ri14] Richerzhagen, Björn; Stingl, Dominik; Hans, Ronny; Gross, Christian; Steinmetz, Ralf: Bypassing the cloud: Peer-assisted event dissemination for augmented reality games. In: 14th IEEE International Conference on Peer-to-Peer Computing. pp. 1–10, 2014.
- [Rü16] Rümmler, Marie-Charlott; Mustafa, Osama; Maercker, Jakob; Peter, Hans-Ulrich; Esefeld, Jan: Measuring the influence of unmanned aerial vehicles on Adélie penguins. *Polar Biology*, 39(7), 2016.
- [SIB12] Sachan, Dr. Vibhav; Imam, Syed; Beg, Mirza: Energy-Efficient Communication Methods in Wireless Sensor Networks: A Critical Review. *International Journal of Computer Applications*, 39, 2012.
- [Wi17] Widodo, Slamet; Pratama, Enno Apriliawan; Pramono, Subuh; Basuki, S Budi: Outdoor propagation modeling for wireless sensor networks 2.4 GHz. In: 2017 IEEE International Conference on Communication, Networks and Satellite (Commnetsat). IEEE, 2017.
- [Wu17] Wu, H.; Wen, X.; Lu, Z.; Pan, Q.: Mobile data offloading under attractor selection in heterogeneous networks. In: 2017 International Symposium on Wireless Communication Systems (ISWCS). 2017.

- [Wu20] Wu, Yin; Guo, Genwei; Tian, Guiyun; Liu, Wenbo: A Model with Leaf Area Index and Trunk Diameter for LoRaWAN Radio Propagation in Eastern China Mixed Forest. 2020.
- [Wy18] Wyckoff, Teal B; Sawyer, Hall; Albeke, Shannon E; Garman, Steven L; Kauffman, Matthew J: Evaluating the influence of energy and residential development on the migratory behavior of mule deer. *Ecosphere*, 9(2), 2018.
- [Xu10] Xu, Jiuqiang; Liu, W.; Lang, Fenggao; Zhang, Yuanyuan; Wang, C.: Distance Measurement Model Based on RSSI in WSN. *Wireless Sensor Networks*, 2, 2010.
- [Xu16] Xu, Jun; Solmaz, Gurkan; Rahmatizadeh, Rouhollah; Turgut, Damla; Boloni, Ladislau: Internet of things applications: Animal monitoring with unmanned aerial vehicle. *arXiv preprint arXiv:1610.05287*, 2016.
- [Yu17] Yu, H.; Cheung, M. H.; Iosifidis, G.; Gao, L.; Tassiulas, L.; Huang, J.: Mobile Data Offloading for Green Wireless Networks. *IEEE Wireless Communications*, 2017.
- [Zh18] Zhou, Huan; Wang, Hui; Li, Xiuhua; Leung, Victor C. M: A Survey on Mobile Data Offloading Technologies. *IEEE access*, 2018.

Exploiting Web Images for Moth Species Classification

Julia Böhlke¹, Dimitri Korsch², Paul Bodesheim³, Joachim Denzler^{4 5 6}

Abstract: Due to shrinking habitats, moth populations are declining rapidly. An automated moth population monitoring tool is needed to support conservationists in making informed decisions for counteracting this trend. A non-invasive tool would involve the automatic classification of images of moths, a fine-grained recognition problem. Currently, the lack of images annotated by experts is the main hindrance to such a classification model. To understand how to achieve acceptable predictive accuracies, we investigate the effect of differently sized datasets and data acquired from the Internet. We find the use of web data immensely beneficial and observe that few images from the evaluation domain are enough to mitigate the domain shift in web data. Our experiments show that counteracting the domain shift may yield a relative reduction of the error rate of over 60%. Lastly, the effect of label noise in web data and proposed filtering techniques are analyzed and evaluated.

Keywords: Web Images; Webly Supervised Learning; Label Noise; Noise Filtering; Species Classification; Fine-grained Recognition; Moth Scanner

1 Introduction

As our environment and climate changes, habitats are destroyed, leading to a decline in plant and animal biodiversity. To effectively combat this loss of species on our planet, monitoring these trends and understanding the reasons is paramount. Moths seem to be especially sensitive to the changes in the environment, as they are affected by an extremely rapid decline in population sizes. While the overall declining trend is undeniable, evidence suggests the situation is complex and heterogeneous for differing species [Wa21, Ha19]. In this work, we analyze different aspects of visual moth species classification to support the continuous monitoring of populations with minimally invasive methods. A clear picture of population changes for specific moth species could help biologists and conservationists make informed decisions to counteract this problem. A moth species classification system is developed as part of the biodiversity monitoring stations in the AMMOD project⁷, where it is referred to as the moth scanner. The hardware setup consists of a light trap that attracts moths using ultraviolet light to illuminate a white screen and a camera regularly taking

¹ Computer Vision Group, Friedrich Schiller University Jena, 07737 Jena, Germany; julia.boehlke@uni-jena.de

² Computer Vision Group, Friedrich Schiller University Jena, 07737 Jena, Germany; dimitri.korsch@uni-jena.de

³ Computer Vision Group, Friedrich Schiller University Jena, 07737 Jena, Germany; paul.bodesheim@uni-jena.de

⁴ Computer Vision Group, Friedrich Schiller University Jena, 07737 Jena, Germany; joachim.denzler@uni-jena.de

⁵ German Aerospace Center (DLR), Institute for Data Science, Mälzerstraße 3, 07745 Jena, Germany

⁶ Michael Stifel Center Jena for Data-Driven and Simulation Science, Ernst-Abbe-Platz 2, 07743 Jena, Germany

⁷ <https://ammod.de>

images of this screen during the night. To automatically analyze the vast amount of recorded images, especially when multiple stations are in operation across the country, the software components of the monitoring system require a classification model for identifying different moth species. Although there are most likely many insects in a single image that need to be localized to classify each individual separately, we only focus on the moth species classification in this paper.

The main challenge for training a good species classifier is the availability of a large training dataset with many example images for each individual species. This especially holds when considering modern deep neural networks with their large number of parameters. In addition, not only the plain images of moths are important but also the appropriate species labels, which can often only be provided by few experts with a corresponding background in taxonomy and systematics. This leads to different annotation conditions as for other visual object recognition tasks, such as the distinction of cars from motorcycles, for which most humans could provide class labels. Note that the previous challenges are relevant for most species classification problems if very similar animals need to be distinguished by small details to obtain the correct label, but they are especially important when considering a certain niche such as moths.

While, in general, citizen science applications or other crowd-sourcing activities are useful to collect additional training data, their benefit is limited when expert knowledge is required for image annotations. An active learning approach is more suited in this situation, where an expert labels images selected by a classification model. With this approach, only a small relevant subset of the raw data is labeled such that the expensive manual effort by experts is reduced. However, an initial small training dataset is needed to build an acceptable classification model that selects the images for labeling [Käl16].

To improve our understanding of the required size of a dataset for acceptable performance, we analyze how the number of images used for training affects the performance of a classification model. We use image search engines to acquire a large dataset from the Internet and analyze different aspects of the use of web data, such as the *domain shift*. A domain shift refers to the differing data distributions in the training and evaluation data, which is expected when acquiring training data from the Internet. We compare a webly-supervised approach, where only web data is used for training, with a semi-supervised approach, where cleanly labeled data and web data together form a merged training dataset.

Because the images retrieved from the Internet using image search engines are weakly labeled through tags and text content surrounding the images, the downloaded data contains label noise. As identified by Krause et al. [Kr16], who showed the potential of noisy data from the Internet for fine-grained recognition, there are two different types of label noise that can be differentiated with the definition of the dataset *domain*. All categories in a fine-grained dataset are sub-categories of a larger umbrella category, which is called the domain of the dataset. In our study, we focus on species identification for moth images, and hence, the domain of the dataset we are using is *moths*. Another example is the CUB200-

2011-dataset [Wal1] containing images of 200 different bird species, for which the domain would be *birds*. Using this definition, we can differentiate between within-domain label noise, called *cross-class noise*, and out-of-domain label noise, also called *cross-domain noise*. Usually, when using images from the Internet, the overlap between the evaluation dataset and web data needs to be eliminated to ensure a fair and robust evaluation. Because the evaluation dataset used in this paper is not publicly available, this aspect of using web data is not relevant here. However, the interested reader is referred to our previous work [Bö21] that addresses this issue.

Lastly, we want to analyze how filtering label noise in a preprocessing step affects classification performance. To handle cross-class noise, Krause et al. [Kr16] removed exact duplicates between different classes in the web data because their true label is ambiguous. We expand on this approach by considering near duplicates with our duplicate detection method. Also, we evaluate our previously proposed methods [Bö21] for cross-domain noise filtering using manually acquired annotations of label noise.

2 Related Work

In this section, we review three topics related to our work: fine-grained classification (Sect. 2.1), using images from the Internet for training a classification model (Sect. 2.2) and identifying duplicate images (Sect. 2.3).

2.1 Fine-grained Classification

The research field of fine-grained classification tackles the problem of distinguishing very similar subcategories with small inter-class variance, such as different moth species, for example. The part-based approaches [GLY19, KBD19] tackle this problem by identifying local regions in the images containing the distinct features such that the classification model can focus on the right regions. On the other hand, global approaches use the entire image for classification and thereby avoid the complex problem of extracting parts from the images. Global approaches may rely on a specific feature pooling method [LRM15, Si17] or a clever pre-training strategy [Cu18], to name a few examples. As this paper focuses on analyzing the impact of web images and dataset size, we employ the straightforward global approach in our experiments to compare different scenarios and filtering methods.

2.2 Images from the Internet and Label Noise Handling

Utilizing the vast visual information of the Internet accessible through image search engines is a data-driven approach to classification. Unfortunately, due to the weak labeling of images on the Internet, this information is tainted by label noise.

To handle label noise, one set of methods applies the concept of multiple instance learning, either learning an additional de-noising network for a fine-grained setup [Ya20] or using a grouping strategy together with attention mechanisms that set the training focus to the non-noisy samples [Zh17, Pe20]. Another approach was proposed by [Kr16] where a web dataset with more categories than the target dataset was used to pre-train a network for fine-grained classification. We expand their approach to handle cross-class noise (Sect. 3.3).

Several approaches rely on a small training set with high-quality labels to build a clean dataset from noisy data. They train an initial model, and iteratively update it by using non-noisy images selected by the model itself [Zh20b, Xu15, CSG13]. These approaches require an initial representative dataset to detect hard-to-classify instances, which increases the visual diversity for a specific class. In contrast, other methods start with a noisy dataset and filter the noise during training or dynamically re-weight noisy samples such that they have less impact during the model update [Re18]. Noisy instances can be either identified using a threshold for the cosine similarity to representations of class centers [Zh20b] or using the loss, which is typically higher for noisy instances, especially at early stages of the training [Zh20a]. Assuming that the predictions of noisy samples change more rapidly for consecutive training epochs compared to clean samples, a large cross-entropy of corresponding class probability vectors can also be an indicator for noise [Li21].

A metric learning-based approach for fine-grained recognition has been proposed by [Cu16], which requires feedback of domain experts during training via a human-in-the-loop concept. In situations where an initial training set with part annotations exists, these can be exploited in a transfer learning approach to detect noisy samples [Xu18] since outliers likely obtain low detection scores when applying both a whole-object detector as well as individual part detectors. In contrast, we focus on automatic noise filtering methods that neither require domain experts in the training loop nor training data with part annotations.

Prominent work on sample selection as a pre-processing step, proposed in frameworks such as co-teaching [Ha18], de-coupling [MSS17], and MentorNet [Ji18] train two networks and, for example, use the disagreement between them to identify the noisy instances. Unfortunately, this is a computationally expensive approach. Further data cleansing approaches have been introduced, which utilize class-wise auto-encoders [ZT19], or a variety of different ensemble methods [Ga16]. In contrast, an unsupervised approach was proposed by [Ni15] who used k-means clustering, an algorithm entirely independent of labels and, therefore, also independent of label noise. They repeatedly cluster the noisy data and use cluster statistics for identifying and relabeling noisy instances. We use a simplified version of this clustering-based idea to filter cross-domain noise as described in Section 3.2.

2.3 Duplicate Detection

To tackle the issue of cross-class noise, Krause et al. [Kr16] removed duplicates among different classes in the web data because their true category is ambiguous. We expand on

this approach by considering *near-duplicates* as well, i.e., image pairs that originate from the same source but differ on a pixel level due to small transformations. Further, near-duplicate detection is important when using publicly available datasets for the evaluation, where overlap between training and testing subsets might occur. The task of near-duplicate detection is non-trivial, especially when efficiency plays a crucial role since the number of comparison operations grows with the number of samples in a dataset.

Originally, efficient duplicate detection was done by comparing hand-crafted features extracted from the images [Ke04, Wa06]. Feature representations of images learned by a convolutional neural network (CNN) offer an alternative to the hand-crafted features, as shown by [BD20]. They used L^2 -normalized feature representations extracted for all images from a CNN pre-trained on the respective training set to compare images. In metric learning approaches, such as the deep ranking method proposed by [Wa14], an embedding of the images in a lower-dimensional space is learned using a task-specific dataset. In the embedded space, similar images are located close to each other, and dissimilar ones are far apart. We employ a variant of the idea proposed by [BD20] utilizing a CNN pre-trained on ImageNet for feature extraction and also use the structural similarity measure (SSIM) [Wa04] discussed further in Section 3.3.

3 Methods

We describe our method of web data acquisition and the two typically occurring variants of label noise in Section 3.1. Two methods for filtering label noise are explained in Sections 3.2 and 3.3, respectively.

3.1 Acquisition of Web Data

Given a small labeled training dataset, also called *seed dataset* in the following, we want to download images for the same classes. The class names, i.e., species names, for this seed dataset are used as keywords in Google Image Search, and the results are downloaded for each class. This process is automated using the Google Cloud Platform services. The dataset of downloaded images will also be referred to as the *augment dataset*.

As mentioned previously, two different types of label noise can be identified in web data. An image of a map, caterpillar, or habitat would be an example of cross-domain noise for the moth species dataset, while an image depicting *Acronicta Aceris* downloaded for the class of *Acronicta Leporina* would be an example of cross-class noise. In the following section, the method for handling cross-domain noise is described, while details on our method for mitigating cross-class noise are given in Section 3.3. The methods described in both sections are used to decide which of the images are instances of label noise and should be discarded.

3.2 Cross-domain Noise Filtering

For cross-domain noise filtering, we want to identify images that do not depict the domain of the seed dataset. The method described in this section was first proposed in our previous work [Bö21]. In theory, features extracted from a CNN of images in the augment set that belong to the same domain as the seed dataset will have smaller distances to features of images from the seed dataset than out-of-domain images. We utilize this observation using a clustering-based approach. We estimate clusters jointly for feature representations of images in the augment and the seed training set. The clusters that contain a certain amount of the seed dataset indicate clusters of images belonging to the domain (*positive clusters*). Images from the augment dataset in clusters with a high proportion of seed data are likely to belong to the same domain as the seed dataset. Augment images in clusters with few to none of the original seed data, on the other hand, are more likely to depict out-of-domain objects, i.e., cross-domain noise. Therefore, we retain images in the positive clusters for training while the remaining images are filtered out.

We distinguish between two different types of positive clusters. A *strong positive cluster* is defined by containing more than $\frac{N}{k}$ samples of the seed dataset, when k clusters are estimated, and N is the total number of images in the seed dataset. The adaptive threshold ensures that at least some strong positive clusters are identified, even when clustering with a large value for k . Preliminary experiments showed that with a visually homogeneous dataset, the features of the seed dataset are too similar and are clustered together in few clusters. In this situation, using images from strong positive clusters only is not enough to identify within-domain images. Therefore, we further define *weak positive clusters* as clusters that are closer to a strong positive cluster than the average pairwise distance between all cluster centers. These clusters are likely to contain images not exactly within the narrow domain of the homogeneous seed dataset but still depict objects from the wider domain the seed dataset defines. These images are valuable in a training set, as they increase the diversity of representations a model can learn for a given class.

Our filtering method does not rely on relatively low levels of label noise to iteratively build an initial model with which non-noisy images are selected. In contrast, the clustering aspect is entirely independent of labels and, therefore, of the level of label noise. Furthermore, with higher levels of label noise, features from noisy images are more likely to be identified as negative clusters. Our previous work showed that, indeed our method is robust to very high levels of label noise [Bö21].

3.3 Cross-class Noise Filtering and Duplicate Detection

Cross-class label noise is arguably the most problematic in a fine-grained domain where only experts can identify wrongly labeled instances. This problem was handled in [Kr16] by filtering out all images that had an exact duplicate in a different class of the augment set.

This approach does not precisely filter cross-class noise but instead mitigates the issue by removing ambiguous images. We expand on this idea by also taking near-duplicates among different classes into account.

We utilize two similarity measures for our near-duplicate detection method. First, the structural similarity index (SSIM) [Wa04] is used, a pixel-based similarity measure taking luminance, structure, and contrast distortions into account. Second, we use the dot product (Dot) of L^2 -normalized feature vectors extracted for the images, which is equivalent to the cosine similarity of the features, which was used by Barz et al. [BD20] for their duplicate detection method. The idea behind using these two similarity measures is that the first (SSIM) will be able to detect small transformations applied to an original image, while the second (Dot) will catch broader transformations, which the SSIM does not pick up.

The similarity measures are used to rank the images in the augment dataset, such that images at the top of the ranking are more likely to have a duplicate in a different class of the augment dataset. Further details on how this ranking is extracted using the similarity measures are found in our previous work [Bö21]. The computation of the similarity measures is omitted for images with exact duplicates identified by comparing MD5 hashes⁸ of the images. The parameter *portion* (p) indirectly specifies how many images are discarded from the top of the ranking in proportion to the number of exact duplicates. When $p = 0.0$, only exact duplicates are discarded in cross-class noise filtering. This parameter *portion* defines some prior knowledge about the level of near duplicates in comparison to exact duplicates.

4 Experiments and Results

The effect of differently sized training datasets and different aspects of the utilization of web images are examined in our experiments. All of them are performed using the EU-Moths dataset described in Section 4.1. In Section 4.2, we analyze the effectiveness of web images and the impact of the domain shift. In Section 4.3, we use manually acquired annotations of cross-domain noise in the downloaded data to evaluate the effect of noise on classification accuracies. Finally, we discuss the effectiveness of filtering strategies in Section 4.4.

4.1 Datasets

The **EU-Moths Dataset**⁹ contains manually taken images of 200 common moth species found in Central Europe collected by the Zoological Research Museum Alexander König, Bonn. The moths were photographed on a rather homogeneous, mostly white background, sometimes together with multiple other insects. The entire dataset consists of approximately 11 images per class. We cropped out the ground truth bounding boxes such that a moth fills

⁸ <https://tools.ietf.org/html/rfc1321>

⁹ http://www.inf-cv.uni-jena.de/eu_moths_dataset.html



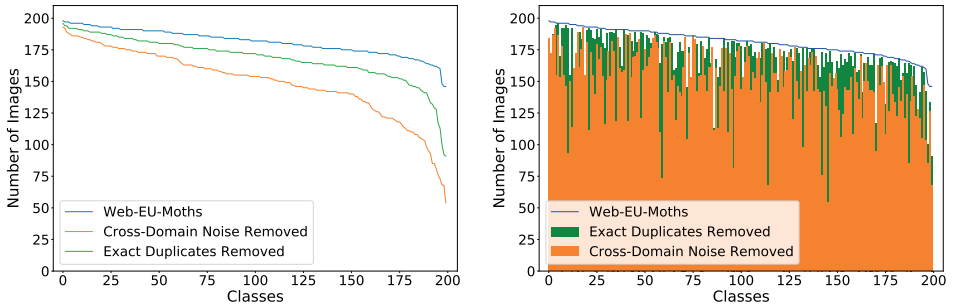
Fig. 1: Example images from the EU-Moths dataset from different classes.

an entire image. Example images from the original dataset are found in Fig. 1. From this dataset, we generated four data splits. Two consist of roughly three images per class in the training and eight in the testing dataset. The other two data splits encompass eight images in the training and three in the testing dataset. As further described in the following, we used these splits to verify the results in two scenarios: (1) an extremely small dataset, with three training images per class only (*three-seed* scenario), and (2) a moderate dataset with eight images per class (*eight-seed* scenario). The four training subsets in these data splits are also referred to as *seed* datasets in the following.

The **Web-EU-Moths Dataset** was compiled by using species names of the 200 moths to query Google Image Search. The links to the images can be found on the dataset website⁹. We downloaded up to 200 images for each class as described in Section 3.1. We call this dataset the *augment dataset* and it consists of 36,274 images in total. The blue line in Fig. 2a shows the sorted distribution of images over the classes in the Web-EU-Moths dataset. This dataset is not balanced because the number of available images downloaded from the Internet varies for different categories. The images in this Web-EU-Moths dataset contain the two discussed types of label noise. Furthermore, the moth images in the Web-EU-Moths dataset cover a larger domain than the narrow domain defined by the EU-Moths dataset of close-up crops of moths photographed on white backgrounds. The images of moths in the Web-EU-Moths dataset were photographed at different angles, on a wider variety of backgrounds, and at different scales. This is the so-called domain shift, and its impact is discussed in the following section.

4.2 The Effect of Webly Annotated Images

Different Learning Approaches Using the seed EU-Moths dataset for training is an example of *supervised* learning, where the training process is supervised by reliable labels. Using web data such as the augment Web-EU-Moths dataset by itself for training is often



(a) Sorted distributions of the Web-EU-Moths dataset, the cleaned version called *Web-EU-Moths dataset where cross-domain noise was identified manually, and the Web-EU-Moths dataset without exact duplicates (cross-class filtering with $p=0.0$).

(b) The number of images per class in the *Web-EU-Moths dataset (Cross-Domain Noise Removed) and dataset where exact duplicates are removed (cross-class noise filtering with $p=0.0$). Some classes are more affected by label noise resulting in class imbalance.

Fig. 2: The data distribution over classes in the Web-EU-Moths dataset and cleaned versions of it, where label noise has been removed

referred to as *webly-supervised* learning, where label noise is an issue. When using a reliably labeled dataset together with web data for training, this is arguably a semi-supervised approach to learning, where some labels are correct and others are not. Therefore, when we augment the EU-Moths seed datasets using data from the Web-EU-Moths dataset for training, we call this a semi-supervised learning approach. In this section, we compare semi-supervised learning with supervised and webly-supervised learning.

All following experiments using these different learning approaches were conducted with an InceptionV3 [Sz16] CNN architecture pre-trained on ImageNet [Ru15]. We repeated each experimental setup four times, each classifier is trained for 30 epochs with an initial learning rate of 1×10^{-4} which was lowered by a factor of 10 after 15 and 25 epochs. We also compare the effect of differently sized seed datasets as described in Section 4.1 for the supervised (training with seed data only) and semi-supervised (training with seed and web data) experiments. The reliably labeled EU-Moths seed datasets used in these experiments either have three or eight images per class in the training dataset. The results in Table 1 are the averaged accuracies for the two splits for each of the two scenarios (Sect. 4.1).

Comparing accuracies for the three-seed with the eight-seed scenario of the supervised learning setting (EU-Moths), we see that the difference of 1,101 training images causes a significant increase from 65.69% to 91.01%. This shows that even relatively few images (compared to modern benchmark datasets) have an immense impact when the initial situation involves an extremely small dataset, such as in the three-seed scenario.

These accuracies are exceeded when using the Web-EU-Moths dataset for training in a webly-supervised learning approach (second row in Table 1). The relative improvement of the webly-supervised learning compared to supervised learning is less pronounced in the

TRAINING DATASET	THREE-SEED			EIGHT-SEED		
	ACC \pm STD	#IMG	%RRE	ACC \pm STD	#IMG	%RRE
EU-MOTHS	65.69% \pm 0.81	552		91.01% \pm 0.40	1,653	
WEB-EU-MOTHS	94.82% \pm 0.27	36,274	84.90%	94.80% \pm 0.58	36,274	42.16%
EU-MOTHS+WEB-EU-MOTHS	96.72% \pm 0.27	36,826	36.68%	97.94% \pm 0.18	37,927	60.38%
EU-MOTHS+*WEB-EU-MOTHS	96.88% \pm 0.32	30,491	4.88%	97.98% \pm 0.14	31,592	1.94%

Tab. 1: This table shows the average accuracies and standard deviations, as well as the number of training images and relative reduction of the error (RRE), compared to the learning approach in the row above. Training with the EU-Moths data is a supervised learning strategy, training on the Web-EU-Moths dataset is considered a webly-supervised approach, and training on the union of both is a semi-supervised approach. The *Web-EU-Moths dataset does not contain any cross-domain noise.

eight-seed scenario than in the three-seed scenario. However, in both cases, training on the large web dataset allows for better performance even though this web data has noisy labels.

In the semi-supervised scenario, the noisy Web-EU-Moths dataset is merged with three and eight images per class from the EU-Moths seed datasets, respectively. The increase of data amounts to 1.52 % in the three-seed scenario and 4.56 % in the eight-seed scenario compared to the Web-EU-Moths dataset. This rather small increase in training data yields an impressive improvement of the classification accuracy compared to the webly-supervised learning of roughly 2 % in the three-seed and 3 % in the eight-seed scenario.

Domain Shift in Web Data As expected, our results show that the amount of data used for training impacts the classification accuracy, but the *type* of data used matters as well, which is especially evident when comparing the results for the different learning strategies in the eight-seed scenario. We find that reasonable accuracies (>90 %) can already be achieved in a supervised learning approach when using a dataset of moderate size (eight-seed scenario). With a substantially larger dataset acquired from the Internet for training, the increase in accuracy is comparably low considering that over twenty times more training data is used. This is surprising, especially when considering the roughly 5 % increase in data used for training in the semi-supervised approach yields an equally large gain in performance. All this suggests that not only the total amount of data used for training but also the type of data plays an important role for the achievable classification accuracy.

These observations are explained by the fact that the training data used in the supervised strategy originates from the same source as the test data (EU-Moths dataset), while the web images have a wide variety of sources. As mentioned earlier, the Web-EU-Moths dataset has a shifted domain compared to the validation subsets of the EU-Moths dataset. The fact that the semi-supervised strategy is superior to the webly-supervised training using relatively little additional data suggests that few images from the same domain as the validation dataset can already counteract this domain shift in the Web-EU-Moths dataset.

	THREE SEED			EIGHT SEED		
	ACC \pm STD	IMAGES	%RRE	ACC \pm STD	IMAGES	%RRE
WEB-EU-MOTHS SUBSET (WEMS)	94.53% \pm 0.25	35,178		95.17% \pm 0.73	35,178	
WEMS + EU-MOTHS	96.66% \pm 0.24	35,730	38.94%	98.46% \pm 0.16	36,831	68.12%
WEMS + MOCK SEED DATASET	94.67% \pm 0.19	35,730	2.56%	95.53% \pm 0.34	36,831	7.45%
EU-MOTHS	65.81% \pm 0.48	552		90.52% \pm 0.51	1,653	
MOCK SEED DATASET	20.87% \pm 0.88	552		53.78% \pm 1.62	1,653	

Tab. 2: Comparing the webly-supervised approach using a subset of the Web-EU-Moths dataset with the two semi-supervised approaches were correctly labeled web images in the mock seed dataset (different domain as the validation data or images from the EU-Moths dataset (the same domain as the validation data) are used as seed datasets for training. The relative reduction of error rate compares the error rates of the semi-supervised approach with the webly-supervised approach in the first row.

To verify this and to make sure the boost in performance does not simply result from the extra data, further experiments were conducted. We compare the effect of additional data that originates from the validation dataset domain with additional data from the web. In these experiments, the exact number of images used for training is controlled. For each of the four EU-Moths seed datasets, we create a *mock* seed dataset with images from the Web-EU-Moths dataset and the same number of images per class to replace the EU-Moths seed datasets in the semi-supervised learning approach. The mock seed datasets were drawn from the pool of images identified manually as belonging to the domain (Sect. 4.3). The subset of images of the Web-EU-Moths dataset not used for the mock datasets is used to train a model, which acts as a baseline for this experiment. Then, the accuracies achieved when using the union of a mock dataset with the Web-EU-Moths subset for training are compared with using the original EU-Moths dataset with the same subset and evaluated on the same test data.

The results in Table 2 show that adding the mock seed data to the Web-EU-Moths Subset yields a smaller gain in performance than adding the EU-Moths seed data for training even though the merged datasets consist of exactly the same number of training images. Especially, the substantial difference in the relative reduction of error rate (RRE) compared to the error rate when using the Web-EU-Moths dataset in the webly-supervised approach underlines the superiority of the semi-supervised approach using data from the evaluation domain. The only explanation for this drastically reduced error compared to the mock datasets is that the few images in the EU-Moths dataset are enough to counteract the domain shift in the web images. This domain shift is the reason why using the large Web-EU-Moths dataset in the webly-supervised approach does not yield higher accuracies. However, our results show that this domain shift can be counteracted with very few images from the source domain resulting in higher predictive accuracies.

In summary, web data can be used to boost accuracies, especially when the existing data is extremely limited, and using a semi-supervised approach with few images from the domain of the validation data is advisable.

4.3 The Effect of Cross-domain Noise

To further understand the impact of web data in a training set on the classifier performance, we need to understand the impact of label noise. We want to compare training using noisy web data with training on clean web data. Since only experts could reliably identify cross-class noise, we limit our focus to cross-domain noise. We build a simple annotation tool to obtain *out-of-domain* vs. *within-domain* binary labels for the images in the Web-EU-Moths dataset. The user clicked only on the out-of-domain samples when looking at a small set of sixteen images in a single panel. Thus, images of maps, habitats, or caterpillars, as well as sketches or close-ups of moths' *heads*, can quickly be labeled as cross-domain noise. Using the binary labels to remove cross-domain noise resulted in the clean *Web-EU-Moths dataset. The distribution of this cleaned dataset is more imbalanced than the original Web-EU-Moths dataset, which is visualized in Fig. 2a. With the manual labeling process, we identified 6,335 images (17.5 %) in the augmented Web-EU-Moths dataset as outside of the moth domain. This level is relatively low, which might be explained, by the fact, that we only downloaded up to two hundred images per class. Also, the species in our EU-Moths dataset are pretty common and well studied, resulting in many images being available across the Internet.

We also use this cleaned version of the Web-EU-Moths dataset in a semi-supervised approach as described in Section 4.2. The results for using this cleaned *Web-EU-Moths dataset together with the EU-Moths training datasets for the two scenarios are found in Table 1. The manually cleaned web images do not yield significant performance gains compared to the simple merge of seed and noisy data. This observation holds independently of the number of images in the seed dataset.

Several studies [Ro17, FP17, Zh16] have shown how deep neural networks are surprisingly robust against moderate levels of label noise in coarse-grained classification problems. Our results indicate that models trained for fine-grained classification problems are also robust against label noise. An explanation was offered by [Zh16] for the robustness of deep neural networks. They claim that highly parameterized models such as deep neural networks (DNNs) have the capacity to learn and generalize from the non-noisy images and use brute force to memorize the noisy samples. This means that the quasi memorized noisy samples in the training dataset do not affect the classification accuracy of the test data, especially when the test set is free of out-of-domain images. Thus, the memorized noise does not influence the prediction of noise-free samples during the application.

The increased class imbalance in the clean *Web-EU-Moths dataset might also explain why the accuracies do not differ more when comparing training with and without label noise. Furthermore, label noise might have a positive regularization effect, which would explain why the predictive accuracy is equivalently high when training with noisy data.

	THREE-SEED			EIGHT-SEED		
	PRECISION	RECALL	F1-SCORE	PRECISION	RECALL	F1-SCORE
CROSS-DOMAIN FILTERS						
$k=50$	0.87446	0.96894	0.919	87.990	96.633	0.921
$k=75$	0.87751	0.95731	0.916	88.088	95.932	0.918
$k=100$	0.86520	0.97431	0.917	88.673	96.167	0.923
$k=125$	0.87696	0.97901	0.925	89.118	96.682	0.927
CROSS-CLASS FILTERS						
$p=0.0$	0.82919	0.93397	0.878	82.919	93.397	0.878
$p=0.01$	0.82965	0.93377	0.879	82.965	93.377	0.879
$p=0.05$	0.83184	0.93336	0.880	83.184	93.336	0.88

Tab. 3: This table shows the precision, recall, and F1-score of different filter setups. Since the cross-class noise filters are entirely independent of the seed dataset, these values are the same for both seed data scenarios.

4.4 The Effect of Filters

In this section, our proposed filtering methods are analyzed. Intuitively, the semi-supervised approach using the *Web-EU-Moths dataset should act as an upper bound for the following setups, where automatically filtered subsets of the augment data are merged with the seed datasets for training. We will discuss the effectiveness of our filtering methods, even though the expected effect on accuracies in our particular application of moth species classification is negligible. Our experiments show, using filtered datasets results in comparable accuracy with far less training data, which reduces training time.

We test the clustering-based cross-domain noise filtering technique with four seed datasets of the EU-Moths dataset, which guide the selection of positive clusters. In addition, the parameter for the number of clusters k in the cross-domain noise filtering method has been set to one of the following four values: 50, 75, 100, or 125. An overview of precision, recall, and F1-score for the different filter settings averaged over two dataset splits for the three-seed and the eight-seed scenario are given in Table 3. The precision refers to the percentage of images retained by the filter that are non-noisy ones. Recall describes the percentage of all non-noisy images, i.e., images depicting moths, retained by a given filter. Since many studies [Ro17, FP17, Zh16] point towards the minimal adverse effect of label noise, it is arguably more critical to retain a large portion of the non-noisy data than to ensure the retained data contains little noise. Therefore, high recall is more important when comparing different hyperparameter settings for the filters.

Overall, we found a high overlap of retained images between different filters using different values of k and differing seed datasets. This indicates the robustness of the proposed filtering method, given different hyperparameter settings and seed datasets. The average intersection-over-union ratio of retained images between all pairs of filters is 94.52 % with

	THREE-SEED			EIGHT-SEED		
	ACC \pm STD	#IMG	%RET	ACC \pm STD	#IMG	%RET
EU-MOTHS+ WEB-EU-MOTHS	96.72% \pm 0.27	36826	100.00%	97.94% \pm 0.18	37927	100.00%
EU-MOTHS+FILTERED WEB-EU-MOTHS						
CROSS-DOMAIN FILTERS						
$\kappa=50$	96.81% \pm 0.18	33,729	91.46%	98.05% \pm 0.36	34,565	90.74%
$\kappa=75$	96.75% \pm 0.14	33,214	90.04%	97.62% \pm 0.36	34,258	89.88%
$\kappa=100$	96.84% \pm 0.16	34,267	92.94%	97.67% \pm 0.56	34,123	89.52%
$\kappa=125$	96.90% \pm 0.31	33,975	92.14%	97.96% \pm 0.11	34,136	89.55%
CROSS-CLASS FILTERS						
$p=0.0$	96.74% \pm 0.32	34,274	92.96%	97.92% \pm 0.46	35,375	92.96%
$p=0.01$	96.85% \pm 0.22	34,248	92.89%	97.76% \pm 0.56	35,349	92.89%
$p=0.05$	96.92% \pm 0.29	34,145	92.61%	97.84% \pm 0.37	35,246	92.61%
COMBINED FILTERS						
$\kappa=50$ $p=0.0$	96.98% \pm 0.28	31,438	85.15%	97.86% \pm 0.48	32,282	84.44%
$\kappa=75$ $p=0.0$	96.96% \pm 0.18	30,958	83.82%	97.85% \pm 0.38	31,974	83.59%
$\kappa=100$ $p=0.0$	96.93% \pm 0.22	31,980	86.64%	98.01% \pm 0.32	31,891	83.36%
$\kappa=125$ $p=0.0$	96.92% \pm 0.10	31,697	85.86%	98.05% \pm 0.37	31,896	83.38%
$\kappa=50$ $p=0.01$	97.00% \pm 0.26	31,432	85.13%	97.98% \pm 0.34	32,276	84.43%
$\kappa=75$ $p=0.01$	96.99% \pm 0.30	30,952	83.81%	98.05% \pm 0.22	31,968	83.58%
$\kappa=100$ $p=0.01$	96.90% \pm 0.36	31,976	86.63%	97.72% \pm 0.50	31,885	83.34%
$\kappa=125$ $p=0.01$	96.87% \pm 0.22	31,692	85.84%	97.85% \pm 0.44	31,890	83.36%
$\kappa=50$ $p=0.05$	97.03% \pm 0.22	31,416	85.09%	98.10% \pm 0.26	32,260	84.38%
$\kappa=75$ $p=0.05$	97.07% \pm 0.31	30,936	83.76%	97.94% \pm 0.20	31,952	83.53%
$\kappa=100$ $p=0.05$	97.00% \pm 0.35	31,960	86.59%	97.96% \pm 0.22	31,869	83.30%
$\kappa=125$ $p=0.05$	96.90% \pm 0.30	31,676	85.80%	98.10% \pm 0.36	31,874	83.32%

Tab. 4: Impact of different filter options on the classification accuracy when training a moth species classifier with a semi-supervised approach using both the EU-Moths dataset and the Web-EU-Moths dataset. The retention rate (ret%) refers to the percentage of web images retained by the filter. The first row shows the results when using the unfiltered web data for comparison

a standard deviation of $\pm 1.52\%$. This means that on average, two filters *agreed* on more than 94% images that should be retained. On the other hand, the overlap was lower for the noisy images with much more variation. Here, the intersection-over-union ratio of discarded images of two noise filters was on average $56.75\% \pm 9.88\%$. This lower level of overlap and higher variation is explained by the fact that the total number of filtered images was low in all setups, making the statistics more sensitive to small differences. Overall, the similar F1-scores also show that the cross-domain filtering techniques are robust to different hyperparameters.

Since evaluating cross-class noise filtering methods would require expert knowledge, it is out of the scope of this work. However, we can assess how well the method for cross-class noise filtering is suited for identifying cross-domain noise. The idea is that out-of-domain

images might be downloaded multiple times in different classes, which means the cross-class noise filtering method based on duplicate detection might be effective for filtering these instances. The lower half in Table 3 reveals that even though the cross-class noise filtering method was not designed to do so, it is surprisingly effective at handling cross-domain noise. This confirms the hypothesis that many noisy images are both out-of-distribution and downloaded multiple times for different classes. As our previous work [Bö21] shows, it is advantageous to set the parameter portion (p), which specifies how many images are removed (Section 3.3), to a low value for this domain. Hence, we evaluate the cross-class noise filtering with the values 0.0, 0.01, and 0.05 for the parameter p . With these low values for p , only a few images are discarded by cross-class noise filtering. When filtering out only exact duplicates ($p = 0.0$), we filter out the minimum cross-class noise. Similar to the cleaned Web-EU-Moths dataset without cross-domain noise, the data distribution of the web dataset where the minimum cross-class noise removed is visualized in Figure 2.

We used the filtered subsets of the Web-EU-Moths dataset to train a model in a semi-supervised manner, where seed datasets of different sizes (three-seed and eight-seed) are merged with the retained images (Table 4). We average the results from two splits of the same size for each of the different scenarios. In the experiments in which cross-domain noise and cross-class noise filters are combined, only those images are retained that both filters consistently identify as non-noisy, i.e., images are discarded if at least one filter identifies it as noise. As expected, there is no clear trend towards a performance boost when using only filtered data for training. However, the classification accuracies remain stable for all filter setups, even when less data is used for training the classifier. This shows that the noise filtering techniques succeed in filtering out irrelevant images and are robust with respect to different hyperparameter settings. Therefore, in situations where the level of noise is higher, our methods are likely to be beneficial in terms of predictive accuracy.

5 Conclusion

Our work shows the benefits of additional training datasets acquired through image search engines from the Internet for learning a classification model. The conducted experiments demonstrate that a semi-supervised learning approach, where web images together with images of the evaluation domain are used for training, can counteract the domain shift in web data. This approach yields higher accuracies than a webly-supervised approach, i.e., training using web data only, and a supervised approach, i.e., training on small seed datasets only. We further found evidence that deep neural networks trained for fine-grained classification tasks are robust to label noise, and removing noise does not significantly improve classification performances in our application. Finally, we studied the effect of noise filtering techniques using ground truth annotations of cross-domain label noise and found that the simple to implement methods we proposed are effective in identifying noisy samples.

Acknowledgements

This work has been funded by the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung, BMBF, Deutschland) via the project "Development of an Automated Multisensor Station for Monitoring of Biodiversity (AMMOD) - Subproject 5: Automated Visual Monitoring and Analysis"(FKZ: 01LC1903E).

Bibliography

- [BD20] Barz, Björn; Denzler, Joachim: Do we train on test data? Purging CIFAR of near-duplicates. *Journal of Imaging*, 6(6):41, 2020.
- [Bö21] Böhlke, Julia; Korsch, Dimitri; Bodesheim, Paul; Denzler, Joachim: Lightweight Filtering of Noisy Web Data: Augmenting Fine-grained Datasets with Selected Internet Images. In: *VISAPP*. 2021.
- [CSG13] Chen, Xinlei; Shrivastava, Abhinav; Gupta, Abhinav; Neil: Extracting visual knowledge from web data. In: *IEEE International Conference on Computer Vision*. pp. 1409–1416, 2013.
- [Cu16] Cui, Yin; Zhou, Feng; Lin, Yuanqing; Belongie, Serge J.: Fine-grained Categorization and Dataset Bootstrapping using Deep Metric Learning with Humans in the Loop. In: *IEEE Conference on Computer Vision and Pattern Recognition*. pp. 1153–1162, 2016.
- [Cu18] Cui, Yin; Song, Yang; Sun, Chen; Howard, Andrew; Belongie, Serge: Large scale fine-grained categorization and domain-specific transfer learning. In: *IEEE Conference on Computer Vision and Pattern Recognition*. pp. 4109–4118, 2018.
- [FP17] Flatow, David; Penner, Daniel: , On the robustness of convnets to training on noisy labels, 2017.
- [Ga16] Garcia, Luís PF; Lorena, Ana C; Matwin, Stan; de Carvalho, André CPLF: Ensembles of label noise filters: a ranking approach. *Data Mining and Knowledge Discovery*, 30(5):1192–1216, 2016.
- [GLY19] Ge, Weifeng; Lin, Xiangru; Yu, Yizhou: Weakly supervised complementary parts models for fine-grained image classification from the bottom up. In: *IEEE Conference on Computer Vision and Pattern Recognition*. pp. 3034–3043, 2019.
- [Ha18] Han, Bo; Yao, Quanming; Yu, Xingrui; Niu, Gang; Xu, Miao; Hu, Weihua; Tsang, Ivor; Sugiyama, Masashi: Co-teaching: Robust training of deep neural networks with extremely noisy labels. In: *Advances in Neural Information Processing Systems (NeurIPS)*. 2018.
- [Ha19] Habel, Jan Christian; Trusch, Robert; Schmitt, Thomas; Ochse, Michael; Ulrich, Werner: Long-term large-scale decline in relative abundances of butterfly and burnet moth species across south-western Germany. *Scientific reports*, 9(1):1–9, 2019.
- [Ji18] Jiang, Lu; Zhou, Zhengyuan; Leung, Thomas; Li, Li-Jia; Fei-Fei, Li: MentorNet: Learning Data-Driven Curriculum for Very Deep Neural Networks on Corrupted Labels. In: *Proceedings of the International Conference on Machine Learning*. 2018.

- [Kä16] Käding, Christoph; Rodner, Erik; Freytag, Alexander; Denzler, Joachim: Active and continuous exploration with deep neural networks and expected model output changes. arXiv preprint arXiv:1612.06129, 2016.
- [KBD19] Korsch, Dimitri; Bodesheim, Paul; Denzler, Joachim: Classification-Specific Parts for Improving Fine-Grained Visual Categorization. In: German Conference on Pattern Recognition. pp. 62–75, 2019.
- [Ke04] Ke, Yan; Sukthankar, Rahul; Huston, Larry; Ke, Yan; Sukthankar, Rahul: Efficient near-duplicate detection and sub-image retrieval. In: ACM Multimedia. volume 4, p. 5, 2004.
- [Kr16] Krause, Jonathan; Sapp, Benjamin; Howard, Andrew; Zhou, Howard; Toshev, Alexander; Duerig, Tom; Philbin, James; Fei-Fei, Li: The unreasonable effectiveness of noisy data for fine-grained recognition. In: European Conference on Computer Vision. pp. 301–320, 2016.
- [Li21] Liu, Huafeng; Zhang, Chuanyi; Yao, Yazhou; Wei, Xiushen; Shen, Fumin; Zhang, Jian; Tang, Zhenmin: Exploiting Web Images for Fine-Grained Visual Recognition by Eliminating Noisy Samples and Utilizing Hard Ones. arXiv preprint arXiv:2101.09412, 2021.
- [LRM15] Lin, Tsung-Yu; RoyChowdhury, Aruni; Maji, Subhansu: Bilinear cnn models for fine-grained visual recognition. In: IEEE International Conference on Computer Vision. pp. 1449–1457, 2015.
- [MSS17] Malach, Eran; Shalev-Shwartz, Shai: Decoupling "when to update" from "how to update". In: Advances in Neural Information Processing Systems (NeurIPS). 2017.
- [Ni15] Nicholson, Bryce; Zhang, Jing; Sheng, Victor S; Wang, Zhiheng: Label noise correction methods. In: IEEE International Conference on Data Science and Advanced Analytics. pp. 1–9, 2015.
- [Pe20] Peng, Xiaojiang; Wang, Kai; Zeng, Zhaoyang; Li, Qing; Yang, Jianfei; Qiao, Yu: Suppressing Mislabeled Data via Grouping and Self-Attention. In: European Conference on Computer Vision. Springer, pp. 786–802, 2020.
- [Re18] Ren, Mengye; Zeng, Wenyuan; Yang, Bin; Urtasun, Raquel: Learning to Reweight Examples for Robust Deep Learning. In: Proceedings of the International Conference on Machine Learning. 2018.
- [Ro17] Rolnick, David; Veit, Andreas; Belongie, Serge; Shavit, Nir: Deep learning is robust to massive label noise. arXiv preprint arXiv:1705.10694, 2017.
- [Ru15] Russakovsky, Olga; Deng, Jia; Su, Hao; Krause, Jonathan; Satheesh, Sanjeev; Ma, Sean; Huang, Zhiheng; Karpathy, Andrej; Khosla, Aditya; Bernstein, Michael; Berg, Alexander C.; Fei-Fei, Li: ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision, 115(3):211–252, 2015.
- [Si17] Simon, Marcel; Gao, Yang; Darrell, Trevor; Denzler, Joachim; Rodner, Erik: Generalized orderless pooling performs implicit salient matching. In: IEEE International Conference on Computer Vision. pp. 4970–4979, 2017.
- [Sz16] Szegedy, Christian; Vanhoucke, Vincent; Ioffe, Sergey; Shlens, Jon; Wojna, Zbigniew: Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 2818–2826, 2016.

- [Wa04] Wang, Zhou; Bovik, Alan C; Sheikh, Hamid R; Simoncelli, Eero P: Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.
- [Wa06] Wang, Bin; Li, Zhiwei; Li, Mingjing; Ma, Wei-Ying: Large-scale duplicate detection for web image search. In: *IEEE International Conference on Multimedia and Expo*. pp. 353–356, 2006.
- [Wa11] Wah, C.; Branson, S.; Welinder, P.; Perona, P.; Belongie, S.: The Caltech-UCSD Birds-200-2011 Dataset. Technical Report CNS-TR-2011-001, California Institute of Technology, 2011.
- [Wa14] Wang, Jiang; Song, Yang; Leung, Thomas; Rosenberg, Chuck; Wang, Jingbin; Philbin, James; Chen, Bo; Wu, Ying: Learning fine-grained image similarity with deep ranking. In: *IEEE Conference on Computer Vision and Pattern Recognition*. pp. 1386–1393, 2014.
- [Wa21] Wagner, David L; Fox, Richard; Salcido, Danielle M; Dyer, Lee A: A window to the world of global insect declines: Moth biodiversity trends are complex and heterogeneous. *Proceedings of the National Academy of Sciences*, 118(2), 2021.
- [Xu15] Xu, Zhe; Huang, Shaoli; Zhang, Ya; Tao, Dacheng: Augmenting Strong Supervision Using Web Data for Fine-Grained Categorization. In: *IEEE International Conference on Computer Vision*. December 2015.
- [Xu18] Xu, Zhe; Huang, Shaoli; Zhang, Ya; Tao, Dacheng: Webly-Supervised Fine-Grained Visual Categorization via Deep Domain Adaptation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(5):1100–1113, 2018.
- [Ya20] Yao, Yazhou; Hua, Xiansheng; Gao, Guanyu; Sun, Zeren; Li, Zhibin; Zhang, Jian: Bridging the web data and fine-grained visual recognition via alleviating label noise and domain mismatch. In: *Proceedings of the 28th ACM International Conference on Multimedia*. pp. 1735–1744, 2020.
- [Zh16] Zhang, Chiyuan; Bengio, Samy; Hardt, Moritz; Recht, Benjamin; Vinyals, Oriol: Understanding deep learning requires rethinking generalization. *arXiv preprint arXiv:1611.03530*, 2016.
- [Zh17] Zhuang, Bohan; Liu, Lingqiao; Li, Yao; Shen, Chunhua; Reid, Ian: Attend in groups: a weakly-supervised deep learning framework for learning from web data. In: *IEEE Conference on Computer Vision and Pattern Recognition*. pp. 1878–1887, 2017.
- [Zh20a] Zhang, Chuanyi; Yao, Yazhou; Shu, Xiangbo; Li, Zechao; Tang, Zhenmin; Wu, Qi: Data-driven meta-set based fine-grained visual recognition. In: *Proceedings of the 28th ACM International Conference on Multimedia*. pp. 2372–2381, 2020.
- [Zh20b] Zhang, Chuanyi; Yao, Yazhou; Zhang, Jiachao; Chen, Jiaxin; Huang, Pu; Zhang, Jian; Tang, Zhenmin: Web-Supervised Network for Fine-Grained Visual Classification. In: *IEEE International Conference on Multimedia and Expo*. pp. 1–6, 2020.
- [ZT19] Zhang, Weining; Tan, Xiaoyang: Combining Outlier Detection and Reconstruction Error Minimization for Label Noise Reduction. In: *IEEE International Conference on Big Data and Smart Computing*. pp. 1–4, 2019.

Automatic Plant Cover Estimation with Convolutional Neural Networks

Matthias Körschens^{1,2}; Paul Bodesheim³; Christine Römermann⁴; Solveig Franziska Bucher⁵; Mirco Migliavacca⁶; Josephine Ulrich⁷; Joachim Denzler⁸

Abstract: Monitoring the responses of plants to environmental changes is essential for plant biodiversity research. This, however, is currently still being done manually by botanists in the field. This work is very laborious, and the data obtained is, though following a standardized method to estimate plant coverage, usually subjective and has a coarse temporal resolution. To remedy these caveats, we investigate approaches using convolutional neural networks (CNNs) to automatically extract the relevant data from images, focusing on plant community composition and species coverages of 9 herbaceous plant species. To this end, we investigate several standard CNN architectures and different pretraining methods. We find that we outperform our previous approach at higher image resolutions using a custom CNN with a mean absolute error of 5.16%. In addition to these investigations, we also conduct an error analysis based on the temporal aspect of the plant cover images. This analysis gives insight into where problems for automatic approaches lie, like occlusion and likely misclassifications caused by temporal changes.

Keywords: Computer Vision; Biodiversity; Deep Learning; Plant Cover

1 Introduction

Environmental changes can have a large impact not only on human life, but also on animals and plants, and can have drastic effects on biodiversity. This gives reason to steadily monitor nature for such changes and their impacts. Plants, for example, are strong indicators of environmental changes like climate change. This can be seen in phenological responses [Ro07, Me06, Bu18, MRP08, Cl12, FF02], and also in changes in plant community compositions [Ro07, Li18, LI09]. However, even more environmental aspects can be

¹ Computer Vision Group, Friedrich Schiller University Jena, 07737 Jena, Germany; matthias.koerschens@uni-jena.de

² Plant Biodiversity Group, Friedrich Schiller University Jena, 07737 Jena, Germany

³ Computer Vision Group, Friedrich Schiller University Jena, 07737 Jena, Germany; paul.bodesheim@uni-jena.de

⁴ Plant Biodiversity Group, Friedrich Schiller University Jena, 07737 Jena, Germany; christine.roemermann@uni-jena.de

⁵ Plant Biodiversity Group, Friedrich Schiller University Jena, 07737 Jena, Germany; solveig.franziska.bucher@uni-jena.de

⁶ Department Biogeochemical Integration, Max Planck Institute for Biogeochemistry, 07701 Jena, Germany; mmigliab@bgc-jena.mpg.de

⁷ Plant Biodiversity Group, Friedrich Schiller University Jena, 07737 Jena, Germany; josephine.ulrich@uni-jena.de

⁸ Computer Vision Group, Friedrich Schiller University Jena, 07737 Jena, Germany; joachim.denzler@uni-jena.de

monitored through plants, like land use [Ge14, AC12] and insect abundance [SZS16]. Therefore, plant community compositions are the focus of a large number of experiments [Li18, Ge14, SZS16, Br18], where the composition is usually recorded at regular time intervals. This is usually done by estimating the plant cover of each species in the predefined vegetation plots. The cover of a plant species describes the percentage of soil area covered by the respective species. In many experiments, this value is estimated at intervals of at least about a week, usually much longer, due to it being a very laborious task, and the estimation itself is usually done by a single biologist. This introduces two problems: a low temporal resolution of recordings of the plant community composition and noisy data caused by human error and subjectivity of the estimations. These problems can be mitigated by automatic image analysis methods. However, while in current times there are many possibilities to collect images and videos of vegetation plots in an automated manner (e.g., [Br16]), performing more complicated analyses automatically like the cover estimation mentioned above is still not possible as we are still missing the methods to do so.

Convolutional neural networks (CNNs) have proven to be a very effective means for image analysis in general and hence could also be utilized as a basis for such methods. However, automatic plant cover prediction is a rather complicated task, as it usually includes large images containing many plant species with only subtle differences, which strongly occlude each other. In addition to this, the annotations for plant cover are relative values (percentages), which is a mostly unexplored kind of annotation in the area of computer vision. In our previous work [Kö20] we laid the groundwork to solve the task of plant cover prediction.

In this work, we will investigate that approach, i.e., the cover prediction via pixel-wise classification using the previously proposed network head, in more detail regarding multiple aspects. Firstly, we aim to investigate how different standard CNN architectures with differing characteristics perform on our task in comparison to the network from [Kö20]. The architectures we compare are a ResNet50 [He16], InceptionV3 [Sz16] and DenseNet121 [Ia14]. Secondly, we investigate whether the addition of information from earlier network layers of the abovementioned networks and an additional increase of the output resolution by using a Feature Pyramid Network [Li17a] can improve the performance in comparison to the networks' standalone performance. Thirdly, as changing the image resolution during training and evaluation can strongly affect the network performance, for example, due to the limitations of the networks' receptive field size, we evaluate the performance of the abovementioned CNN architectures on different image resolutions. Fourthly, as pretraining in a similar domain to the target task usually results in better network performances than out-of-domain pretraining [Cu18], we aim to find out how large the benefits of this method are for plant cover prediction. To this end, we compare the previously mentioned configurations in two different pretraining settings. For our domain-similar ("in-domain") pretraining setting, we constructed a dedicated plant image dataset using images from the GBIF website⁹, here referred to as the GBIF dataset. In our out-of-domain setting, we utilize freely available off-the-shelf Imagenet [Ru15] pretraining weights. Lastly, to explain

⁹ Global Biodiversity Information Facility, <https://www.gbif.org/>

the errors made by the networks during the plant cover prediction, we conduct a temporal analysis of the prediction error of a selected network based on the week of recording of the respective images.

In the following, we will elaborate on related work in the area of automatic plant analysis (section 2), followed by an introduction of the datasets and methods we use in section 3 and section 4, respectively. After that, we present the results of our experiments (section 5) and end the paper with a conclusion and an outlook on possible future work.

2 Related Work

2.1 Plant Identification

In the last years, there have been several CNN-based approaches developed for the identification of plant species from images [YR16, Ba17, Le15, GYA17]. Barré et al. [Ba17] developed a rather simple CNN architecture called LeafNet, whose aim is to determine the plant species from leaf images. Yalcin et al. [YR16] utilize a pretrained CNN with 11 layers to classify images of agricultural plants. However, simple identification of plant species via images is not the only area where Deep Learning approaches like CNNs are applied. Other tasks include the detection of fruits via segmentations [Ga19], counting of fruits of agricultural plants via regression networks [Lu17, Xi19] and plant disease prediction [Ch20]. A more comprehensive list of approaches can be found, for example, in the survey by Kamilaris et al. [KPB18].

2.2 Plant Cover Prediction

While the number of approaches for problems concerning plant identification is rather abundant, only two approaches try to solve the problem of cover prediction via images. The first one is by Kattenborn et al. [Ka20], who tried to determine the cover percentages of woody species from UAV-based remote sensing data. For their analysis, they utilized a simple custom CNN with eight layers to calculate the cover percentages for several herb, shrub and woody species. In contrast to the data used in this paper, there was not only no occlusion to be taken into account, but the annotations were delineations in the images. In addition to that, the woody plant species investigated mostly had heterogeneous appearances, making automatic discrimination between them relatively easy compared to the much more similar herbaceous species found in the InsectArmageddon dataset [UI20, Kö20] we use. The second approach in the area of plant cover prediction is our recent work presented in [Kö20], which we will build on in this paper. In [Kö20] we utilized a novel custom network architecture with 12 layers and a block similar to those of the Inception network [Sz16] for feature aggregation. The respective experiments were also done on the InsectArmageddon dataset from Ulrich et al. [UI20]. In contrast to our previous work, we will not utilize a



Fig. 1: Example images from our GBIF dataset. The plant species shown are from left to right: *Achillea millefolium* (Yarrow), *Centaurea jacea* (brown knapweed), *Plantago lanceolata* (ribwort plantain), *Trifolium pratense* (red clover), *Scorzoneroideides autumnalis* (autumn hawkbit) and *Grasses*, which are not differentiated into different species.

custom network here but evaluate the performance of several standard CNN architectures in conjunction with Feature Pyramid Networks (FPNs) [Li17a]. To this end, we utilize the network head proposed in [Kö20] and hence adopt the pixel-wise classification approach. This way, we aim to determine how well these architectures are suited for this task and if the shortcut connections and higher output resolution added by the FPN can be beneficial for plant cover prediction. In addition to this, we also investigate different image resolutions as well as different types of pretraining, as they can strongly affect the performance of CNNs.

3 Datasets

In our experiments, we utilize two datasets. The first one is used for pretraining and will be referred to as GBIF dataset, as it contains plant images taken from the Global Biodiversity Information Facility (GBIF)¹⁰. The second dataset is the so-called InsectArmageddon dataset introduced in [U120] and [Kö20]. These datasets will be explained in subsection 3.1 and subsection 3.2, respectively.

3.1 The GBIF Dataset

The GBIF dataset is a classification dataset we designed specifically for pretraining models in preparation for using them for cover prediction on the InsectArmageddon dataset. Hence, it contains mostly the same classes as the latter, namely the plant species *Achillea millefolium* (Ach_mil), *Centaurea jacea* (Cen_jac), *Lotus corniculatus* (Lot_cor), *Medicago lupulina* (Med_lup), *Plantago lanceolata* (Pla_lan), *Scorzoneroideides autumnalis* (Sco_aut) and *Trifolium pratense* (Tri_pra), as well as *Festuca Tourn. ex L.*, which represents the collective class of grasses in the InsectArmageddon dataset, in which the individuals are not further identified to the species-level.

¹⁰ <https://www.gbif.org/>



Fig. 2: Example images from the InsectArmageddon dataset. We can see temporal changes in plant coverages in the experimental units of the Ecotron, where different plant species have been grown from seed.

This dataset contains 7,200 images with 750 training and 150 validation images per class and is therefore also balanced. The images were selected randomly from the complete sets of human observations of the respective plant species [GB20]. A selection of example images from the dataset is shown in Figure 1.

3.2 The InsectArmageddon Dataset

The InsectArmageddon dataset was generated during the eponymous project¹¹ in 2018, in which the effect of invertebrate density on plant composition and growth were investigated [UI20]. In this experiment, 24 so-called EcoUnits, i.e., enclosed boxes with a base area of about $1.5\text{ m} \times 1.5\text{ m}$ containing small, closed ecosystems with different experimental treatments, were equipped with two cameras each, collecting images from a height of about 2 m over a duration of 18 weeks. During this time, plants of 9 different species grew in the EcoUnits, resulting in large variations of the plants in the images as seen in Figure 2. The plant species in this dataset are the same as in section 3.1 with the addition of dead biomass, referred to as *Dead litter*, which was introduced due to lack of visual distinguishability in the images. For more information on the InsectArmageddon experiment itself or the image collection process, we would like to refer to [UI20, ET18, Tü17] and [Kö20], respectively.

The images themselves contain the plant species in largely different distributions, partially caused by the different treatments. Therefore, the contents of the images can differ strongly. In addition to this, the dataset is heavily imbalanced. On the one hand, multiple plant species are either very abundant or have large leaf sizes, each leading to large cover percentages. On the other hand, there are also plant species with tiny leaves, leading to excessively small

¹¹ https://www.idiv.de/en/research/platforms_and_networks/idiv_ecotron/experiments/insect_armageddon.html

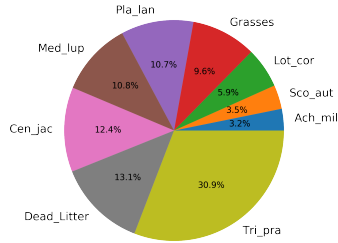


Fig. 3: The distribution of plant species in the InsectArmageddon dataset, displayed as fractions of the total sum of cover percentages over the complete dataset. *Trifolium pratense* is the most abundant plant in the dataset, while *Achillea millefolium* represents the least abundant species.

cover percentages. This distribution is visualized in Figure 3, which shows that the most prevalent plant is *Trifolium pratense*, representing about a third of the complete dataset. In contrast, *Achillea millefolium* can be found on the other end of the distribution with only about 3% cover.

The dataset contains a large number of difficulties for automatic image analysis methods. Due to technical reasons, the zoom levels and image quality are inconsistent within and across the EcoUnits. However, one of the biggest challenges in the dataset is the occlusion caused by massive mutual overlaps of the plants, which is inherent to plant communities, irrespective of the number of co-occurring species. This makes a correct plant cover prediction very difficult.

While images from the EcoUnits have been collected daily, due to the necessity of annotating them, we can only utilize one image per week per camera for our investigations. This leads to only 682 images with annotations, which are vectors containing the estimated plant cover for each plant species of the dataset.

Plant Cover Annotations. The plant cover is defined as the area of ground covered by each plant species. This value is typically estimated by a biologist directly in the field and can thus be highly subjective. Therefore, the annotations of the InsectArmageddon dataset, which were provided by Ulrich et al. [UI20], can be considered noisy. In addition to this, as an accurate percentage estimation is practically infeasible for humans, the estimated plant cover is usually done with a certain quantization, which in the case of this dataset is the so-called Schmidt-scale [Pf97]. The latter includes the percentages 0, 0.5, 1, 3, 5, 8, 10, 15, 20, 25, 30, 40, 50, 60, 70, 75, 80, 90 and 100 percent. The quantization process introduces additional label noise. It should be noted that, due to disregarding the occlusion during the estimation, the sum of the plant cover percentages can exceed 100%.

4 Method

We apply the same approach as in [Kö20]: we view the plant cover prediction problem as a pixel-wise classification problem, where these classifications can then be aggregated into a final cover prediction. One additional advantage of this approach is that the pixel-wise classifications can be viewed as segmentation maps, serving as a way for the user to verify the correctness of the result. Taking this approach requires us to identify the plant species at each location in the images before aggregating them into a single cover value for each species. Figure 2 shows that the plants are, in part, very similar to each other, and, due to the relatively large area captured in the single images, rather small compared to the whole image. In order to still be able to identify the plants correctly and be able to take into account their small and essential features, we need to process the images in a relatively high resolution. Often, when high-resolution images are used during CNN training, the images are divided into smaller parts, which are then used for training (e.g., [He17]). This approach, however, is not possible in our case, as the numerical plant cover annotations are only valid for the complete image. Therefore, we are required to train on high-resolution images directly.

While this was done already in our previous work [Kö20] with a resolution of 672x336 px, here we test the effect of training on even higher resolutions. Hence, we evaluate our method on images with resolutions of 768x384 and 1536x768 px, respectively. To receive a better insight on the effect of different network architectures on the task of plant cover prediction, we also measure the performance of standard classification network architectures in conjunction with a Feature Pyramid Network (FPN) [Li17a]. The networks we will compare are ResNet50 [He16], DenseNet121 [Ja14], and InceptionV3 [Sz16]. ResNet50 was selected as it is used for comparison and as backbone in almost all possible settings and tasks, from simple classification [He16] over object detection [Li17b] to instance segmentation [He17]. InceptionV3, in turn, is commonly used in fine-grained classification tasks (e.g., [Cu18]), and DenseNet121 was chosen due to its particular network structure, which incorporates information from earlier layers directly in the decision process in later layers via long-range skip connections. As these networks only have output feature maps of small sizes, we also apply the FPN, which is usually utilized in tasks like object detection (e.g., [Re15, Li17b]) and semantic segmentation (e.g., [He17]), and could therefore be a sensible choice due to the similarity of the underlying task. The FPN increases the size of the output feature maps by incrementally fusing feature maps from the end of the network with higher resolution ones generated in earlier layers of the network. Hence, this structure also enables us to generate high-resolution output feature maps when joined with the abovementioned standard networks. This also allows us to evaluate whether generating higher resolution class outputs results in a better cover prediction, and to receive information on the usefulness of features from earlier layers.

As mentioned above, we also want to quantify the effect of in-domain and out-of-domain transfer learning. To this end, we use, for the in-domain pretraining, the previously introduced GBIF dataset, as due to the specifically selected plant species it is much more similar to our

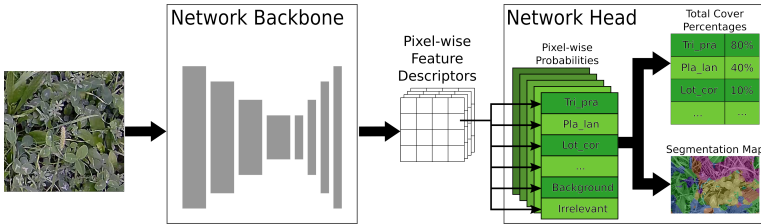


Fig. 4: The basic structure of the network. The image is input in the network backbone, which extracts pixel-wise feature descriptors. These are then used by the network head to calculate the pixel-wise probabilities for each plant, which are afterwards aggregated into the total cover percentages. The generated pixel-wise probabilities can also be interpreted as a segmentation map, so that the predictions of the network can also be interpreted by the user.

target task than, for example, the widely used ImageNet [Ru15] dataset containing merely general objects, which we use as out-of-domain pretraining for comparison.

4.1 Basic Network Structure

Our network consists of two main components: a feature extractor backbone and a network head. The backbone consists of one of the abovementioned standard classification networks potentially in conjunction with a Feature Pyramid Network to increase the output resolution. It hence generates a pixel-wise feature map, which is then used as input for the network head. The latter comprises a calculation model, explained in detail in section 4.2, which initially determines the pixel-wise probabilities of each plant species and afterward aggregates the predicted values into the total cover percentage of each species. As already mentioned, the generated class or probability map can also be used as a segmentation map for the user to confirm the correctness of the results. An overview of this network structure is shown in Figure 4.

4.2 Calculation Model

The calculation model represents the computations taking place in the network head to lastly generate the cover predictions. This calculation is separated into two parts: pixel-wise class predictions and aggregations.

Class Prediction. As the pixel-wise predictions are largely dependent on the actual cover calculation, we will first introduce the intuitive formula to calculate the cover for a single plant species:

$$cover^{plant} = \frac{A^{plant}}{A^{plant} + A^{uncovered}}, \quad (1)$$

where A^{plant} is the area covered by the single plant species and $A^{uncovered}$ is the area not covered by it. Under this premise, there are several things to take into account for the pixel-wise class prediction. First, as occlusion is usually ignored for calculating the plant cover, it has to be possible to predict multiple different plant species in the same location. Second, the areas in which no plants exist also have to be taken into account in the calculation. Third, as seen in the example images in Figure 2, there are areas in the image, which are irrelevant for cover calculation (e.g., the walls of the EcoUnits) and should thus be ignored. While the existence of multiple plants in a single pixel does not have to be mutually exclusive, plants, soil and irrelevant pixels cannot be predicted in the same location and therefore have to be mutually exclusive. To be able to calculate the plant cover accurately, we also include these premises in our calculation model. As these factors are already included in the calculation model from our previous work [Kö20], we also utilize this model here with slight modifications.

We define $P_{x,y}^p$ as the probability of a plant p existing at location x, y . We also define $P_{x,y}^{bio}$, $P_{x,y}^{bg}$, and $P_{x,y}^{irr}$ as the probability of any plant, background (uncovered soil), and irrelevant data being at location x, y , respectively. Note that, as mentioned above, the single plant probabilities $P_{x,y}^p$ do not necessarily add up to one, as they are not mutually exclusive, while the three probabilities $P_{x,y}^{bio}$, $P_{x,y}^{bg}$ and $P_{x,y}^{irr}$ do.

As $P_{x,y}^{bio}$ is naturally dependent on the probabilities $P_{x,y}^p$, we can set them in relation, as in [Kö20]:

$$P_{x,y}^{bio} = 1 - (P_{x,y}^{bg} + P_{x,y}^{irr}) = 1 - \frac{\kappa}{\kappa + \sum_p P_{x,y}^p} = \frac{\sum_p P_{x,y}^p}{\kappa + \sum_p P_{x,y}^p}, \quad (2)$$

where κ is a threshold value, which determines if a pixel is considered a plant pixel or not. While κ is a tunable hyperparameter in [Kö20], we use it as a learnable parameter in our experiments.

Aggregation. Now, to aggregate the predicted probabilities into the plant-wise cover percentages, we can calculate the areas

$$(A^{bg}, A^{irr}, A^{bio}) = \left(\sum_{x,y} P_{x,y}^{bg}, \sum_{x,y} P_{x,y}^{irr}, \sum_{x,y} P_{x,y}^{bio} \right), \quad (3)$$

$$A^{total} = A^{bio} + A^{bg} + A^{irr}. \quad (4)$$

Thus, we can finally calculate the cover values for each plant p :

$$cover^p = \frac{\sum_{x,y} P_{x,y}^p}{A^{total} - A^{irr}} = \frac{\sum_{x,y} P_{x,y}^p}{A^{bio} + A^{bg}} \quad (5)$$

While in [Kö20] we also applied a vegetation percentage, i.e., the percentage of ground covered by plants in general, as an auxiliary value, we found that this is not necessary when using pretrained networks, which in turn removes the need for additional annotations.

5 Experiments

In our experiments, we compare the performance of the previously introduced network architectures with each other using two different resolutions. This comparison is done first using in-domain pretraining for each of the networks and then repeated using ImageNet pretraining. At the end of this section, we provide an error analysis based on the experiment duration of the InsectArmageddon dataset. The metrics we will use in our experiments for performance evaluation are the mean absolute error (MAE) and the mean scaled absolute error (MSAE) as defined in [Kö20]. As the MAE is heavily influenced by the average cover percentages of each plant species, we introduced the MSAE, which is the mean absolute error divided by the species-wise plant cover mean, to allow for a fairer comparison between the different species. For the exact values used to scale the errors, we would like to refer the reader to [Kö20].

5.1 Comparison of Different Networks

In this experiment different network architectures were utilized as backbone and their performance is compared. Three different architectures were used: ResNet50 [He16], InceptionV3 [Sz16] and DenseNet121 [Ia14]. We applied these networks in two different versions: with and without an FPN. In the latter case, we utilized an FPN [Li17a] with depth 512 to increase spatial resolution and hence to be able to better predict the classes pixel-wise. We used the P2 layer of the FPN, i.e., the second-to-last layer of the FPN, resulting in a final resolution of $\frac{1}{4}$ th of the original input image resolution.

Experimental Setup. In the following experiments, pretraining on the GBIF dataset (see section 3.1) was utilized, whereafter the main training on the InsectArmageddon (section 3.2) dataset took place.

The pretraining on GBIF was done on the base network and the FPN, if used. All networks were initialized with off-the-shelf ImageNet weights provided by the Keras framework [Ch15]. During the pretraining we used an image resolution of 448×448 px for ResNet50 and DenseNet121, and 427×427 px for InceptionV3 due to its different network structure. Each network was trained using a batch size of 12 for 20 epochs with an initial learning rate of 0.01, decaying by a factor of 0.1 at epoch 10 and 15. To improve the adaptability of the learned weights on the InsectArmageddon dataset, we also utilized several data augmentations to improve generalization and adaptation to the target dataset: random

Tab. 1: The numerical results of the experiment using several standard networks with pretraining on the GBIF dataset. Top results are marked in bold font. MAE and MSAE represent the *mean absolute error* and the *mean scaled absolute error*, respectively.

Image Resolution	768 × 384		1536 × 768	
	MAE	MSAE	MAE	MSAE
ResNet50 [He16]	5.49%	0.528	5.22%	0.503
ResNet50 [He16] + FPN [Li17a]	5.35%	0.512	5.18%	0.496
InceptionV3 [Sz16]	5.51%	0.531	5.30%	0.506
InceptionV3 [Sz16] + FPN [Li17a]	5.51%	0.528	5.27%	0.503
DenseNet121 [Ia14]	5.37%	0.522	5.18%	0.495
DenseNet121 [Ia14] + FPN [Li17a]	5.30%	0.515	5.16%	0.494

rotations, random crops, and random horizontal flipping. The optimizer during training is SGD with a momentum of 0.9, and the loss is the standard categorical cross-entropy.

During training on the InsectArmageddon dataset, we exchange the classification head of the network from the pretraining step with the plant cover prediction head explained in section 4.2, which we then train in conjunction with the FPN, if used. The rest of the network is frozen during training, as in previous experiments not shown here we found that fully fine-tuning the network does either not change the results or is even detrimental to them. We also found that for this training, a batch size of 1 is sufficient, and larger batch sizes also do not change the results. We assume this is the case due to the calculation performed in the network head. The latter can be interpreted as performing many single classifications with strongly changing and weakly defined target values. This, in turn, massively regularizes the network to a point at which bigger batches and additional regularization do not improve the results in any way. The networks are trained using the Adam [KB15] optimizer for faster convergence for 40 epochs starting with a learning rate of 0.001 (0.00001 in case of the network containing an FPN), decaying with a factor of 0.1 at epoch 20 and 30. We apply the mean absolute error as loss during the training process. Only horizontal flips are applied as data augmentation, as we found that the effect of most augmentations was either negligible or detrimental to the results. All experiments are performed in a special 12-fold cross-validation, i.e., during each fold, we select two EcoUnits, whose images are used for testing, while the images of the remaining units are used for training. For generating the experimental results below, the results of these folds are averaged.

Results. The results of the GBIF-pretraining experiments can be seen in Table 1. It is visible that the best results concerning the mean absolute error (MAE) and mean scaled absolute error (MSAE) were achieved with a DenseNet121 with FPN. We can also note that the DenseNet outperforms the two other networks in almost every single setting, which can likely be attributed to its structure, that directly incorporates information from earlier layers in the network for the final prediction. Similarly, the ResNet50 outperforms InceptionV3 in all settings, which leads to the assumption that the structure of the Inception modules

might be detrimental in this task. From these results, we can also observe that using a higher resolution is strictly better than using the lower one, which means that, even though the 768×384 px resolution is already quite high, there is still more relevant information contained in the images, which can only be utilized when using an even higher image resolution. In the experimental data, we can also see that the results improve upon using the FPN for some networks more than for others. The network profiting most from the FPN is the ResNet50, which seems to produce useful information for identification at multiple levels of the FPN and therefore receives a substantial performance boost. The DenseNet, as mentioned before, already includes low-level features by default in its network structure, which is why the FPN does not add much novel information and therefore only improves results slightly. While the InceptionV3 does not have long-range skip connections, the improvements by addition of the FPN are also only minor. While the reason for this is not entirely apparent, this could be caused by the differing network structure as compared to the ResNet50 and DenseNet121, which might not benefit as much from the FPN. The top result of the DenseNet121 outperforms the network proposed in [Kö20] in terms of mean absolute error (5.16% vs. 5.30%) and mean scaled absolute error (0.494 vs. 0.500). However, it should also be noted that in this work, we completely abstain from using the auxiliary vegetation-percent annotations, which were used in [Kö20] in addition to the plant cover annotations, and the amount of training necessary here is also much less, while still outperforming the network from [Kö20]. Additionally, it should be noted that, while the differences in errors are comparably small, visual inspection suggests that this is likely primarily due to improvements in the detection of less abundant plants, which have a rather small influence on the total error value.

5.2 Comparison with ImageNet Pretraining

To evaluate the effect of the additional in-domain pretraining using the GBIF dataset in section 5.1 we conduct an ablation study and compare it with a standard ImageNet pretraining. It should be noted that pretrained ImageNet weights were only available for the base classification networks without FPN, which is then randomly initialized. The pretraining on the GBIF dataset in section 5.1 has been performed on the base network including the FPN.

Experimental Setup. In these experiments, the setup is very similar to the one in the experiments before. However, instead of pretraining the network on GBIF, we merely load off-the-shelf ImageNet weights for the base networks, which are provided by the Keras framework [Ch15]. During training, in case of not using an FPN, we again only train the last layer using the same scheme as described above with an initial learning rate of 0.001. In the cases where an FPN is used, as its weights are not initialized, we train the FPN and the cover prediction head jointly with a learning rate of 0.00001.

Tab. 2: The numerical results of the experiment using several standard networks with only ImageNet pretraining. Top results are marked in bold font. MAE and MSAE represent the *mean absolute error* and the *mean scaled absolute error*, respectively.

Image Resolution	768 × 384		1536 × 768	
Metric	MAE	MSAE	MAE	MSAE
ResNet50 [He16]	5.55%	0.537	5.32%	0.510
ResNet50 [He16] + FPN [Li17a]	5.37%	0.521	5.30%	0.510
InceptionV3 [Sz16]	5.74%	0.546	5.34%	0.505
InceptionV3 [Sz16] + FPN [Li17a]	5.62%	0.535	5.33%	0.506
DenseNet121 [Ia14]	5.39%	0.515	5.17%	0.491
DenseNet121 [Ia14] + FPN [Li17a]	5.31%	0.510	5.20%	0.494

Results. The results of our experiments with off-the-shelf ImageNet weights can be seen in Table 2. We can see that the top results are generated again by the DenseNet121, with an MAE of 5.17% and an MSAE of 0.491. These results differ only slightly from the previous experiments. This difference likely results from the structure of the DenseNet, which directly incorporates features from earlier layers into top-level predictions, which are beneficial for high-resolution plant predictions. Here, we can also see that the FPN appears to be beneficial in all cases for the lower image resolution. For the higher resolutions, we notice that the performances of the networks with FPN are very similar to the ones of the standalone base networks in all cases. This leads to the assumption that the learned lower-level features incorporated by the FPN are not very relevant on this resolution anymore, and higher-resolution ones become more crucial. Compared to the GBIF pretraining experiments, we can see that if we only consider the base networks without FPN, the GBIF pretraining is beneficial in almost all cases, with only a slight deviation from this pattern for DenseNet121.

5.3 Error Analysis

To finally also analyze the output and errors of our method, we take a look at the error of the DenseNet121 pretrained on GBIF with FPN for an input image resolution of 1536 × 784 px from section 5.1 plotted over the weeks recorded in the dataset. As mentioned above, the images contain time series of images collected weekly over a range of 18 weeks, with the plants growing progressively over time while dying towards the end of the experiment. As these changing conditions can hugely affect the outcome of the automatic analysis, we observe the error of our method over the single weeks. Here, we will only focus on the MSAE. An overview of the MSAE changing over time can be seen in Figure 5, in which the change of the average plant cover percentage summed over all plants is also shown for comparison. We see that in the first few weeks, where not many plants have grown, the error is small but starts rising strongly around week 4, which coincides with the plant growth. Following this sudden rise, the error stays mostly consistent, also coinciding with the cover percentages. This high error can probably be attributed to the large amount of occlusion taking place in the images. It is visible in the graph that the latter peaks at 200%,

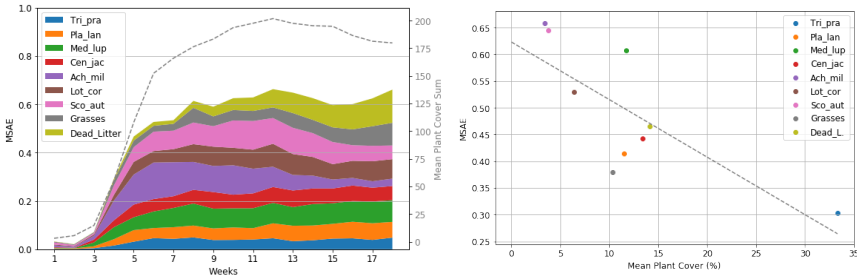


Fig. 5: Left: The mean scaled absolute errors (MSAE) averaged over all plant species in the dataset in dependence on the week of the image recording. For reference, the plant cover sum averaged over all weeks is shown as dashed line. Right: The relationship between the species-wise mean plant cover and the mean scaled absolute error (MSAE). We can see that the values are not or even negatively correlated, i.e., as opposed to the MAE the MSAE does not depend strongly on the plant abundances (correlation of $p = -0.774$; $R^2 = 0.600$).

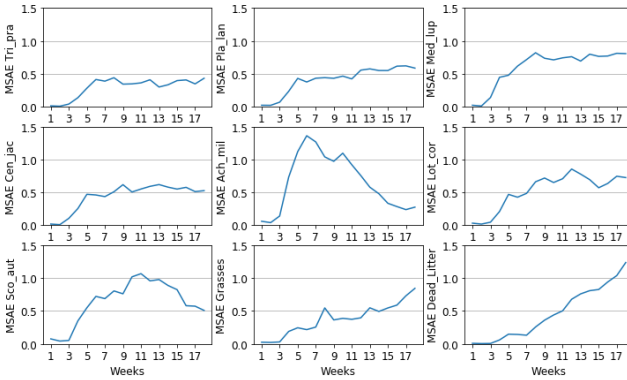


Fig. 6: The mean scaled absolute error in dependence on the week of the image recording.

meaning about half of the plant area used for estimating the ground truth values is occluded, making a correct estimation for the network difficult. Towards the end of the experiment, we can see that the error does not coincide with the changes in cover percentages anymore, as the former is rising while the latter are dropping. This can be explained when taking a look at the plant-wise errors in Figure 6. There we can see that the error for most plants is decreasing, as they are dying towards the end of the experiment. This process, i.e., the death of plants by either age (senescence) or due to the experimental setup, leads to plants being counted not as their original class anymore, but as the *Dead litter* class, where we see a rise in the error. This creates the assumption that the network cannot really identify the *Dead litter* correctly and possibly still assigns the respective pixels to its original class. From this, we can assume that the network could not work out a suitable discrimination between *Dead litter* and the other classes. This is understandable, as the point in time at which a plant starts

to be considered *Dead litter* can be quite arbitrary and subjective. While its label changes over the duration of one week, the plants might still bear significant resemblances to their previous label. In summary, we can therefore say that most of the errors are likely caused by the amount of occlusion in the images and the problem of discriminating between the regular plant classes and *Dead litter*. This, however, could also be caused by the non-existence of the *Dead litter* class in the pretraining dataset.

6 Conclusions and Future Work

In this work, we found that utilizing standard CNNs can yield similar or slightly better results than using, for example, our previously proposed network from [Kö20]. We also found that the usage of FPNs is in general beneficial, while the ResNet50 benefits most from them. It should also be noted that FPNs increase the output resolution of the resulting segmentation map, which, in turn, leads to better interpretability of the results compared to their non-FPN counterparts. We also found that using higher image resolutions is always beneficial in our setting. Therefore, collecting images with high resolutions should be prioritized to improve prediction accuracy. Pretraining the network on a more related dataset rather than using off-the-shelf ImageNet weights proved beneficial for all the networks with only few exceptions when using DenseNet121. Hence an in-domain pretraining should, in most cases, be preferred to an out-of-domain one, if available. Also, with pretraining, we can achieve similar or better results to [Kö20] with only training a part of the network on the final task, saving time and hardware capacities for training on the plant cover prediction task. From our week-based error analysis, we were able to conclude that some parts of the errors are most likely caused by the massive occlusions in the images, as well as the network's inability to differentiate between *Dead litter* pixels and its preceding plant species before its death.

In future work, we aim to tackle the abovementioned problems. For occlusion, some approaches in the area of amodal segmentation [LM16] might be suitable. For a better differentiation between plants, other pretraining datasets like iNaturalist [Va18] or even pretraining the network on segmentation datasets like MSCOCO [Li14] to prime it for pixel-wise classification might be a possibility.

Acknowledgement

Matthias Körschens thanks the Carl Zeiss Foundation for the financial support. We would also like to thank Alban Gebler and the iDiv for providing the data for our investigations.

Bibliography

- [AC12] Aggemyr, Elsa; Cousins, Sara AO: Landscape structure and land use history influence changes in island plant composition after 100 years. *Journal of Biogeography*, 39(9):1645–1656, 2012.

- [Ba17] Barré, Pierre; Stöver, Ben C; Müller, Kai F; Steinhage, Volker: LeafNet: A computer vision system for automatic plant species identification. *Ecological Informatics*, 40:50–56, 2017.
- [Br16] Brown, Tim B; Hultine, Kevin R; Steltzer, Heidi et al.: Using phenocams to monitor our changing Earth: toward a global phenocam network. *Frontiers in Ecology and the Environment*, 14(2):84–93, 2016.
- [Br18] Bruelheide, Helge; Dengler, Jürgen; Purschke, Oliver; Lenoir, Jonathan; Jiménez-Alfaro, Borja; Hennekens, Stephan M; Botta-Dukát, Zoltán; Chytrý, Milan; Field, Richard; Jansen, Florian et al.: Global trait–environment relationships of plant communities. *Nature Ecology & Evolution*, 2(12):1906–1917, 2018.
- [Bu18] Bucher, Solveig Franziska; König, Patrizia; Menzel, Annette; Migliavacca, Mirco; Ewald, Jörg; Römermann, Christine: Traits and climate are associated with first flowering day in herbaceous species along elevational gradients. *Ecology and Evolution*, 8(2):1147–1158, 2018.
- [Ch15] Chollet, François et al.: , Keras. <https://keras.io>, 2015.
- [Ch20] Chen, Junde; Chen, Jinxiu; Zhang, Defu; Sun, Yuandong; Nanekaran, YA: Using deep transfer learning for image-based plant disease identification. *Computers and Electronics in Agriculture*, 173:105393, 2020.
- [Cl12] Cleland, Elsa E; Allen, Jenica M; Crimmins, Theresa M; Dunne, Jennifer A; Pau, Stephanie; Travers, Steven E; Zavaleta, Erika S; Wolkovich, Elizabeth M: Phenological tracking enables positive species responses to climate change. *Ecology*, 93(8):1765–1771, 2012.
- [Cu18] Cui, Yin; Song, Yang; Sun, Chen; Howard, Andrew; Belongie, Serge: Large scale fine-grained categorization and domain-specific transfer learning. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp. 4109–4118, 2018.
- [ET18] Eisenhauer, Nico; Türke, Manfred: From climate chambers to biodiversity chambers. *Frontiers in Ecology and the Environment*, 16(3):136–137, 2018.
- [FF02] Fitter, AH; Fitter, RSR: Rapid changes in flowering time in British plants. *Science*, 296(5573):1689–1691, 2002.
- [Ga19] Ganesh, P; Volle, K; Burks, TF; Mehta, SS: Deep Orange: Mask R-CNN based Orange Detection and Segmentation. *IFAC-PapersOnLine*, 52(30):70–75, 2019.
- [GB20] GBIF.org: , GBIF Occurrence Downloads, 13 May 2020. <https://doi.org/10.15468/dl.xg9y85>, <https://doi.org/10.15468/dl.zgbmn2>, <https://doi.org/10.15468/dl.cm6hqj>, <https://doi.org/10.15468/dl.fez33g>, <https://doi.org/10.15468/dl.f8pqjw>, <https://doi.org/10.15468/dl.qbmyb2>, <https://doi.org/10.15468/dl.fc2hqk>, <https://doi.org/10.15468/dl.sq5d6f>.
- [Ge14] Gerstner, Katharina; Dormann, Carsten F; Stein, Anke; Manceur, Ameer M; Seppelt, Ralf: EDITOR’S CHOICE: REVIEW: Effects of land use on plant diversity—A global meta-analysis. *Journal of Applied Ecology*, 51(6):1690–1700, 2014.
- [GYA17] Ghazi, Mostafa Mehdipour; Yanikoglu, Berrin; Aptoula, Erchan: Plant identification using deep neural networks via optimization of transfer learning parameters. *Neurocomputing*, 235:228–235, 2017.

- [He16] He, Kaiming; Zhang, Xiangyu; Ren, Shaoqing; Sun, Jian: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. IEEE, pp. 770–778, 2016.
- [He17] He, Kaiming; Gkioxari, Georgia; Dollár, Piotr; Girshick, Ross: Mask r-cnn. In: Proceedings of the IEEE international conference on computer vision. IEEE, pp. 2961–2969, 2017.
- [Ia14] Iandola, Forrest; Moskewicz, Matt; Karayev, Sergey; Girshick, Ross; Darrell, Trevor; Keutzer, Kurt: Densenet: Implementing efficient convnet descriptor pyramids. arXiv preprint arXiv:1404.1869, 2014.
- [Ka20] Kattenborn, Teja; Eichel, Jana; Wisler, Susan; Burrows, Larry; Fassnacht, Fabian E; Schmidtlein, Sebastian: Convolutional Neural Networks accurately predict cover fractions of plant species and communities in Unmanned Aerial Vehicle imagery. *Remote Sensing in Ecology and Conservation*, 2020.
- [KB15] Kingma, Diederik P.; Ba, Jimmy: Adam: A Method for Stochastic Optimization. CoRR, abs/1412.6980, 2015.
- [Kö20] Körschens, Matthias; Bodesheim, Paul; Römermann, Christine; Bucher, Solveig Franziska; Ulrich, Josephine; Denzler, Joachim: Towards Confirmable Automated Plant Cover Determination. In: *European Conference on Computer Vision*. Springer, pp. 312–329, 2020.
- [KPB18] Kamilaris, Andreas; Prenafeta-Boldú, Francesc X: Deep learning in agriculture: A survey. *Computers and electronics in agriculture*, 147:70–90, 2018.
- [Le15] Lee, Sue Han; Chan, Chee Seng; Wilkin, Paul; Remagnino, Paolo: Deep-plant: Plant identification with convolutional neural networks. In: 2015 IEEE international conference on image processing (ICIP). IEEE, pp. 452–456, 2015.
- [Li14] Lin, Tsung-Yi; Maire, Michael; Belongie, Serge; Hays, James; Perona, Pietro; Ramanan, Deva; Dollár, Piotr; Zitnick, C Lawrence: Microsoft coco: Common objects in context. In: *European conference on computer vision*. Springer, pp. 740–755, 2014.
- [Li17a] Lin, Tsung-Yi; Dollár, Piotr; Girshick, Ross; He, Kaiming; Hariharan, Bharath; Belongie, Serge: Feature pyramid networks for object detection. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp. 2117–2125, 2017.
- [Li17b] Lin, Tsung-Yi; Goyal, Priya; Girshick, Ross; He, Kaiming; Dollár, Piotr: Focal loss for dense object detection. In: *Proceedings of the IEEE international conference on computer vision*. pp. 2980–2988, 2017.
- [Li18] Liu, Huiying; Mi, Zhaorong; Lin, LI; Wang, Yonghui; Zhang, Zhenhua; Zhang, Fawei; Wang, Hao; Liu, Lingli; Zhu, Biao; Cao, Guangmin et al.: Shifting plant species composition in response to climate change stabilizes grassland primary production. *Proceedings of the National Academy of Sciences*, 115(16):4051–4056, 2018.
- [LI09] Lloret, Francisco; Peñuelas, Josep; Prieto, Patricia; Llorens, L; Estiarte, Marc: Plant community changes induced by experimental climate change: seedling and adult species composition. *Perspectives in Plant Ecology, Evolution and Systematics*, 11(1):53–63, 2009.
- [LM16] Li, Ke; Malik, Jitendra: Amodal instance segmentation. In: *European Conference on Computer Vision*. Springer, pp. 677–693, 2016.

- [Lu17] Lu, Hao; Cao, Zhiguo; Xiao, Yang; Zhuang, Bohan; Shen, Chunhua: TasselNet: counting maize tassels in the wild via local counts regression network. *Plant methods*, 13(1):79, 2017.
- [Me06] Menzel, Annette; Sparks, Tim H; Estrella, Nicole; Koch, Elisabeth; Aasa, Anto; Ahas, Rein; Alm-Kübler, Kerstin; Bissolli, Peter; Braslavská, Ol'ga; Briede, Agrita et al.: European phenological response to climate change matches the warming pattern. *Global change biology*, 12(10):1969–1976, 2006.
- [MRP08] Miller-Rushing, Abraham J; Primack, Richard B: Global warming and flowering times in Thoreau's Concord: a community perspective. *Ecology*, 89(2):332–341, 2008.
- [Pf97] Pfadenhauer, J.: *Vegetationsökologie - ein Skriptum*. IHW-Verlag, Eching, 2. verbesserte und erweiterte auflage edition, 1997.
- [Re15] Ren, Shaoqing; He, Kaiming; Girshick, Ross; Sun, Jian: Faster r-cnn: Towards real-time object detection with region proposal networks. *arXiv preprint arXiv:1506.01497*, 2015.
- [Ro07] Rosenzweig, Cynthia; Casassa, Gino; Karoly, David J et al.: Assessment of observed changes and responses in natural and managed systems. *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, pp. 79–131, 2007.
- [Ru15] Russakovsky, Olga; Deng, Jia; Su, Hao et al.: Imagenet large scale visual recognition challenge. *International journal of computer vision*, 115(3):211–252, 2015.
- [Sz16] Szegedy, Christian; Vanhoucke, Vincent; Ioffe, Sergey; Shlens, Jon; Wojna, Zbigniew: Rethinking the inception architecture for computer vision. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. IEEE, pp. 2818–2826, 2016.
- [SZS16] Souza, Lara; Zelikova, Tamara Jane; Sanders, Nathan J: Bottom-up and top-down effects on plant communities: nutrients limit productivity, but insects determine diversity and composition. *Oikos*, 125(4):566–575, 2016.
- [Tü17] Türke, Manfred; Feldmann, Reinhart; Fürst, Bernd; Hartmann, Henrik; Herrmann, Martina; Klotz, Stefan; Mathias, Georg; Meldau, Stefan; Ottenbreit, Markus; Reth, Sascha et al.: Multitrophische Biodiversitätsmanipulation unter kontrollierten Umweltbedingungen im iDiv Ecotron. In: *Lysimetertagung*. pp. 107–114, 2017.
- [UI20] Ulrich, Josephine; Bucher, Solveig Franziska; Eisenhauer, Nico; Schmidt, Anja; Türke, Manfred; Gebler, Alban; Barry, Kathryn; Lange, Markus; Römermann, Christine: Invertebrate decline leads to shifts in plant species abundance and phenology. *Frontiers in plant science*, 11:1410, 2020.
- [Va18] Van Horn, Grant; Mac Aodha, Oisín; Song, Yang; Cui, Yin; Sun, Chen; Shepard, Alex; Adam, Hartwig; Perona, Pietro; Belongie, Serge: The inaturalist species classification and detection dataset. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. IEEE, pp. 8769–8778, 2018.
- [Xi19] Xiong, Haipeng; Cao, Zhiguo; Lu, Hao; Madec, Simon; Liu, Liang; Shen, Chunhua: TasselNetv2: in-field counting of wheat spikes with context-augmented local regression networks. *Plant Methods*, 15(1):150, 2019.
- [YR16] Yalcin, Hulya; Razavi, Salar: Plant classification using convolutional neural networks. In: *2016 Fifth International Conference on Agro-Geoinformatics (Agro-Geoinformatics)*. IEEE, pp. 1–5, 2016.

DiversityNaviKey, a Progressive Web Application for interactive diagnosis and identification


Dagmar Triebel¹, Ariane Grunz¹, Stefan Seifert¹, Anton Link¹ & Gerhard Rambold²


Abstract: DiversityNaviKey is designed as a diagnostic tool in the field of biology and related sciences to identify organisms as well as to interactively select other entities and objects related to research based on a set of predefined properties. It allows queries on structured sources of descriptive data (trait data) to diagnose groups of objects based on combinations of optionally modified descriptor-states or values that are selected consecutively during the diagnostic process. The Web App is implemented as a single-page application and contains the entire presentation logic to dynamically change a pre-generated HTML page in the browser. The content data is accessed via a web service as JSON packages. DiversityNaviKey is a progressive web application that uses caching mechanisms of browsers, such as Service Worker and IndexedDB. Thus, the main tasks are also available in offline mode. The current set up uses the SNSB technical infrastructure and data pipelines with PostgreSQL Cache Database and the data management tool DiversityDescriptions as backend. The exemplary data sources are from various domains, two of them are large datasets from the long-term projects DEEMY and LIAS.

Keywords: Diagnosis, identification, PostgreSQL Cache DB, RESTful web service, single-page application, vue.js, offline use, use cases DEEMY and LIAS

1 Introduction

Early web-based identification keys were created by and for fauna, flora and mycota experts and ambitious naturalists [NV10, HRM10]. They are embedded as simple statically tagged HTML web sites in content management systems or in wiki-based community networks. Examples are the early LIAS keys for multi-access identification of lichens [Ra97] and the more than 1.100 dichotomous or table-oriented identification keys of Offene Naturführer (ON) [Of21, Ha10]. To allow a more responsive presentation of text and multimedia content on web servers, various types of browser-based applications have been set up with simple query capabilities on an underlying data source. In most approaches, the underlying backend structure is adapted to a defined domain-specific data schema, the requirements in a specific scientific field and often limited on the identification of a single organism group in a given geographical region (see FRIDA approach for lichens of Italy [NM20]).

¹ Staatliche Naturwissenschaftliche Sammlungen Bayerns, IT Center, Menzinger Straße 67, 80638 München, triebel@snsb.de,  <https://orcid.org/0000-0003-1980-3148>

² Universität Bayreuth, Dept. of Mycology, Universitätsstraße 30, 95440 Bayreuth, gerhard.rambold@uni-bayreuth.de,  <https://orcid.org/0000-0002-9473-3250>

Apart from the approaches mentioned above, there are a number of non-domain specific tools for advanced queries of highly structured data, either as output from backend databases based on a generic data model or by accessing data files structured according to the DELTA data coding and exchange standard [CLM10, Da80, DEL21a]. These types of generic interactive selection tools (web-based or not) are designed to support the identification of biological objects, mainly organisms, but also suitable for e.g. the diagnosis of plant and animal diseases, environmental damages as well as the recommendation of species for medical or agricultural and forestry use. This can also be done from large data sources with categorical or measurement data and different types of textual data. Binary data objects such as multimedia files may be included as additional resources, e.g. for defining the object representations and traits. Such frontend tools with various functionalities and mostly used by taxonomists, have been built on Perl, PHP, Java, and ASP [Ra14, Un10]. Examples include the NaviKey Java applet [Na21a, Na21b], Open DELTA [Op21], XPER 2, the XPER 3 platform [Xp21a, Xp21b, Xp21c], WEBiKEY [ADC16] and the commercial LUCID platform [Lu21].

2 General background on structured trait data and (bio-)diversity items (objects)

There are a number of theoretical approaches that address the general problem of handling and structuring morpho-anatomical and ecological trait data (= descriptive data) to finally be interactively selected for diagnostic purposes. Within the last decade, they focussed mainly on evolutionary aspects and data integration based on semantic web and recognised ontologies (e.g. [Da17, Vo17, Wa14]).

Authors like Diederich [Di97], Hagedorn [Ha08], Rambold et al. [Ra19] and Harjes et al. [Ha20] discussed the conceptual and logical structuring of descriptions, descriptors, the definition of basic properties and the creation of concepts to organise hierarchies, dependencies, concatenations and other types of relationships between biological and non-biological objects and their properties.

These concepts explain ‘descriptions’ of an item (= entity, = object) through a triple structure, namely ‘item-descriptor-state’ or ‘item-descriptor-value’ or ‘item-character-character state’ or ‘entity-attribute-value’. The descriptors (= characters, attributes, properties) are not limited to morphological characters and functional traits, but can also be physiological, ecological, behavioural and survey parameters, molecular descriptors and descriptors of data content schemes. The descriptor-states or values that make up the item/description can be generalized to one of several data types (categorical, numeric, textual and other more specific ones). This term ‘backbone’ is used in the following for the DiversityNaviKey functions for the interactive selection of trait data and for DiversityDescriptions functions to store highly structured trait data [DD21].

In general, natural objects or items, especially from the field of biodiversity and geosciences, are circumscribed by various knowledge organization systems, such as terminologies, ontologies, categories, and classifications at various levels. Systems for categorising organisms are traditional taxonomies based on coding by taxon names to classify them according their nomenclatural status (rules of the International Codes of Nomenclature ICN, ICZN, ICNP) and to circumscribe them according one or more scientifically recognised taxonomic concepts (see, e.g. [Pu05]). Taxonomic concepts in a broad sense include those applied in molecular biology, e.g., that of exact sequence variants (ESVs) and operational taxonomic units (OTUs), concepts as applied in ecology, e.g. that of ecological functional units (EFUs) or that of ecological guilds, and, e.g., those used by citizen scientists with regionally applied vernacular names for species. Other biological objects or groups of objects may be named or categorized with codes, e.g. environmental samples, products from laboratory work or vegetation units. The circumscription of objects or items is a matter of research and taxonomic opinion. The same applies to (phylogenetic) classification aspects and the use of ecological and environmental vocabulary.

Against this background, the diagnosis and identification of such objects or entities requires highly structured backend data sources. This is further complicated by the fact that there are a multitude of recognized vocabularies, terminologies and ontologies describing natural objects and items to be identified as well as characters or attributes of natural objects addressing their overall diversity. Linking data source items, descriptors and states with synonymous terms and recognised terminologies and ontologies needs to be organised, e.g., as part of an early data curation measure.

2.1. Accessing highly structured backend data sources

Data sources (datasets and databases) in biology vary in terms of data structure and content. There are tables and matrices from research studies with a high number of structured descriptors or matrices with few characters but numerous categorical states, or of a format with textual or numeric data types. The tables vary from small to large and the data content is often incomplete and leaves room for interpretation. In consequence, each backend data source can be considered a unique source with unique data structure, ontology and content. The direct web access to such highly structured backend data sources for interactive querying is challenging. It can only be successful and convenient if original data is reorganised and provided either via a flattened ‘cache’ database with well-documented generic data model or via standardised XML, JSON and CSV files. Such files might be standardized according to the TDWG Standard for Descriptive Data SDD (for SDD see [SD21], [Ha05], GFBio standards [GF21b] and GFBio type 3 data [GF21c]) or structured according to the DELTA data format and standard [DEL21a] as consumed by the DiversityNaviKey predecessor software, called NaviKey [Na21a].

The current set up for the new interactive DiversityNaviKey selection tool manages the access to highly structured data sources from various biological and environmental

domains via a web service and a PostgreSQL cache database as described in chapter 3. The highly structured data sources are primarily managed in MS SQL database installations of DiversityDescriptions with features to import and export SDD and DELTA files (see chapter 3.6). The ‘flattened’ PostgreSQL cache database for web data access is organised by DiversityDescriptions data pipelines (see chapter 3.4, 3.6). Several use case datasets are described below (see chapter 3.7).

2.2. Constructing queries for interactive selection

Constructing queries for interactive selection of the various types of descriptors has to fulfil special requirements [Ha08, Ha10]. For certain types of industrial products, it can be assumed that the characters of individuals of a series are uniquely defined and identical to other individuals and, accordingly, the specification of a query is fairly unambiguous. In such ‘ideal cases’, the predefined specifications of descriptor categories can be linked by AND (and more rarely OR) operators and the numeric values exactly defined and queried by ‘=’ or ‘≠’. In text search, fuzzy search logic is usually not required. While industrial products as objects can often be fully represented according to a given scheme of descriptors and complete data entries, biological items or objects from other disciplines of natural sciences in backend data sources are often incompletely represented due to deficiencies in the available field or laboratory data and proprietary schemes. These reasons make the diagnosis and interactive selection process for most types of items from the biological domain challenging. The categorical states of individual characters, for example, can be often better linked by OR (and more rarely AND) operators. This addresses a more flexible and subjective user view on descriptor-states or values. For numeric characters/descriptors, it is also advisable to distinguish upper/lower ranges of values from their extreme values.

In addition to functions for identifying biological objects, there should also be functions for text search that follow the logic of fuzzy search. Finally, it is desirable to have two query modes, namely a retrieval mode and a diagnosis mode, which differ according to how missing data is handled.

In general, the processes ‘diagnosis’ and ‘identification’ have to be supported with maximum user guidance and flexibility. For instance, it must be possible to save the set of criteria in order to be able to document all or part of the process. Furthermore, each step of an interactive selection should be flexible enough to allow the query conditions to be changed.

The software application described below is primarily aimed at professionals interested in object identification. Depending on the available data sources (see some use cases below), these may be flora, mycota and fauna experts, curators in (culture) collections, or experts e.g. in taxonomy, ecology, plant pathology, pedology and mineralogy.

3 Technical description of DiversityNaviKey identification software application and services

The web app DiversityNaviKey (DNK) is realised to identify organisms or other objects that are part of or related to bio- and geosciences research by means of predefined characters or descriptors.

A descriptor can be seen as a feature of an object with one or more states and values. Predefined descriptor-states describe objects and groups of objects or entities by descriptor-states as being values. DNK web app allows interactive queries on those structured sources of descriptive data, where the user consecutively defines descriptors and states as values either by selecting predefined options (categorical type) or by entering the values (textual and numerical types). An example is described in the section ‘Data types and interactive query description’.

DiversityNaviKey is a reimplement of the Java applet and offline usable application NaviKey [Na21a] (in the latter version also integrated in the Diversity Navigator database client [DN21]). Primarily, DNK is designed as a Diversity Workbench (DWB) tool [DW21a] that uses DWB data pipelines as step up at the SNSB IT Center. This is described in detail under ‘Diversity Workbench data pipelines with DiversityDescriptions as the backend data source’.

3.1 Progressive Web Application

DNK is designed and implemented as a ‘Progressive Web Application’ (PWA). A PWA is a web application providing a native app-like experience through the browser itself. It combines web page and native app capabilities and, unlike a native app, is independent of operating systems or devices. The pages are responsive and adjust to the respective end device (desktop, tablet, smartphone). Modern browsers enable PWA functions with technologies like Service Workers and IndexedDB, opening up a new way for writing app-like web applications by supporting offline use and installation without an app store.

3.2 DiversityNaviKey architecture

The architecture of DiversityNaviKey is shown in Fig. 1. On the one hand there is the frontend with the web browser running the DiversityNaviKey PWA, on the other hand, there is the backend with a web service that delivers the data from a database to the web app and the installed web application that contains the entire view and logic of the DNK. The web app not only presents the view, but also contains the functionality of the search and filter methods. A web service gets the requested data from the database and sends it to the web app in JSON format. In an initial installation and implementation of DNK at the SNSB IT Center, the data comes from existing data sources managed via DiversityDescriptions. This software has the functionality to pre-process and transfer the

data into a PostgreSQL CacheDB. A detailed description of this process is given in chapter 3.6. The web service requests the data and converts the response into JSON packages that can be used by the DNK app. It works as a service for accessing different data sources of the CacheDB. All data sources defined in the web service are automatically read out by the DNK and provided in the database selection dialog of the web app. The web service is explained in the ‘Backend’ section. The data format and structure as well as the minimal requirements of the CacheDB data sources will be provided along with the source code (autumn 2021, information under [DNK21]). The DNK web app is designed as a generic tool, where the web service may be replaced by other web services fulfilling the interface.

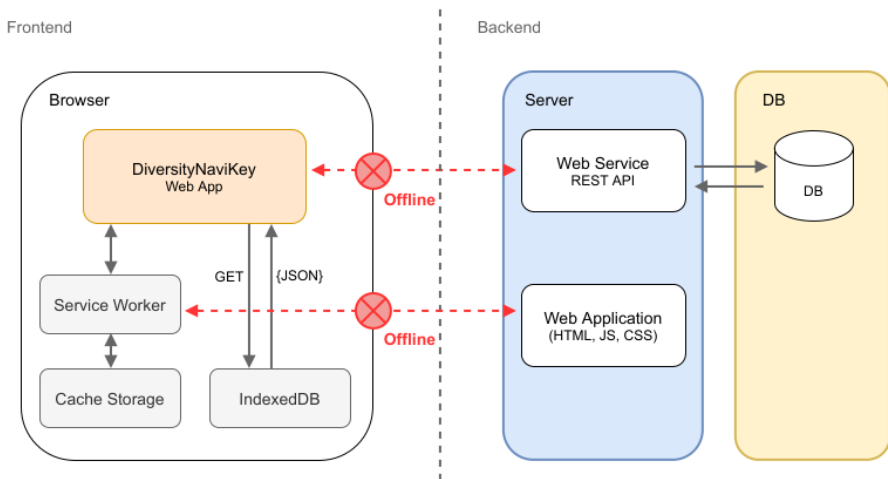


Fig. 1 DiversityNaviKey: Frontend (web browser) with Progressive Web Application (PWA) and backend (server) with web service and web application

3.3 Frontend

The DiversityNaviKey web app is a single-page application (SPA). SPAs are web applications that require only the loading of a single, pre-generated HTML page. This page contains the entire presentation logic for dynamically changing the page in the browser. The first time the DNK is called up in a browser, a request is sent to the server web application. The server responds with all the information of the app (HTML, JS, CSS). DNK includes a service worker that is installed the first time it is called. The service worker checks for updates of the app and enables offline use of the app by caching it (Fig. 1). The content data is retrieved via a web service and stored in the browsers IndexedDB.

Framework Vue.js

Various JavaScript frameworks exist for the design and creation of user interfaces for single-page applications (SPAs), such as Vue.js, React and Angular. For the development of the SPA DiversityNaviKey, the Vue.js framework and its standard tooling Vue CLI are used [Vu21a].

The Vue.js framework is an open source front-end JavaScript framework. It is very popular (185k stars on GitHub <https://github.com/vuejs/vue>, as of June 2021) and was originally written by Evan You and the open-source community. Vue.js was developed by taking the most popular ideas of frameworks like Angular or React. It is lightweight, flexible and easy to get started with. Vue.js is not backed by a huge corporation like Angular (Google) and React (Facebook), but is entirely community-driven.

To provide the features of a Progressive Web Application (PWA), the plugin `cli-plugin-pwa` was added. Based on the Vue core library, DiversityNaviKey uses the Vue Router [Vu21b], Vuetify [Vu21c] and Vuex [Vu21d] libraries. The Vue Router library is the official router for Vue.js, used to navigate between the different Vue components of the page. The Vuetify library is a complete UI Library developed according to the Material Design Specification [Ma21]. The Vuex library is a state management library. It serves as a global singleton, a central data store for the shared data of all components. It is reactive and has rules for state changes to ensure consistency of data.

Data access and offline support

After starting DiversityNaviKey, a data source can be selected from a list of available data sources. The current setup on the server infrastructure of the SNSB IT Center provides six exemplary data sources (Fig. 2), two of them, LIAS light [LI21] and DEEMY [DEE21], are described below. This list of data sources is queried via a web service. The architecture of the REST web service is explained below. It provides all data for a selected data source as JSON packages.

Modern browsers support Service Worker and IndexedDB to enable offline support of PWAs. DiversityNaviKey uses a Service Worker for caching the application's static assets and IndexedDB for caching the content data received via the backend data sources (Fig. 1). Since all logic and functionality is implemented in the PWA, the functions of the application are also available in offline mode. Users can choose between different data loading modes in DNK. In online mode, the web app always loads the data via the web service so that the data is always up to date. The online-first mode offers the possibility to access the browser's data cache after a failed data connection. Conversely, the offline mode accesses the cache, which speeds up the loading of data at the beginning and avoids repeated downloading of data over the network. In offline-first mode, the remote data is accessed when no data is available in the cache. The limitations and dependencies of the browser are discussed in chapter 4.

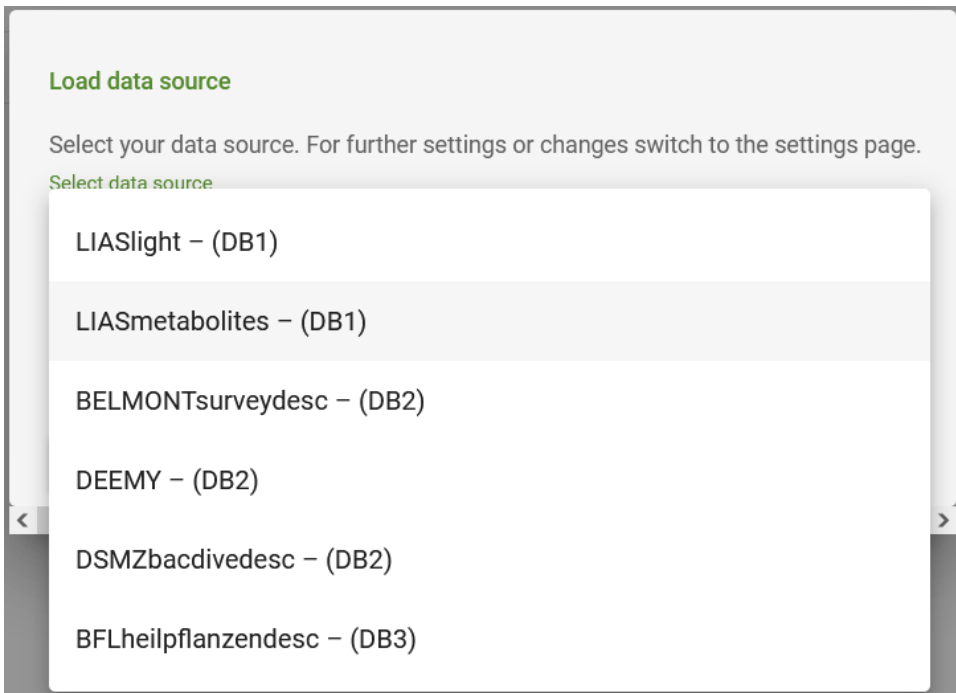


Fig. 2 DiversityNaviKey: User interface with picklist of available data sources (screenshot of test installation)

3.4 Backend

Web Service

The single page app needs a web-based endpoint from which it can retrieve data. This endpoint is provided by a REST web service. REST (representational state transfer) means a stateless service that uses the HTTP protocol to deliver data to the calling web application. The resources available in the web services are addressed by URLs and the data is delivered in JSON. The implemented web service is the connection between a database and the web application. It translates the data into the required JSON format. The database backend is a PostgreSQL installation. The tables in PostgreSQL are preformatted to minimize the query overhead of the web service. This optimization is part of the data transfer from the DiversityDescriptions database to the PostgreSQL database. The web service is used as a defined interface for the web application to retrieve the data from the PostgreSQL cache database. As the web application contains all the functionality the web service is reduced to deliver the data only. It has no capabilities to query or filter the data.

Since the web service serves only ‘quasi’ static read-only data, the service’s responses are extensively cached to reduce response times. The responses follow the json:api recommendations in such a way that all data is stored in a JSON ‘data’ dictionary and the pagination links in the ‘links’ dictionary. The pagination was implemented to avoid sending too large data packages to the application, which could hinder the web app’s responsiveness. All responses are transparently compressed over the HTTP transport. This is important when the web app is used without a stable internet connection and when internet bandwidth is limited. The web service endpoints are structured hierarchically to provide meaningful URLs. The endpoints are documented according to the OpenAPI specification.

The web service is implemented in Python using the Flask framework as an uWSGI service. It can be deployed on almost all current operating systems and can be used standalone or behind a web server such as Apache or Nginx.

DB

The database in the backend is a PostgreSQL database that is only used as a cache database (CacheDB). The CacheDB is not used for data management. In the DNK installation in the SNSB IT Center (Fig. 3), this CacheDB is filled by the DiversityDescriptions software. This data pipeline is explained in detail in chapter 3. The minimum requirements and the database scheme will be published and made available together with the source code of DNK (autumn 2021, information under [DNK21]).

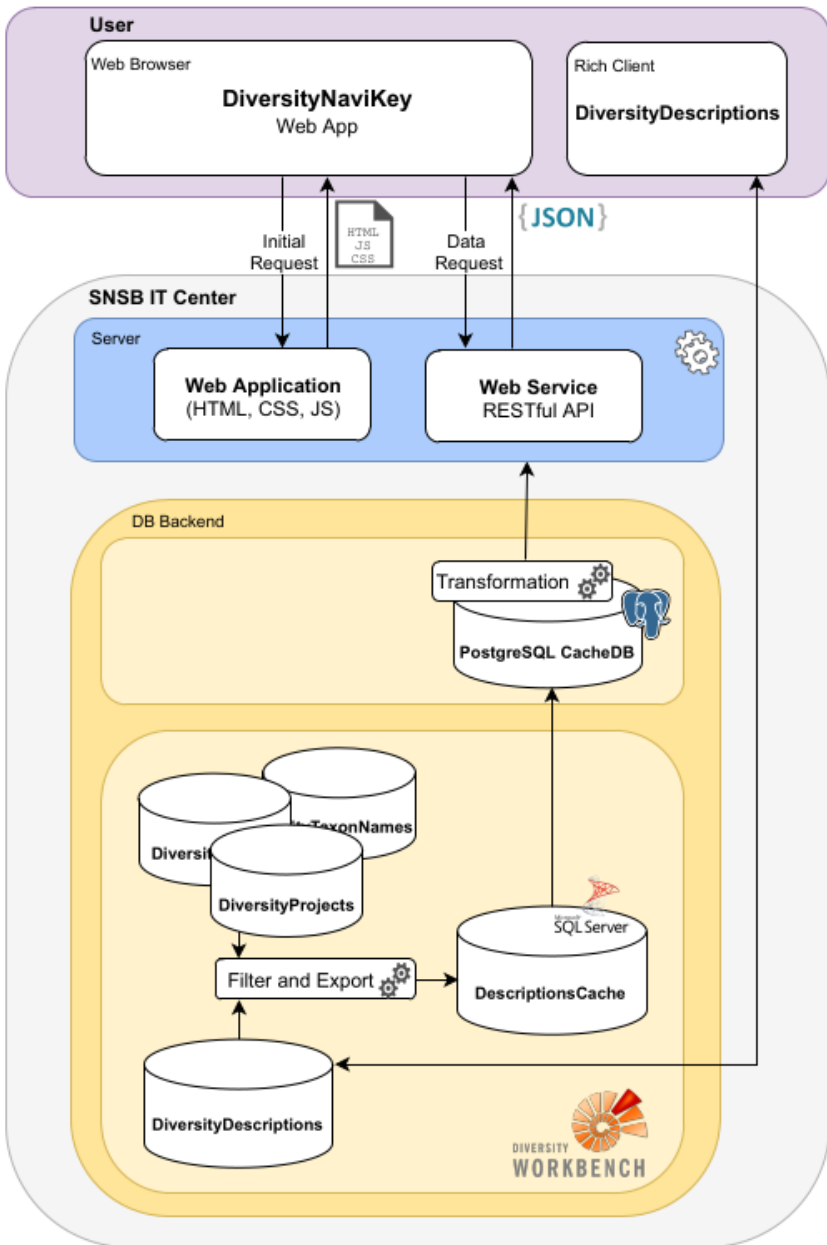


Fig. 3 DiversityNaviKey: Frontend at user side and backend with data processing at the SNSB IT Center: Web App, Web Application, Web Service and DWB Databases

3.5 Data types and interactive query description

Data types

Objects or groups of objects are described with a descriptor-state-value triple. The descriptor can mean a feature, property, symptom, etc. The descriptor can be categorical, numeric and textual. The states are the different possible manifestation of the descriptor and stand for categorical values.

Interactive query

The web app DiversityNaviKey implements and offers various sorting and filtering options. The application does not use a backend for this filtering process. In order to optimize the query and enable a fast and efficient search within the frontend, the data stored in the CacheDB and retrieved via the web service is structured by pre-processing. For each descriptor-state pair, a list with all associated objects is stored. On the one hand, this can be used like an index to speed up the search, on the other hand it is suitable for a filtering approach, the so-called faceted search/navigation. This faceted navigation guides the user through his search. In each step, only the currently available options are displayed, and those that would lead to dead ends are eliminated. This minimises the likelihood of null results. Most website implementations with such multi-select filters concatenate different facets conjunctively (AND), while values within a given facet are applied disjunctively (OR). This query mode is the default query in DNK.

Fig. 4 shows an example of an interactive query using the LIAS light dataset (a description of this data set follows below). In the first step shown, the descriptor 'secondary metabolites compound' was selected. A dialog opens to select the states, in this example 'gyrophoric acid' and 'stictic acid'. The second screen shows all the criteria selected so far. By default, the DNK links the different descriptors, like 'ascospores septa kind' and 'substrate kind' with AND, i.e. all descriptors must match the specified state values, while the states within a descriptor, such as 'stictic acid' and 'gyrophoric acid' are linked with OR, i.e. one of the two must match. The example query results in seven remaining matches (see Fig. 5). These results can be listed and each of the results can be selected to see all the detailed information.

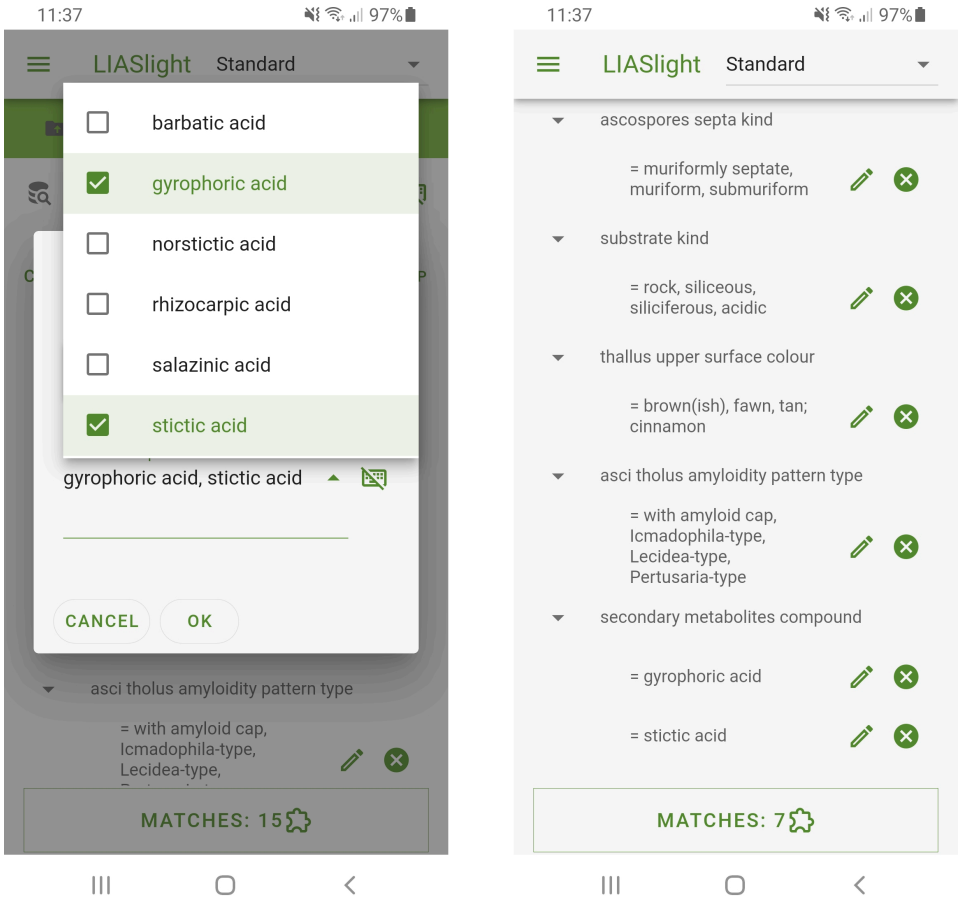


Fig. 4 DiversityNaviKey: Query user interface with LIAS data (screenshot of test installation, query for mobiles)

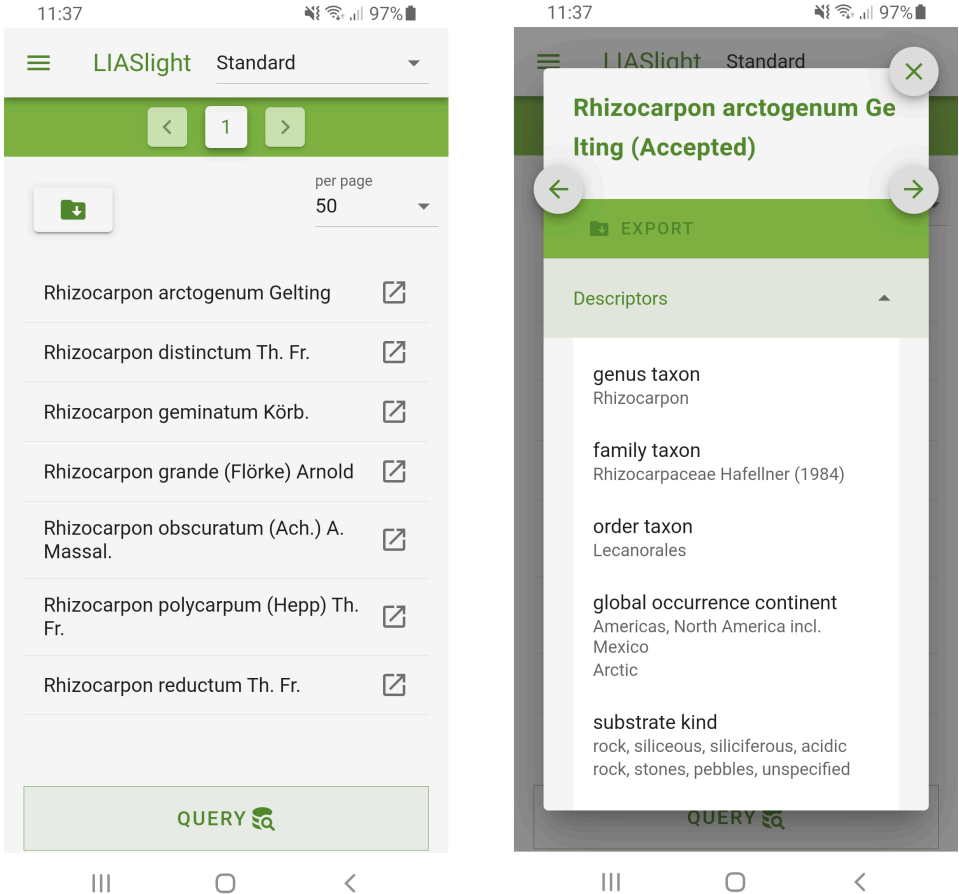


Fig. 5 DiversityNaviKey: Result user interface with LIAS data (screenshot of test installation, result for mobiles)

As this kind of search does not cover all user requirements, DNK allows the settings for the AND/OR group to be changed. This opens up further query and retrieval options, such as an even more restrictive search by combining the values/states within a descriptor/facet conjunctively (AND) or extending the search for more results by applying the descriptors disjunctively (OR). Furthermore, the combination of different AND/OR groups within a single search is possible in the extended query mode.

This type of search and filter approach covers the requirements of biologists who, on the one hand, want to identify natural objects with unclear and fuzzy features and, on the other hand, rely on descriptive data sets with incomplete and subjective data entries that are the result of interpretation by the data creators.

Selected query criteria can be saved in JSON format and reused as pre-sets. In particular, limiting the dataset to a subset that can be worked with effectively (e.g. a specific set of taxa for a specific region, country or continent) is essential in this context.

DiversityNaviKey separates a ‘query mode’, which treats missing data as a given feature, from a ‘diagnostic mode’, which treats missing values as given in case of doubt. This ‘tolerance’ option has already been implemented in local diagnostic applications such as Intkey of the CSIRO DELTA package [Da80, CLM10].

Result list and descriptive information

The matches or results from a given query are output as a list of matching object names in alphabetical order (Fig. 5). In the biology domain with a focus on organism identification, these objects are often represented by taxon names for which additional information may be displayed as part of the result, e.g. currently accepted taxon names. The description data are generated directly from the records. They are represented in tabular form (desktop version) or list view (mobile version). In a later version, additional information could be well presented by linking to internet sources, as is planned for the multimedia object information.

Multilingualism

DiversityNaviKey even supports a multilingual approach. In the LIAS light use case, content data is provided in more than 20 different languages. Switching between languages also affects labels and explanatory texts on the user interface. Multilingualism also includes the possibility to provide descriptors and descriptor values of a specific language in different levels of linguistic and content precision, from detailed to informal, and to select them depending on the context in the application.

3.6 Diversity Workbench data pipelines with DiversityDescriptions as the backend data source

Both the web service and the web application DiversityNaviKey are installed in their prototypic version at the SNSB IT Center (Fig. 3). The backend data sources come from a PostgreSQL database with a simple ‘flattened’ data structure, i.e. the DiversityDescriptions cache database. This database is a direct backend and is dynamically generated by data export, filtering and transformation mechanisms from one or more DiversityDescriptions installations [DD21].

DiversityDescriptions (DWB-DD) itself is a relational MS SQL database application with a rich client. It is a component of the Diversity Workbench (DWB) database framework [DW21a], which is set up to manage research and collection data in study projects. DWB-DD can be installed as part of a DWB environment, but can also be used as a stand-alone application. The ten DWB core modules have rich clients adapted for interaction with users from biology and related sciences.

All modules, except DWB-DD, are based on a specific digital object domain (e.g. specimen and occurrence data, taxon data, agent data etc.; DWB – Software [DW21b]). DWB-DD, on the other hand, is generic and built on a model corresponding to the Entity Attribute Value Model (EAV) [En21], an ‘open schema’. The design and ER data model of DWB-DD have been published (see under ‘DWB – Information models’ [DW21c]).

The DWB-DD rich client has functions for organising research data from environmental studies [Ha20] and can be used for curating schema elements [Ra18, Ra19]. It is also a tool for taxonomists to manage descriptive data from monographic studies and for internal diagnostic purposes. There are functions to import and export SDD- and DELTA-structured digital objects [Ha20]. The client is capable of managing character hierarchies and dependencies, combining object descriptions (e.g. specimen descriptions) into a summary description (“taxon or species description”) and is ready to link knowledge organization systems, e.g. taxonomies available as web services or as part of DWB cloud services [DW21d]. Identity and rights management can be set on user request. In summary, DWB-DD is well suited to serve as a back-end data management system to enrich and optimize data sources in preparation for data transformation to DiversityNaviKey.

3.7 Two use cases of biodiversity datasets from long-term projects

The prototype version of DiversityNaviKey has been tested and realised with several DNK use cases and access to six backend data sources. The content of the datasets is shortly addressed on the introductory web page of DNK [DNK21]. LIAS light [LI21] (Fig. 4, 5) and DEEMY [DEE21], are by far the largest among them. These data originate from two long-term projects with continuous data curation by domain experts [Ra97, RA97, Ra14, Ra16, Tr10]. They include two sets of data on morphology, secondary metabolism, geoinformation and ecology of mutualistic associations: lichens and ectomycorrhizae. Both data sources are comprehensive, with DEEMY being mainly based on original data from field and laboratory analysis, and LIAS light on the compilation of data from reference literature. LIAS light contains more than 11,000 species descriptions and DEEMY descriptions of 550 species or taxonomically unassigned fungal-tree associations. The data are highly structured, with the LIAS data source containing more than 70 descriptors and more than 2,900 predefined values, and DEEMY providing 500 descriptors with more than 1,300 categorical values. In addition to categorical descriptors, both data sources also contain textual and quantitative descriptors. DEEMY uses its own image resources, LIAS uses external resources.

The data is managed in DiversityDescriptions, quality-controlled and curated with respect to current standard vocabularies, taxonomy and classifications, and dynamically made available for use with DiversityNaviKey (Fig. 3). Both data sources are not only established for diagnostic purposes, but also provide a range of non-diagnostic characters for other purposes.

The DNK use cases LIAS light and DEEMY will be made accessible via their respective web sites.

4 Discussion and future

A Progressive Web App like DiversityNaviKey is a web application that provides an app-like experience through the browser. Unlike native apps, it is independent of operating systems and devices, but depends on the browser's capabilities of supporting PWA features such as Service Worker and IndexedDB. Popular modern browsers (Firefox, Edge, Chrome, Safari) support the Service Worker API (see <https://caniuse.com/serviceworkers>; June 2021), which enables offline support for the PWAs. DNK uses the browser's Service Worker for caching the web application (HTML, JS, CSS). The browser's IndexedDB is used to cache the content data from the web service. If a browser does not support the IndexedDB for data storage, e.g. if private browsing mode is selected, DNK can still be used in online mode. Since browsers have their own tools to calculate and keep track of memory and remove data if necessary, in rare cases the browser itself may delete cached data. In such cases, the user must go online before the web app can be used in offline mode again. The descriptive information of all items in DNK may be a large dataset, e.g., LIAS light with approximately 11,000 taxon descriptions. Therefore, DNK does not load and cache this data by default, but the user is offered an alternative option to do so.

DNK is implemented by using the popular open source Vue.js framework, which is supported by a strong open-source community. Only highly supported and maintained libraries (Vue Router, Vuex, Vuetify) are used as extensions to the core Vue.js framework to ensure future support. The DNK web app is designed as a generic tool that provides the possibility to integrate different data sources via the web service. There is also the possibility to extend the DNK web app to replace the web service with other web services that fulfil the interface. The DNK software architecture is designed to support future extension with new search and filtering algorithms.

Upcoming DiversityNaviKey versions will address character and state dependence shortcomings and improve the display of linked web-based multimedia resources. Future plans also include software enhancements to facilitate data export from the web app in CSV format. This could support ideas to use DNK for simple species monitoring approaches to record the occurrence of taxa at a given spatial location and export search result files. For such purposes, DNK could be an alternative solution for the DiversityMobile app [DM21, Ja09], but only in its basic functions. DiversityMobile was a successful tool for species monitoring projects over years but is now outdated due to the underlying Windows Phone platform. With data export in CSV format the user might also get the option of data exchange with Xper 2 and the Xper 3 platform [Xp21a, Xp21b, Xp21c], with DELTA IntKey [DEL21b] and the Open DELTA network [Op21].

The Java web interface NaviKey [Na21a] was the predecessor of DiversityNaviKey. Since 2007 it has been used by a number of projects worldwide [Na21b] with a PostgreSQL database as backend. Because of security reasons, Java applet technology was generally abandoned. With the new web app DiversityNaviKey, all these projects might have a chance of revival, provided the backend data source is adapted.

DiversityNaviKey is not primarily designed as an identification tool for Citizen Science projects that focus on educational and science communication aspects. Its target groups are scientists and other professionals. DNK structures and queries data and makes existing (published) knowledge of taxonomists or other bio- and geodiversity experts usable for (scientific) diagnosis and identification. Some recent approaches for identification keys in citizen science projects are implemented as PWAs [Ba20]. However, most are realised as native apps for a single user community. In these cases the targets are primarily those appealing to citizens, such as butterflies, birds, mammals and flowering plants. In addition, most tools are optimized on processing and analysing multimedia data (images, sounds). Some implement and use computer vision and machine learning techniques for (semi-) automatic identification approaches [F121, WM18, WM19]. Future collaboration to compare the technical design and data of the diagnosis tool implementations for citizens on the one hand and DiversityNaviKey for professionals and scientists on the other hand could be promising.

An important vision for the future may be that individual biodiversity researchers or research groups use the option to manage research projects with descriptive data in their own local DWB-DD installation. Alternatively they use DWB-DD installations hosted by a recognized domain-specific data repository or installations offered by a recognized technical service provider, e.g., GFBio e.V. (keyword: Software as a Service, SaaS). At some point, researchers decide to make their data available for open access via data publishing services (e.g., as EML-SDD structured data files) and – in parallel – to make the data openly and freely available via DiversityNaviKey diagnostic and identification services. Those services could be well integrated into the application layer of cloud computing architectures.

DiversityNaviKey outside the Diversity Workbench in a comprehensive data infrastructure

The DiversityNaviKey PWA with the web service is technically completely independent from the Diversity Workbench database suite (see chapter 3 and Fig. 1, 3). The components can be dockerised to allow an easy deployment of static data sources. In this way DNK might become one of the envisaged “practical tools” of the NFDI4BioDiversity [G120, NF21]. NFDI4BioDiversity is one of the consortia of the evolving German National Research Data Infrastructure (NFDI). As part of the application layer of the NFDI Research Data Commons (NFDI-RDC) [G120] it might be a useful completion to the GFBio tool for analysis and visualisation (VAT) [Be17, GF21d]. Similar to what the GFBio data portal and VAT tool already do for ABCD-XML-structured data, a DiversityNaviKey service might support the usage of SDD-structured XML trait data as already now available via the GFBio data portal [GF21c].

In its envisaged extension DNK related services will transform the SDD-XML files to suit the database schema of a cloud-based DNK PostgreSQL cache database and DNK web services. With this realisation, the data can be handled by a DiversityNaviKey web app for identification and diagnostic purposes.

The DiversityNaviKey programming code with web application and web service including documentation of the underlying DiversityDescriptions cache database (PostgreSQL) as backend will be available under <https://github.com/snsb>. The DiversityNaviKey app with exemplary data sources will soon be available on a server of the SNSB IT Center under <https://divnavikey.snsb.info>.

The LIAS and DEEMY long term web projects will directly integrate DiversityNaviKey services.

5 Acknowledgements

The work is partly granted by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG), LIS infrastructure programme with the project ‘German Federation of Biological Data (GFBio, <https://www.gfbio.org>)’, funding reference TR 290/7-3. Support for Diversity Workbench software engineering and concept development is received from the Bavarian State Ministry of Science and the Arts (StMWK), the DFG project SFB 1357, no. 391977956 and from the European ERDF fund/ ETC-INTEREG V program, action no. 216. We would like to thank Eva Haiduk, Dieter Neubacher and Markus Weiss (all Munich, Germany) for stimulating discussions and software testing.

6 References

- [ADC16] Attigala L.; De Silva N.I.; Clark L.G.: Simple Web-based interactive key development software (WEBiKEY) and an example key for Kuruna (Poaceae: Bambusoideae). *Appl Plant Sci.*, 4(4): apps.1500128, 2016.
- [Ba20] Batsaikhan, A.; Hachinger, S.; Kurtz, W.; Heller, H.; Frank, A.: Application of Modern Web Technologies to the Citizen Science Project BAYSICS on Climate Research and Science Communication. *Sustainability*, 12(18): 7748, 2020. <https://doi.org/10.3390/su12187748>
- [Be17] Beilschmidt C.; Dröner J.; Mattig M.; Schmidt M.; Authmann C.; Niamir A.; Hickler T.; Seeger B.: VAT: A Scientific Toolbox for Interactive Geodata Exploration. *Datenbank-Spektrum*, 17(3): 233-243, 2017. <https://doi.org/10.1007/s13222-017-0266-5>

- [CLM10] Coleman, C.; Lowry, J.; Macfarlane, T.: DELTA for Beginners. An introduction into the taxonomy software package DELTA. *ZooKeys*, 45: 1–75, 2010. <https://doi.org/10.3897/zookeys.45.263>
- [Da17] Dahdul, W; Balhoff, J.; Lapp, H.; Uyeda, J.; Vision, T.: Enabling machine-actionable semantics for comparative analyses of trait evolution. Zenodo, 2017. <https://doi.org/10.5281/zenodo.885538>
- [Da80] Dallwitz, M.J.: A General System for Coding Taxonomic Descriptions. *Taxon*, 29 (1): 41–46, 1980. <https://doi.org/10.2307/1219595>. JSTOR 1219595. S2CID 85981894.
- [DD21] DWB – DiversityDescriptions (DWB-DD), <https://diversityworkbench.net/Portal/DiversityDescriptions>, accessed: 2021-05-06.
- [DEE21] DEEMY – An Information System for Characterization and Determination of Ectomycorrhizae, <http://www.deemy.de/>, (see also re3data repository entry under <http://doi.org/10.17616/R3ZF5V>), accessed: 2021-05-07.
- [DEL21a] DELTA (taxonomy) data format, [https://en.wikipedia.org/wiki/DELTA_\(taxonomy\)](https://en.wikipedia.org/wiki/DELTA_(taxonomy)), accessed: 2021-05-07.
- [DEL21b] DELTA Intkey, <https://www.delta-intkey.com/>, accessed: 2021-05-07.
- [Di97] Diederich, J.: Basic properties for biological databases: Character development and support. *Mathl. Comput. Modelling*, 25(10): 109-127, 1997. [https://doi.org/10.1016/S0895-7177\(97\)00078-2](https://doi.org/10.1016/S0895-7177(97)00078-2)
- [DM21] DWB – DiversityMobile (DWB-DM), <http://www.diversitymobile.net>, accessed: 2021-05-07.
- [DN21] DN – Diversity Navigator, <http://www.diversitynavigator.net/>, accessed: 2021-05-06.
- [DNK21] DiversityNaviKey (DNK), <https://diversityworkbench.net/Portal/DiversityNaviKey>, accessed: 2021-07-07.
- [DW21a] DWB – Diversity Workbench, <https://diversityworkbench.net>, accessed: 2021-05-06.
- [DW21b] DWB – Software: <https://diversityworkbench.net/Portal/Software>, accessed: 2021-05-06.
- [DW21c] DWB – Information Models, https://diversityworkbench.net/Portal/Information_models, accessed: 2021-05-06.
- [DW21d] DWB cloud services, https://diversityworkbench.net/Portal/DWB_cloud_services, accessed: 2021-05-06.
- [En21] Entity attribute value model (EAV), https://en.m.wikipedia.org/wiki/Entity%E2%80%93attribute%E2%80%93value_model, accessed: 2021-05-05.
- [FI21] Flora Incognita, <https://floraincognita.com>, accessed: 2021-05-07.
- [GF21a] GFBio – German Federation for Biological Data, <https://www.gfbio.org>, Accessed: 2021-05-05.
- [GF21b] GFBio relevant standards, protocols and formats, https://gfbio.biowikifarm.net/wiki/Data_exchange_standards_protocols_and_formats









relevant_for_the_collection_data_domain_within_the_GFBio_network,
accessed: 2021-05-07.

- [GF21c] GFBio type 3 data mit SDD und EML, https://gfbio.biowikifarm.net/wiki/Major_Types_of_Biological_Data#Type_3:_Environmental_Biological_and_Ecological_Data, accessed: 2021-05-07.
- [GF21d] GFBio VAT Tool, <https://www.gfbio.org/de/visualize>, accessed: 2021-05-05.
- [GI20] Glöckner, F.O.; Diepenbroek, M.; Felden, J.; Güntsch, A.; Stoye, J.; Overmann, J.; Wimmers, K.; Kostadinov, I.; Yahyapour, R.; Müller, W.; Scholz, U.; Triebel, D.; Frenzel, M.; Gemeinholzer, B.; Goesmann, A.; König-Ries, B.; Bonn, A.; Seeger, B.: NFD14BioDiversity – A Consortium for the National Research Data Infrastructure (NFDI). Zenodo, 2020. <https://doi.org/10.5281/zenodo.3943645>
- [Ha05] Hagedorn, G.; Thiele, K.; Morris, R.; Heidorn, P.B.: Structured Descriptive Data (SDD) w3c-xml-schema, Version 1.0. Biodiversity Information Standards (TDWG) <http://www.tdwg.org/standards/116> (version 1.1 available), 2005. accessed: 2021-05-05.
- [Ha08] Hagedorn, G.: Structuring descriptive data of organisms – Requirement analysis and information models. Dissertation, Universität Bayreuth, pp. 1-417, 2007/2008.
- [Ha10] Hagedorn, G.; Weber, G.; Plank, A.; Giurgiu, M.; Homodi, A.; Veja, C.; Schmidt, G.; Mihnev, P.; Roujinov, M.; Triebel, D.; Morris, R.A.; Zelazny, B.; Van Spronsen, E.; Schalk, P.; Kittl, C.; Brandner, R.; Martellos, S.; Nimis, P.L.: An online authoring and publishing platform for field guides and identification tools. In (Nimis, P.L.; Vignes Lebbe, R. (eds.)): Tools for identifying biodiversity: progress and problems, Paris, pp.13-18, 2010.
- [Ha20] Harjes, J.; Link, A.; Weibulat, T.; Triebel, D.; Rambold, G.: FAIR digital objects in environmental and life sciences should comprise workflow operation design data and method information for repeatability of study setups and reproducibility of results. Database, 2020 (Article ID baaa059): 1–20, 2020. <https://doi.org/10.1093/database/baaa059>
- [HRM10] Hagedorn, G.; Rambold, G.; Martellos, S.: Types of identification keys. In (Nimis, P.L.; Vignes Lebbe, R. (eds.)): Tools for identifying biodiversity: progress and problems, Paris, pp. 59-64, 2010.
- [Ja09] Jablonski, S.; Kehl, A.; Neubacher, D.; Poschlod, P.; Rambold, G.; Schneider, T.; Triebel, D.; Volz, B.; Weiss, M.: DiversityMobile – Mobile data retrieval platform for biodiversity research projects. In (Fischer, S.; Maehle, E.; Reischuk, R. (eds.)): Informatik 2009 – Im Focus das Leben. Beiträge der 39. Jahrestagung der Gesellschaft für Informatik e. V. (GI). GI-Edition: Lecture Notes in Informatics (LNI), Proceedings 154: 27, pp.610–624, 2009.
- [LI21] LIAS light – A Database for Rapid Identification of Lichens, <https://liaslight.lias.net/>, see also re3data repository entry under <http://doi.org/10.17616/R3134N>, accessed: 2021-05-05.
- [Lu21] Lucid, <https://www.lucidcentral.org/>, accessed: 2021-05-05.
- [Ma21] Material Design specification, <https://material.io/design>, accessed: 2021-05-05.

- [Na21a] NaviKey, <http://www.navikey.net/>, accessed: 2021-04-05.
- [Na21b] NaviKey – Webprojects, <http://www.navikey.net/webprojects.html>, accessed: 2021-04-05.
- [NF21] NFDI4BioDiversity – A Consortium for the National Research Data Infrastructure (NFDI), <https://www.nfdi4biodiversity.org/>, accessed: 2021-05-05.
- [NM20] Nimis, P.L.; Martellos, S.: Towards a digital key to the lichens of Italy. *Symbiosis*, 82: 149-155, 2020. <https://doi.org/10.1007/s13199-020-00714-8>
- [NV10] Nimis, P.L.; Vignes Lebbe, R. (eds.): Tools for identifying biodiversity: progress and problems. Proceedings of the International Congress, Paris, September 20-22, 2010, pp.1-480, 2010.
- [Of21] Offene Naturführer – Bestimmungsschlüssel, https://offene-naturfuehrer.de/web/Kategorie:Alle_Bestimmungsschlüssel, accessed: 2021-04-05.
- [Op21] Open DELTA, [https://downloads.ala.org.au/p/Open%20DELTA%20\(Atlas%20of%20Living%20Australia](https://downloads.ala.org.au/p/Open%20DELTA%20(Atlas%20of%20Living%20Australia)), accessed: 2021-04-05.
- [Pu05] Pullan, M.R.; Armstrong, K.E.; Paterson, T.; Cannon, A.; Kennedy, J.B.; Watson M.F.; McDonald, S.; Raguenaud, C.: The Prometheus Description Model: An examination of the taxonomic description-building process and its representation. *Taxon*, 54(3): 751–765, 2005. <https://doi.org/10.2307/25065431>
- [Ra97] Rambold, G.: LIAS – the concept of an identification system for lichenized and lichenicolous ascomycetes. In (Türk, R. & Zorer, R. (eds.)): *Progress and problems in lichenology in the Nineties -IAL 3. Biblioth. Lichenol.* 68: 67-72, 1997.
- [RA97] Rambold, G.; Agerer, R.: DEEMY – the concept of a characterization and determination system for ectomycorrhizae. *Mycorrhiza* 7: 113-116, 1997.
- [Ra14] Rambold, G.; Elix, J.A.; Heindl-Tenhunen, B.; Köhler, T.; Nash III, T.H.; Neubacher, D.; Reichert, W.; Zedda, L.; Triebel, D.: LIAS light – Towards the ten thousand species milestone. *MycKeys*, 8: 11–16, 2014. <https://doi.org/10.3897/mycokeys.8.6605>
- [Ra16] Rambold, G.; Zedda, L.; Coyle, J.; Peršoh, D.; Köhler, T.; Triebel, D.: Geographic heat maps of lichen traits derived by combining LIAS light description and GBIF occurrence data, provided on a new platform. *Biodiversity and Conservation* 25(13): 2743–2751, 2016. <https://doi.org/10.1007/s10531-016-1199-2>
- [Ra18] Rambold, G.; Yilmaz, P.; Harjes, J.; Link, A.; Glöckner, F.O.; Triebel, D.: MOD-CO schema – a conceptual schema for processing sample data in meta’omics research (version 1.0), http://mod-co.net/wiki/MOD-CO_Schema_Reference, 2018. accessed: 2021-05-05.
- [Ra19] Rambold, G.; Yilmaz, P.; Harjes, J.; Klaster, S.; Sanz, V.; Link, A.; Glöckner, F. O.; Triebel, D.: Meta-omics data and collection objects (MOD-CO): A conceptual schema and data model for processing sample data in meta-omics research. *Database*, 2019 (Article ID baz002): 1–13, 2019. <https://doi.org/10.1093/database/baz002>

- [SD21] SDD – Structured Descriptive Data, <https://github.com/tdwg/sdd>, accessed: 2021-05-07.
- [Tr10] Triebel, D.; Neubacher, D.; Weiss, M.; Heindl-Tenhunen, B.; Nash, T. H. III; Rambold, G.: Integrated biodiversity data networks for lichenology – data flows and challenges. In (Nash, T.H. III; Geiser, L.; McCune, B.; Triebel, D.; Tomescu, A.M.F.; Sanders, W.B. (eds.)): *Biology of lichens – symbiosis, ecology, environmental monitoring, systematics and cyber applications*. *Biblioth. Lichenol.* 105: 47–56, 2010.
- [Un10] Ung, V.; Dubus, G.; Zaragüeta-Bagils, R.; Vignes Lebbe, R.: Xper2: Introducing e-taxonomy. *Bioinformatics* 26(5), 2010. <https://doi.org/703-4.10.1093/bioinformatics/btp715>
- [Vo17] Vogt, L.: Assessing similarity: on homology, characters and the need for a semantic approach to non-evolutionary comparative homology. *Cladistics* 33(5): 513–539, 2017. <https://doi.org/10.1111/cla.12179>
- [Vu21a] Vue.js Framework, <https://vuejs.org/>, accessed: 2021-05-07.
- [Vu21b] Vue Router, <https://router.vuejs.org/>, accessed: 2021-05-07.
- [Vu21c] Vuetify, <https://vuetifyjs.com/en/>, accessed: 2021-05-07.
- [Vu21d] Vuex, <https://vuex.vuejs.org/>, accessed: 2021-05-07.
- [Wa14] Walls RL, Deck J, Guralnick R, Baskauf S, Beaman R, Blum S, et al.: Semantics in Support of Biodiversity Knowledge Discovery: An Introduction to the Biological Collections Ontology and Related Ontologies. *PLoS ONE* 9(3): e89606, 2014. <https://doi.org/10.1371/journal.pone.0089606>
- [WM18] Wäldchen, J.; Mäder, P.: Machine learning for image based species identification. *Methods in Ecology and Evolution* 9(8): 2216–2225, 2018. <https://doi.org/10.1111/2041-210X.13075>
- [WM19] Wäldchen, J.; Mäder, P.: Flora Incognita – how artificial intelligence revolutionizes plant identification. *Biol. Unserer Zeit* 49: 99–101, 2019. <https://doi.org/10.1002/biuz.201970211>
- [Xp21a] xper Software Wiki, <http://wiki.xper3.fr/doku.php>, accessed: 2021-05-07.
- [Xp21b] xper2 Desktop software, <https://infosyslab.fr/?q=en/resources/software/xper2>, accessed: 2021-05-07.
- [Xp21c] xper3 Collaborative descriptive data platform, <http://xper3.fr/>, accessed: 2021-05-07.

A pragmatic approach to concept-based annotation of scientific names in biodiversity and environmental research data

Andreas Kohlbecker ¹, Anton Güntsch ², Norbert Kilian ³, Wolf-Henning Kusber ⁴, Katja Luther ⁵, Andreas Müller ⁶, Eckhard von Raab-Straube ⁷ and Walter Berendsohn ⁸

Abstract: With the increasing amount of interdisciplinary and international networks dedicated to long-term persistence and interoperability of research data, the demand for semantic linking of environmental research data has grown. Data related to organisms frequently inherit a major obstacle. Organisms often are ambiguously identified by using only the scientific name, which is not a precise identifier for the taxonomic concept that is implicitly being used. Here we describe a robust taxon concept definition that allows deducing a set of rules for semi-automatically managing concepts. These rules define specific taxonomic operations as transition points at which new taxon concepts emerge from former concepts. Implemented into the business logic of taxon management systems, these rules can assure the stability of taxon concepts so that environmental data sets can be reliably annotated with the corresponding persistent identifier. Our approach limits the risk that referenced taxon concepts are modified unnoticed.

Keywords: Annotation, biodiversity, environmental data, identifier, linked open data, scientific name, semantics, taxonomy, taxonomic concept.

1 From names to concepts

Due to the increasing availability of powerful network technologies and the international standardization of access protocols and data formats, ever-larger volumes of biodiversity-related research data are being made available on the Internet for the reproducibility and subsequent use of research results. Important data sources include

¹ BGBM Berlin, FU-Berlin,* A.Kohlbecker@bgbm.org,  <https://orcid.org/0000-0003-1046-8750>

² BGBM Berlin, FU-Berlin,* A.Guentsch@bgbm.org,  <https://orcid.org/0000-0002-4325-4030>

³ BGBM Berlin, FU-Berlin,* N.Kilian@bgbm.org,  <http://orcid.org/0000-0003-0890-4373>

⁴ BGBM Berlin, FU-Berlin,* W.H.Kusber@bgbm.org,  <https://orcid.org/0000-0003-4543-5764>

⁵ BGBM Berlin, FU-Berlin,* K.Luther@bgbm.org,  <https://orcid.org/0000-0002-3404-4680>

⁶ BGBM Berlin, FU-Berlin,* A.Mueller@bgbm.org,  <https://orcid.org/0000-0003-1609-7882>

⁷ BGBM Berlin, FU-Berlin,* E.Raab-Straube@bgbm.org,  <https://orcid.org/0000-0001-6561-4490>

⁸ BGBM Berlin, FU-Berlin,* W.Berendsohn@bgbm.org,  <http://orcid.org/0000-0001-9945-7606>

*Botanic Garden and Botanical Museum Berlin, Freie Universität Berlin, Germany, Department of Biodiversity Informatics, Königin-Luise-Str.6-8, D-14195 Berlin

international portals, such as GBIF (<https://www.gbif.org>), which provide open and free access to thousands of decentrally operated biodiversity databases but also centers with a focus on data generated by projects of national research funding agencies [DGG14].

The use of standards is an important prerequisite for the interoperability of data from disparate sources, but they are not sufficient if data fields are described semantically but the content continues to be transmitted as texts that cannot be unambiguously assigned to a semantics [GGH20]. In the biodiversity data domain, this problem is particularly evident in the use of species names.

A scientific name in these data mostly denotes a taxon concept at the ranks of species or below (subspecies, variety). According to [GB03] taxon concepts are defined by the “set of explicit or implicit criteria, which allow deciding whether a particular element (specimen, observation or lower taxon) belongs to a taxon or not.” Conversely, a taxon concept can be seen as the set of all individual organisms that fulfill all descriptive features covered by its circumscription and that are not fully covered by the circumscriptions of another taxon of the same rank. Changing knowledge and different options on the specific characteristics of a taxon concept may cause specific individual organisms to no longer belong to that concept per its definition. That is, the concept has changed even if its scientific name remains the same (rules of biological nomenclature immutably bind the scientific name to a single element).

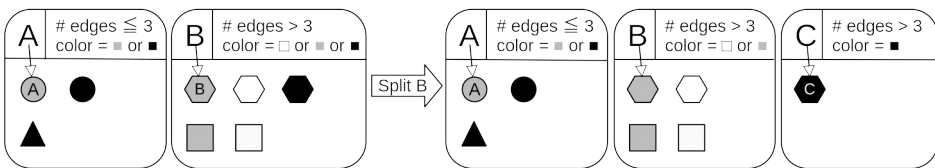


Fig. 1: Uppercase single letters (A, B, C) are scientific names denoting taxon concepts (rounded boxes). Names are immutably bound to a single element per nomenclature (arrow). Each concept is defined by the circumscription consisting of the number (#) of edges and colors of the contained elements (colored shapes). The taxonomic operation “split” applied on “B” creates two new concepts named “B” and “C”. Even if bearing the same name “B”, the earlier and later taxon concepts are significantly different.

The uninformed use of names as the key element for the integration and automated inference of species-related research data from disparate sources is therefore highly problematic.

One approach to overcoming this uncertainty is to annotate scientific names in research data with stable and persistent references to taxonomic concepts, which are used to represent taxa as semantic resources independent from the scientific name [FP06].

The definition of conditions that may lead to concept changes and associated reassignment of identifiers is non-trivial [GB10] and must be communicated to the users

of the associated services in an appropriate form. Within the framework of the NFDI4BioDiversity Initiative [GI20], services for the annotation of botanical research data are being created using the Euro+Med Plantbase (<http://europlusmed.org/>) as a proof of concept, and practicable annotation workflows for data providers are being defined. Based on this, an extension of the approach to other organism groups is planned.

2 Related work and differentiation

Others have presented approaches for expressing taxon concepts and their relationships. This includes the following works: The “Linked Taxonomic Knowledge” (LTK) aims at preserving and presenting the change in taxonomic knowledge for linked data [CT16]. With TAXREF Michel et al. [MG17] present a semantic model for linking with remote LOD (Linked Open Data) sets, may they represent nomenclatural or taxonomic information. The Taxon Concept Schema [Ta06] is an XML standard that was conceived to allow the representation of taxonomic concepts as defined in published taxonomic classifications, revisions and, databases.

While all these works are on the representation of taxon concepts and their interrelations, provenance, and history, none of them focuses on a robust definition of taxon concepts which prevents taxa from being changed unnoticed due to changes in the underlying data.

Here we describe an approach to overcome this limitation. It relies on a robust taxon concept definition which makes it possible to identify taxonomic operations as transition points at which new taxon concepts emerge. This is the base for defining rules for managing taxon concepts semi-automatically. Managing the formation of new taxon concepts has always been a manual step, for which expert editors of taxonomic projects are responsible. They are required to be aware of the changes in the taxon circumscription, which will cause new taxon concepts to emerge and to subsequently create a new taxon concept with a new persistent identifier.

A subset of these rules describes operations in which the emergence of a new taxon concept is indisputable so that a new taxon concept with an identifier has to be created. Other cases are not mandatory, and the decision whether to create a new taxon concept or not is the editor's responsibility. Even if automatic management of taxon concept transitions is not possible in the latter cases, a system that implements these rules can notify and assist editors in these ambiguous situations.

The aim and scope of this work are purely on annotating data with persistent identifiers for taxon concepts that are stabilized by a rule-based system. The semantic representation of taxon relations and background information, reasons, and provenance are explicitly out of the scope.

3 Concept-based annotation

The approach of concept-based annotation as described here relies on four key elements, (1) a robust and stringent definition of taxon concepts on the base of the taxonomically relevant data associated with scientific names, (2) reliable, unique, and persistent identifiers for all taxon concepts, (3) semi-automatic workflow support for the creation of new taxon concepts including the relations between precursors and successors when concepts are changing, and (4) means for swift matching of names in environmental research data as accurately as possible to presumably used taxon concepts.

We use classifications and synonymies to delimit taxonomic concepts rather than descriptive information, which is notoriously difficult to compare and can lead to extremely high numbers of possible concepts in the system which are difficult to handle in practice. To define manageable taxon concepts we take advantage of the fact that scientific organisms follow strict nomenclatural rules and are based on (ideally) unequivocal conserved specimens called nomenclatural ‘types’. The type specimens are available in research collections for comparison. If several type specimens (and thus names) fall into the circumscription of a single species, the oldest name automatically becomes the accepted name (principle of priority), while the others are treated as synonyms. In the hierarchical classification system, child taxa (subspecies, varieties) are analogously defined and named. Thus, a taxonomic concept can be defined as the list of all names included. The type specimens of these names define a set of criteria (morphological and molecular characters, location, collection date, etc) that form the delimitation of the implicitly circumscribed taxon concept. This definition of taxon concepts is fully compliant with the definitions in [GB03] (see also chapter 1).

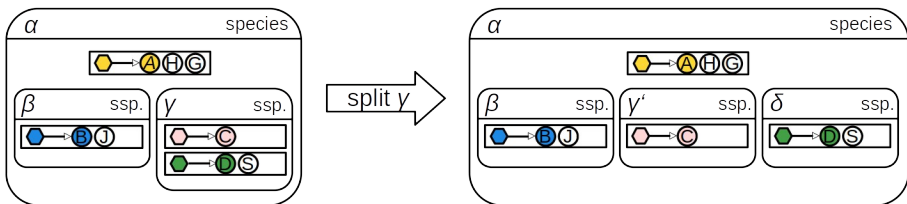


Fig. 2: Taxon concept (rounded boxes) definition by classification and homotypic groups (rectangles) of names (circle) that are assigned to the same type specimen (hexagon). The oldest name associated (arrow) with the type specimen is the accepted name. In the case of concept α the name A is the oldest and thus name-giving. Other names are synonyms, for example, H and G are synonyms to A. The concept β is defined by the homotypic group of name B. Concept α is defined by the homotypic groups contained within α and β and γ since the subspecies are contained in A per classification hierarchy. Splitting the homotypic group containing name D from concept γ creates two new taxon concepts γ' and δ . The name D becomes the accepted name for δ . The concept α remains unchanged since it still contains the same set of homotypic groups.

The border of the circumscription of a single taxon may be fuzzy but is, at the same time, clearly delimited by the names (and inherent characteristics) included in the concepts of the contiguous taxa in the same classification. This way, the concept becomes relatively robust and comparable in comparison to concepts based on explicitly defined feature descriptions which may be extended or even changed frequently.

Taxon concepts used as a reference system for annotations must never be changed. Once type specimens are added or removed from the set associated with a taxon concept. A new concept and corresponding relationships, like for example “split from” or “merged into”, between the predecessor and successor concepts have to be created. This allows data inference depending on the type of relationship in consuming systems. A suitable approach to express these relations semantically in Linked Open Data has been described in [CT16].

To assure stable interrelated taxon concepts, taxonomists would have to be aware in advance of the consequences of editing processes, that might alter a concept. Prior to the operation the predecessor concept need to be cloned, and taxon relations corresponding to the operation needs to be created, between both. Finally, a new persistent identifier needs to be assigned to a new concept. This manual operation is error-prone and the obligation to do so is easily forgotten.

To perform these steps we, therefore, suggest a semi-automatic mechanism, which is triggered by specific taxonomic operations. Depending on the type of operation the process of creating a new taxon and establishing appropriate relations can happen with (semi-automatic) or without (automatic) user intervention:

- Automatic: Split taxon (Fig. 1 & 2), combine taxa, move synonym to other accepted taxon, swap heterotypic synonym with accepted taxon
- Semi-automatic: New child taxon, new heterotypic synonym.

In the latter cases, the emergence of a new concept is not implicit, therefore systems ought to delegate the decision to the expert user. Practically this can be implemented into the business logic of taxonomic data management systems. Service methods that perform the above operations in a transactional context would either automatically execute sub-processes to manage the involved taxon concepts and relations, or would in advance trigger actions in the UI to poll for the user’s decision.

For linking scientific names in research data to taxon concepts, persistent identifiers are crucial. These must also be unique and stable over an adequately long time, that is, as long as the identifier is being used, it must stay resolvable. We decided on HTTP URIs as identifiers, since these are intrinsically resolvable. The organizational and management efforts to guarantee the continued availability of the linked resources are well understood and these are directly actionable in many digital media. Moreover, HTTP URIs are recommended building blocks for Linked Open Data networks [HAV14].

4 Application

We have chosen Euro+Med PlantBase [EM21, GR05] as a pilot project for developing a pragmatic approach to taxonomic concept management. E+M provides a dynamic taxonomic checklist of all native and introduced vascular plants of Europe, the Mediterranean region, and the Caucasus. The E+M PlantBase is continuously updated by an international network of taxonomic and regional editors and advisers. As soon as new information on the taxonomy, nomenclature, the geographic distribution, or the status of taxa becomes available, those data are being integrated into the current taxonomic treatments of E+M. Since E+M is the taxonomic reference system for many regional floras and checklists and as it is part of EU-Nomen (<http://www.eu-nomen.eu>) which is mentioned as the preferred taxonomic backbone in the EU's INSPIRE directive [IN13], E+M constitutes an excellent use case for concept-based annotations. At present, E+M contains 46,356 accepted names at generic, specific, and subspecific ranks, and over 7100 explicitly created taxon relationships. Technically E+M PlantBase is based on the EDIT Platform for Cybertaxonomy, a comprehensive system for managing taxonomic information including concept relationships [GL10], which furthermore allows the implementation of the beforehand mentioned semi-automatic taxon concept management workflow. In this regard, E+M has significant advantages over other projects like the World Flora Online (<http://www.worldfloraonline.org>) or "The Plant List" (<http://www.theplantlist.org>). Even if these are having a much broader geographic scope, E+M also fulfills the requirement for a taxonomic backbone in NFDI4BioDiversity.

The Pilot will be developed in the context of the NFDI4BioDiversity initiative that is dedicated to make the variety of biodiversity data from a multitude of sources available and to unify and harmonize the disparate and heterogeneous data. This also involves that data labeled with scientific names only is annotated with a reference to the taxon concept which has originally been used implicitly in the creation of the data. This concept is not necessarily the concept currently accepted by a checklist like E+M. The annotation will either be applied at the individual environmental data set or if possible at the taxonomic checklist the dataset refers to.

For these tasks web services are needed, as well as web UIs to allow users of systems such as PANGAEA® (<https://pangaea.de/>) to manually match scientific names to taxon concepts. Expected complications are that multiple taxa exist for a name or that the taxon originally implicitly used in the creation of the dataset cannot be found. The services need to support further filtering criteria to facilitate exact matching or to decide that none of the provided concepts is suitable. These criteria may include for example the citation of the context in which the name is used, the so-called sec. reference [GB03], the geographic context of a concept and the time range of its validity.

For cases when the taxon concept cannot be ascertained we propose to annotate data with identifiers that primarily refer to the scientific name, but which also include all taxon concepts this name was involved in. Such identifier can be similar to the "simple

nominal entity” as suggested by Chawuthai et al. [CT16] or something that we preliminary named “name-related concepts identifier”. The latter, in short, refers to a chronologically and semantically linked list of all taxon concepts this name was involved in in a given dynamic taxonomic checklist like Euro+Med. This implicitly includes historic concepts which never have been preserved in data and which are thus lost. This option represents progress as compared to a pure name-based annotation of problematic data sets, which fulfills many of the requirements for the annotation with taxon concepts, such as stable identifiers that resolve to stable information. Even if the referenced information is a dynamic list of taxon concepts related to name in a specific checklist or globally (depends on the semantics of the identifier being used), it represents the best match possible based on the data currently available. At the same time, it reflects the uncertainty of the data and is open for subsequently refined annotation once more data becomes available, like for example when further historic or current concepts have been added to a checklist.

5 Outlook

For data networks like NFDI4BioDiversity it is important to be able to semantically link environmental research data to taxon concepts provided by taxonomic research projects like the E+M checklist, which may use updated and thus differing taxonomic concepts. The here presented conception for the annotation of environmental data has the potential to encourage the usage of taxon concept identifiers directly in the environmental data at the time of their generation. This would relieve the need for error-prone retrospective annotation and thus makes inference on the data more reliable.

Literature

- [CT16] Chawuthai, R. et al.: Presenting and Preserving the Change in Taxonomic Knowledge for Linked Data, *Semantic Web*. 7, pp. 589-616, 2016.
- [DGG14] Diepenbroek, M.; Glöckner, F.; Grobe, P.; Güntsch, A.; Huber, R.; König-Ries, B.; Kostadinov, I.; Nieschulze, J.; Seeger, B.; Tolksdorf, R.; Triebel, D.: Towards an Integrated Biodiversity and Ecological Research Data Management and Archiving Platform: The German Federation for the Curation of Biological Data (GFBio). In Plödereeder, E.; Grunskel, L.; Schneider, E.; Ull, D. (eds): *Informatik 2014 – Big Data Komplexität meistern. GI-Edition: Lecture Notes in Informatics (LNI) – Proceedings 232*. Köllen Verlag, Bonn, 1711-1724, 2014.
- [EM21] Euro+Med: Euro+Med PlantBase - the information resource for Euro-Mediterranean plant diversity. – Published at <http://www.europlusmed.org>, 2006+ [continuously updated] accessed 26 Apr 2021.
- [FP06] Franz, N.; Peet, R.; Weakley, A.: On the use of taxonomic concepts in support of biodiversity research and taxonomy, *The New taxonomy, systematics association special volume series 74*, vol. 76, 2006. <https://doi.org/10.1201/9781420008562>

- [GB03] Geoffroy, M.; Berendsohn, W. G.: The concept problem in taxonomy: importance, components, approaches. In Berendsohn, W.G. (ed.): *MoReTax - Handling factual Information Linked to Taxonomic Concepts in Biology*. BfN Schriftenreihe für Vegetationskunde 39. Federal Agency for Nature Conservation, Bonn, 5-14, 2003. https://www.researchgate.net/publication/285456910_The_concept_problem_in_taxonomy_Importance_components_approaches
- [GB10] Güntsch, A.; Berendsohn, W. G.; Geoffroy, M.: Versioning and the use of GUIDs for PESI. A Pan-European Species-directories Infrastructure (PESI). Report D5.5. 2010. <https://doi.org/10.13140/RG.2.1.1036.9440>
- [GL10] Güntsch, A.; Luther, K.; Müller, A.; Kohlbecker, A.; Plitzner, P.; Kilian, N.; Henning, T.; Mering, S. von; Raab-Straube, E. von; Kusber, W.-H.; Berendsohn, W. G.: The Platform for Cybertaxonomy: Standards, services and tools, *Biodiversity Information Science and Standards*, vol. 3, p. e37177, 2019. <https://doi.org/10.3897/biss.3.37177>
- [GGH20] Groom, Q.; Güntsch, A.; Huybrechts, P.; Kearney, N.; Leachman, S.; Nicolson, N.; Page, R. D. M.; Shorthouse, D. P.; Thessen, A. E.; Haston, E.: People are essential to linking biodiversity data. *Database (Oxford)*, 2020, baaa072, 2020. <https://doi.org/10.1093/database/baaa072>
- [GR05] Greuter, W.; Raab-Straube, E. von: Euro+Med Notulae 1. *Willdenowia* 35, 223-239, 2005.
- [HAV14] Hyland, B.; Atemezing, G.; Villazón-Terrazas, B.: Best practices for publishing Linked Data, <https://www.w3.org/TR/ld-bp/>, accessed 09 Jan 2014.
- [IN13] European Commission: INSPIRE Data Specification for the spatial data theme Species Distribution. D2.8.III.19. Technical Guidelines, 1-125, 2013. https://inspire.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_SD_v3.0.pdf
- [MG17] Michel, F. et al.: A Model to Represent Nomenclatural and Taxonomic Information as Linked Data. [S4BioDiv@ISWC, pp. 1-12, 2017](https://doi.org/10.1007/978-3-319-54364-5_1)
- [GI20] Glöckner, F. O.; Diepenbroek, M.; Felden, J.; Güntsch, A.; Stoye, J.; Overmann, J.; Wimmers, K.; Kostadinov, I.; Yahyapour, R.; Müller, W.; Scholz, U.; Triebel, D.; Frenzel, M.; Gemeinholzer, B.; Goesmann, A.; König-Ries, B.; Bonn, A.; Seeger, B. Jul. 2020: NFDI4BioDiversity - A Consortium for the National Research Data Infrastructure (NFDI). <https://doi.org/10.5281/zenodo.3943645>
- [Ta06] Taxonomic Names and Concepts Interest Group: Taxonomic Concept Transfer Schema (TCS), version 1.01. *Biodiversity Information Standards (TDWG)*, 2006 <http://www.tdwg.org/standards/117>

Minimizing the Annotation Effort for Detecting Wildlife in Camera Trap Images with Active Learning

Daphne Auer,¹ Paul Bodesheim,² Christian Fiderer^{3, 4}, Marco Heurich^{5, 6, 7}, and Joachim Denzler^{8, 9, 10}

Abstract: Analyzing camera trap images is a challenging task due to complex scene structures at different locations, heavy occlusions, and varying sizes of animals. One particular problem is the large fraction of images only showing background scenes, which are recorded when a motion detector gets triggered by signals other than animal movements. To identify these background images automatically, an active learning approach is used to train binary classifiers with small amounts of labeled data, keeping the annotation effort of humans minimal. By training classifiers for single sites or small sets of camera traps, we follow a region-based approach and particularly focus on distinct models for daytime and nighttime images. Our approach is evaluated on camera trap images from the Bavarian Forest National Park. Comparable or even superior performances to publicly available detectors trained with millions of labeled images are achieved while requiring significantly smaller amounts of annotated training images.

Keywords: Active Learning; Wildlife Monitoring; Camera Trap Images

1 Introduction

The world experiences an accelerated loss of biodiversity that is unprecedented in the history of life [Ce15]. The major drivers of this development are habitat loss, direct persecution, climatic change, and invasive species. Therefore, the monitoring of animal populations is crucial as an early warning system and for evaluating the success of conservation measures [Se19]. Thus, national parks have installed networks of camera traps to observe

¹ Computer Vision Group, Friedrich Schiller University Jena, 07737 Jena, Germany; daphne.auer@uni-jena.de

² Computer Vision Group, Friedrich Schiller University Jena, 07737 Jena, Germany; paul.bodesheim@uni-jena.de

³ Bavarian Forest National Park Germany, Visitor Management and National Park Monitoring, Freyunger Str. 2, 94481 Grafenau, Germany; christian.fiderer@npv-bw.bayern.de

⁴ Wildlife Ecology and Wildlife Management, Faculty of Environment and Natural Resources, University of Freiburg, 79106 Freiburg, Germany

⁵ Bavarian Forest National Park Germany, Visitor Management and National Park Monitoring, Freyunger Str. 2, 94481 Grafenau, Germany; marco.heurich@npv-bw.bayern.de

⁶ Inland Norway University of Applied Science, Institute for forest and wildlife management, Campus Evenstad, NO-2480 Koppang, Norway

⁷ Wildlife Ecology and Wildlife Management, Faculty of Environment and Natural Resources, University of Freiburg, 79106 Freiburg, Germany

⁸ Computer Vision Group, Friedrich Schiller University Jena, 07737 Jena, Germany; joachim.denzler@uni-jena.de

⁹ German Aerospace Center (DLR), Institute for Data Science, Mälzerstraße 3, Jena, Germany

¹⁰ Michael Stifel Center Jena for Data-Driven and Simulation Science, Ernst-Abbe-Platz 2, Jena, Germany

animals like deer, roe deer, and wild boars but also smaller ones like hares and squirrels to document their appearance. This leads to thousands of images already taken in a short period, which are currently often analyzed and annotated by experts manually. Thus, an automatic species classification system would greatly enhance subsequent studies on varying animal appearances.

Furthermore, many images from camera traps do not even contain an animal but only show the background scene instead. This is often due to rapidly changing lighting or weather conditions, as well as moving plants like grasses or branches of trees caused by wind. Hence, it is not only a tedious task for humans to perform the species identification but also cumbersome to identify and filter out those background images. Our work, therefore, focuses on the automatic separation of images with and without animals to spare experts to go through uninformative data. In addition, this binary classification task serves as preprocessing step for an automatic species classification that can be applied afterwards, since the classifier does not have to recognize images showing background only anymore and can focus on the details to distinguish different animals.

To keep the manual labeling effort minimal, the task of automatically identifying background images from camera traps is tackled with an active learning approach. Starting the learning process with an initial model only trained with a small labeled training dataset, an active learning strategy automatically selects additional images that shall improve the recognition performance when annotated and incorporated in the learning process of the classifier. The aim is that only the most informative examples are chosen such that few additional annotations lead to the most significant performance gains compared to the initial model. In this work, we use the EMOC approach [FRD14] as a selection criterion in the active learning process because it has already proven to be effective in related monitoring tasks [Käl16a].

For evaluation, our approach is applied to image data from the Bavarian Forest National Park in Germany¹¹. Our experiments reveal that region-specific classification models learned with only a few annotated images have the potential to outperform existing models for detecting wildlife in camera trap images, which are trained with millions of annotated images, such as the MegaDetector [BMY19] or the approach of Tabak et al. [Ta19]. Furthermore, we argue that learning separate classifiers for daytime and nighttime images is beneficial because individual models can then focus on the actual classification task compared to a single model handling both scenarios internally.

2 Related Work

Many approaches are used to automatically analyze images from camera traps, mostly adopting deep learning models for tasks like species identification and counting [VSV17, No18, Sc20, SMF20, Ta19]. Some work specifically focuses on the effect of varying sizes

¹¹ <https://www.nationalpark-bayerischer-wald.bayern.de/english/index.htm>

of the training dataset on the final species identification accuracy and showed that there is a logarithmic relationship between the number of training samples and the performance of the classifier [SMF20, Ta19]. Furthermore, the problem of class imbalance in the training data is specifically investigated with very deep network architectures [VSV17], and solutions based on transfer learning have been proposed [Wi19]. Interestingly, the authors of [Wi19] also applied their model in a live online citizen science project.

An important aspect of species identification systems is their generalization to new locations. Several studies found that most common approaches struggle when applied in previously unseen environments leading to poor performance due to the domain shift [BVHP18, Sc20]. This supports the idea of learning new, region-specific models for detection, thereby only requesting a few annotations with an active learning strategy.

As mentioned previously, we focus on the distinction between images with and without animals. This task has been tackled by [No18] with a binary classifier as well, using thousands of annotated example images. Only recently, filtering background images from camera traps in embedded systems has been proposed [Cu21]. In their study, the authors found that the performance largely depends on the number of available images for training and that neural network models with higher input resolution perform better. In contrast, our approach keeps the annotation effort minimal and uses as few labeled data as possible.

Active learning for animal detection has been previously applied [Ke19], but for drone images rather than camera trap images. However, they also propose a web-based annotation tool called AIDE that requests feedback from experts following the human-in-the-loop concept to improve a species classifier over time [KTM20]. While they use active learning for species identification in drone images, we use active learning to improve the detection of animals in camera trap images with various backgrounds and conditions.

A popular approach for analyzing camera trap images is the MegaDetector [BMY19] that has been trained on millions of example images from camera traps in North America and East Africa using the standard Faster R-CNN [Re15] as a deep object detection model. This generic detector is trained for three scenarios, i.e., it can localize animals, humans, and vehicles. Background images can be identified when even the highest detection scores for an image are below a certain threshold. Typically, a species classifier is applied for the detected animals afterwards. The MegaDetector [BMY19] is used as a baseline in our experiments.

Another detection approach that also allows for identifying background images has been proposed by Tabak et al. [Ta19]. They also provide a software package called Machine Learning for Wildlife Image Classification (MLWIC) which is used for comparison to our methods.

3 Methods

The aim of this work is to train models that separate background images from images showing an animal with only a few annotated images, which can be realized using the concept of active learning. Within the detection task, it needs to be addressed that background images and images showing an animal are very similar for a single camera trap, while the numerous backgrounds of all camera traps vary a lot, but they all belong to one common background class. That is why the task of separating images with and without animals is tackled by training region-specific models focussing on particular environments containing only a few backgrounds and perspectives.

In this work, models are trained from scratch using similar data of the region where the detector will be applied later on. Therefore, Gaussian processes are used instead of deep classifiers, since Gaussian processes already achieve good performance on very small labeled datasets, while deep classifiers require large annotated datasets. Furthermore, closed-form solutions exist for model updates of Gaussian process classifiers that are beneficial when updating the model during active learning. To also exploit recent advances in deep learning for creating meaningful image representations, the fixed CNN Inception-v3 [Sz16] with pre-trained weights from ImageNet [De09] is used as a black box feature extractor.

The Gaussian process framework and the basic idea of active learning are summarized in the following sections. Also, the sample selection strategy used for active learning is explained, namely the expected model output change (EMOC) of the Gaussian process classifier.

3.1 Gaussian Processes

The Gaussian process framework offers a probabilistic formulation for both regression and classification tasks [RW06]. The advantage is that one can directly infer uncertainties associated with the estimations of the model. These uncertainties have already been proven to be useful for active learning [Fr12, Ro17] but we will use the EMOC principle together with Gaussian processes as explained in Sect. 3.3.

The advantage of closed-form solutions for Gaussian process inference can be exploited when using label regression for classification, a common choice for visual recognition tasks [Ro17]. In this case, binary classification is achieved by taking the sign of the regression estimate when using labels $y \in \{-1, 1\}$ for the training data. Typically, labels of N training samples are collected in a vector \mathbf{y} . For the features \mathbf{x} of the training images, one computes pairwise similarities using a kernel function κ to obtain a kernel matrix \mathbf{K} of size $N \times N$. A common choice for two feature vectors \mathbf{x} and \mathbf{x}' is the Gaussian kernel with hyperparameter σ :

$$\kappa(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{2\sigma^2}\right) \quad (1)$$

such that small distances in the feature space lead to high similarity values. For predicting the class of a test sample \mathbf{x}^* , it is necessary to obtain the similarities to the N training samples with the kernel function in a vector \mathbf{k}_* of length N . This vector is then used to compute the regression estimate y^* for \mathbf{x}^* via:

$$y^* = \mathbf{k}_*^T (\mathbf{K} - \sigma_n^2 \mathbf{I})^{-1} \mathbf{y} \quad (2)$$

with a regularization term $\sigma_n^2 \mathbf{I}$ that adds a small constant σ_n^2 to the main diagonal of the kernel matrix \mathbf{K} . Due to this closed-form solution, efficient updates of the model are possible with only a few algebraic operations when additional data need to be incorporated.

3.2 Active Learning

Active learning (AL) aims at improving an initial classification model trained with few data by iteratively updating the model with additional images, which are chosen using a certain strategy. In every iteration, one image is selected out of a pool of unlabeled images first. Then, it gets labeled by an expert to afterwards update the model with the additional image. This process continues until either every image is labeled, or the budget for annotation is reached. In many applications such as the analysis of camera trap images, a large amount of unlabeled data exists to choose from. The goal is to only query and annotate the meaningful images that lead to the most significant improvements of the classifier, which is guided by the above-mentioned AL strategy.

However, for experimental evaluations, a completely annotated dataset is typically used to simulate the AL pipeline in a fully automatic manner. Hence, for each image that is treated as unlabeled, the existing label will not be considered unless the image is selected by the AL strategy. Then the image label is revealed by an oracle that plays the role of the human annotator such that the labeled query image can be used to update the model. For the incorporation of additional labeled samples, we adapt a framework proposed by [Kä16b].

3.3 EMOC for Active Learning

Several strategies exist to choose an unlabeled image for additional annotation that is then used to update the model. However, this work focuses on the expected model output change (EMOC) which was already successfully applied for wildlife detection in camera trap images [Kä16a]. EMOC calculates a score for every image that tells us how much the model predictions would change if this additional image is used for extending the model.

Given an image represented by its feature vector \mathbf{x}' and a model f , the change of the model output $\Delta f(\mathbf{x}')$ after updating it with \mathbf{x}' can be calculated with the following formula:

$$\Delta f(\mathbf{x}') = \sum_{y' \in \mathcal{Y}} \left(\frac{1}{|\mathcal{Q} \cup \mathcal{U}|} \sum_{\mathbf{x} \in \mathcal{Q} \cup \mathcal{U}} \mathcal{L}(f(\mathbf{x}), f'(\mathbf{x})) \right) p(y'|f(\mathbf{x}')) \quad (3)$$

Here, f' refers to the updated model and the data consist of labeled images \mathfrak{L} as well as unlabeled images \mathfrak{U} . Furthermore, \mathcal{L} is a loss function measuring the difference in model outputs $f(\mathbf{x})$ and $f'(\mathbf{x})$ before and after the model update for each image $\mathbf{x} \in \mathfrak{L} \cup \mathfrak{U}$. Since we do not know the label of the selected image \mathbf{x}' , we calculate the average loss (term inside large parentheses) for each possible annotation $y' \in \mathcal{Y}$ of \mathbf{x}' and average the results with weights determined by the likelihood for the label y' of the image \mathbf{x}' .

In each iteration of the model updating process, EMOC needs to be calculated following Eq. (3) as an AL score for every unlabeled image. The unlabeled sample with the largest score is then selected for annotation and the model update. In our work, Gaussian process models are used as explained in Sect. 3.1.

4 Dataset

The methods are applied to camera trap images from the Bavarian Forest National Park, that were gathered in the years 2018 and 2019 within the project “*Neue Wege zu einem grenzüberschreitenden Rotwildmanagement in Zeiten des Klimawandels*”¹². The collected data contains roughly 286,000 images from 100 camera traps, showing forest scenes from different perspectives day and night. Within the data, 31 native animal species could be identified by experts. The camera traps were mounted to tree trunks 60 to 70 cm above the ground and recorded a series of five images each time when motion was detected by its pyroelectric infrared sensor. However, these are still individual images taken with some seconds in between, so the camera traps did not provide continuous video clips which would have allowed for better integration of temporal context. Hence, we focus on processing the single images independently. Besides animals, there are sometimes also humans in the images, and sometimes images of one camera trap may show different forest scenes because of location changes during the operation time period.

Visibility of Animals Occlusions by vegetation, poor lighting conditions, and the distance between a camera trap and the animal mainly affect the visibility of animals in camera trap images. Sometimes, only an animal’s head or leg can be seen especially if it enters the camera’s field of view at the moment the image is taken. Various perspectives and backgrounds complicate the task of detecting animals in those images additionally. The animals that shall be detected vary strongly in size, ranging from larger animals like deer to smaller animals such as hares. Example images are shown in Fig. 1 which also demonstrate, that the detection can be challenging for humans as well.

Label Noise The dataset was labeled by experts with the intention to count the number of animals and define their species. As mentioned above, the camera traps took a sequence

¹² <https://www.nationalpark-bayerischer-wald.bayern.de/forschung/projekte/rotwildprojekt.htm>



Fig. 1: Appearing animals in camera trap images from the Bavarian Forest National Park. Already for humans, the separation of images with and without animals can be a difficult task in several situations.

of five images in a short time interval when detecting motion. Thus, the same animal(s) can be seen during a sequence and therefore the sequence needed only one annotation. Unfortunately, this label cannot be mapped to all five frames, since an animal often does not appear in the last frames anymore, leading to wrong annotations. Instead, the label is assigned to the first frame only, which reduces the dataset to about 20% of all images. Thereby, it is not guaranteed that this assignment is always correct, which explains the label noise within the data. This is why the following results need to be interpreted with caution.

Note that the following labeling is used: an image without animals is treated as a background image, including those showing humans only. This may challenge the classifier later on since all living animals belong to one class but humans need to be explicitly recognized to combine them in a class with background scene images.

Class Imbalance The ratio between images with and without animals is quite unbalanced since 68.6% of the images in the dataset do not contain an animal. The imbalance between the two classes also appears in most of the single camera trap subsets. Therefore, balanced test sets are used for evaluation, i.e., they contain the same number of testing images for each class. Of course, the class imbalance is also reflected in the set of images that the AL approach uses for drawing queries, but a reasonable selection strategy might be able to also choose selective images from the underrepresented class.

5 Experiments

For the task of separating images with and without animals, the performance of general detectors is compared with the performance of our models, which use active learning (AL) to minimize human annotation efforts. The main idea is to train one detector for a single camera trap or for a small subset of camera traps following a region-specific approach, which focuses on the corresponding scene conditions. Thereby, the number of different backgrounds is heavily reduced for a single subset, such that a model does not need to handle many backgrounds and can therefore focus on the small differences between images with and without animals. Nevertheless, the appearance of a camera trap's background still varies regarding weather and lighting conditions, as well as for different seasons.

In Sect. 5.3, experiments show that active learning with EMOC leads to competitive performance on the given task and dataset. Afterwards, those models are compared to the MegaDetector [BMY19] and the MLWIC package [Ta19] in Sect. 5.3. In Sect. 5.4 the focus lies on camera trap stations, for which these existing tools perform poorly, and it is shown that the AL approach can achieve better separations for these special cases due to well-calibrated models with a very small amount of image annotations. The dataset split for evaluating the different methods is outlined in Sect. 5.1, and the experimental setup is described in Sect. 5.1.

5.1 Dataset Preparation for Performance Comparisons

All experiments are conducted on the dataset from the Bavarian Forest National Park. As mentioned before regarding the region-specific approach, the dataset is split into small subsets, as described in Sect. 5.1.1. It also needs to be considered that each camera trap recorded a different number of images, and different ratios of class imbalance exist among the different stations, which is described in Sect. 4. In order to allow for an unbiased comparison of different approaches with respect to the class imbalance, balanced test sets of a certain size are defined in Sect. 5.1.2. Finally, camera traps that took only very few images have to be treated differently (see Sect. 5.1.3).

5.1.1 Separating Day and Night

First, the dataset is split into 100 subsets, each belonging to a single camera trap, since this work focuses on the region-specific approach. In addition, each of the subsets is separated into colored images taken during the day and grayscale images mostly taken at night, where flash is used to illuminate the scenes. The automatic separation of colored and grayscale images is easy since the corresponding color channels have characteristic values. Thus, one daytime subset and one nighttime subset are obtained for each camera trap, containing images that are illuminated either naturally or artificially by the flash. To now observe the effect of this separation on the model performance, the results for two approaches will be compared: one approach treats the daytime and nighttime subsets of a camera trap independently and two separate models are trained (dt-model and nt-model), while the other approach ignores the split and one model is trained for both daytime and nighttime images (d&nt-model). Given this distinction, a comparable setup with respect to the test sets needs to be created for the three scenarios (dt, nt, d&nt).

5.1.2 Uniform Data Splits for Balanced Test Sets

To define a uniform and consistent rule for choosing test and training data for each of the three scenarios (dt, nt, and d&nt), the variety within the numerous stations needs to be considered.

Therefore, a balanced test set for each subset is created, i.e., every test set contains the same number of images with and without animals to avoid a class imbalance in the test set. This prevents misleading accuracy values that are based on a good accuracy for the prevalent class in the data set but rather neglect the performance for the less representative one.

dt- and nt-Models After splitting the data for each camera trap into subsets for day and night, the size of the subsets ranges from only 7 images up to 3,428. Especially the class imbalance, i.e., the ratio of the two classes, differs between 97:4 to 44:42 and 6:2,291. To find general solutions for a balanced test set in each experiment, the focus is on the underrepresented class with b images. Per class, $n = \text{floor}(b/3)$ images are randomly chosen for testing, but at least five if available. For the initial training, only three images are randomly chosen per class. All remaining data serve as a pool for the AL query process.

d&nt-Models To ensure the comparability between a d&nt-model and its associated dt- and nt-model, the test set of the two models are merged for each station. Since their test sets are class-balanced, the test set for each d&nt-model also contains as many images with as without animals. The same is done with images that were chosen for the initial training and therefore six images per class images are used for the initial model for d&nt-models.

In general, it is also important to consider whether enough images are available to allow for a proper learning setup. Especially stations with only a few images either during day or night are problematic. Therefore, image subsets are merged, if they do not provide a reasonable amount of images, as described in the following section.

5.1.3 Merging Small Subsets

In the following experiments, the goal is to train a detector with only 156 train images for each subset. When splitting the subsets into their training set and test set, it turns out that 72 out of 100 dt-subsets, as well as 72 out of 100 nt-subsets, and 43 of 100 d&nt-subsets do not contain 156 images for training. Within each scenario those subsets having not enough images are merged, to generate a setup that can be applied to all experiments in Sect. 5.2. The new subgroups are created by randomly combining too small subsets until their training set includes at least 156 images.

As a result, 100 dt-subsets are merged to 48 dt-subsets, 100 nt-subsets are merged to 49 nt-subsets and 100 d&nt-subsets are merged to 73 d&nt-models. In each merging process, also the test sets are merged to a new one, instead of generating a new test set. Hence, the test sets of all d&nt-models still contain exactly the images of the test sets from the dt-models and nt-models, since the merging did not affect the test image selection. Note that models based on merged subsets need to handle more perspectives and are therefore more general, but training them is also more challenging.

5.2 Experimental Setup

After generating consistent subsets, the experimental setup can be applied to each subset within each scenario as described in the following. In every experiment, the model performance is analyzed regarding the initial accuracy, the random baseline, active learning with EMOC, and training with all available data. The initial accuracy results by training the model for the first time with only three images per class for dt- and nt-subsets and six images per class for d&nt-subsets, respectively. Afterwards, the AL process is simulated by querying 150 single images from the remaining subset iteratively and updating the model with the queried images. Choosing these images only randomly leads to a random baseline. This is compared to our AL strategy with the EMOC sample selection criterion. All experiments are repeated ten times, changing the images for the initial training randomly in each run. For comparison, another model does not only use 150 additional images for training but all data of the corresponding subset except the test set. Its performance reflects the possible enhancement using more data, and also indicates the influence of label noise and class imbalance on the training process.

Both the random selection and the selection made by EMOC generate an updated model in every querying step, which can be evaluated regarding its performance and saved afterwards. Thus, it may be possible to achieve higher accuracies with a model that was trained with fewer queried images. The main reason is the limited number of images per class in the pool of unlabeled images from which EMOC selects the queries. If there are only images of one class left after a certain amount of queries, then EMOC is also only able to select further images from the remaining class. This leads to an increasing class imbalance for the training set of the classification model. In the evaluation, the highest accuracy achieved during the AL process is therefore also reported in order to demonstrate the capacity of the approach. Automatically determining an optimal number of queries is an unsolved problem so far and needs further investigation.

For comparison, the MegaDetector [BMY19] and the MLWIC package [Ta19] are evaluated on the same test sets as the models of the AL approach. Thus, all results are affected by the label noise discussed in Sect. 4.

5.3 Overall Performance

The evaluation of the numerous models can be handled by inspecting single models regarding their overall classification accuracy, i.e., the number of correctly classified images divided by the number of all images in the test set. However, since every subset has individual parameters like the number of possible training images for the learning process or the ratio between the two classes, it is statistically more interesting to observe the performance on the whole dataset in the first place. This will be done for all models within the scenarios day, night, and day&night. Since the test sets of d&nt-models consist of the test sets of dt-models and the test sets of nt-models, the comparability of the accuracies is given.

	InitAcc	Our active learning approach				TrainAll	Baselines methods	
		Highest accuracy Random	EMOC	Final accuracy Random	EMOC		MegaDetector [BMY19]	MLWIC [Ta19]
Only daytime images, dt-models	67.32% ±1.44%	84.62% ±0.42%	85.53% ±0.47%	79.62% ±0.34%	81.24% ±0.43%	81.76%	82.61%	67.84%
Only nighttime images, nt-models	65.25% ±1.37%	81.00% ±0.42%	81.46% ±0.43%	75.94% ±0.47%	77.01% ±0.59%	76.63%	78.15%	61.16%
All test images, dt-models + nt-models	66.25% ±0.66%	82.75% ±0.23%	83.43% ±0.33%	77.72% ±0.18%	79.06% ±0.33%	79.11%	80.31%	64.39%
All test images, d&nt-models	66.95% ±0.83%	80.76% ±0.51%	81.50% ±0.36%	76.03% ±0.66%	77.46% ±0.38%	78.85%	80.31%	64.39%
Enhancement due to daytime-nighttime separation	-0.70%	+1.99%	+1.93%	+1.69%	+1.60%	+0.26%	-	-

Tab. 1: Per-image accuracies averaged over the test images of all region subsets showing the differences between the combination of dt- and nt-models, the d&nt-models, and the baseline methods. Accuracies from the AL experiments are presented as follows: the initial accuracies obtained from a model trained on a small initial training set (InitAcc), the highest accuracy achieved during the AL process (Highest accuracy), and the final accuracy after 150 queries (Final accuracy). For comparison, the accuracies of the models learned with all training images are also reported (TrainAll). As baseline methods, the MegaDetector [BMY19] and the MLWIC package [Ta19] have been evaluated on the same test data.

Evaluation Methods Two evaluation methods are used to analyze the overall performance. One method is to treat all test sets in a scenario as a single, huge test set and evaluate the predicted labels for the corresponding images, which will be referred to as per-image accuracy. This accuracy reflects how many test images are classified correctly, which equals the probability that the prediction for a single image is correct. Another method is to evaluate how well the models perform on average for a specific region, i.e., for a single camera trap station or a subset of stations obtained from the merging of small stations. The resulting values are referred to as per-region accuracies in the following and can be interpreted as the average percentage of regions for which the predictions are correct.

For obtaining the per-region accuracies, every region subset gets the same weight when computing average values. In contrast, the per-image accuracy reflects a weighted mean using the sizes of the individual test sets as weights. As defined in Sect. 5.1.2, the size of a test set indicates the availability of the less represented class. Training a model for a binary decision with a few exemplary images for this class often leads to biased models with bad performance. By considering the per-image accuracy, the performance of models that have been trained with strongly unbalanced data is less significant. Hence, mainly the per-image accuracies are discussed since the influence of class imbalance is mitigated by considering all test images as one huge test set for the evaluation.

	InitAcc	Our active learning approach				TrainAll	Baselines methods	
		Highest accuracy		Final accuracy			MegaDetector	MLWIC
		Random	EMOC	Random	EMOC		[BMY19]	[Ta19]
Only daytime images, dt-models	68.08% ±1.57%	83.16% ±0.50%	84.23% ±0.38%	76.90% ±0.63%	78.17% ±0.46%	78.51%	84.74%	67.53%
Only nighttime images, nt-models	64.98% ±1.10%	80.09% ±0.60%	80.54% ±0.44%	73.93% ±0.48%	75.03% ±0.41%	74.91%	79.67%	59.61%
All test images, dt-models + nt-models	66.51% ±0.78%	81.61% ±0.25%	82.37% ±0.28%	75.40% ±0.30%	76.58% ±0.18%	76.69%	82.20%	63.53%
All test images, d&nt-models	66.94% ±0.96%	79.65% ±0.43%	80.36% ±0.32%	74.05% ±0.54%	75.40% ±0.44%	76.70%	81.71%	63.58%
Enhancement due to daytime-nighttime separation	-0.43%	+1.96%	+2.01%	+1.35%	+1.18%	-0.01%	-	-

Tab. 2: Per-region accuracies averaged over the different region subsets. Organization is the same as in Table 1, see corresponding caption explanations of the individual columns.

Separation of Daytime and Nighttime Images First, the impact of learning separate models for daytime and nighttime images is investigated. The per-image accuracies reported in Table 1 show that dt-models (first row) perform better than nt-models (second row). Although the baseline methods do not explicitly distinguish between daytime and nighttime images, they also perform better on daytime images. Thus, the distinction between images with and without animals seems to be more challenging for nighttime images. When evaluating the predictions of the individual dt-models and nt-models across all images of the test sets (third row), a clear improvement of the accuracies can be seen compared to training single d&nt-models that are able to handle both daytime and nighttime images (fourth row). The differences are indicated in the fifth row of Table 1.

It is important to note that the test sets used for the evaluations in the third and fourth row are identical, which allows for a fair comparison. This is also the reason why the per-image accuracies of the baseline methods are the same in the third and fourth row since the same baseline models are evaluated on the same test set in both lines, while for the AL experiments different models are evaluated. However, the per-region accuracies also differ for the baseline methods as shown in Table 2, because the average is taken across different region subsets of varying size but with the same weight. For the per-region accuracies, similar behavior of the individual dt- and nt-models can be observed as for the per-image accuracies, although the differences in the fifth row are less pronounced.

Active Learning Results for the AL experiments with the proposed EMOC strategy can be observed in the left parts of Tables 1 and 2. The corresponding columns are listed in order of the number of training images that have been used for learning the models. In the first column (InitAcc), the initial accuracies are reported, which denote the model performances

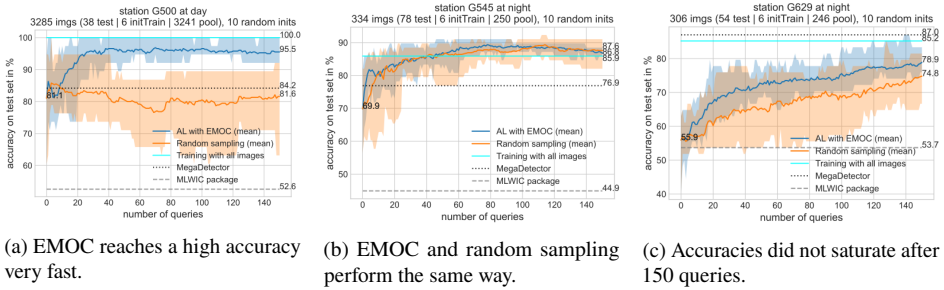


Fig. 2: Visualizations of the active learning process for three example models. Training with a random sample selection is shown in orange, the learning process with EMOC as active learning strategy is shown in blue. Each line represents the mean accuracy over ten runs.

when using three training images per class for dt- and nt-models, or six training images per class for d&nt-images. Accuracies from the AL approach are shown in the following columns. While the final accuracies belong to models trained with 150 additional images, the highest accuracies are often achieved with fewer queried samples due to the reasons mentioned in Sect. 5.2. As can be seen in both tables, the final accuracies of EMOC are often very similar to the accuracies achieved by models trained with all images except those of the test set (TrainAll). Hence, it is not necessary to annotate all camera trap images for training the models in order to achieve good performance, up to 156 labeled images per region are often sufficient. Incorporating more annotated images might lead to learning biased models due to class imbalance in the training data.

EMOC vs. Random Sampling The proposed EMOC strategy is compared with a random sample selection regarding the final accuracy and the highest accuracy reached during the training process. One observation in Tables 1 and 2 is that active learning with EMOC leads to training better models in all cases compared to random sampling of images. This is consistent with the findings of previous work [Käl16a, FRD14] and shows that EMOC automatically selects meaningful images. The two different selection strategies can also be compared based on the visualizations for the AL process of three example models in Fig. 2. The optimal learning process is shown in Fig. 2a, where the pool of images is more than 20 times larger than the number of queried images. Applying the EMOC criterion results in a steep ascent in performance since it chooses meaningful images out of that large pool, in contrast to a random selection. Sometimes, there is no decisive difference between the two sampling methods like in Fig. 2b. At least, it shows that the performance for all initial runs using EMOC results in similar final accuracies, while sampling randomly often leads to more different final accuracies in each run. Some models for merged subsets, which include images from many perspectives, need more than 150 queries to reach the saturation of the training process, as demonstrated in Fig. 2c. Nevertheless, models trained with images selected by EMOC perform better than those using randomly selected images.

Comparison with MLWIC package One baseline method is provided by Tabak et al. in the MLWIC package [Ta19]. From the accuracies in Tables 1 and 2, a superior performance of the AL models can be clearly observed. Interestingly, even the initial models that are only trained with either three images per class (dt- and nt-models) or six images per class (d&nt-models) often already achieve higher accuracies, and selecting additional samples via AL even enlarges this gap. This comparison shows that applying out-of-the-box methods to an own dataset does not work in general and often leads to poor classification performance, even if these existing methods are trained with thousands of example images. The reasons are the different characteristics of the datasets that are used for training and in the application, with varying challenges and scene conditions. In contrast, better models can already be obtained using a handful of annotated images from the corresponding application domain or region, thus requiring only minimal annotation efforts.

Comparison with MegaDetector The MegaDetector [BMY19] performs better than the method of Tabak et al. [Ta19] in all scenarios regarding both the per-image accuracies (Table 1) and the per-region accuracies (Table 2). For the latter, images of more regions are classified correctly by the MegaDetector than by the models trained with our AL approach, on average. This is not surprising, since the MegaDetector has been trained with millions of annotated images. In contrast, the AL models only use about 150 to 160 annotated training images. Given this large difference in the amount of training data, the difference in per-region performance is rather low. Even more encouraging, the per-image accuracies obtained from models trained with EMOC are comparable to those of the MegaDetector when individual daytime and nighttime models are considered for the AL approach (first three rows in Table 1). Concerning the highest accuracies achieved by one of our models during the AL process, they almost identical accuracies as the MegaDetector in the aforementioned scenarios. This shows the potential for minimizing the annotation efforts for humans and keep high classification performances using an AL strategy with a clever selection of images.

5.4 Investigating Selected Region Subsets

The comparison of the AL approach with the MegaDetector [BMY19] in the previous section has shown that models trained with AL are able to achieve comparable average accuracies across all region subsets while requiring significantly less annotation effort. However, we do not state that our approach is better suited for all possible regions and camera trap scenes since it would be difficult to compete with an existing model trained on millions of images in general. In contrast, we argue that existing methods might struggle in particular regions when encountering certain scene conditions or properties of individual camera traps. Thus, models trained with AL and a small amount of additionally labeled images are one possible solution to cope with difficult or unique situations.

Therefore, ten difficult region subsets and ten easy region subsets are selected for each scenario (only daytime images, only nighttime images, both types of images) to compare

	EMOC Highest accuracy	EMOC Final accuracy	MegaDetector [BMY19]
<i>Ten difficult region subsets w.r.t.</i>			
Only daytime images, dt-models (564 test images 16 camera traps)	81.67% 78.71%	80.53% 76.86%	68.62% 71.25%
Only nighttime images, nt-models (481 test images 13 camera traps)	77.28% 72.89%	75.41% 70.75%	59.88% 54.90%
Both daytime and nighttime images, d&nt-models (673 test images 10 camera traps)	70.76% 70.32%	69.05% 67.68%	63.60% 62.66%
<i>Ten easy region subsets w.r.t.</i>			
Only daytime images, dt-models (386 test images 17 camera traps)	83.19% 79.98%	81.01% 76.83%	93.01% 94.10%
Only nighttime images, nt-models (439 test images 17 camera traps)	83.01% 76.86%	81.09% 75.00%	92.71% 92.39%
Both daytime and nighttime images, d&nt-models (543 test images 11 camera traps)	81.33% 80.19%	79.65% 77.83%	93.00% 93.85%

Tab. 3: Per-image accuracies *resp. per-region accuracies* for selected region subsets. According to the performance of the MegaDetector [BMY19], ten difficult and ten easy region subsets were chosen.

the performance of the MegaDetector with the performance of the AL approach. Here, the terms “difficult” and “easy” reference the level of difficulty for the MegaDetector as a baseline method, and do not indicate the complexity factor of the underlying task itself. The per-image accuracies and per-region accuracies are shown in Table 3.

As expected, the models trained with only a few annotated images and AL do not reach the high accuracies of the MegaDetector for the easy subsets, which are the ones where the MegaDetector performs best. These subsets contain images of situations that can be well represented by considering large amounts of images to model the still complex distribution of common scene conditions. While a large amount of training data helps to model the most likely scene conditions, the derived models such as the MegaDetector cannot capture the rare special cases in the tails of the overall data distribution, since these are underrepresented and require special attention. Thus, when looking at the region subsets that are most difficult for the MegaDetector, it can be obtained that the models trained with AL are beneficial and lead to better class separations. The reason is that these models are specifically trained for the special cases of certain region subsets using annotated images of the corresponding application domain, in contrast to generic tools like the MegaDetector. As a main result, the experiments show that one can not simply apply existing tools like the MegaDetector to an individual dataset and expect a satisfying performance for the task at hand. In contrast, it turns out that spending some effort annotating few example images for an application in a particular region and training more specific models is beneficial.

6 Conclusions and Future Work

Out-of-the-box tools for animal detection in camera trap images, like the MegaDetector [BMY19] or MLWIC [Ta19], do not perform well on images of every camera trap. One reason is that these models were trained on images from other regions like North America covering different environments and scene conditions. With active learning, an efficient method was presented to train specialized models for single camera traps, reaching competitive average performance compared to general models that have been trained with millions of images. Only 156 images were needed for training a specialized model with active learning, and the annotation effort for experts can be kept minimal. The concept allows to include new situations and sites of a camera trap as well with low annotation effort.

The investigation of single camera trap images showed, that learning region-specific models leads to improved recognition accuracies especially in regions where existing general models struggle due to difficult circumstances. Therefore, also the separation of daytime and nighttime images followed by learning individual models for these scenarios led to clear enhancements of the model performances. Another observation was that having too few images of the less represented class in the training set likely leads to low model performance since the model has only images of one class left for selecting further training data at some point. To utilize more image data of the less represented class, it is helpful to have one annotation per image instead of annotation for the whole image sequence. This would also allow for a more comprehensive analysis and for exploiting more semantic information given by the additional image labels.

The separation of images with and without animals serves as a pre-processing step for the following species classification, where empty images do not need to be identified, and the species classifier can focus mainly on animal features. In further investigations on species classification, active learning will be applied as well to choose meaningful images for training. Other investigations on an early stopping strategy for active learning would be interesting to avoid learning biased models due to class imbalance in the data. It is a direct consequence of our analyses, where it was shown that the highest model accuracies are sometimes achieved with less than 150 queried images, which have been used to obtain the final accuracies in the active learning experiments. This would also reduce the annotation efforts for humans even more.

Acknowledgements

This work has been funded by the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung, BMBF, Deutschland) via the project "Development of an Automated Multisensor Station for Monitoring of Biodiversity (AMMOD) - Subproject 5: Automated Visual Monitoring and Analysis"(FKZ: 01LC1903E).

Bibliography

- [BMY19] Beery, Sara; Morris, Dan; Yang, Siyu: Efficient Pipeline for Camera Trap Image Review. In: KDD Workshop on Data Mining and AI for Conservation. 2019.
- [BVHP18] Beery, Sara; Van Horn, Grant; Perona, Pietro: Recognition in Terra Incognita. In: European Conference on Computer Vision (ECCV). pp. 472–489, 2018.
- [Ce15] Ceballos, Gerardo; Ehrlich, Paul R.; Barnosky, Anthony D.; García, Andrés; Pringle, Robert M.; Palmer, Todd M.: Accelerated modern human-induced species losses: Entering the sixth mass extinction. *Science Advances*, 1(5), 2015.
- [Cu21] Cunha, Fagner; dos Santos, Eulanda M.; Barreto, Raimundo; Colonna, Juan G.: Filtering Empty Camera Trap Images in Embedded Systems. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2021. Accepted for publication, arxiv preprint: <https://arxiv.org/abs/2104.08859>.
- [De09] Deng, Jia; Dong, Wei; Socher, Richard; Li, Li-Jia; Li, Kai; Fei-Fei, Li: ImageNet: A large-scale hierarchical image database. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 248–255, 2009.
- [Fr12] Freytag, Alexander; Rodner, Erik; Bodesheim, Paul; Denzler, Joachim: Rapid Uncertainty Computation with Gaussian Processes and Histogram Intersection Kernels. In: Asian Conference on Computer Vision (ACCV). pp. 511–524, 2012.
- [FRD14] Freytag, Alexander; Rodner, Erik; Denzler, Joachim: Selecting Influential Examples: Active Learning with Expected Model Output Changes. In: European Conference on Computer Vision (ECCV). pp. 562–577, 2014.
- [Kä16a] Käding, Christoph; Freytag, Alexander; Rodner, Erik; Perino, Andrea; Denzler, Joachim: Large-scale Active Learning with Approximated Expected Model Output Changes. In: German Conference on Pattern Recognition (GCPR). pp. 179–191, 2016.
- [Kä16b] Käding, Christoph; Rodner, Erik; Freytag, Alexander; Denzler, Joachim: Watch, Ask, Learn, and Improve: A Lifelong Learning Cycle for Visual Recognition. In: European Symposium on Artificial Neural Networks (ESANN). pp. 381–386, 2016.
- [Ke19] Kellenberger, Benjamin; Marcos, Diego; Lobry, Sylvain; Tuia, Devis: Half a Percent of Labels is Enough: Efficient Animal Detection in UAV Imagery Using Deep CNNs and Active Learning. *IEEE Transactions on Geoscience and Remote Sensing*, 57(12):9524–9533, 2019.
- [KTM20] Kellenberger, Benjamin; Tuia, Devis; Morris, Dan: AIDE: Accelerating image-based ecological surveys with interactive machine learning. *Methods in Ecology and Evolution*, 11:1716–1727, 2020.
- [No18] Norouzzadeh, Mohammad Sadegh; Nguyen, Anh; Kosmala, Margaret; Swanson, Alexandra; Palmer, Meredith S.; Packer, Craig; Clune, Jeff: Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *Proceedings of the National Academy of Sciences*, 115(25):E5716–E5725, 2018.
- [Re15] Ren, Shaoqing; He, Kaiming; Girshick, Ross; Sun, Jian: Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. In: *Advances in Neural Information Processing Systems (NIPS)*. 2015.

- [Ro17] Rodner, Erik; Freytag, Alexander; Bodesheim, Paul; Fröhlich, Björn; Denzler, Joachim: Large-Scale Gaussian Process Inference with Generalized Histogram Intersection Kernels for Visual Recognition Tasks. *International Journal of Computer Vision (IJCV)*, pp. 253–280, 2017.
- [RW06] Rasmussen, Carl Edward; Williams, Christopher K. I.: *Gaussian Processes for Machine Learning*. Adaptive Computation and Machine Learning. The MIT Press, 2006.
- [Sc20] Schneider, Stefan; Greenberg, Saul; Taylor, Graham W.; Kremer, Stefan C.: Three critical factors affecting automated image species recognition performance for camera traps. *Ecology and Evolution*, 10(7):3503–3517, 2020.
- [Se19] Serrouya, Robert; Seip, Dale R.; Hervieux, Dave; McLellan, Bruce N.; McNay, R. Scott; Steenweg, Robin; Heard, Doug C.; Hebblewhite, Mark; Gillingham, Michael; Boutin, Stan: Saving endangered species using adaptive management. *Proceedings of the National Academy of Sciences*, 116(13):6181–6186, 2019.
- [SMF20] Shahinfar, Saleh; Meek, Paul; Falzon, Greg: “How many images do I need?” Understanding how sample size per class affects deep learning model performance metrics for balanced designs in autonomous wildlife monitoring. *Ecological Informatics*, 57, 2020.
- [Sz16] Szegedy, Christian; Vanhoucke, Vincent; Ioffe, Sergey; Shlens, Jon; Wojna, Zbigniew: Rethinking the Inception Architecture for Computer Vision. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 2818–2826, 2016.
- [Ta19] Tabak, Michael A.; Norouzzadeh, Mohammad S.; Wolfson, David W.; Sweeney, Steven J.; Vercauteren, Kurt C.; Snow, Nathan P.; Halseth, Joseph M.; Di Salvo, Paul A.; Lewis, Jesse S.; White, Michael D.; Teton, Ben; Beasley, James C.; Schlichting, Peter E.; Boughton, Raoul K.; Wight, Bethany; Newkirk, Eric S.; Ivan, Jacob S.; Odell, Eric A.; Brook, Ryan K.; Lukacs, Paul M.; Moeller, Anna K.; Mandeville, Elizabeth G.; Clune, Jeff; Miller, Ryan S.: Machine Learning to Classify Animal Species in Camera Trap Images: Applications in Ecology. *Methods in Ecology and Evolution*, 10(4):585–590, 2019.
- [VSV17] Villa, Alexander Gomez; Salazar, Augusto; Vargas, Francisco: Towards automatic wild animal monitoring: Identification of animal species in camera-trap images using very deep convolutional neural networks. *Ecological Informatics*, 41:24–32, 2017.
- [Wi19] Willi, Marco; Pitman, Ross T.; Cardoso, Anabelle W.; Locke, Christina; Swanson, Alexandra; Boyer, Amy; Veldhuis, Marten; Fortson, Lucy: Identifying animal species in camera trap images using deep learning and citizen science. *Methods in Ecology and Evolution*, 10:80–91, 2019.

ISTMINER: Interactive Spatiotemporal Co-occurrence Pattern Extraction: A Biodiversity case study

Dina Sharafeldeen ¹, Mohamed Bakli ², Alsayed Algergawy ³, Birgitta König-Ries ⁴

Abstract: In recent years, the exponential growth of spatiotemporal data has led to an increasing need for new interactive methods for accessing and analyzing this data. In the biodiversity domain, species co-occurrence models are critical to gain a mechanistic understanding of the processes underlying biodiversity and supporting its maintenance. This paper introduces a new framework that allows users to explore species occurrences datasets at different spatial and temporal periods to extract co-occurrence patterns. As a real-world case study, we conducted several experiments on a subset of the Global Biodiversity Information Facility (GBIF) occurrences dataset to extract species co-occurrence patterns interactively. For better understanding, these co-occurrence patterns are visualized in a map view and as a graph. Also, the user can export these patterns in CSV format for further use. For many queries, runtimes are in a range that allows for interaction already. Further optimizations are on our research agenda.

Keywords: Spatiotemporal data mining; Co-occurrence patterns; Biodiversity data mining

1 Introduction

Biodiversity research aims to study biologic diversity at different levels of scale ranging from genes to ecosystems [In15]. The biologist Edward O. Wilson, known as the “father of biodiversity” wrote: “It is reckless to suppose that biodiversity can be diminished indefinitely without threatening humanity itself” [Wi99]. Considering the rapid loss in biodiversity we evidence right now [Br19], understanding the mechanisms behind biodiversity is crucial for supporting its conservation.

Biodiversity research is a multidisciplinary, challenging research area that has experienced a tremendous increase in data over the last years. New data are being observed in different places all the time [Sh98]. Species occurrences are one of the most frequently available

¹ Friedrich Schiller University Jena, Heinz-Nixdorf Chair for Distributed Information Systems, Germany
dina.sharafeldeen@uni-jena.de

² Université libre de Bruxelles, Computer & Decision Engineering Department, Brussels, Belgium
mohamed.bakli@ulb.be

³ Friedrich Schiller University Jena, Heinz-Nixdorf Chair for Distributed Information Systems, Germany
alsayed.algergawy@uni-jena.de

⁴ Friedrich Schiller University Jena, Heinz-Nixdorf Chair for Distributed Information Systems, Germany
birgitta.koenig-ries@uni-jena.de

types of biodiversity data. The main attributes of a species occurrence are a taxonomic name, geographic location, and date (i.e., "what," "where," and "when"). In this context, two species are considered to have a *spatiotemporal co-occurrence pattern*, if these two species frequently co-occur with each other within a spatial range and a temporal window. These co-occurrence patterns are of interest for a number of important questions ranging from the identification of suitable indicator species to monitor to the investigation of predator-prey or symbiotic relationships.

There is a growing success in the biodiversity community in mobilizing occurrences data. For example, the Global Biodiversity Information Facility (GBIF)⁵ is an international organisation that focuses on making biodiversity data collected by a large number of contributors, available. Data available through the GBIF portal are primarily occurrence data on plants, animals, fungi, and microbes, and scientific names data. GBIF hosts more than one and a half billion occurrence records. In order to take advantage of this wealth of data to improve understanding of biodiversity and support decision-making, the research community is in need of advanced interactive spatiotemporal mining tools [Ga21].

Spatiotemporal co-occurrence pattern mining refers to discovering the subsets of event types whose instances frequently co-locate in a spatial range and coincide in a temporal window. Previous spatiotemporal co-occurrence pattern extraction algorithms are either apriori-based or pattern growth-based approaches. In the apriori-based approach [Ay14], the number of candidate patterns can grow exponentially with the number of event types. The pattern growth-based approaches [HAA16] show better performance compared to the apriori-based ones. All of these approaches are not sufficiently dealing with big data sets interactively with user preferences. In other words, they are not able to answer questions like, "what are the objects that co-occurred from 2010 till 2015, where they co-occurred within 10 days within 50m as a spatial range?".

Concerning species co-occurrence patterns, some research using data mining was conducted. Multivariate spatial associations were used in [FOM14] for grouping species into disjoint sets with similar co-association values. Also, a data mining method has been developed in [Si16] to assist ecologists in assessing occurrences patterns for species in plant communities. A semantic data mining framework has been recently developed [SAK19] to extract hidden species co-occurrence patterns. However, these methods do not provide interactive facilities to support data exploration. We believe that such facilities are crucial in our context. A user may, for instance, be interested in co-occurrence patterns of insects and a certain plant. She/he may then want to explore different temporal (within a day, a season, a year) and spatial (within 10 m, 100 m, 1 km) resolutions, restrict the exploration to a certain country or region, look at bees only instead of all insects, or a family of plants instead of a particular species. With this, they will gradually gain a deeper understanding of relevant patterns.

To this end, we propose an *interactive spatiotemporal co-occurrence extraction framework*, called *ISTMINER*, in a three-tier architecture. The first tier is responsible for data storage and

⁵ <https://www.gbif.org>

indexing. The second tier is the main processing engine of ISTMINER. It performs candidate generation and filtering in order to extract the spatiotemporal co-occurrence patterns based on the prevalence measures. Also, in this layer, semantic enrichment is conducted for the knowledge graph construction. The third tier comprises all the interactive features provided to the user. Due to the computationally expensive nature of spatiotemporal operations required for co-occurrences mining, we exploit MobilityDB [Zi19] as an infrastructure for efficient spatiotemporal indices and operations. It is an open-source⁶ spatiotemporal database management system, where it provides types and operations on top of PostgreSQL and PostGIS. It has features of high performance as it distributes indices and parallel query execution over the machine cores. Moreover, it supports big data management [BSZ20, BSZ19] on a cluster of machines.

Our proposed framework can be applied to any dataset that contains information that describes an object, when, and where it is observed (OBJECT, TIME, LOCATION). In this paper, we take the biodiversity domain as a case study. We exploit our proposed framework to extract spatiotemporal species co-occurrence patterns. Our contributions are as follows. 1- Designing and constructing an interactive dynamic query to extract spatiotemporal co-occurrence patterns. 2. Reducing the search space for the patterns extraction because the candidates are generated after the filter operations. 3. As a result of reducing the search space, the run-time is improved relative to traditional spatiotemporal co-occurrence patterns extraction methods.

The rest of the paper is organized as follows. Section 2 introduces the concepts of spatiotemporal co-occurrence patterns extraction. We present our proposed framework in Section 3. Section 4 shows our experiments on a subset of the GBIF occurrences data set to extract spatiotemporal species co-occurrence patterns. We discuss future work issues and present our conclusions in the last section.

2 Spatiotemporal Co-occurrence Patterns

Discovering implicit spatiotemporal patterns from datasets is important for many scientific domains including astronomy, ecology, meteorology, public health, agricultural sciences [Er04]. These patterns can be used for modeling various scientific phenomena (e.g., tornadoes, propagation of epidemics, clouds). Moreover, they can be utilized for performing large-scale verification of current knowledge, as well as the prediction of unknown spatiotemporal relationships (e.g., predicting the spread of epidemics [Ku04]). The growing amount of data being generated and collected from various scientific sources makes the data-driven knowledge discovery process very challenging to researchers in these fields.

In this section, we project the spatiotemporal co-occurrence patterns concepts on GBIF's occurrences dataset for a better explanation. GBIF data is de-normalized and is stored in

⁶ <https://github.com/MobilityDB/MobilityDB>

a single table. GBIF occurrence records are stored in a simple column-based data format (Darwin Core) [Wi12]. Darwin Core makes data easy to enter and view, but it is potentially challenging to extract spatiotemporal relationships between entities.

In the context of the spatiotemporal co-occurrence pattern concepts, the GBIF occurrence records are the spatiotemporal instances, denoted by \mathbb{I} . Table 1 shows an excerpt of GBIF occurrence instances where objects are observed in different places at different times.

Tab. 1: An excerpt of a GBIF occurrence dataset [GB20]

<i>gbifid</i>	<i>taxonrank</i>	<i>kingdom</i>	<i>scientificname</i>	<i>eventdate</i>	<i>latitude</i>	<i>longitude</i>	<i>countrycode</i>	<i>stateprovince</i>
49652757	SPECIES	Animalia	Ectropis crepuscularia	2003-04-13	53.07632	10.33553	DE	Ni
1314754881	GENUS	Bacteria	Methylobacterium	2008-01-01	48.38	9.52	DE	Swabian Jura
1500777636	SPECIES	Chromista	Ceratium lineatum	2002-09-12	39	0.16	ES	Valencia
580793	SPECIES	Fungi	Myriospora scabrida	2003-08-07	69.24917	20.39393	NO	Troms
1580401	SPECIES	Plantae	Achillea clypeolata Sibth.	2001-06-06	41.05	26.083333	GR	Nom. Evrou
178565421	SPECIES	Protozoa	Ceratiomyxa fruticulosa	2003-07-02	42.9	3.03	FR	Pyrénées-Orientales

A **feature type** is a non-spatiotemporal attribute of an instance that identifies the properties of the instance like the *countrycode*, *scientificName* attributes. A feature type is denoted as f_i , and the set of all features is denoted by $\mathbb{F} = \{f_1, f_2, \dots, f_m\}$. The set of all instances of a particular feature type f_i is denoted as $\mathbb{I}f_i$.

A **spatiotemporal co-occurrence pattern** is a subset of all feature types, whose instances frequently co-occur within a time window and spatial range. A pattern is denoted as P , where $P = \{f_{i1}, \dots, f_{ik}\}$ and $P \subset \mathbb{F}$.

The number of feature types in a pattern will be referred as the *cardinality* of the pattern. A k -cardinality pattern refers to a k -subset of all feature types. The minimum cardinality of the pattern is 2 ($k \geq 2$).

The pattern instance (denoted by $PInst$) of P is a unique incident of a spatiotemporal co-occurrence (spatial and temporal overlap) among the instances of the all feature types in P .

Prevalence measure is used for determining the interestingness of a pattern. The prevalence measure of a feature type f_i in the pattern P is calculated as

$$pr(P, f_i) = |P_{f_i}| / |\mathbb{I}f_i|$$

where $|P_{f_i}|$ denotes the number of unique instances of feature type f_i participating in the pattern instances of P , and $|\mathbb{I}f_i|$ is the total number of instances of feature type f_i .

3 ISTMNER Framework Overview

As mentioned earlier, ISTMNER is a three-tier architecture for spatiotemporal co-occurrence patterns extraction. Figure 1 shows the architecture overview of our proposed framework.

After the user uploads the dataset, the storage & indexing layer is responsible for performing data preprocessing steps as preparation for data indexing. The core layer is responsible for candidate generation and filtering to extract the spatiotemporal co-occurrence patterns and hence conducting semantic enrichments on the extracted patterns. As an example of these semantic enrichments, currently, we are working on verifying the extracted patterns by connecting to Wikidata, which is a step towards knowledge graph construction. We focus in this paper mainly on the spatiotemporal co-occurrence patterns extraction process. Therefore, we developed the ISTMINER tool. The interactive features layer contains the features provided by the ISTMINER tool⁷. These features keep the user involved in the whole process, starting from providing the dataset to exploring and exporting the extracted patterns in different available formats like map view, CSV view, and graph view. Meanwhile, the user participates in the extraction process by providing her/his preferences through the interactive query construction. The following sections will explain the three layers with the help of our use case, aiming to extract spatiotemporal species co-occurrence patterns from the GBIF data.

3.1 Storage & Indexing

In this layer, after the user upload the GBIF occurrences dataset, the spatiotemporal instances are stored after data cleaning from multiple aspects, as described in the following section. Then, the relevant attributes are selected (e.g., *eventDate*, *latitude*, *longitude*). For efficiently managing the location and time operations, MobilityDB is exploited as a PostgreSQL extension to use the appropriate spatiotemporal type. Finally, data types are indexed to speed up the filtering process.

Data Cleaning. In this step, input data is cleaned by removing duplicates and inconsistency in data. We have encountered problems such as duplicate values, the inconsistency of data types, missing quotes, and null values for important fields. Some of these problems have been fixed, but others have been filtered out, especially those without location data and/or time.

Features Selection. Usually, the datasets contain tens of features. Therefore, the main features fall under the three labels: “what”, “where”, and “when” are selected. In the GBIF dataset, the label “what” identifies what to select and filter, such as the taxonomic level, kingdom level, class, order, etc. The other two labels are used to construct a spatiotemporal type which is then used in the spatiotemporal operations.

Spatiotemporal Types Selection. To represent the location and time of an object, we use the *tpoint* type of MobilityDB which represents the geometry point over time. The *tpoint* is a temporal extension of the PostGIS geometry/geography point. It is used by an extensive set of operations such as topological functions (e.g., intersects, contains), spatiotemporal

⁷ All resources of the paper are available under the link: <https://github.com/fusion-jena/ISTMINER>

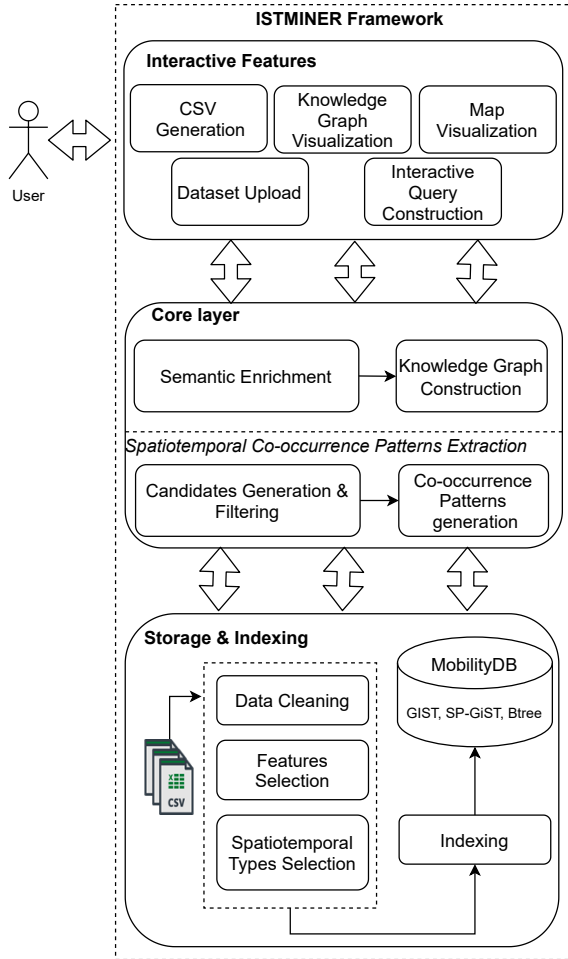


Fig. 1: Architecture Overview of ISTMINER

bounding operators (e.g., , @>), distance, temporal and range predicates, lifting operations, etc. The value of the *point* is stored as follows: *point@time*. The point is constructed using the *latitude* and *longitude* attributes and the *eventDate* attribute represents the time ((*latitude*, *longitude*)@*eventDate*). The schema of the database table is as follows:

GBIF: <*GBifId*: *bigint*, *STPoint*: *tpoint(geometry)*, *countrycode*: *text*, *stateProvince*: *text*, *taxonRank*: *text*, *kingdom*: *text*, *phylum*: *text*, *class*: *text*, *order*: *text*, *family*: *text*, *genus*: *text*, *scientificname*: *text*>

Indexing. After the data types are created, the appropriate indices are built on top of them. The B-tree index is used for all basic data types and the spatiotemporal GiST (which is an R-tree) and SP-GiST (which is an Oct-tree) indices are used for the MobilityDB tpoint type. The two spatiotemporal indices are used to perform a quick range filter before executing the expensive operations (i.e., distance, intersects).

3.2 Spatiotemporal Co-occurrence Patterns Extraction

Starting from this step, the user interactively participates in the extraction process by providing her/his preferences through the interactive query construction facility in the ISTMINER tool. We show an example of how the species spatiotemporal co-occurrence patterns is extracted from the GBIF occurrences dataset. We will explain based on one of the question examples,

”Find the co-occurrence between species in the Animalia kingdom in Germany within a distance of 200m and an interval of 10 days during the period from 01 July, 2019 to 30 July, 2019”.

The spatiotemporal co-occurrence patterns extraction is outlined in Algorithm 1. The two main phases in this algorithm are the *”Candidates Generation and Filtering”* and *”Spatiotemporal Co-occurrence Patterns Generation”*. This algorithm starts with reading the given dataset provided as an input. Normally, the set of all instances (\mathbb{I}) and the set of all features (\mathbb{F}) are explored, where (CP) is used as intermediate temporary storage. The improvement in our proposed algorithm is that the search space is filtered by the user preferences before the candidates generation phase. Then, the 2-cardinality spatiotemporal co-occurrence patterns are discovered, which occur within the (*distance within*, and the *temporal within*) values. Finally, the prevalent spatiotemporal co-occurrence patterns are characterized by the prevalence measure threshold (p_{th}) where the *prevalence measure* for the pattern must be greater than or equal to the *prevalence measure threshold* p_{th} . The *prevalence measure* is calculated as mentioned before in section 2.

Algorithm 1: Spatiotemporal Co-occurrence Patterns Extraction Algorithm

Input: A spatiotemporal dataset, FilterPeriod, Dwithin, Twithin, prevalence threshold (p_{th})

Output: The set of prevalent spatiotemporal co-occurrence patterns FP

function ISTMINER()

```

1  |   $\langle \mathbb{I}, \mathbb{F} \rangle \leftarrow ReadDataset();$ 
2  |   $CP \leftarrow \mathbb{F};$ 
3  |   $FP = GenerateAndFilterCandidates(CP, FilterParameters);$ 
4  |   $FP = GenerateSTCo-occurrencePattern(FP, Swithin, Twithin, p_{th});$ 
5  |  return FP;
```

We add filters for the base type based on the non-spatiotemporal parameters such as taxonomic and kingdom levels. For the spatiotemporal parameters (i.e., time period or distance), we take care of adding the appropriate spatial or temporal bounding filters and operators. To construct dynamic queries/query plans based on the user parameters, the dynamic query construction goes through two phases:

Phase 1: Candidates Generation & Filtering

In this phase, the non-spatiotemporal parameters' filters exploit the presence of the indices to be used by the PostgreSQL planner. For instance, we apply the overlapping operator (`&&`) to perform a range filter whether it is temporal, spatial, or both. This step is done before executing the next phase which contains the expensive operations.

Once the dynamic query is created, it is sent to PostgreSQL for planning and execution. This query part aims at filtering the records at answering the complete query. The SQL query is constructed based on the user selection as follows:

```

1 SELECT T1.scientificname As Object1Name, T2.scientificname As Object2Name
2     ,T1.STPoint As Object1STPoint, T2.STPoint As Object2STPoint
3 FROM GBIF T1, GBIF T2 WHERE T1.GBifId < T2.GBifId
4 AND T1.countrycode='DE' AND T2.countrycode='DE'
5 AND T1.taxonRank = 'SPECIES' AND T1.kingdom = 'Animalia'
6 AND T2.taxonRank = 'SPECIES' AND T2.kingdom = 'Animalia'
7 AND T1.STPoint && period('2019-07-01','2019-07-30')
8 AND T2.STPoint && period('2019-07-01','2019-07-30')

```

The query joins the GBIF table with itself to generate the candidate instances. The join predicate (line 3) ensures that no same pair is included in the results. The instances are filtered (Lines 4-6) by country code and then by descriptive attributes about species (e.g., taxonomic level, kingdom level). As shown in lines 7-8, the overlaps predicate (`&&`) is used to perform a temporal range filter, where it checks the temporal overlapping between the bounding box of the STPoint attribute and a given time period. The execution plan of the above query is as follows:

```

1 -> Nested Loop
2 -> Join Filter: (t1.gbifid < t2.gbifid)
3 -> Bitmap Heap Scan on GBIF t1
4   -> Filter: ((taxonrank='..') AND (kingdom='..'))
5   -> Bitmap Index Scan on countrycode_btree_idx
6     -> Index Cond: (countrycode='..')
7   -> Bitmap Index Scan on stpoint_spgist_idx
8     -> Index Cond: (stpoint && period)
9 -> Materialize
10 -> Bitmap Heap Scan on GBIF t2
11   -> Filter: ((taxonrank='..') AND (kingdom='..'))
12   -> Bitmap Index Scan on countrycode_btree_idx
13     -> Index Cond: (countrycode = '..')
14   -> Bitmap Index Scan on stpoint_spgist_idx
15     -> Index Cond: (stpoint && period)

```

The plan shows how the query is executed in MobilityDB. The planner uses a two-step plan. The first step (Lines 5-8) performs an index scan to prune data that will not contribute to

the results. The B-tree index of the countrycode attribute and the SP-GiST index of the STPoint attribute are invoked here to find the candidate rows matching the predicates. Then, the second step, the planner (Lines 3-4) fetches the data from the table itself and applies additional filters for the taxonRank and kingdom attributes. As the query is a self-join and in order not to re-scan the whole table again, the planner creates a Materialize node (Line 10) for saving an in-memory copy of the rows generated. Thus, the query is executed faster.

Phase 2: Spatiotemporal Co-occurrence Pattern Extraction

After the non-spatiotemporal parameters' filters are exploited, the spatiotemporal co-occurrence patterns are extracted. Afterward, the co-occurrence rules are generated. Table 2 shows a sample of the generated co-occurrence rules with their prevalence values. As an example, the following query retrieves the spatiotemporal co-occurrence patterns that co-occur within a distance of 200 meters, during an interval of 10 days, and with a *prevalence measure* greater than or equal to 0.001.

```

1 WITH TEMP AS(
2 SELECT Object1Name, Object2Name, count(*) As InteractionsCount
3 FROM Phase1
4 WHERE
5 Object1STPoint::period && expandTemporal(Object2STPoint,interval '10 day')::period
6 AND ST_DWithin(Object1STPoint::geometry, Object2STPoint::geometry, 200)

7 GROUP BY Object1Name, Object2Name
8 )
9 SELECT Object1Name, Object2Name
10 FROM TEMP
11 WHERE InteractionsCount /(SELECT sum(InteractionsCount) FROM TEMP) > 0.001

```

The query plan is as follows:

```

1 -> CTE Scan on temp
2 -> Filter: (((interactionscount)::numeric / $2) > 0.001)
3 -> CTE temp
4 -> GroupAggregate
5 -> Group Key: phase1.object1name, phase1.object2name
6 -> Sort:
7 -> Sort Key: phase1.object1name, phase1.object2name
8 -> Gather
9 -> Workers Planned: 8
10 -> Parallel Seq Scan on phase1
11 -> Filter: ((... && ...) AND st_dwithin(...))

```

A CTE query (Lines 1-7) aggregates the count of the interactions for each pair of the pattern instances. The function `expandTemporal` expands the temporal dimension of the STPoint attribute and then the overlapping predicate checks the temporal boundaries for the two arguments. In line 5, the function `ST_DWithin` returns true if the Euclidean distance between the two points is within 200 meters. Finally, the main query returns the patterns with a prevalence greater than or equal 0.001. In line 2, the participation calculation filter is performed to generate the prevalent spatiotemporal patterns. As shown in the CTE plan (Line 9), the planner distributes the workload on the available eight workers of the machine. The results are then collected, sorted, and aggregated (Line 4-7) to get the count of interactions.

Tab. 2: Part of the generated co-occurrence rules

co-occurrence rule	Prevalence
Calidris alpina (Linnaeus, 1758) <i>Co-occur</i> Plectrophenax nivalis (Linnaeus, 1758)	0.08241
Corvus corone (Linnaeus, 1758) <i>Co-occur</i> Oenanthe oenanthe (Linnaeus, 1758)	0.00683
Periparus ater (Linnaeus, 1758) <i>Co-occur</i> Dendrocopos major (Linnaeus, 1758)	0.00331
Sternaspis Otto, 1821 <i>Co-occur</i> Spiophanes Grube, 1860	0.00754
Ophelina Örsted, 1843 <i>Co-occur</i> Maldanidae	0.00332
Ophelina Örsted, 1843 <i>Co-occur</i> Aricidea Webster, 1879	0.03029
Aricidea Webster, 1879 <i>Co-occur</i> Maldanidae	0.03169
Accipiter nisus (Linnaeus, 1758) <i>Co-occur</i> Oenanthe oenanthe (Linnaeus, 1758)	0.03229
Accipiter nisus (Linnaeus, 1758) <i>Co-occur</i> Corvus corone Linnaeus, 1758	0.03607

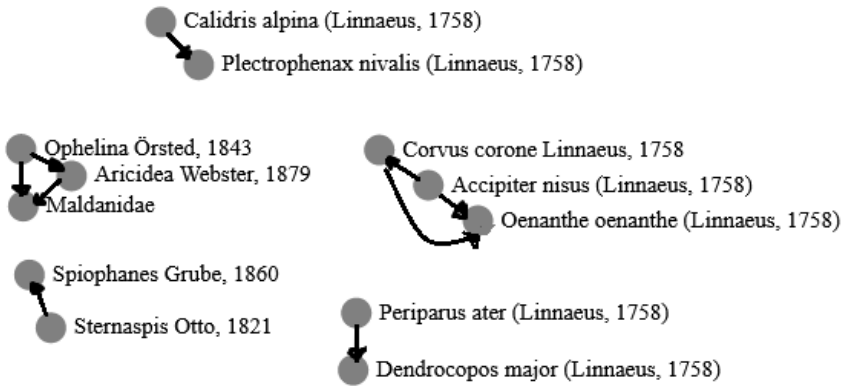


Fig. 2: Part of the generated species co-occurrence graph

3.3 Interactive Features

For getting the user more involved in the spatiotemporal co-occurrence patterns extraction process, we developed ISTMINER tool as shown in figure 3. Using ISTMINER, the user is involved starting from data uploading, data exploration till the end of the species co-occurrence patterns extraction. After the extraction of the species co-occurrence patterns, the user can export them and show them in different formats.

- **Interactive Query Construction:** While the user is selecting his/her preferences (i.e., taxon level, spatiotemporal window), the query is constructed automatically.
- **CSV Generation:** As an initial output format, the user can export the output as a CSV format to be used in further analysis. Furthermore, the extracted species

co-occurrence patterns are visualized as a graph for better understanding. Figure 2 visualizes a part of the generated graph.

- **Map Visualization:** While the user is interactively extracting the species co-occurrence patterns, the species co-occurrences are visualized in a map view where the user can have a good overview of the output. As shown in figure3, the species co-occurrences are highlighted based on the user selection.

ISTMINER: Interactive Spatio-temporal miner Using MobilityDB

Connected to ISTMINER Settings ▾

Object 1	Object 2	Period	Co-Occur Within
Taxonomic level SPECIES ▾ kingdom level Bacteria ▾	Taxonomic level SPECIES ▾ kingdom level Plantae ▾	From 07/01/2009 To 10/31/2009	Distance (in meters) 200 Interval (in days) 10

Co-Occur

```

SELECT * FROM observations T1, observations T2
WHERE T1.gbifid < T2.gbifid
AND T1.taxonrank = 'SPECIES' AND T1.kingdom = 'Bacteria'
AND T2.taxonrank = 'SPECIES' AND T2.kingdom = 'Plantae'
AND T1.tpoint && PERIOD('2009-07-01', '2009-10-31')
AND T2.tpoint && PERIOD('2009-07-01', '2009-10-31')
AND T1.tpoint::Period && expandTemporal(T2.tpoint, interval '10 day')::Period
AND ST_dWithin(T1.tpoint::geometry, T2.tpoint::geometry, 200);
  
```

Show all rows Show specific row
 object1_location object2_location

object1_name: Focidula hypoleuca (Pallas, 1764)
 object1_date: 2009-10-02 00:00:00+02
 object2_name: Larus marinus Linnaeus, 1758
 object2_date: 2009-10-03 00:00:00+02

Fig. 3: ISTMiner: Spatiotemporal co-occurrence patterns extraction

- **RDF Generation:** Currently the user can export the species co-occurrence patterns in CSV format. We are working on constructing the species co-occurrences as a knowledge graph in RDF format, where the nodes represent objects and the edges are the co-occur relationships.

4 Experimental Evaluation

In this section, we report on preliminary experiments that show the run time of the proposed model with different datasets in different queries. The experiment is conducted on a single-node machine running PostgreSQL, PostGIS, and MobilityDB over Ubuntu 20.04. It is equipped with 6 cores per socket (2 sockets and 2 threads per core) Intel i7-7800X CPU @ 3.50GHz processor, 1TB SSD, and 32GB RAM.

We downloaded occurrences records from the GBIF portal that met the following filter criteria [GB20] *present status, including coordinates, year range (2000, 2020), continent Europe*⁸ On the downloaded dataset, we followed the steps mentioned in the framework section. This resulted in a dataset with around 143 Million records and a size of roughly 78 GB. As a basis for determining spatial distances, afterwards, the data records are projected based on the spatial reference system of Europe⁹ to have accurate spatial operations. Then, the spatiotemporal types are selected and the data are indexed properly. Table 3 shows the characteristics of the data sets. It shows the number of occurrence records retrieved by our query in the subset of European countries that we used in the remainder of the evaluation.

Tab. 3: Country code and number of occurrence records

Country code	NO	DK	IE	FR	DE	PL	GR	GB	RO	SI
	Norway	Denmark	Ireland	France	Germany	Poland	Greece	United Kingdom	Romania	Slovenia
Count	1320197	991367	966553	870258	626947	210144	170913	168003	126824	126824

To gain an insight into the run-time performance of our approach, we used different query templates with different filters on the different country codes depicted in Table 3. Let's assume that the user want to study the species co-occurrences between insects and plants or co-occurrences between animals with other animals in different families. The following query template shows the main attributes we could build the different queries on.

“Find the co-occurrence between species from {Family} in {class} in {order} in {country code(CC)} within a distance of {Dwithin}m and an interval of {Twithin} days during the period {FilterPeriod(FP)}”

Biological fieldwork is mostly conducted in the growing season; thus, there are more records for summer than for other seasons. Therefore, we select July in our first set of example queries. A second set of queries examines the time interval between begin of January to end of March. We conducted different experiments with different settings as follows:

```

1 Q1:FilterPeriod(FP)=(from 01 July, 2019 to 30 July, 2019), kingdom1 = 'Animalia', class1='
    Insecta', kingdom2= 'Plantae', Dwithin= 15 m, Twithin= 10 days
2 Q2:FilterPeriod(FP)=(from 01 July, 2019 to 30 July, 2019), kingdom1 = 'Animalia' ,class1='
    Insecta', kingdom2= 'Plantae', Dwithin= 5 m, Twithin= 2 days
3 Q3:FilterPeriod(FP)=(from 01 Jan, 2019 to 30 March, 2019), kingdom1 = 'Animalia' ,kingdom2
    = 'Animalia', Dwithin= 15 m, Twithin= 10 days
4 Q4:FilterPeriod(FP)=(from 01 Jan, 2019 to 30 March, 2019), kingdom1 = 'Animalia' ,kingdom2
    = 'Animalia', Dwithin= 5 m, Twithin= 2 days

```

The traditional candidates generation process in co-occurrence patterns mining generates the Cartesian product of the number of the records in each dataset. For example, the number of records in the Norway (NO) dataset is 1320197, this would result in a Cartesian product with roughly $1.7 * 10^{12}$ records. In our proposed algorithm, the number of generated candidates after filtering is only 7309576. Table 4 shows the number of candidates for our example

⁸ Due to incomplete location hierarchies in GBIF this will not return all European datasets. There exist datasets that contain country information only; these do not qualify for our query.

⁹ <https://epsg.io/3035>

Tab. 4: Number of candidates after filtering in (Q1-Q4)

Country code	NO	DK	IE	FR	DE	PL	GR	GB	RO	SI
<i>Q1</i>	1792852	11655	463386	18276515	23594117	68162	76818	395766	0	134928
<i>Q2</i>	1792852	11655	463386	18276515	23594117	68162	76818	395766	0	134928
<i>Q3</i>	146689	36481	767376	105288121	28686736	6400	265225	5044516	44890000	784
<i>Q4</i>	146689	36481	767376	105288121	28686736	6400	265225	5044516	44890000	784

countries and time intervals. Due to the way the filter query is formulated, Queries 1 and 2 and 3 and 4, respectively, result in the same candidate set. Table 5 shows the number of the extracted species co-occurrence patterns from the four different queries with a prevalence value of 0.001.

Tab. 5: Number of extracted co-occurrence patterns

Country code	NO	DK	IE	FR	DE	PL	GR	GB	RO	SI
<i>Q1</i>	27	21	0	2	6	79	35	24	0	34
<i>Q2</i>	18	7	0	2	1	21	31	4	0	32
<i>Q3</i>	8	0	16	1	9	27	23	0	19	30
<i>Q4</i>	8	0	16	0	4	27	13	0	17	30

As we can see from table 5 and table 6, the run-time is proportional to the number of generated candidates in each query, and the proposed filtering is very efficient in reducing the number of candidates which improves the run-time.

Tab. 6: Runtime in secs in (Q1-Q4)

Country code	NO	DK	IE	FR	DE	PL	GR	GB	RO	SI
<i>Q1</i>	3.87	2.45	2.67	9.80	14.75	2.14	3.06	2.92	1.99	3.63
<i>Q2</i>	3.57	2.47	2.7	9.63	11.55	2.44	2.52	2.54	1.26	0.27
<i>Q3</i>	4.66	3.62	4.18	13.84	16.43	4.18	4.33	4.84	6.81	0.21
<i>Q4</i>	3.58	3.74	3.99	15.71	15.17	3.67	3.84	5.57	6.66	0.21

5 Conclusion and Future Work

This paper presents ISTMINER, an interactive spatiotemporal co-occurrence patterns extraction framework where the user is involved in the whole extraction process. Experiments were conducted on the GBIF occurrences dataset to extract species co-occurrence patterns. These co-occurrence patterns can be exported as RDF format as a step forward to the biodiversity knowledge graph. Comparing to traditional spatiotemporal co-occurrence extraction methods, our proposed framework shows high responsive performance. As a future work, we aim to evaluate the prototype with biodiversity researchers to determine which settings of the parameters (e.g., prevalence) result in biologically meaningful output. Thus, ISTMINER is a tool that effectively supports their work.

References

- [Ay14] Aydin, Berkay; Kempton, Dustin; Akkineni, Vijay; Gopavaram, Shaktidhar Reddy; Pillai, Karthik Ganesan; Angryk, Rafal: Spatiotemporal indexing techniques for efficiently mining spatiotemporal co-occurrence patterns. In: 2014 IEEE international conference on big data (Big Data). IEEE, pp. 1–10, 2014.
- [Br19] Brondizio, Eduardo S; Settele, Josef; Díaz, S; Ngo, Hien T et al.: Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. IPBES secretariat, Bonn, 2019.
- [BSZ19] Bakli, Mohamed; Sakr, Mahmoud; Zimanyi, Esteban: Distributed Moving Object Data Management in MobilityDB. In: Proceedings of the 8th ACM SIGSPATIAL International Workshop on Analytics for Big Geospatial Data. BigSpatial '19, ACM, New York, NY, USA, pp. 1–10, 2019.
- [BSZ20] Bakli, Mohamed; Sakr, Mahmoud; Zimányi, Esteban: Distributed Spatiotemporal Trajectory Query Processing in SQL. In: Proceedings of the 28th International Conference on Advances in Geographic Information Systems. SIGSPATIAL '20, p. 87–98, 2020.
- [Er04] Erwig, Martin: Toward spatio-temporal patterns. In: Spatio-temporal databases, pp. 29–53. Springer, 2004.
- [FOM14] Flügge, Anton J; Olhede, Sofia C; Murrell, David J: A method to detect subcommunities from multivariate spatial associations. *Methods in Ecology and Evolution*, 5(11):1214–1224, 2014.
- [Ga21] Gadelha Jr, Luiz MR; de Siracusa, Pedro C; Dalcin, Eduardo Couto; da Silva, Luís Alexandre Estevão; Augusto, Douglas A; Krempser, Eduardo; Affe, Helen Michelle; Costa, Raquel Lopes; Mondelli, Maria Luiza; Meirelles, Pedro Milet et al.: A survey of biodiversity informatics: Concepts, practices, and challenges. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 11(1):e1394, 2021.
- [GB20] GBIF Occurrence Download. <https://doi.org/10.15468/d1.rv4rgh>, 2020. Downloaded: 2020-11-23.
- [HAA16] Hamdi, Shah Muhammad; Aydin, Berkay; Angryk, Rafal A: A pattern growth-based approach for mining spatiotemporal co-occurrence patterns. In: 2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW). IEEE, pp. 1125–1132, 2016.
- [In15] Inthasone, Somsack; Pasquier, Nicolas; Tettamanzi, Andrea GB; da Costa Pereira, Célia: *Biodiversity and Environment Data Mining*. 2015.
- [Ku04] Kuhn, Katrin; Campbell-Lendrum, Diarmid; Haines, Andy; Cox, J: Using climate to predict infectious disease outbreaks: A review. World Health Organization, 2004.
- [SAK19] Sharafeldeen, Dina; Algergawy, Alsayed; König-Ries, Birgitta: Towards Knowledge Graph Construction using Semantic Data Mining. In: The 21st International Conference on Information Integration and Web-based Applications & Services (In press). 2019.
- [Sh98] Shah, Anup: Why is biodiversity important? Who cares? *Global Issues*, 1998.
- [Si16] Silva, Luís Alexandre Estevão; Siqueira, Marinez Ferreira; dos Santos Pinto, Flávia; Barros, Felipe Sodré M; Zimbão, Geraldo; Souza, Jano Moreira: Applying data mining techniques for spatial distribution analysis of plant species co-occurrences. *Expert Systems with Applications*, 43:250–260, 2016.

- [Wi99] Wilson, Edward O: The diversity of life. WW Norton & Company, 1999.
- [Wi12] Wieczorek, John; Bloom, David; Guralnick, Robert; Blum, Stan; Döring, Markus; Giovanni, Renato; Robertson, Tim; Vieglais, David: Darwin Core: an evolving community-developed biodiversity data standard. *PloS one*, 7(1):e29715, 2012.
- [Zi19] Zimányi, E.; Sakr, M.; Lesuisse, A.; Bakli, M.: MobilityDB: A Mainstream Moving Object Database System. In: *Proc. of the 16th International Symposium on Spatial and Temporal Databases, SSTD 2019*. ACM, Vienna, Austria, pp. 206–209, 2019.

Towards a Semantic Toolbox for Reproducible Knowledge Graph Generation in the Biodiversity Domain - How to Make the Most out of Biodiversity Data

Samira Babalou ¹, David Schellenberger Costa ², Jens Kattge ³, Christine Römermann ⁴, Birgitta König-Ries ⁵

Abstract: Knowledge Graphs are widely regarded as one of the most promising ways to manage and link information in the age of Big Data. Their broad uptake is still hindered by the large effort required to create and maintain them, though. In this paper, we propose the creation of a semantic toolbox that will support data owners in transforming their databases into reproducible, dynamically extendable knowledge graphs that can be integrated and jointly used. We showcase the need, potential benefits and first steps towards the solution in our example domain, biodiversity research.

Keywords: Knowledge Graph; Dynamic; Reproducible; Biodiversity

1 Introduction

Knowledge Graphs (KGs) are widely regarded as one of the most promising ways to manage and link information in the age of Big Data. While the notion of KG is somewhat vague, for this paper, the definition provided by Nickel et al. [Ni15] is useful. According to them, KGs are graph-structured knowledge bases that store factual information in form of relationships between entities. These graphs consist of nodes and edges connecting them. Nodes represent entities, i.e., concepts and real-world things. Edges represent relationships among entities. Both nodes and edges have unambiguous labels. Typically, KGs are encoded

¹ Heinz-Nixdorf Chair for Distributed Information Systems, Institute for Computer Science, Friedrich Schiller University Jena, Germany; German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig samira.babalou@uni-jena.de

² Systematic Botany and Functional Biodiversity, Institute of Biology, University of Leipzig, Johannisallee 21-23, 04103 Leipzig, Germany; German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig david.schellenberger.costa@uni-leipzig.de

³ Max Planck Institute for Biogeochemistry, Hans Knöll Str. 10, 07745 Jena, Germany; German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig jkattge@bgc-jena.mpg.de

⁴ Institute of Ecology and Evolution with Herbarium Haussknecht and Botanical Garden, Department of Plant Biodiversity, Friedrich Schiller University Jena, D-07743 Jena; German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig christine.roemermann@uni-jena.de

⁵ Heinz-Nixdorf Chair for Distributed Information Systems, Institute for Computer Science, Friedrich Schiller University Jena, Germany; German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig birgitta.koenig-ries@uni-jena.de

as (subject, predicate, object)-triples. Nodes in the KG are subjects and objects, and relationships are predicates. One popular data model to use is RDF (Resource Description Framework) [La98], the standard data model of the Semantic Web. The triple structure allows for intuitive visualisations of KGs. Because of the underlying formal model, there are other important ways KGs can be used: First, with the help of unique identifiers, KGs can be interlinked with general-purpose KGs like WikiData [VK14] or with other domain-specific KGs. They are thus an excellent tool for data integration. Second, many KGs offer SPARQL endpoints enabling complex, declarative queries. Third, new knowledge can be inferred, using reasoners like Pellet [Si07] or HermiT [G114]. In addition, machine learning or data mining can be applied. One could infer, e.g., that two nodes represent the same real-world entity based on the similarity of the subgraphs surrounding them.

The potential of KGs has also been recognized for our example domain, biodiversity science. This discipline explores the diversity of life on earth, including diversity of genes, species, and ecosystems. Conserving biodiversity is a prerequisite to reaching most of the sustainable development goals. For this, a better understanding of biodiversity and the underlying mechanisms is needed. For many questions, this understanding can only be gained based on large, integrated data. Sachs et al. [Sa19] state: “To unite our data silos in biodiversity science, we need agreement and adoption of a data modelling framework. A knowledge graph built using RDF, supported by an identity broker such as WikiData, has the potential to link data and change the way biodiversity science is conducted.”

However, to date, few biodiversity KGs exist (cf. [Pa19, Pe19]). We believe, that – as in other domains – one major roadblock is the large effort and high semantic web expertise still needed to create them. To ameliorate the situation, we propose the development of a toolbox that will allow data owners who are domain experts but not necessarily semantic web experts to transform their databases into interlinkable KGs. A strong focus of our work is on making this creation reproducible. With this, it will not only be possible to effortlessly add new entries to a KG when new data becomes available, but also to track the provenance of each entry in the KG, a prerequisite for their usage in research. Finally, the recorded workflows can serve as templates for other databases, easing the creation of further KGs.

In particular, in this paper, we present the workflow of iKNOW, a planned semantic toolbox for KG creation and evolution in the biodiversity domain. The creation of this toolbox is the main aim of the iKNOW project ⁶ within the German Centre for Integrative Biodiversity Research (iDiv) ⁷.

In the future the application of this workflow will facilitate a greater alignment between data and expertise in the biodiversity domain. The resulting KG serves as discovery tool for biodiversity data and can provide a robust infrastructure for managing biodiversity knowledge. The ultimate goal is to make biodiversity data more accessible and visible in the world.

⁶ <https://fusion.cs.uni-jena.de/fusion/projects/iknow/>

⁷ <https://www.idiv.de/en/index.html>

The remainder of this paper is structured as follows: First, we will provide an overview of the state of the art with respect to biodiversity knowledge graphs and knowledge graph creation. We will then provide some insights into the context of our research and the sample datasources we work with, give an overview of our proposed approach and conclude with an outlook to future work.

2 State of the Art

Biodiversity KGs were first proposed by Page et al. [Pa16]. By now, first, still rather isolated examples of such KGs exist: Ozymandias [Pa19] combines scholarly data about the Australian fauna from different sources, including the Atlas of Living Australia, the Biodiversity Heritage Library, ORCID, and WikiData, and links to external KGs. OpenBiodiv [Pe19], another KG in the biodiversity domain, generates a KG from scientific literature (around 5000 scholarly articles). It integrates knowledge extracted from biodiversity publications in a graph database using the OpenBiodiv-O ontology [Se18] and an RDF version of the Global Biodiversity Information Facility (GBIF) ⁸ taxonomic backbone. All of these existing approaches focus on KG creation from publications. In contrast, the major novelty of iKNOW is to generate the biodiversity KG from tabular datasets. Tabular data such as CSV files is one of the most commonly used input formats in data analytics pipelines used by biodiversity scientists.

To understand the semantic structure and meaning of tabular data, the Semantic Web Challenge on Tabular Data to Knowledge Graph Matching [Ji20] is held annually. The aim of this challenge is to automatically detect matching WikiData entries for individual cell values, columns, and relationships across columns. This is to some extent relevant to our work, but our setting is different and in a sense easier in that we have additional knowledge sources including often rich metadata and human expertise that can be leveraged. Nevertheless, the approaches developed in the SemTab Challenge can inform parts of our workflow.

Only a few approaches to dynamic KG generation have been proposed so far (cf. [Ji15, He17, Ch17]). In this group of studies, once an entity is mapped to the KG's entity (through a mapping matrix), the mapped matrix will be updated. In this way, the construction method will be updated once one entity is added. But, the final KG is static for the given input data. This does not quite match our requirements. DBPedia [Au07] and WikiData [VK14] aim to establish dynamic, cross-domain knowledge graphs. They serve as valuable hubs for our KG. However, their creation methods can be applied partly only to our problem. In particular, in contrast to those approaches, we aim for automatically captured provenance information that is sufficiently fine-grained to allow recreation of the KG.

While this ability to recreate or reproduce should be standard for any scientific endeavor, to date, this is not a given with existing KGs or KG creation: One of the main points of critic

⁸ <https://www.gbif.org/what-is-gbif>

of the Google Knowledge Graph, e.g., is exactly its lack of provenance information. While DBPedia [Au07] and WikiData [VK14] require provenance information, they, too, do not quite achieve what we need here: On the one hand, despite all efforts, there are serious quality problems with the provenance information in WikiData. According to [Pi17] only 61% of the provenance information resolved to relevant and authoritative data sources. In our context even more important is a second drawback of, e.g., the WikiData provenance information: While this information tells *where* a piece of information stems from, it does not describe *how* it has been processed to be included in WikiData. With this, it is possible to understand and verify information but not possible to reproduce the KG creation.

Unfortunately, the same is true for many non-KG related biodiversity data integration endeavors. Like many other disciplines [Ba16], biodiversity research faces a reproducibility crisis [SKR21]. Despite growing awareness of the problem and a rising availability of data and code that publications are based upon, often, recreating published results is still not possible with the information provided. A frequent reason for this is, that, e.g., some data cleaning and preprocessing steps are not documented at all or not in sufficient detail.

There is a substantial body of work though on workflow and provenance management in general (e.g., [Pi19]), on the level of individual statements in a KG, in particular [HP18], and on tracking provenance in the workflows that lead to the creation of a KG, e.g. for Jupyter Notebooks [SKR18a] or in R, e.g., <https://cran.r-project.org/web/packages/rdtLite>.

Important input for our approach can also be found in attempts to create data and workflows following the FAIR principles [Wi16], that is making data and workflows findable, accessible, interoperable and reproducible. For the biodiversity domain, relevant proposals exist, e.g., from the collections community [LKH20, Ha20]. If these approaches gain traction, the data produced (and described) with them would be ideally suited as input to the KG creation proposed in iKNOW.

For creating a toolbox, that is supporting users in finding the right tools for the tasks at hand and combining them in the most efficient way, another field of research is relevant, namely semantic description, discovery and composition of software components. Providing such descriptions will enable dataowners (and a supporting system) to find and combine suitable tools to transform their data into KGs. Here, we can build on an extensive body of work (see [KEA18] for a recent review).

3 Biodiversity Datasources

Our work is motivated by the wealth of well-curated data sources available at iDiv, the German Centre for Integrative Biodiversity Research, that can be transformed into formalized KGs with limited effort.

Examples of such datasources are:

- TRY [Ka11], the plant trait database (founded in 2007), provides unprecedented plant trait data coverage under an open access data policy and is the main plant trait database used by the research community worldwide.
- The sPlot project [Br19] collates vegetation plots worldwide to explore global patterns of taxonomic, functional, and phylogenetic diversity at the plant community level.
- PhenObs [No21] ⁹ is a global network of botanical gardens, and aims to discover the impacts of climate change on the phenology of herbaceous plant species.

The examples already show the wide variety of datasources, but also obvious connections across them making it worthwhile to aim for integration. The PlantHub strategic project ¹⁰ aims to harmonize these data sources and leverage their joint power to answer important scientific questions, to showcase the fascinating variety of biodiversity to the public, and to integrate the data in national and international developments including NFDI4BioDiversity ¹¹. The iKNOW project will extend this work towards the creation of a toolbox for reproducible KG creation.

The TRY database stores functional traits of plant species. These traits evolved driven by environmental constraints and represent the adaptations of plants to their current environment. SPlot is a database of vegetation records, which in turn can be linked to environmental conditions. Therefore, joining those databases allows for inferences about local climate at sPlot vegetation record locations, but also on the relationship between plant functional traits and their environment on a global scale. PhenObs collects data on the phenology of herbaceous species, i.e. start of flowering or fruiting in botanical gardens together with their functional traits. Therewith this data source complements existing data from sPlot and TRY: PhenObs provides data on phenology and traits for plant populations measured in-situ, thereby including the aspect of intraspecific trait variability.

4 Proposed Approach

All of the databases described above are built by integrating and carefully curating data originating from a variety of heterogeneous sources. Therefore, all of the databases have well-established workflows and tools for data integration, cleaning, and harmonization. These will form the basis of our toolbox. We consider these processes in the pre-processing phase in our architecture.

Figure 1 shows the high-level view of the iKNOW procedure to provide a reproducible KG that can dynamically evolve over time. In a first step, data sources need to be pre-processed

⁹ <https://www.plantbiodiv.uni-jena.de/PhenObs>

¹⁰ <https://www.lw.uni-leipzig.de/en/institute-for-biology/abteilungen/systematic-botany-and-functional-biodiversity/research/big-data/>

¹¹ <https://www.nfdi4biodiversity.org/>

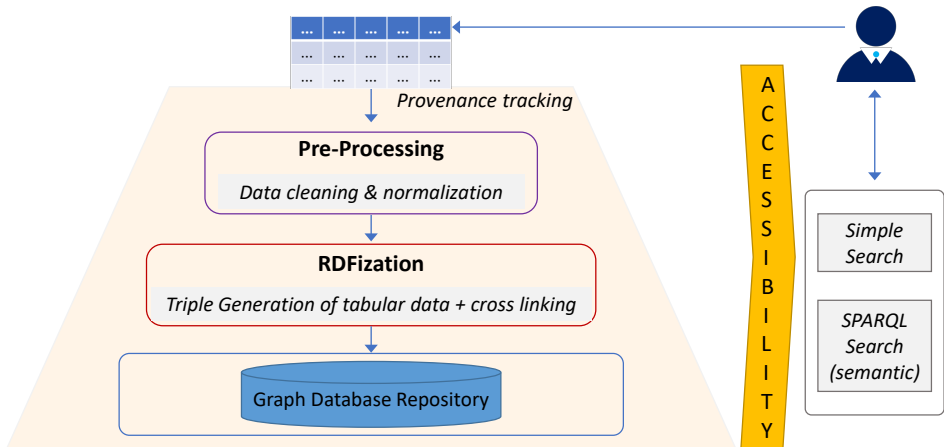


Fig. 1: High-level architecture of iKNOW.

to ensure that the data are of high-quality and harmonized. As mentioned above, in our toolbox, we will offer examples of tools and workflows used by the existing data sources with extensive human- (and in the long run also machine-) readable descriptions. We will also include other widely used tools in the toolbox. Owners of datasources that aim to create a KG from their source can explore the toolbox and adapt the existing workflows to their needs.

Once the data is cleaned, we will provide a set of tools to generate the RDF triples from the tabular data. Depending on the structure of the data and the amount of structured information it provides, different tools can be used in this step. We will aim to leverage and adapt existing tools for automatic mapping. In particular, we aim to adapt JenTab [AS20] a participant in the SemTab Challenge to the biodiversity domain. Since non of the mentioned biodiversity databases is a KG yet, we provide the cross-links to the domain-specific KGs. The output of this step will be stored and maintained in a graph database repository such as Blazegraph ¹².

The goal of iKNOW is to make the usage of KG easy for domain experts. Retrieving knowledge in our KG is possible via either the SPARQL endpoint or a free-text search functionality (for users with little or no prior knowledge of SPARQL).

Both the existing and the newly developed tools will be described in both human and machine understandable ways, so that discovery and (semi-)automatic composition will be possible in the future. We will ensure that these descriptions are compatible with those developed in the framework of NFDI4Biodiversity.

Provenance Tracking & Reproducibility: According to the provenance life-cycle [Mo08],

¹² <https://github.com/blazegraph/>

once the process and execution are documented and stored, the provenance of data results can be retrieved via querying over the provenance store. To this end, we plan to adopt one of the recent works on provenance capture in the R language (e.g., `rdtLite`¹³), or provenance tracking via Jupyter Notebooks [SKR18b] towards semantic provenance management. This has several advantages: First, by capturing the provenance of the source code, input, and the results, other scientists can reproduce the result. Second, the captured provenance allows to replay the workflow with altered input data (e.g., when new data becomes available in the data sources). Thus, keeping KGs up to date and synchronized with the sources becomes effortless. And finally, by making transparent how a KG was created (open-access code, data, result, and process), we increase trust in the information provided and support open data and open science.

5 Outlook

iKNOW is a planned semantic-based toolbox for the dynamic building Knowledge Graph in the biodiversity domain. It is a codebase for the automatic transformation of tabular biodiversity data into RDF statements. From the biodiversity research perspective, this transformation contributes to the creation of the Linked Open Data (LOD) dataset by interlinking to cross-domain and specific-domain KGs. The open accessibility of the source code and provenance tracking will ensure the reproducibility of all processes and results. Thus, this is an important steps towards strengthening the establishment of open science practices in the biodiversity domain. The generated KGs will contribute to a further understanding of biodiversity and its mechanisms.

From the computer science perspective, iKNOW will contribute to the development of tools for dynamic, low-effort creation of reproducible knowledge graphs.

Acknowledgments

The work described in this paper is conducted in the iKNOW Flexpool project of iDiv, the German Centre for Integrative Biodiversity Research, funded by DFG (Project number 202548816).

Bibliography

- [AS20] Abdelmageed, Nora; Schindler, Sirko; Jentab: Matching tabular data to knowledge graphs. In: The 19th International Semantic Web Conference (ISWC). 2020.
- [Au07] Auer, Sören; Bizer, Christian; Kobilarov, Georgi; Lehmann, Jens; Cyganiak, Richard; Ives, Zachary: Dbpedia: A nucleus for a web of open data. In: The semantic web, pp. 722–735. Springer, 2007.

¹³ <https://cran.r-project.org/web/packages/rdtLite>

- [Ba16] Baker, Monya: 1,500 scientists lift the lid on reproducibility. *Nature News*, 533(7604):452, 2016.
- [Br19] Bruelheide, Helge; Dengler, Jürgen; Jiménez-Alfaro, Borja; Purschke, Oliver; Hennekens, Stephan M; Chytrý, Milan; Pillar, Valério D; Jansen, Florian; Kattge, Jens; Sandel, Brody et al.: sPlot—A new tool for global vegetation analyses. *Journal of Vegetation Science*, 30(2):161–186, 2019.
- [Ch17] Chang, Liang; Zhu, Manli; Gu, Tianlong; Bin, Chenzhong; Qian, Junyan; Zhang, Ji: Knowledge graph embedding by dynamic translation. *IEEE access*, 5:20898–20907, 2017.
- [Gl14] Glimm, Birte; Horrocks, Ian; Motik, Boris; Stoilos, Giorgos; Wang, Zhe: Hermit: an OWL 2 reasoner. *Journal of Automated Reasoning*, 53(3):245–269, 2014.
- [Ha20] Harjes, Janno; Link, Anton; Weibulat, Tanja; Triebel, Dagmar; Rambold, Gerhard: FAIR digital objects in environmental and life sciences should comprise workflow operation design data and method information for repeatability of study setups and reproducibility of results. *Database*, 2020, 2020.
- [He17] He, He; Balakrishnan, Anusha; Eric, Mihail; Liang, Percy: Learning symmetric collaborative dialogue agents with dynamic knowledge graph embeddings. *arXiv preprint arXiv:1704.07130*, 2017.
- [HP18] Hertling, Sven; Paulheim, Heiko: Provision and usage of provenance data in the WebIsA-LOD Knowledge Graph. In: *CEUR Workshop Proceedings*. volume 2317. RWTH, p. 6, 2018.
- [Ji15] Ji, Guoliang; He, Shizhu; Xu, Liheng; Liu, Kang; Zhao, Jun: Knowledge graph embedding via dynamic mapping matrix. In: *Proceedings of the 53rd annual meeting of the association for computational linguistics and the 7th international joint conference on natural language processing (volume 1: Long papers)*. pp. 687–696, 2015.
- [Ji20] Jiménez-Ruiz, Ernesto; Hassanzadeh, Oktie; Efthymiou, Vasilis; Chen, Jiaoyan; Srinivas, Kavitha; Semtab 2019: Resources to benchmark tabular data to knowledge graph matching systems. In: *European Semantic Web Conference*. Springer, pp. 514–530, 2020.
- [Ka11] Kattge, Jens; Diaz, Sandra; Lavorel, Sandra; Prentice, I Colin; Leadley, Paul; Bönsch, Gerhard; Garnier, Eric; Westoby, Mark; Reich, Peter B; Wright, Ian J et al.: TRY—a global database of plant traits. *Global change biology*, 17(9):2905–2935, 2011.
- [KEA18] Kurniawan, Kabul; Ekaputra, Fajar J; Aryan, Peb R: Semantic service description and compositions: A systematic literature review. In: *2018 2nd International Conference on Informatics and Computational Sciences (ICICoS)*. IEEE, pp. 1–6, 2018.
- [La98] Lassila, Ora; Swick, Ralph R et al.: Resource description framework (RDF) model and syntax specification. 1998.
- [LKH20] Lannom, Larry; Koureas, Dimitris; Hardisty, Alex R: FAIR data and services in biodiversity science and geoscience. *Data Intelligence*, 2(1-2):122–130, 2020.
- [Mo08] Moreau, Luc; Groth, Paul; Miles, Simon; Vazquez-Salceda, Javier; Ibbotson, John; Jiang, Sheng; Munroe, Steve; Rana, Omer; Schreiber, Andreas; Tan, Victor et al.: The provenance of electronic data. *Communications of the ACM*, 51(4):52–58, 2008.

- [Ni15] Nickel, Maximilian; Murphy, Kevin; Tresp, Volker; Gabrilovich, Evgeniy: A review of relational machine learning for knowledge graphs. *Proceedings of the IEEE*, 104(1):11–33, 2015.
- [No21] Nordt, Birgit; Hensen, Isabell; Bucher, Solveig Franziska; Freiberg, Martin; Primack, Richard B; Stevens, Albert-Dieter; Bonn, Aletta; Wirth, Christian; Jakubka, Desiree; Plos, Carolin et al.: The PhenObs initiative: A standardised protocol for monitoring phenological responses to climate change using herbaceous plant species in botanical gardens. *Functional Ecology*, 35(4):821–834, 2021.
- [Pa16] Page, Roderic: Towards a biodiversity knowledge graph. *Research Ideas and Outcomes*, 2, 2016.
- [Pa19] Page, Roderic DM: Ozymandias: a biodiversity knowledge graph. *PeerJ*, 7:e6739, 2019.
- [Pe19] Penev, Lyubomir; Dimitrova, Mariya; Senderov, Viktor; Zhelezov, Georgi; Georgiev, Teodor; Stoev, Pavel; Simov, Kiril: Openbiodiv: a knowledge graph for literature-extracted linked open data in biodiversity science. *Publications*, 7(2):38, 2019.
- [Pi17] Piscopo, Alessandro; Kaffee, Lucie-Aimée; Phethean, Chris; Simperl, Elena: Provenance information in a collaborative knowledge graph: an evaluation of Wikidata external references. In: *International semantic web conference*. Springer, pp. 542–558, 2017.
- [Pi19] Pimentel, João Felipe; Freire, Juliana; Murta, Leonardo; Braganholo, Vanessa: A survey on collecting, managing, and analyzing provenance from scripts. *ACM Computing Surveys (CSUR)*, 52(3):1–38, 2019.
- [Sa19] Sachs, Joel; Page, Roderic; Baskauf, Steven J; Pender, Jocelyn; Lujan-Toro, Beatriz; Macklin, James; Comspon, Zacchaeus: Training and hackathon on building biodiversity knowledge graphs. *Research Ideas and Outcomes*, 5:e36152, 2019.
- [Se18] Senderov, Viktor; Simov, Kiril; Franz, Nico; Stoev, Pavel; Catapano, Terry; Agosti, Donat; Sautter, Guido; Morris, Robert A; Penev, Lyubomir: OpenBiodiv-O: ontology of the OpenBiodiv knowledge management system. *Journal of biomedical semantics*, 9(1):1–15, 2018.
- [Si07] Sirin, Evren; Parsia, Bijan; Grau, Bernardo Cuenca; Kalyanpur, Aditya; Katz, Yarden; Pellet: A practical owl-dl reasoner. *Journal of Web Semantics*, 5(2):51–53, 2007.
- [SKR18a] Samuel, Sheeba; König-Ries, Birgitta: ProvBook: Provenance-based Semantic Enrichment of Interactive Notebooks for Reproducibility. In: *International Semantic Web Conference (P&D/Industry/BlueSky)*. 2018.
- [SKR18b] Samuel, Sheeba; König-Ries, Birgitta: ProvBook: Provenance-based Semantic Enrichment of Interactive Notebooks for Reproducibility. In: *International Semantic Web Conference (P&D/Industry/BlueSky)*. 2018.
- [SKR21] Samuel, Sheeba; König-Ries, Birgitta: Understanding experiments and research practices for reproducibility: an exploratory study. *PeerJ*, 9:e11140, 2021.
- [VK14] Vrandečić, Denny; Krötzsch, Markus: Wikidata: a free collaborative knowledgebase. *Communications of the ACM*, 57(10):78–85, 2014.

- [Wi16] Wilkinson, Mark D; Dumontier, Michel; Aalbersberg, IJsbrand Jan; Appleton, Gabrielle; Axton, Myles; Baak, Arie; Blomberg, Niklas; Boiten, Jan-Willem; da Silva Santos, Luiz Bonino; Bourne, Philip E et al.: The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, 3(1):1–9, 2016.

Workshop: Energie- und Ressourceneffizienz von Rechenzentren (E2R21)

Softwarearchitektur für die Bewertung des Ressourcenbedarfes in Rechenzentren durch ganzheitliche Ressourceneffizienz

Fernando Penaherrera¹, Alexandra Pehlken², Ralph Hintemann³, Simon Hinterholzer⁴, und Björn Koch⁵


Abstract: Der steigende Energiebedarf der Rechenzentren stellt eine Herausforderung bei der Erreichung der Klimaschutzziele sowie der Realisierung der Energiewende dar. Auch die sogenannte graue Energie – also die Energie, die für Herstellung, Transport, Lagerung und Entsorgung der Komponenten benötigt wird – kann ganz erhebliche Anteile am Gesamtenergiebedarf ausmachen. Dieser Artikel stellt eine selbstentwickelte Modellstruktur vor, welche die graue Energie in Rechenzentren erstmalig berechnet.


Keywords: Softwarearchitektur; graue Energie; Life Cycle Assessment; Rechenzentrum; Materialeffizienz; Recycling

1 Einleitung

Rechenzentren sind aus dem Alltag nicht mehr wegzudenken. Streamingdienste, Automatisierung oder Echtzeitanalysen wie z.B. der Verkehrsdichte sind nur einige von vielen Themen, die wir heute selbstverständlich nutzen. Im Hintergrund laufen die Rechenzentren, die aufgrund der hohen Rechenleistung immer mehr mit ihrem Energiehunger zu kämpfen haben.

Wie eine Reihe von Studien ermittelt hat, sind die ca. 50.000 Rechenzentren [Hi17] mit mehr als 16 TWh für knapp 2,9% des Strombedarfs in Deutschland verantwortlich, obwohl die Effizienz der Rechenzentrumsinfrastrukturen um 21 % gestiegen ist [Hi20], [HH20], [BSB08]. Der steigende Energiebedarf von Rechenzentren wurde in zahlreichen Studien nicht nur für Deutschland, sondern auch für Europa, sowie weltweit dokumentiert [Hi14], [Pr14], [Co14]. Ein Grund hierfür ist insbesondere der stark steigende Bedarf,

¹ OFFIS – Institut für Informatik, Escherweg 2, 26121 Oldenburg, fernandoandres.penaherreravaca@offis.de,  <https://orcid.org/0000-0003-3726-0028>

² OFFIS – Institut für Informatik, Escherweg 2, 26121 Oldenburg, pehlken@offis.de,  <https://orcid.org/0000-0003-1798-8679>

³ Borderstep Institut für Innovation und Nachhaltigkeit, Clayallee 323, D-14169 Berlin, hintemann@borderstep.de,

⁴ Borderstep Institut für Innovation und Nachhaltigkeit, Clayallee 323, D-14169 Berlin, hinterholzer@borderstep.de,

⁵ COAST – Universität Oldenburg, Ammerländerheerstr. 114-118, 26129 Oldenburg, bjoern.koch@uol.de

große Datenmengen zu speichern und zu verarbeiten. Im Jahr 2020 betrug der Anteil des Energiebedarfs der Speichersysteme am Gesamtbedarf ca. 33% des Energiebedarfs der Server [Mo20]. Bei den im Konsortium von TEMPRO (Total Energy Management for Professional Data Centers), einem deutschen Projekt, beteiligten Rechenzentrumsbetreibern liegt der Energiebedarf der Speichersysteme aktuell bei ca. 30% des Gesamtstrombedarfs.

Der steigende Energiebedarf der Rechenzentren stellt eine Herausforderung bei der Erreichung der Klimaschutzziele sowie der Realisierung der Energiewende dar. Verstärkt wird diese Herausforderung noch durch zwei Aspekte. Zum einen ist insbesondere im Bereich der IKT-Produkte mit ihren sehr kurzen Lebenszyklen nicht nur der Energiebedarf in der Nutzungsphase relevant [Ga12]. Auch die sogenannte graue Energie – also die Energie, die für Herstellung, Transport, Lagerung und Entsorgung der Komponenten benötigt wird – kann ganz erhebliche Anteile am Gesamtenergiebedarf ausmachen [FH14]. Dieser Aspekt gewinnt zunehmend an Bedeutung, da die Materialintensität der Rechenzentren steigt. Verbunden mit einem deutlichen Trend zu größeren Rechenzentren ist ein erheblicher Zuwachs von IT-Infrastrukturen wie Server, Speicher und Netzwerkkomponenten zu verzeichnen. Allein die Zahl der Server hat zwischen 2010 und 2018 um 15% zugenommen und wird bis 2025 noch weiter ansteigen [Mo20]. Außerdem gibt es bei Rechenzentren vielfältige Möglichkeiten in die Energienetze der Zukunft integriert zu werden, wie z.B. Abwärmenutzung, Bereitstellung von Regelenergie durch Notstromaggregate, Demand Side Management, usw. Diese Möglichkeiten werden jedoch bis heute praktisch nicht genutzt.

Neben einem immer weiter ansteigenden Energieverbrauch rückt allerdings auch der Ressourcenhunger von Rechenzentren in den Vordergrund. Schon bevor ein Rechenzentrum eine kWh durch Datenverarbeitung oder Speicherung verbraucht hat, ist im Aufbau eines Rechenzentrums sehr viel graue Energie enthalten. Das Forschungsprojekt TEMPRO hat die graue Energie in Rechenzentrum zum ersten Mal empirisch ermittelt und analysiert [Pe20]. Dieser Beitrag baut auf den Erkenntnissen des Forschungsprojektes auf und stellt eine Software-Architektur vor, welche als Bewertungstool für die graue Energie in Rechenzentren dient.

Kapitel 2 zeigt eine Zusammenfassung der Zusammensetzung eines Rechenzentrums. Kapitel 3 beschreibt das vorgeschlagene Softwaretool zur Berechnung von Ressourcenbedarfen eines Rechenzentrums, basierend auf einer dreistufigen Architektur. Kapitel 4 zeigt Ergebnisse der durchgeführten Berechnungen. Kapitel 5 gibt einen Überblick über die Vorteile und Grenzen dieses Tools und stellt zukünftige Arbeiten vor.

2 Allgemeine Zusammensetzung von Rechenzentren

Zur Vergleichbarkeit der massenspezifischen rohstofflichen Zusammensetzung eines Rechenzentrums sollen zur Kategorisierung eines Rechenzentrums Faktoren der

physischen Größe angewendet werden. Die technische Betrachtung der Größe ist über die Zahl der Racks, Zahl der Server, Zahl der Prozessoren, des Stromverbrauchs, die Fläche, Rechenleistung, Speicherleistung oder der Art der Anwendung zu bestimmen [HFS10].

Die höchste Rohstoffrelevanz des Rechenzentrumsinventars liegt in den Geräten, die einen erhöhten Anteil kritischer Rohstoffe aufweisen. Der Anteil und die Vielfalt kritischer Rohstoffe sind vor allem im IKT-Equipment besonders hoch, dem funktionell bedeutendsten Bestandteil des Rechenzentrumsinventars. Dies ist vor allem auf den Anteil der verbauten elektronischen Schaltungen in Form von Leiterplatten zurückzuführen, der in Servern einen Anteil von etwa 15% ausmacht [Pe20]. Der Rohstoffeinsatz aus Computersteuerung und LCD-Bildschirmanzeigen in Geräten der Energieversorgung, Klimatisierung und weiterer Infrastruktur ist vernachlässigbar klein. Abb. 1 repräsentiert die Konfiguration der Komponenten eines Rechenzentrums.

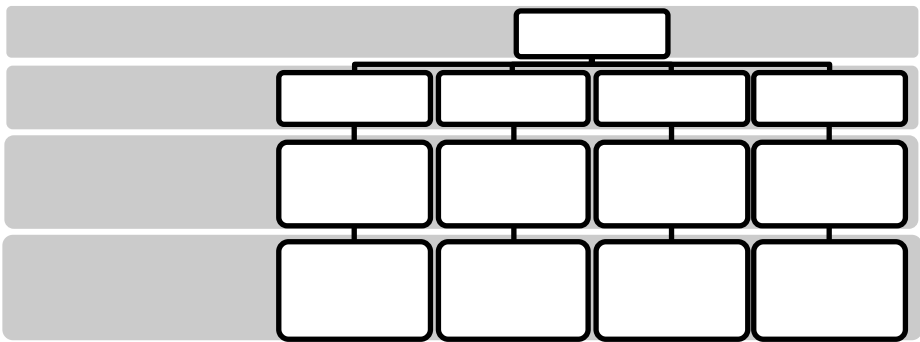


Abb. 1: Komponenten eines Professionellen Rechenzentrums für die Fallstudie.

Den größten Teil der Leiterplatten im IT-Equipment machen, mit einem Anteil von etwa 80 %, die Hauptplatinen (Mainboards) aus. Zudem werden in Servern Leiterplatten in Form von Adapterplatinen mit verschiedenen Anwendungen eingesetzt, die der Hauptplatinen zugeschaltet werden. Auf den Hauptplatinen der Server sind in der Regel Speichermodule und CPU(-Chips) angebracht. Server enthalten abhängig von ihrer Konfiguration weitere Geräte, die als Bauteile der Server betrachtet werden. Diese Geräte, Laufwerke und Netzteile enthalten ebenfalls Platinen, die eine hohe Konzentration an kritischen Materialien aufweisen [Pe20], [Sz18], [HFS10]. Zudem sind auch die Geräte der Netzwerktechnik, wie Switches mit Hauptplatinen ausgestattet. Die Zusammensetzung aller Platinen wird ausgehend von der Zusammensetzung von Hauptplatinen, wie sie in Stand-PCs (zumeist in Nutzung als Clients, seltener als Server) verbaut sind, festgelegt. Da die Platinen die Komponenten im Rechenzentrum darstellen, die die meisten kritischen Rohstoffe mit zum Teil hohen Energiebedarfen in der Gewinnung darstellen, liegt der Fokus dieses Ansatzes auf den Leiterplatten.

Mit Hilfe des Geräteclusters wird die anwendungsbezogene Zusammenstellung des Rechenzentrums-Equipments rohstofflich aufgeschlüsselt und so der Rohstoffbezug hergestellt. Zu beachten ist hierbei, dass die rohstoffliche Aufschlüsselung der Komponenten nur über das Recycling zu erfassen war, da keinerlei Herstellerdaten über

den Inhalt (also der Rohstoffe) verfügbar sind. Daher ist der Stand der Technik aus dem Recycling mit dem Stand der Rechenzentren von vor etwa 2-5 Jahren (je nach Lebensdauer der Komponenten) vergleichbar und schränkt das Modell etwas ein. Die Klimatisierung wurde lediglich über Luftkühlung betrachtet. Zur Differenzierung der individuellen Rechenzentren wird nach dem Hauptkriterium der Serveranzahl unterschieden. Das Cluster dient als Auswertungsgrundlage des Rohstoffpotenzials. Aus dem Untersuchungsmaterial geht das Gewicht der einzelnen Bauteilgruppen, Bauteile und Komponenten hervor. Aus der Datenanalyse geht der Rohstoffgehalt hervor und die damit einhergehenden Informationen über die Rohstofffakten dieser Teile.

3 Softwarearchitektur zur Ermittlung der grauen Energie in Rechenzentren

Ziel der Software-Architektur ist es ein Instrument zu erstellen, mit dessen Hilfe eine ganzheitliche Bewertung des gesamten Material- und Energiebedarfs für den Bau und Betrieb eines Rechenzentrums ermöglicht werden kann. Hierbei werden neben dem Energiebedarf in der Betriebsphase und dem Materialverbrauch sowie dem damit zusammenhängenden Primärenergieverbrauch für die Herstellung der Komponenten des Rechenzentrums ebenso die Einsparpotentiale durch gezieltes Recycling der Komponenten nach der Betriebsphase in die Berechnungen einbezogen. Um all diese Aspekte bewerten zu können, müssen zahlreiche Indikatoren berücksichtigt werden, die den gesamten Energie- und Materialbedarf widerspiegeln und die potenziellen Einsparungen aufgrund von Recycling angemessen darstellen können.

3.1 Methode Life Cycle Assessment und Datenverfügbarkeit

Lebenszyklusanalyse (LCA) ist eine wissenschaftliche Methode, um unter anderem die Erschöpfung natürlicher Ressourcen und ihre Auswirkungen auf das Ökosystem zu bewerten. Durch die Aufstellung der Komponenten werden Modelle erstellt, die Prozessketten von Produkten innerhalb eines Systems darstellen, welche von der Extraktion von Rohstoffen über die Herstellung, den Transport sowie den Betrieb und eine eventuelle Entsorgung reichen. Die Analyse der Flüsse in diesen Modellen ermöglicht die Beurteilung der Umweltauswirkungen [CWP11], [IS06].

In der Regel werden bei elektronischen Geräten selten eine Bill-of-Materials (BoM) vom Hersteller an den Nutzer übermittelt. Ebenso gibt es meistens eine sehr große Diskrepanz zwischen den Informationen, die aus dem BoM abgerufen oder vom Hersteller empfangen werden können. Daher erscheint eine präzise und belastbare LCA auf Grundlage dieser Daten in der Regel nicht möglich. Aus diesem Grund wurden bereits im Projekt TEMPRO eigene Materialanalyse von ausgewählten IKT-Produkten durch Demontage, Probenahme und Laboranalyse erstellt und in mehreren relationalen Datenbanken gesammelt [Pe20], [Sz18], [HFS10]. Diese Laborergebnisse ermöglichen die Beurteilung der

Materialzusammensetzung von Rechenzentrenkomponenten und sind einmalig in der Bewertung von Rechenzentren, da eigene Daten mit Hilfe eines professionellen und lizenzierten Recyclers generiert wurden. Zusätzlich wurden Informationen zu den einzelnen Rohstoffen, wie z.B. die Kritikalität, gesammelt und in getrennten Datenbanken gespeichert [Eu20]. Aktuelle Warenwirtschaftssysteme aus dem Betrieb haben keinen Bezug zum Rohstoffinhalt (z.B. Gramm Gold pro Leiterplatte) und haben auch keinen Bezug zur Kritikalität der Rohstoffe.

In dem vorgestellten Ansatz werden zusätzlich externe Datenbanken eingebunden, um noch bestehende Datenlücken zu schließen. So enthält z.B. dieecoinvent 3.4 Datenbank [We16] die benötigten Informationen zu bereits vorhandenen Prozessen, Komponenten, Energie- und Materialressourcen vorrangig aus dem Herstellungsprozess und der Nutzungsphase. Das Recycling von Serverhardware durch den Anbieter ist meist bereits in Ausschreibungen enthalten, und entfällt nicht zwingend auf den Endkunden. Deswegen wurde das Recycling der Komponenten über eine für diesen Fall spezialisierte Recyclinganlage betrachtet, die eine Monocharge Rechenzentrumschrott im Labor analysierte und uns die Daten bereitstellte. Diese Daten wurden in das Modell integriert. Das Recycling von sogenannten Elektroschrott ist derzeit auf europäischer Ebene gut dokumentiert (Stichwort WEEE) und in Kombination mit den Labordaten in das Modell eingeflossen.

Zur Berechnung der Umweltauswirkungen wurde die open-source Software openLCA ausgewählt, da der Zugriff und die Manipulation der verschiedenen Modellkomponenten über externe Skripts ermöglicht wird [Gr21]. Durch diese Möglichkeit über externe Skripte auf die einzelnen Modellkomponenten zugreifen zu können wird die Anpassung bestehender und die Erstellung neuer Modelle mit aktualisierten Informationen vereinfacht und somit eine schnellere Berechnung und Ergebnisdarstellung ermöglicht. Mit diesem Modell lassen sich zum Beispiel die bereits in der Gewinnung der Rohstoffe inklusive der Produktion der Bauteile emittierten CO₂-Emissionen anzeigen (graue Energie), bevor es überhaupt zum Betrieb kommt.

3.2 Vorgeschlagene Softwarearchitektur

Die vorgeschlagene Softwarearchitektur basiert auf einer klassischen dreistufigen Architektur (three-tier architecture), in der die drei Hauptebenen der Anwendung - die Präsentationsebene (der Benutzerschicht bzw. die GUI), die Anwendungslogik (Anwendungsschicht) und die persistente Speicherung von Daten (die Datenbankschicht) - voneinander getrennt werden, wie Abb. 2 zeigt. Diese Architektur bietet mehrere Vorteile: Zentraler Zugriff auf Datenbanken und Hardware, Zentralsicherheit und Zwischenspeicherung, reduzierter Netzwerkverkehr und Last auf Kundenkonsolen sowie Skalierbarkeit. In diesem Fall wird neben den bereits oben genannten Datenbanken in der Datenbankanwendung und in openLCA in der Anwendungsebene eine webbasierte Anwendung für die Eingabe von Inventarinformationen von Rechenzentren als Präsentationsebene implementiert. Abb. 2 schematisiert diesen Lösungsansatz.

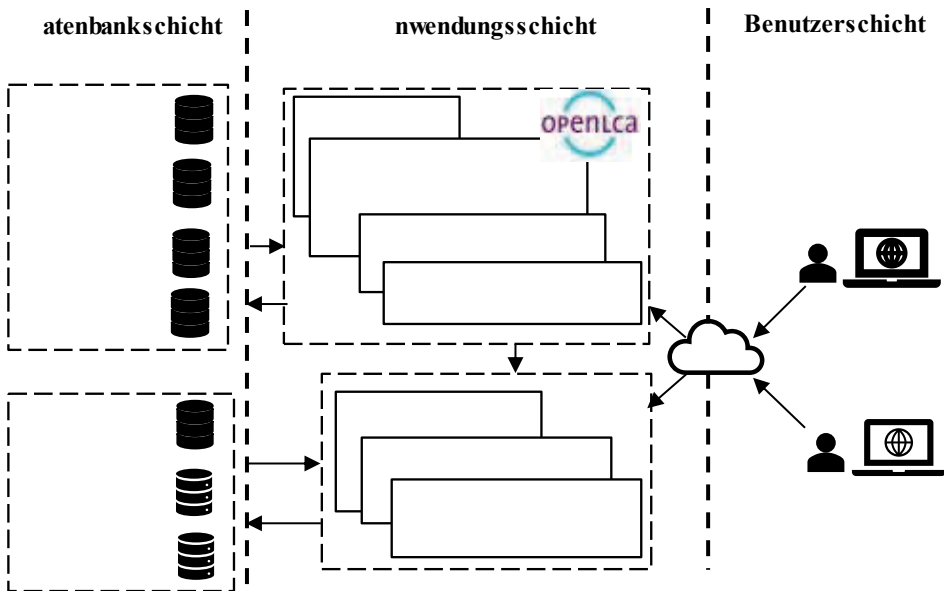


Abb. 2. Vorgeschlagene Softwarearchitektur.

In der Datenbankebene befinden sich verschiedene Datenbanken verteilt auf die für die Prozesse und deren Modellierung benötigte Daten, wie die BoMs der einzelnen Komponenten, die Inventarlisten einzelner Rechenzentren, Daten zu den Recyclingprozessen, generelle Rohstoffinformationen, die ecoinvent-Datenbank, Modelle aus dem TEMPRO Projekt und weitere Bewertungsmethoden.

Die Anwendungsebene besteht aus verschiedenen Modulen, die eine Modellerstellung, Berechnung und Speicherung von Ergebnissen ermöglichen, um die Berechnung von Umweltauswirkungen zur Bewertung von Bedarf von Material- und Energieressourcen durchzuführen. Die Modellierungsanwendungen greifen hierbei auf die verschiedenen Datenbanken in der Datenbankebene zu, um Produktsystemmodelle zu erstellen, welche dann mit Hilfe von openLCA analysiert werden sollen. Somit können die Umweltauswirkungen einzelner Produktsysteme unter Berücksichtigung verschiedener Berechnungsmethoden berechnet und auch analysiert werden. Es handelt sich hier nicht um eine herkömmliche LCA, da neue (eigene) Datenbanken eingebunden werden, die wesentlich auf die Rohstoffinhalte fokussieren. Gängige Rechenzentren verschiedener Größe lassen sich damit leichter über den modularen Aufbau abbilden,

Die Möglichkeit über ein Scripting dieser Modellierungsprozesse in openLCA ermöglicht eine Automatisierung der verschiedenen Schritte, so dass das gesamte Inventar eines Rechenzentrums für jede Komponente ohne manuelle Eingriffe analysiert werden kann. Elemente einer Datenbank (Flows, Prozesse, Produktsysteme, Berechnungsmethoden) werden als Objekte importiert und mit den Informationen in der Datenbankschicht

verknüpft. Dies ermöglicht die Manipulation von Objekten und Automatisierung der Erstellung von Modellen und Berechnungen.

Die Implementation der Präsentationsebene als Webanwendung ermöglicht dem Benutzer eine vereinfachte Interaktion mit dem System. So wird über eine grafische Ausgabe der Ergebnisse eine schnelle Bewertung des gesamtheitlichen Energie- und Ressourcenverbrauchs eines Rechenzentrums ermöglicht.

Generell ergibt sich die Möglichkeit, firmeneigene Inventarinformationen für benutzerspezifische Fallstudien einzuführen. Hierbei können Informationen verwendet

Bewertung der bekanntesten Indikatoren für Umweltauswirkungen wie Treibhausgasemissionen oder kumulativer Energieaufwand zur Verfügung stellen. Die häufigsten Key-Performance-Indikatoren für Rechenzentren, wie beispielsweise die Power-Usage-Effectiveness oder der Rechenzentren-Infrastruktureffizienz, können ebenfalls berücksichtigt werden, indem die Informationen aus vorhandenen Werkzeugen integriert werden.

4 Ergebnisse und Diskussion der Bewertung des Materialressourcenverbrauchs in Rechenzentren

Die in Abb. 2 beschriebene Architektur ermöglicht es, eine Analyse der Umweltauswirkungen von Rechenzentren und seiner Komponenten durchzuführen und stellt die Fallstudie dieser Veröffentlichung dar. Die Ergebnisse aus der in Kapitel 3 vorgestellten Softwarearchitektur wurden an einem realen laufenden professionellen Rechenzentrum mit Luftkühlung generiert, und mit einer maximalen Leistung von 100kW, in den entsprechenden Datenbanken berechnet und gespeichert.

Zusätzliche Berechnungsmethoden zum Einfluss der Materialkritikalität wurden integriert. [KPP19] präsentieren Methoden, um die Ressourcennutzung mit EU-Kritikalitätsaspekten zu bewerten [Eu20]. Diese Indikatoren zeigen die relative Bedeutung kritischer Rohstoffe auf den gesamten Verbrauch von Materialressourcen für verschiedene Komponenten. Diese Entwicklung zeigt die Skalierbarkeit der Architektur. Abb. 3 repräsentiert eine generelle Übersicht der Ergebnisse dieser Berechnungen aus der Fallstudie für Geräte in Rechenzentren.

Bei der Betrachtung der unterschiedlichen Teile eines Gerätes in einem Rechenzentrum muss der jeweilige Einfluss der verschiedenen Teile auf die Auswirkungen des Gesamtressourcenverbrauchs bewertet werden, um das Recyclingpotenzial zu bewerten. Die Integration von genaueren Recyclingmodellen für die Komponenten eines Rechenzentrums ermöglicht die Beurteilung der Vorteile des Recyclings. Das Recycling von vorrangig kritischen Rohstoffen kann den Druck auf die Lieferkette dieser Rohstoffe verringern und somit dann langfristig Lieferkettenprobleme vermeiden, da dort recycelt wird, wo die Komponenten anfallen und ebenso Bergbau der primären Rohstoffe vermieden werden kann.

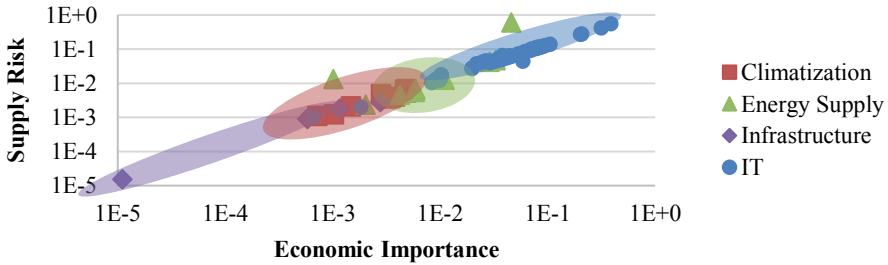


Abb. 3. Kritikalität gewichtete Erschöpfung für verschiedene Geräte eines Rechenzentrums. Zur Masse normiert.

Die resultierenden Berechnungen mit Hilfe der vorgestellten Architektur erlauben eine Abschätzung des Materialgehalts und des kumulierten Energiebedarfs für ein Rechenzentrum auf Basis der Bestandsinformationen. Abb. 4 zeigt daher die Komponenten eines Rechenzentrums auf, welche den meisten Energieaufwand generieren und kann daher Verbesserungspotentiale anzeigen.

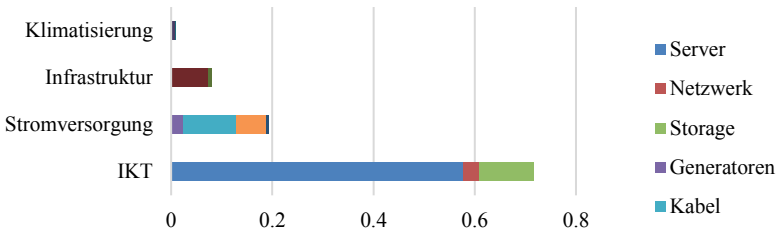


Abb. 4. Kumulierte Energieaufwand (MJ-eq) eines Rechenzentrums, normiert.

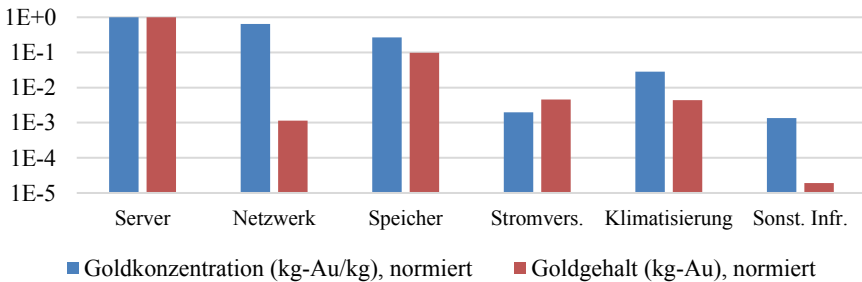


Abb. 5. Ergebnisse für Gehalt von Gold in einem Rechenzentrum.

Durch das Wissen über den spezifischen Goldgehalt in verschiedenen Produktgruppen wie in Abb. 5 gezeigt, kann man beim Recycling gezielt und strategisch die Komponenten für das Recycling aufkonzentrieren, die einen wertvollen Rohstoff (hier als Beispiel Gold) enthalten und unterstützen daher ein produktzentrisches Recycling. Durch die gute Datenbasis aus der eigenen Analyse der Rechenzentrenkomponenten aus der

Recyclinganlage sind hier alle Rohstoffe erfassbar, die derzeit in der Recyclingbranche dieser Rohstoffe kann hierdurch sowohl absolut wie in Abb. 5 als auch quantitativ dargestellt werden.

5 Ausblick

Mit dieser vorgeschlagenen selbstentwickelten Architektur ist es durch den modularen Aufbau möglich, Modelle für eine Vielzahl von Rechenzentrumskonfigurationen zu erstellen. Basierend auf den Ergebnissen aus der Analyse der Materialzusammensetzung von Geräten aus dem Projekt TEMPRO, können die unterschiedlichen Umweltauswirkungen eines Rechenzentrums schnell erfasst werden. In Zukunft können auch weitere Informationen zu verschiedenen Rechenzentrumskonfigurationen gesammelt und die vorhandenen Modelle in den Datenbanken verbessert werden. Ziel ist es, Rechenzentrumsbesitzern und politischen Entscheidungsträgern ein Entscheidungstool zur Verfügung zu stellen, welches den Energie- und Materialbedarf von Rechenzentrumskomponenten bewertet und Informationen über kritische Materialflüsse während der Lebensdauer dieser Komponenten liefert. Durch die Lebensdauerbetrachtung kann auch vorhergesagt werden, wieviel und welche Rohstoffe im Jahr x anfallen können. Somit wird der Rohstoffanteil aus Recyclingmaterialien vorhersagbarer, was sich positiv auf den Lieferkettendruck einzelner Rohstoffe auswirken kann.

Angesichts der kurzen Lebensdauer von IKT-Geräten in einem Rechenzentrum, welche in der Regel auf 2-5 Jahre ausgelegt ist, könnten die Informationen auf dem Gerät schnell veraltet sein. Die Einspeisung der Informationen in Datenbanken und der Entwicklung von Anwendungen für ein schnelles Update der Modelle, erleichtert die Aktualisierung von Informationen, sodass die Qualität der Ergebnisse verbessert wird.

Open-Source Software ermöglicht die Zugänglichkeit solcher Tools, allerdings ist eine ordnungsgemäße Dokumentation erforderlich, um den Strukturen der mit dieser Software entwickelten Modelle zu folgen. Es besteht die zusätzliche Notwendigkeit des Zugriffs auf einen Webserver zur Umsetzung der Ergebnisvisualisierung, und weitere Webdienste sind für die Interaktion mit den Benutzern erforderlich. Mit dem Stand 2021 ist dies noch nicht umgesetzt, aber es ist zu erwarten, dass ähnliche Ansätze in naher Zukunft folgen werden.

Diese Fallstudie kann derzeit noch nicht den Stand der Technik in der aktuellen Rechenzentrumsbranche darstellen, da das Recycling dem Stand der Technik aufgrund der Lebensdauer der Komponenten immer hinterherhinkt. Wünschenswert wäre eine Übergabe der Bill-of-Materials einzelner Komponenten direkt an den Recycler. Dies würde die Modelle erheblich verbessern, ist aber derzeit noch nicht gängige Praxis. Auch fokussiert die Klimatisierung derzeit nur die Luftkühlung und nicht die Wasserkühlung und schränkt das Modell somit noch weiter ein.

Weitere Arbeiten fokussieren sich auf die ordnungsgemäße Entwicklung einer Webanwendung, die mit den Benutzern und den Ergebnisdatenbanken ordnungsgemäß

interagiert, sodass die Informationen visualisiert werden können. Auch ein Reporting auf EU-Ebene kann mit diesem Modell in Zukunft ermöglicht werden.

6 Literaturverzeichnis

- [BSB08] Behrendt, F.; Schafer, M.; Belusa, T.: Konzeptstudie zur Energie- und Ressourceneffizienz im Betrieb von Rechenzentren. In IZE TU Berlin, Universitätsverlag, Berlin, 2008.
- [Co14] Cook, G. et al.: Clicking clean: how companies are creating the green internet. In Greenpeace Inc., Washington, DC, 2014; S. 19.
- [CWP11] Chomkamsri, K.; Wolf, M.-A.; Pant, R.: International reference life cycle data system (ILCD) handbook: Review schemes for life cycle assessment. In Towards life cycle sustainability management, 2011; S. 107–117.
- [Eu20] European Commission: *Cloud Computing* 2020
- [FH14] Fichter, K.; Hintemann, R.: Beyond Energy: The Quantities of Materials Present in the Equipment of Data Centers. In Journal of Industrial Ecology, 2014, 18; S. 846–858.
- [Ga12] Garnier, C. et al.: Data centre life cycle assessment guidelines. In The Green Grid, white paper, 2012, 45; v2.
- [Gr21] GreenDelta: openLCA. GreenDelta, 2021.
- [HFS10] Hintemann, R.; Fichter, K.; Stobbe, L.: Materialbestand der Rechenzentren in Deutschland - Eine Bestandsaufnahme zur Ermittlung von Ressourcen- und Energieeinsatz, 2010.
- [HH20] Hintemann, R.; Hinterholzer, S.: Rechenzentren in Europa – Chancen für eine Nachhaltige Digitalisierung, Berlin, 2020.
- [Hi14] Hintemann, R.: The Future of Data Center Energy Demand The Impact of the Changing Structure of Data Centers. Unpublished, 2014.
- [Hi17] Hintemann, R.: 3 Internetgestützte Innovationsprozesse, 2017.
- [Hi20] Hintemann, R.: Energiebedarf der Rechenzentren steigt trotz Corona weiter an, 2020.
- [IS06] ISO: ISO 10668:2006 *Energy-efficient Data Centers – Requirements and Guidelines*, 2006.
- [KPP19] Koch, B.; Peñaherrera, F.; Pehlken, A.: Criticality and LCA – Building comparison values to show the impact of criticality on LCA. In European Journal of Sustainable Development, 2019, 8; S. 304.
- [Mo20] Montevocchi, F. et al.: Energy-efficient Cloud Computing Technologies and Policies for an Eco-friendly Cloud Market. Final Study Report, Vienna, 2020.
- [Pe20] Pehlken, A. et al.: Abschlussbericht Verbundprojekt TEMPRO, 2020.
- [Pr14] Prakash, S. et al.: Study on the practical application of the new framework methodology for measuring the environmental impact of ICT-cost/benefit analysis. In European Commission, Brussels, Studie, 2014.
- [Sz18] Szczepaniak, K.: Das Rohstoffpotenzial von Rechenzentren: Quantifizierung kritischer Rohstoffe im Rechenzentrumsinventar zur Ermittlung des Grauen-Energie-Verbrauchs. Masterarbeit TUHH, Hamburg, Germany, 2018.
- [We16] Wernet, G. et al.: The ecoinvent database version 3 (part I): overview and methodology. In The International Journal of Life Cycle Assessment, 2016, 21; S. 1218–1230.

Mehr Leistung, mehr Strom – und wie weiter?

Was heute im Rechenzentrum schon machbar ist und morgen möglich wird

Karsten Kutzer¹

Abstract: Rechenzentren werden in den kommenden Jahren immer mehr im Spannungsfeld zweier Entwicklungen stehen: Auf der einen Seite werden die Anforderungen an Energieeffizienz und die Suche nach Einsparpotentialen - bedingt durch den Klimawandel und damit zusammenhängende regulatorische Veränderungen und steigende Energiekosten - immer wichtiger. Auf der anderen Seite nähert sich die Halbleiterfertigung immer mehr den physikalischen Grenzen an und vergrößert dazu massiv die Anzahl von Transistoren auf den Chips. Das führt zu stark wachsendem Stromverbrauch von Systemen bei gleichzeitiger Erhöhung der Energiedichte. Dieser Vortrag beschreibt Möglichkeiten der IT-Infrastruktur, um diesen Herausforderungen zu begegnen, und die Auswirkungen auf die Planung und den Betrieb von Rechenzentren. An konkreten Beispielen werden die Machbarkeit heute aufgezeigt und „Best Practices“ abgeleitet. Darüber hinaus werden die notwendige Integration innerhalb der IT und in das Rechenzentrum beleuchtet und wichtige Trends für die ganzheitliche Planung heute und den Betrieb morgen erläutert.

Keywords: Rechenzentrum, Energieeffizienz, Kühlungsinfrastruktur, Wasserkühlung.

1 Kühlung von Serversystemen

Jedes Watt an Strom, das ein Server verbraucht, wird in Wärme umgesetzt und muss gekühlt werden. Das ist ein zunehmend essentieller Punkt im Systemdesign.

1.1 Steigender Strombedarf

Der Strombedarf von Serversystemen ist in den letzten Jahren stark gewachsen. Vor 10 Jahren verbrauchte ein zwei-Steckel Server noch maximal um die 400 Watt, heute planen wir mit knapp 800 Watt. Verantwortlich für den stark gestiegenen Strombedarf ist vor allem die enorm angestiegene „Thermal Design Power“ (TDP) der Prozessoren. Die in den letzten Jahren immer öfter für spezielle Rechenaufgaben eingesetzten Graphical Processing Units (GPUs) tragen mit einer zusätzlichen TDP von bereits heute jeweils bis zu 400 Watt dazu bei, dass sich diese Entwicklung auch in Zukunft nochmal weiter beschleunigen wird.

¹ Lenovo Global Technology Germany GmbH, ISG, Meitnerstraße 9 , 70563 Stuttgart, kkutzer@lenovo.com

1.2 Folgen für die Kühlungsinfrastruktur

Immer leistungsstärkere Lüfter, deren Strombedarf inzwischen bei einem Server unter Last leicht 80 Watt erreicht, übertragen die Wärme über großflächige Kühlkörper in die Luft des Rechenzentrums. Aus der Luft wird die Wärmeenergie über Klimageräte in Wasser abgegeben. Dieses Wasser muss vorher über Kompressionskälteanlagen abgekühlt werden. Dieser gesamte Kühlungsprozess ist sehr energieintensiv und die Spirale immer höherer Stromverbräuche setzt sich so in der Kühlungsinfrastruktur des Rechenzentrums fort. Auch an dieser Stelle steigt somit der Stromverbrauch massiv an.

Es gibt verschiedene Ansätze, den Kühlungsprozess zu optimieren. Der Einsatz von Rack-basierten Kühlsystemen zum Beispiel vermeidet die Erwärmung der Luft im Rechenzentrum, benötigt allerdings immer noch heruntergekühltes Wasser. Noch effizienter ist es, einen möglichst großen Anteil der Wärme direkt an ein besser wärmeleitendes Medium abzuführen, wofür es unterschiedliche Ansätze gibt – von Kühlung im Ölbad bis hin zur direkten Wasserkühlung des kompletten Servers.

Die direkte Warmwasserkühlung hat sich inzwischen etabliert und ist beispielsweise am Leibniz Rechenzentrum (LRZ) in Garching bei München seit mehr als 10 Jahren bei mehreren tausend Servern sehr erfolgreich im Einsatz. Ein großer Vorteil dieses Ansatzes ist, dass ein großer Anteil der Wärme an das Wasser abgeführt wird und auch sehr warmes Wasser ($> 45^{\circ}\text{C}$) zur Kühlung eingesetzt werden kann. Das eröffnet weitere Energieeinsparpotentiale auf Ebene des Rechenzentrums durch „Freie Kühlung“ (bei der das Herunterkühlen von Kühlwasser entfällt) oder die Weiternutzung der Wärmeenergie, was nachhaltig die Energiebilanz von Rechenzentren weiter verbessern kann.

2 Konsequenzen für Rechenzentren

Energieeffizienz ist ein immer wichtigeres Ziel beim Betrieb von Rechenzentren. Die Optimierung der Kühlungsinfrastruktur ist dabei die größte Stellschraube und eine direkte Warmwasserkühlung von Servern der etablierte und erfolgsversprechende Ansatz. Folgende Empfehlungen ergeben sich daraus für den Bau und Betrieb von Rechenzentren:

- Neben dem traditionellen, mit Kompressionskälte betriebenen Kühlkreislauf sollte im Rechenzentrum auch ein Warmwasserkreislauf umgesetzt werden. Für den Warmwasserkreislauf sollte mit Temperaturen geplant werden, die eine Weiternutzung der Wärmeenergie oder zumindest „Freie Kühlung“ erlauben.
- Um Ansätze für die Weiternutzung von Abwärme eines Rechenzentrums bei Neu- oder Umbaus zu erkennen und zu optimieren ist die enge interdisziplinäre Zusammenarbeit von Fachleuten aus der System- und Gebäudetechnik schon in einer sehr frühen Planungsphase erforderlich. Für den energieeffizienten Betrieb ist eine tiefgehende Integration von Rechner- und Gebäudeleittechnik eine wichtige Voraussetzung.

Nachhaltigkeit und Abwärmenutzung an HPC-Zentren

Norbert Conrad¹, Michael Resch²

Abstract: Managementsysteme helfen Rechenzentren, Ziele wie die Minimierung von Umweltauswirkungen zu erreichen. Abwärmenutzung trägt dabei wesentlich zur Verbesserung der Ökobilanz von Rechenzentren bei.

Keywords: Abwärmenutzung; HPC; Managementsysteme; Nachhaltigkeit

Die aktuelle Nachhaltigkeits- und Klimadiskussion betrifft HPC-Zentren, da sie einen erheblichen Energiebedarf haben. Aufgabe von Managementsystemen ist es, bestimmte Ziele einer Organisation durch Maßnahmen, Organisationsstrukturen und Prozessabläufe in einem kontinuierlichen Verbesserungsprozess (KVP) zielgerichtet zu erreichen. Neben dem bekanntesten standardisierten Managementsystem für Qualität ISO 9001 gibt es auch das Umweltmanagementsystem ISO 14001 und Energiemanagementsystem ISO 50001. Das Europäische Umweltmanagementsystem EMAS [Um21] geht über die ISO 14001 noch hinaus und beinhaltet u.a. auch Teile eines Energiemanagementsystems. Der Blaue Engel ist ein Umweltlabel für Produkte und Dienstleistungen. Anders als Managementsysteme macht er keine Vorgaben über das „wie“, sondern nur zum „was“. Da der Blaue Engel produktspezifische Anforderungen hat, können diese wesentlich konkreter sein, als die abstrakten Ziele der Managementsysteme, die die jeweilige Organisation selbst konkretisieren muss. Der für Rechenzentren relevante „Blauer Engel für energieeffizienten Rechenzentrumsbetrieb“ DE UZ-161 [RA21] verlangt unter anderem die Nutzung von Ökostrom, Maximalwerte für die Power Usage Effectiveness PUE und begrenzte Leistungsaufnahme von Servern im Idle-Zustand.

Die korrekte Umsetzung der Systeme kann von einem unabhängigen Gutachter überprüft und zertifiziert werden. Das HLRS ist seit 2019 nach EMAS, ISO 14001 und ISO 50001 zertifiziert. Den Blauen Engel für energieeffizienten Rechenzentrumsbetrieb hat das HLRS 2020 erhalten. Zertifizierungen wirken nach innen (Bestätigung, Motivation) und nach außen (Ansehen, Vorbild). Die Außenwirkung des Blauen Engels geht aufgrund seines sehr hohen Bekanntheitsgrades weit über die der Zertifikate von Managementsystemen hinaus.

Die Managementsysteme umfassen alle Bereiche nachhaltigen Handelns. Lebenswegbetrachtungen sind aufgrund der Datenlage bei Rechnern bisher nicht möglich. Der Energiebedarf der Rechner wird über die Betrachtung der Total Cost of Ownership (TCO) bei der Beschaffung minimiert. Beim Energiebedarf der Infrastruktur eines Zentrums ist die Kühlung

¹ HLRS der Universität Stuttgart, Nobelstraße 19, 70569 Stuttgart, Deutschland, conrad@hlrs.de

² HLRS der Universität Stuttgart, Nobelstraße 19, 70569 Stuttgart, Deutschland, resch@hlrs.de

entscheidend. Neben hocheffizienten Kühlsystemen ist die Nutzung der Abwärme, z.B. zur Heizung, eine wichtige Möglichkeit, die Umweltauswirkungen zu minimieren.

Für die Abwärmenutzung sind drei Größen entscheidend: die Leistung, das Temperaturniveau und die Entfernung zum Nutzer der Abwärme. Die Abwärmemengen von HPC-Zentren gehen oft über den Heizbedarf der eigenen Gebäude hinaus. Dann sind externe Verbraucher zu suchen und mit Fernwärmeleitungen anzubinden. Wenn Fernwärmeleitungen neu errichtet werden müssen, spielen die Kosten dafür oft eine dominierende Rolle für die Wirtschaftlichkeit. Sofern die nötige Heiztemperatur höher als die Kühlwassertemperatur ist, sind Wärmepumpen erforderlich, die Investitions- und Betriebskosten verursachen.

Ein weiterer Aspekt ist die Verfügbarkeit von Wärme-Quellen und -Senken. Rechenzentren können die Abwärme nicht ununterbrochen 7 x 24 h zur Verfügung stellen. Unterbrechungen wegen Wartungsarbeiten oder Ausfällen müssen berücksichtigt werden. Dies wird in der Regel eine Ersatzheizung nötig machen. Da Heizenergie nur in der kalten Jahreszeit benötigt wird, muss trotz Abwärmenutzung ein Kühlsystem vorhanden sein, das die gesamte Abwärme auch im Sommer dauerhaft an die Umgebung abgeben kann.

Schließlich ist noch der Zeithorizont von Investitionsentscheidungen zu berücksichtigen. Gebäude und ihre Infrastruktur sind auf Jahrzehnte angelegt. Speziell HPC wird zum großen Teil in öffentlichen Forschungseinrichtungen betrieben und ist nicht selten projektfinanziert. Die führt zu weiteren Risiken für langfristige Investitionen. Sofern die Abwärme außerhalb der eigenen Institution (z.B. Universität) genutzt werden soll, bringt der Verkauf von Wärme weitere organisatorische und rechtliche Herausforderungen.

Das HLRS betreibt derzeit einen HPC-Rechner mit ca. 3 MW Leistung. Die Kühlung erfolgt so weit wie möglich mit freier Kühlung, der Rest mit Kompressionskälte. Maximal 100 kW werden im Winter zur Gebäudeheizung (Büros) des HLRS mit einer Wärmepumpe verwendet. Ab 2026 ist ein neues HPC-System mit ca. 10 MW und ein neues RZ-Gebäude geplant. Die Abwärme soll zur Heizung des Campus Vaihingen der Universität Stuttgart verwendet werden. Hierzu beginnen die Ende 2021 die Untersuchungen möglicher Konzepte.

Zusammenfassend lässt sich feststellen, dass das am HLRS eingeführte Umweltmanagementsystem die Nachhaltigkeitsbestrebungen systematisiert und zu einem integrierten Bestandteil des Zentrums gemacht hat. Die Nutzung der Rechnerabwärme ist mit sehr vielen Herausforderungen im Detail verbunden. Sie ist wegen der möglichen Einsparungen an Energie und klimaschädlichen Emissionen trotzdem ein sehr lohnendes Ziel.

Literaturverzeichnis

[RA21] Blauer Engel für energieeffizienten Rechenzentrumsbetrieb. <https://www.blauer-engel.de/de/produktwelt/elektrogeraete/rechenzentren>, abgerufen 2021-07-27.

[Um21] EMAS. <https://www.emas.de>, abgerufen 2021-07-27.

Abwärmenutzung eines warmwassergekühlten Hochleistungsrechners mit einer Wärmepumpe

Frederik Feike,¹ Frank Dammel,¹ Peter Stephan,¹ Andreas Wolf², Christian Bischof²

Keywords: Warmwasserkühlung; Abwärmenutzung; Hochleistungsrechner; Fernwärme

Zusammenfassung

Der fortschreitende Ausbau von Rechenzentrumsinfrastruktur in Deutschland führt zu einem rasanten Anstieg der in diesem Bereich benötigten elektrischen Energie. Die verbrauchte elektrische Energie wird dabei nahezu vollständig in Wärme umgesetzt, welche durch den Einsatz von zusätzlicher Energie an die Umgebung abgeführt werden muss. Stattdessen sollte diese Wärme als Ressource angesehen werden, mit der umliegende Gebäude beheizt werden können.

An der TU Darmstadt wurde ein Konzept entwickelt, die Abwärme des "Lichtenberg II"-Hochleistungsrechners zu nutzen, indem sie in den Rücklauf des Fernwärmenetzes eingespeist wird ([1], siehe Abb. 1). Durch die installierte Warmwasserkühlung steht die Abwärme bei einer Temperatur von 45°C zu Verfügung. Über eine Wärmepumpe wird das Temperaturniveau angehoben auf 60°C bis 75°C, sodass die Abwärme über einen Wärmeübertrager in das Fernwärmenetz eingespeist werden kann. Wenn kein Wärmebedarf im Netz vorhanden ist, kann die Abwärme das gesamte Jahr aufgrund der hohen Kühlwassertemperaturen mittels freier Kühlung an die Umgebung abgeführt werden. Neben den eingesparten fossilen Brennstoffen im Heizwerk des Fernwärmenetzes wird somit auch elektrische Energie einer ansonsten erforderlichen Kompressionskältemaschine eingespart.

Dieses Konzept wurde umgesetzt und ist Anfang 2021 in Betrieb gegangen. Dabei wurde zur Analyse und Optimierung des Betriebs ein Messkonzept implementiert, welches an den relevanten Stellen Wärmeströme und Temperaturen, sowie die elektrische Leistung der Wärmepumpe misst.

Durch die hohe Temperatur der Rechnerabwärme konnte ein effizienter Betrieb der Wärmepumpe mit einer mittleren Leistungszahl von 5 erreicht werden. Zur Verbesserung

¹ Technische Universität Darmstadt, Institut für Technische Thermodynamik, Alarich-Weiß Straße 10, 64287 Darmstadt, feike@ttd.tu-darmstadt.de, dammel@ttd.tu-darmstadt.de, pstephan@ttd.tu-darmstadt.de

² Technische Universität Darmstadt, Hochschulrechenzentrum, Alexanderstr. 2, 64283 Darmstadt, andreas.wolf@tu-darmstadt.de, christian.bischof@tu-darmstadt.de

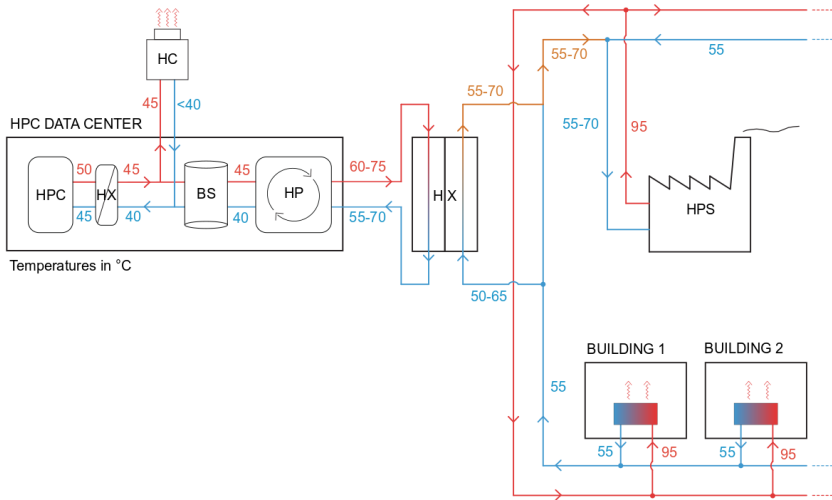


Abb. 1: Schema der Hochleistungsrechner-Abwärmenutzung [1]

der Effizienz wird nach weiteren Optimierungspotentialen gesucht. Beispielsweise wird gemeinsam mit dem Fernwärme-Betreiber eine Erhöhung des Massenstroms im Netz an der Einspeisestelle angestrebt. Darüber hinaus soll gemeinsam mit dem Hersteller des Hochleistungsrechners die Ausbeute der direkten Warmwasserkühlung erhöht werden.

In der Konzeptentwicklung wurden Energie- und CO₂-Einsparungen berechnet, die durch die bisher vorhandenen Messdaten als realistisch eingeschätzt werden können. So ist von einer jährlichen Verringerung des Erdgasbedarfs für die Wärmeerzeugung von ca. 2200 MWh sowie einer resultierenden Reduktion von ca. 700t CO₂-Emissionen auszugehen.

Diese Aktivitäten zur Energieoptimierung sind Teil des BMWi-geförderten Projekts "En-Eff:Stadt Campus Lichtwiese"³, das eine energieeffiziente Weiterentwicklung des Campus Lichtwiese der TU Darmstadt durch intelligente Systemvernetzung verfolgt. Dabei arbeiten die Projektpartner - entsprechend ausgewiesene Fachgebiete mit dem Dezernat "Baumanagement und Technischer Betrieb"- in enger Kooperation mit dem Hochschulrechenzentrum zusammen.

Literatur

- [1] F. Dammel; J. Oltmanns; P. Stephan. „Nutzung der Abwärme des Hochleistungsrechners am Campus Lichtwiese der TU Darmstadt“. In: *EuroHeat&Power* (48. Jg (2019), Heft 1-2 2019). Hrsg. von VDE VERLAG GMBH, S. 28–31.

³ https://www.intern.tu-darmstadt.de/verwaltung/dez_v/vf_technischer_betrieb/energiemanagement/eneff_stadt_campus_lichtwiese/index.de.jsp

VSC-3 – a summary of the experiences with an HPC system immersed in oil

Jan Zabloudil¹, Ernst Haunschmid², Herbert Störi³

Abstract: The electric energy consumed by computer systems is fully converted into heat, which needs to be removed from the system. As the internal temperature of processor chips is typically 80°C in operation and the temperature of the outside air hardly exceeds 45°C, a physicist would not easily understand, why heat pumps, consuming extra energy, are needed to remove heat to the outside. Traditional computer centers nevertheless consume copious amounts of extra electric energy to operate heat pumps. Consequently, various options for the efficient cooling of computers have been considered for several decades. Immersion cooling is an intriguingly simple option. You just find the right fluid and submerge the whole system in it. Then you only cool the fluid, which should be simple. There are, however, challenges to be met, which will be discussed.

After several years of experiments with a small system, VSC-3 was procured as a full-scale immersion cooled system. Already at installation, several issues needed to be addressed, mostly concerning the compatibility of various submerged components with the fluid. Nevertheless, with some adaptations, the system got up and running. Due to the very simple hardware surrounding the nodes, the system had a very good price/performance ratio and, from a point of view of energy efficiency, the whole installation surpassed the expectations with an average PoE of 1.03. Additionally, roughly 10% of the energy consumption of a conventional system were saved by the absence of cooling fans.

There were, however, serious downsides. First and foremost, the maintenance of VSC-3 was challenging to say the least. While compute nodes had hardly any problem, many other components suffered from the prolonged contact with hot fluid. Cables became stiff and some wicked oil outside the containers. Legions of InfiniBand switches failed for reasons, that will probably never be fully understood. Also, the containers and the oil installation were prototypes, leading to leakage.

Ultimately, the question of the future of immersion cooled systems remains. Let's assume, engineering problems can be solved without destroying the economic appeal of immersion cooling. The energy efficiency aspect will become more important in a greener future. The main problem, which remains to be solved, is the compatibility of computing components with the immersion fluid.

Looking at VSC-3 as our main system for several years, we must however say, it served us very well. During it's 6.5 years it delivered more than 1.25 billion core-hours to more than 700 projects.

Keywords: HPC; Immersion Cooling; Energy Efficiency

¹ University of Natural Resources and Life Sciences & VSC Research Center TU Wien, Vienna, Austria
jan.zabloudil@boku.ac.at

² VSC Research Center & TU.it, TU Wien, Vienna, Austria ernst.haunschmid@tuwien.ac.at

³ VSC Research Center, TU Wien, Vienna, Austria herbert.stoeri@tuwien.ac.at

Workshop:
Umweltinformationssysteme – Ein
Beitrag zur digitalen nachhaltigen
Transformation der Gesellschaft
(UISDigi-Trans)

Der UBA-Umweltatlas als Werkzeug des Umweltbundesamtes für die Öffentlichkeitsinformation

Das Beispiel „Reaktiver Stickstoff“

Stefan von Andrian-Werburg¹, Maria Theuring²

Abstract: Der UBA-Umweltatlas ist eines neues digitales Werkzeug des Umweltbundesamtes (UBA) für die Information der Öffentlichkeit über umweltpolitische Querschnittsthemen. In diesem 2020 veröffentlichten interaktiven und multimedialen Angebot werden, gestützt auf statistische Daten und Geo-Daten sowie qualitative Informationen, komplexe Umweltthemen einer breiten Öffentlichkeit verständlich und schutzgutübergreifend präsentiert und vermittelt. Der UBA-Umweltatlas unterstützt damit das UBA bei der Erfüllung einer seiner Kernaufgaben, die in der Information der Öffentlichkeit über den Zustand der Umwelt und den Umweltschutz im Sinne des Umweltinformationsgesetzes (UIG) besteht.

Keywords: Öffentlichkeitsinformation, Umweltinformationsgesetz, Reaktiver Stickstoff

1 Einleitung

„Die informationspflichtigen Stellen [des Bundes] unterrichten die Öffentlichkeit in angemessenem Umfang aktiv und systematisch über die Umwelt. In diesem Rahmen verbreiten sie Umweltinformationen, die für ihre Aufgaben von Bedeutung sind und über die sie verfügen. [...] Die Verbreitung von Umweltinformationen soll in für die Öffentlichkeit verständlicher Darstellung und leicht zugänglichen Formaten erfolgen.“ (§10 Abs. 1 und 3 UIG³)

Der Zugang zu Umweltinformationen, die Beteiligung der Öffentlichkeit an Entscheidungsverfahren in Umweltangelegenheiten sowie der Zugang zu Gerichten in Umweltangelegenheiten sind gemäß der Aarhus-Konvention⁴ notwendige Voraussetzungen, um Bürger*innen ein Leben in einer gesunden und dem Wohlbefinden zuträglichen Umwelt und ein aktives Engagement für einen wirksamen Umweltschutz zu ermöglichen. Zur Umsetzung der Aarhus-Konvention und der EU-Umweltinformationsrichtlinie (2003/4/EG) wurde in Deutschland das

¹ Bosch & Partner GmbH, Arbeitsbereich Umweltinformation und Klimafolgenanpassung, Pettenkoflerstraße 24, D-80336 München, s.andrian@boschpartner.de

² Kopfarbyte UG, Simplonstraße 44, 10245 Berlin, maria@kopfarbyte.com

³ Umweltinformationsgesetz in der Fassung der Bekanntmachung vom 27. Oktober 2014 (BGBl. I S. 1643), das zuletzt durch Artikel 2 des Gesetzes vom 25. Februar 2021 (BGBl. I S. 306) geändert worden ist (UIG)

⁴ UNECE-Übereinkommen vom 25.06.1998 über den Zugang zu Informationen, die Öffentlichkeitsbeteiligung an Entscheidungsverfahren und den Zugang zu Gerichten in Umweltangelegenheiten (Aarhus-Konvention)

Umweltinformationsgesetz (UIG) novelliert (SZS20). Damit wurde der rechtliche Rahmen zum einen für den freien Zugang zu Umweltinformationen bei informationspflichtigen Stellen sowie zum anderen für die aktive Verbreitung von Umweltinformationen durch diese Stellen angepasst beziehungsweise neu geschaffen (§1 Abs. 1 UIG).

Aus der gesetzliche Vorgabe zur Verbreitung von Umweltinformationen leitet sich für das Umweltbundesamt (UBA) eine seiner Kernaufgaben ab – die Öffentlichkeit aktiv und systematisch über den Zustand der Umwelt und aktuelle Umweltveränderungen in Deutschland zu informieren. Das UBA kommt diesem Auftrag in vielfältiger Weise nach, beispielsweise über eine öffentlich zugängliche Fachbibliothek, über Veranstaltungen, über digitale Fachinformationssysteme sowie vielfältige digitale Medien und Print-Medien. Von besonderer Bedeutung ist angesichts der Nutzung des Internets als hauptsächliche Kommunikations- und Wissensplattform in großen Teilen der Gesellschaft die digitale Aufbereitung und Bereitstellung von Umweltinformationen über Online-Formate, die über die UBA-Website www.umweltbundesamt.de zugänglich sind. Als Erweiterung des Online-Informationsangebots der UBA-Website ließ das UBA in einem FuE-Vorhaben ab dem Jahr 2016 den UBA-Umweltatlas entwickeln. Das Vorhaben umfasste die inhaltliche, gestalterische und technische Konzeption des Atlas sowie die Erarbeitung und Umsetzung der zwei thematischen Pilot-Bausteine „Bauen, Wohnen, Haushalte“ sowie „Reaktiver Stickstoff“, auf den im Folgenden beispielhaft eingegangen wird.

2 Grundlegendes zum UBA-Umweltatlas

2.1 Aufgabe und Funktion

Der vorrangige Zweck des UBA-Umweltatlas ist die Vermittlung von Wissen zu ausgewählten Schwerpunktthemen an die Öffentlichkeit im Sinne des UIG, also an interessierte Bürger*innen sowie Multiplikator*innen, Journalist*innen, Lehrende und politisch Verantwortliche. Damit soll er zur Umweltbewusstseinsbildung beitragen und darüber hinaus Anregungen zum eigenen Handeln bieten, indem er den Nutzenden wichtige umweltpolitische Botschaften zum jeweiligen Thema verständlich vermittelt.

Der UBA-Umweltatlas ermöglicht es den Interessierten erstens, sich einen **Überblick** über das jeweilige Schwerpunktthema zu verschaffen. Umweltthemen beziehungsweise -probleme haben häufig vielschichtige Verursacherkomplexe und betreffen eine Vielzahl an Umweltbestandteilen. In der Folge können Maßnahmen der jeweiligen Akteur*innen an unterschiedlichen Stellen ansetzen.

Um die Vielschichtigkeit von Umweltthemen zu verdeutlichen, sei an dieser Stelle das Beispiel reaktiver Stickstoff angeführt: Reaktive Stickstoffverbindungen entstehen im Wesentlichen in der Landwirtschaft, in Energiewirtschaft und Industrie, im

Verkehrssektor sowie bei der Abwasserbehandlung. Reaktive Stickstoffverbindungen sind in der Umwelt sehr wandelbar, ihre Auswirkungen betreffen alle Umweltbestandteile von der Atmosphäre über Boden und Wasser bis zur biologischen Vielfalt und zur menschlichen Gesundheit. Ebenso vielfältig sind die Handlungsmöglichkeiten, um mit reaktivem Stickstoff und seinen Auswirkungen umzugehen.

Auf der UBA-Website führte diese Vielschichtigkeit dazu, dass reaktiver Stickstoff in seinen unterschiedlichen Erscheinungsformen auf zahlreichen verschiedenen Unterseiten angesprochen wird. Ein geeignetes Online-Format für einen Überblick über dieses Querschnittsthema, wie er beispielsweise in der Publikation „Reaktiver Stickstoff in Deutschland“ (GF15) gegeben wird, fehlte bisher aber. Um diese Lücke zu füllen, ist der UBA-Umweltatlas als eine integrierte Informationsplattform konzipiert, die Informationen themen- und schutzgutübergreifend darstellt und vermittelt.

Zweitens bieten die Bausteine des UBA-Umweltatlas den Interessierten auch einen ersten, aber dennoch fundierten **Einstieg** in das jeweilige Schwerpunktthema. Die Nutzenden können sich step-by-step und je nach Interessenlage intensiv in das Schwerpunktthema insgesamt oder in einzelne seiner Teilbereiche vertiefen und sich damit auseinandersetzen. Die Informationen werden ihnen dafür verständlich erläutert und anhand von interaktiven Karten, Infografiken und Diagrammen sowie durch Bilder und Videos anschaulich vermittelt.

Drittens fungiert der UBA-Umweltatlas als **Türöffner** zum ausführlicheren und differenzierteren Informationsangebot auf der UBA-Website sowie zu weiterführenden Online-Angeboten anderer Quellen außerhalb des UBA. Personen, die ein über den thematischen Einstieg hinausgehendes Interesse am behandelten Schwerpunktthema haben, werden über Verweise und Links in die Lage versetzt, gezielt vertiefende Informationen aufzusuchen.

2.2 Vertikale und horizontale Struktur der Bausteine

Um den Nutzenden einen einfachen Einstieg und einen informativen Besuch des UBA-Umweltatlas und seiner Bausteine zu ermöglichen, verwendet der Atlas eine klare und harmonisierte Grundstruktur, die sich auf unterschiedliche geartete Themenstellungen anwenden lässt. Die Inhalte der Bausteine sind in der Regel in Anlehnung an das DPSIR-Modell (Drivers, Pressures, State, Impact, Responses) der Europäischen Umweltagentur (EUA99) gegliedert. Diese Kategorien bilden die horizontale Struktur der Bausteine (siehe Abb. 1), in der das gewählte Schwerpunktthema mit seinen relevanten Ursachenfaktoren, deren Wirkungen auf die verschiedenen Umweltbestandteile sowie den Handlungsmöglichkeiten von Politik und Bürger*innen dargestellt wird. Um die damit verbundene große thematische Breite zu fassen, sind die Bausteine vertikal in eine hierarchische Struktur gegliedert (siehe Abb. 1):

- Rubriken (beispielsweise Verursacher, Wirkungen, Politisches Handeln, Persönliches Handeln),

- Themenfelder (beispielsweise Landwirtschaft, Verkehr; Boden, Luft und Atmosphäre, Grundwasser und Oberflächenwasser etc.) sowie
- Slides (beispielsweise Was ist reaktiver Stickstoff? Wie verändert der Mensch den natürlichen Stickstoffkreislauf?).

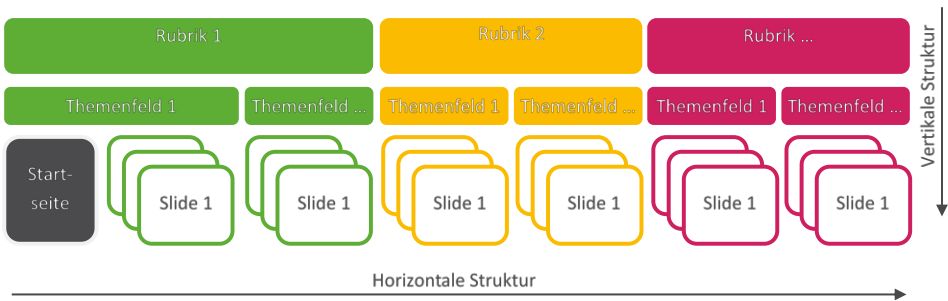


Abb. 1: Vertikale und horizontale Struktur innerhalb der Bausteine (ABE21)

Die Slides sind das Grundelement für Struktur und Gestaltung des UBA-Umweltatlas. Bei einer Slide handelt es sich um eine einzelne Seite, auf der mithilfe eines Mediums – der UBA-Umweltatlas verwendet die Medientypen interaktives Diagramm, interaktive Karten oder Infografiken, Kartenbilder sowie Videos und Bilder – sowie einem kurzen erläuternden Text jeweils eine Botschaft zu einem thematischen Aspekt des Themenfelds präsentiert wird (siehe Abb. 2). Die Slides sind innerhalb eines Themenfelds in einer stringenten Reihenfolge angeordnet, können aber auch jeweils für sich allein gelesen werden. Den Abschluss eines Themenfelds bildet jeweils eine Schluss-Slide, die keine eigene Abbildung enthält, sondern auf thematisch verwandte Publikationen und Dokumente verweist sowie Verlinkungen zu weiterführenden Internet-Informationen auf der UBA-Website oder auf externen Websites anbietet.

The screenshot shows a web page from the UBA (Umweltbundesamt) titled "Was ist reaktiver Stickstoff?". The page layout includes a top navigation bar with categories like "Umwelt Bundesamt", "Start", "Das UBA", "Themen", "Presse", "Publikationen", "Tipps", and "Dates". Below this is a secondary navigation bar with "Reaktiver Stickstoff" selected, and sub-categories: "Einführung", "Verursacher", "Wirkungen", "Politisches Handeln", and "Mein Handeln". A "Navigation" button is on the right.

The main content area features a slide with the following elements:

- Text Box (Left):** Contains the title "Was ist reaktiver Stickstoff?" and a detailed paragraph explaining that reactive nitrogen is the most abundant form of molecular atmospheric nitrogen, which is fixed into various compounds. It lists examples like nitric oxide (NO_x), nitrous oxide (N_2O), ammonia (NH_3), and nitrate (NO_3^-), and notes that these compounds are highly reactive and can be harmful to the environment and health. It also mentions that reactive nitrogen is a key component of many biological molecules like proteins and DNA.
- Diagram (Center):** A collection of seven circular icons representing different nitrogen compounds: NO_x (Nitrogenoxide), N_2O (Lachgas), NH_3 (Ammoniak), NH_4^+ (Ammonium), NO_2^- (Nitrit), NO_3^- (Nitrat), and N_{org} (Stickstoff gebundene Substanz).
- Medium Box (Right):** A red-bordered box containing the word "Medium".

The footer of the slide includes the UBA logo and name, the tagline "Für Mensch und Umwelt", and social media icons for Facebook, Twitter, YouTube, and Instagram.

Abb. 2: Eine Slide und deren Inhalte – Beispiel aus dem Baustein „Reaktiver Stickstoff“ (UBA-Umweltatlas „Reaktiver Stickstoff“ – Was ist reaktiver Stickstoff?)

3 Verwendung und Präsentation von Umweltinformationen

Der UBA-Umweltatlas kann grundsätzlich alle in §2 Abs. 3 UIG benannten Arten von Umweltinformationen nutzen und darstellen. Der Baustein „Reaktiver Stickstoff“ enthält beispielsweise Umweltinformationen über

- den Zustand von Boden, Grund- und Oberflächenwasser, Luft und Atmosphäre etc.,
- die Emissionen von reaktivem Stickstoff aus Landwirtschaft, Energiewirtschaft, Verkehr und Abwasserbehandlung in die genannten Umweltbestandteile,
- das politische Konzept zum integrierten Umgang mit reaktivem Stickstoff, über Gesetze und Verordnungen zur Reinhaltung von Luft und Gewässern sowie über Handlungsmöglichkeiten der Bürger*innen,
- die deutsche Berichterstattung atmosphärischer Emissionen als Vertragsstaat der Klimarahmenkonvention sowie der Genfer Luftreinhaltekonvention oder die Berichterstattung zur EU-Nitratrichtlinie (91/676/EWG),
- die Ergebnisse von Kosten-Nutzen-Analysen zum Einsatz von reaktivem Stickstoff oder zur Aufbereitung von Trinkwasser sowie

- die möglichen Gefährdungen der menschlichen Gesundheit infolge der übermäßigen Konzentration von reaktivem Stickstoff in den verschiedenen Umweltbestandteilen.

Mit den unterschiedlichen Arten von Umweltinformationen, die im UBA-Umwelatlas präsentiert werden, sind unterschiedliche Anforderungen an deren Präsentation und damit an die Verwendung der Medien in den Slides verbunden. Wesentlich für die Verwendung eines Mediums ist die Unterscheidung in quantitative und qualitative Umweltinformationen.

Quantitative Umweltinformationen über den Zustand von Umwelt und Gesundheit, über das Ausmaß und die Entwicklung von Emissionen oder aus Kosten-Nutzen-Analysen können in Medientypen präsentiert werden, die sich für die zeitlich und räumlich differenzierte Darstellung von statistischen bzw. von Geo-Daten eignen. Dies sind

- interaktive Diagramme,
- interaktive Kartenanwendungen (siehe Abb. 3),
- Kartenbilder, Kartenvergleiche und Stopmotion-Animation.

In den Baustein „Reaktiver Stickstoff“ sind vielfältige quantitative Umweltinformationen eingeflossen, beispielsweise in interaktive Diagramme und Kartenanwendungen zu den Emissionen der verschiedenen Verursacherebereiche.

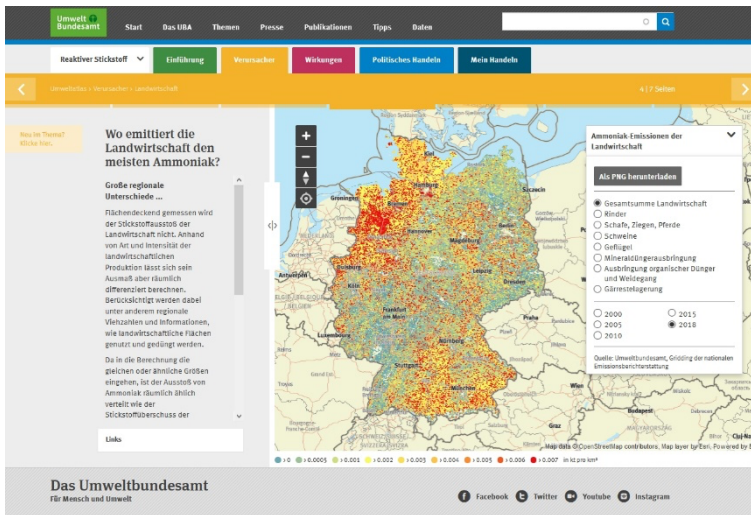


Abb. 3: Beispiel einer interaktiven Kartenanwendung (UBA-Umwelatlas „Reaktiver Stickstoff“ – [Wo emittiert die Landwirtschaft den meisten Ammoniak?](#))

Qualitative Umweltinformationen lassen sich in aller Regel nicht in zahlenbasierten Medientypen darstellen. Das gilt für Informationen etwa zum politischen Handeln, beispielsweise Informationen zu Rechts- und Verwaltungsvorschriften oder zu politischen Strategien, ebenso wie für Informationen zum persönlichen Handeln mit Hinweisen für ein umweltschonendes Konsumverhalten. Qualitative Umweltinformationen bringen vielmehr die Anforderung mit sich, Sachverhalte verständlich und grafisch anschaulich aufzubereiten. Dies trifft mit Blick auf das UIG insbesondere für komplexere Zusammenhänge zu, die auch interessierten Personen nicht ohne Weiteres bekannt oder nachvollziehbar sind. In der Regel sind hierfür folgende Medientypen des UBA-Umweltatlas geeignet:

- interaktive Infografiken (siehe Abb. 4),
- Videoanimationen,
- Zeitstrahlen,
- interaktive Spiele.

Beispiele im Baustein „Reaktiver Stickstoff“ sind etwa die interaktiven Infografiken zur Erläuterung des Stickstoffkreislaufs oder zu den natürlichen und anthropogenen Einträgen von reaktivem Stickstoff in die Atmosphäre sowie die Videoanimation der Wirkmechanismen von Lachgas zum Abbau der stratosphärischen Ozonschicht.

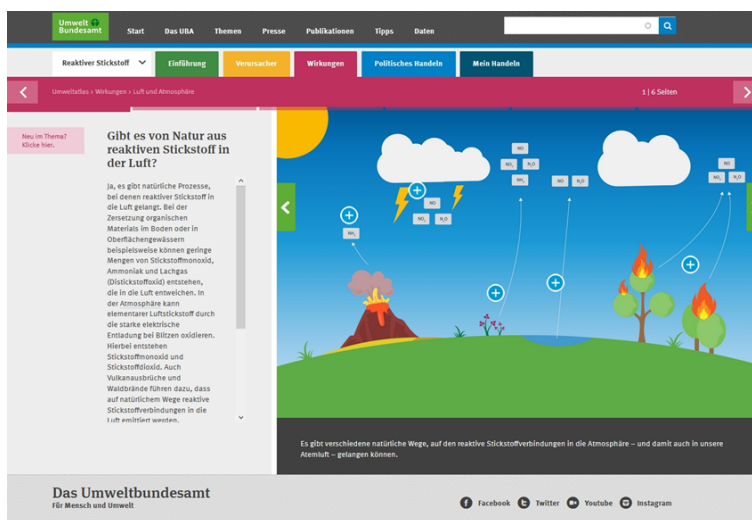


Abb. 4: Beispiel einer interaktiven Infografik
(UBA-Umweltatlas „Reaktiver Stickstoff“ – [Gibt es von Natur aus reaktiven Stickstoff in der Luft?](#))

4 Technische Umsetzung

4.1 Technischer und gestalterischer Rahmen

Der UBA-Umweltatlas ist integraler Bestandteil der UBA-Website und berücksichtigt insofern deren Grundprinzipien und Rahmenbedingungen. Hierzu gehören die Einhaltung des Datenschutzes in Übereinstimmung mit der Datenschutz-Grundverordnung (DSGVO), Barrierefreiheit, Responsivität und Suchmaschinenoptimierung (SEO).

Hinsichtlich des Backend ist mit der Integration des UBA-Umweltatlas die Einbindung in das mit der Plattform DRUPAL implementierte Content Management System (CMS) der UBA-Website verbunden. Dadurch ist es möglich, dass Erweiterungen und Aktualisierungen des UBA-Umweltatlas grundsätzlich von UBA-Mitarbeitenden selbst geleistet und dass erstellte Medien auch an anderer Stelle der UBA-Website über eine Referenzierung verwendet werden können. Die Inhalte des UBA-Umweltatlas lassen sich über das CMS mit wenig Aufwand pflegen und fortschreiben, Programmierungen und Programmierkenntnisse sind dafür in der Regel nicht erforderlich. Das UBA kann den UBA-Umweltatlas unter diesen Bedingungen dauerhaft als Informationsplattform anbieten und regelmäßig fortschreiben. Zum anderen ermöglicht die Art seiner Umsetzung im CMS einen modularen Aufbau des UBA-Umweltatlas, der bei Bedarf um zusätzliche Themenbausteine erweitert werden kann.

Das Frontend des UBA-Umweltatlas beachtet die gestalterischen Vorgaben hinsichtlich der darzustellenden Standardelemente der UBA-Website sowie des Corporate Designs, das für das UBA im Corporate Design-Leitfaden (UBA18) festgelegt ist.

4.2 Realisierung im UBA-CMS

Die technische Umsetzung der UBA-Website basiert auf der Verwendung sogenannter Inhaltstypen. Die Inhaltstypen dienen dazu, die Inhalte der einzelnen Seiten im Backend des CMS strukturiert für die Darstellung auf der Website abzulegen. Die Eingabe erfolgt über definierte Formulare. Sogenannte Templates steuern die Darstellung im Frontend.

Die technische Umsetzung des UBA-Umweltatlas ist nahtlos in dieses System integriert. Zur Eingabe der Inhalte für den UBA-Umweltatlas bestehen im Backend die zwei eigens programmierten Inhaltstypen „Umweltatlas“ und „Umweltatlas-Karte“. Der Inhaltstyp „Umweltatlas“ ist dabei flexibel angelegt und kann auf unterschiedliche Darstellungserfordernisse reagieren. Dazu bindet der Inhaltstyp auch die bereits vorhandenen Inhaltstypen „Bild“, „Video“, „Interaktiv Diagramm“ und „Interaktiv Tool“ aus dem CMS der UBA-Website ein. Das bedeutet, dass bestehende Inhalte aus dem CMS in den UBA-Umweltatlas integriert werden können. Umgekehrt können aber auch für den UBA-Umweltatlas erstellte Inhalte, einschließlich von Inhalten des Typs „Umweltatlas-Karte“, an anderer Stelle auf der UBA-Website verwendet werden. [ABE21]

Über den **Inhaltstyp „Umweltatlas“** lassen sich alle Informationen, die für das Befüllen einer konkreten Slide erforderlich sind, in einem einzigen Formular im Backend einfügen. Dieses Formular ermöglicht die Anlage neuer Slides sowie die Pflege und Ergänzung bestehender Slides. Die Darstellung im Frontend erfolgt gemäß mehrerer für den UBA-Umweltatlas entwickelter Templates. Das zu nutzende Layout ist dabei im Backend auswählbar und kann bei Bedarf angepasst werden. [ABE21]

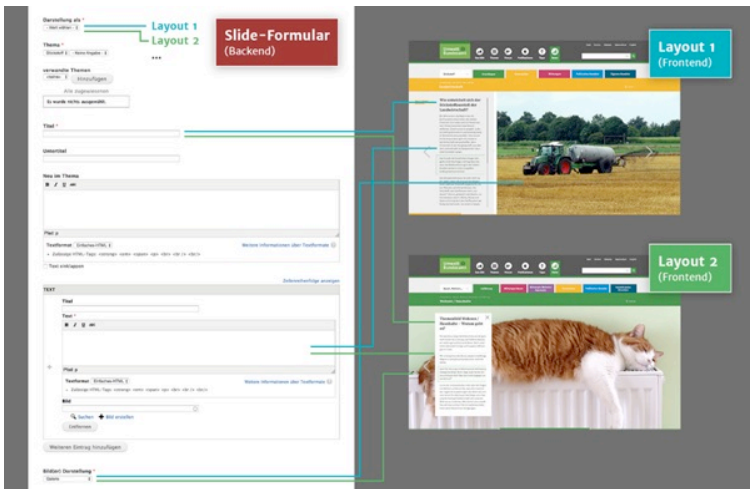


Abb. 5: Zusammenhang von Backend und Frontend [ABE21]

Für den UBA-Umweltatlas steht außerdem der eigens entwickelte **Inhaltstyp „Umweltatlas-Karte“** zur Verfügung, mit dem interaktive Kartenanwendungen über das CMS umgesetzt werden. Dieser Inhaltstyp dient der adäquaten Abbildung von räumlichen Daten aus mehreren definierten Quellen. Die regelmäßige Verfügbarmachung und Anbindung der Daten für den UBA-Umweltatlas beispielsweise für Zwecke der Aktualisierung wurde durch die Entwicklung eigener Schnittstellen gelöst, die bei einer Beibehaltung der bestehenden Datenstruktur langfristig für die Datenanbindung genutzt werden können. Mithilfe dieser Schnittstellen können der Datahub des UBA, der IÖR-Monitor⁵ sowie Shape-Files als Datenquellen für interaktive Kartenanwendungen verwendet werden. Im CMS lassen sich Einstellungen unter anderem zur Darstellung und zur Filterung der eingelesenen Daten vornehmen und in einem mehrstufigen Prozess individuelle Karten erstellen [ABE21]. Besonderer Fokus wurde bei der Entwicklung des Inhaltstyps auf die Performanz sowie die Möglichkeiten der Darstellung von Daten auf der Karte gesetzt. Die über die Schnittstellen erhaltenen Daten werden dafür vorgefiltert, in das PBF-Format („Protocolbuffer Binary Format“) umgewandelt, browserbasiert ausgeliefert und entsprechend der Filteroptionen abgebildet. Um eine größtmögliche

⁵ Monitor der Siedlungs- und Freiraumentwicklung (IÖR-Monitor): <https://www.ioer-monitor.de/>

Datenvielfalt präsentieren zu können, wurde auf neueste Techniken zurückgegriffen – etwa das dynamische Laden von Vector-Tiles oder die Verwendung von WebGL – um die Datenvisualisierung individuell an die jeweilige Zoomstufe anpassen zu können.

5 Fazit und Ausblick

Mit der Entwicklung des UBA-Umweltatlas ist es dem UBA gelungen, seine Website um eine attraktive Informationsplattform mit vielfältigen interaktiven Möglichkeiten zu ergänzen, mit der der breiten Öffentlichkeit auch komplexere umweltpolitische Schwerpunktthemen auf der Basis von statistischen Daten und Geo-Daten verständlich vermittelt werden können. Die Berichterstattung des UBA, also die Verbreitung von Umweltinformationen im Sinne des UIG durch das UBA, konnte dadurch in ihrer Qualität weiter verbessert werden. Im August 2020 ging der UBA-Umweltatlas⁶ mit seinem ersten Baustein „Reaktiver Stickstoff“ online, der zweite Baustein „Bauen, Wohnen, Haushalte“ folgte im Mai 2021. Derzeit wird das Schwerpunktthema „Umwelt und Landwirtschaft“ für die Präsentation als dritter Baustein des UBA-Umweltatlas ausgearbeitet. Perspektivisch soll der UBA-Umweltatlas Bausteine zu fünf bis zehn umweltpolitischen Schwerpunktthemen umfassen.

6 Literaturverzeichnis

- [ABE21] Andrian-Werburg, S. v.; Buth, M.; Eichhorn, D.; Meinel, G.; Theuring, M.: Thematischer Umweltatlas: Medienwirksame Aufbereitung von Umweltinformationen durch Verknüpfung von räumlichen Darstellungen, Umweltdaten und Fachbeiträgen. UBA Texte 10/2021 Umweltbundesamt, Dessau-Roßlau, 2021.
- [EUA99]: EUA, Europäische Umweltagentur (Hrsg.): Environmental indicators - Typology and overview. Technical report No 25, EUA, Kopenhagen, 1999.
- [GF15] Geupel, M.; Frommer, J.: Reaktiver Stickstoff in Deutschland. Ursachen, Wirkungen, Maßnahmen. Umweltbundesamt, Dessau-Roßlau, 2015.
- [SZS20] Stracke, K.; Zschiesche, M.; Schomerus, T.; Habigt, L.; Konopatzki, P.; Lütkemeyer, E. M.; Saiger, A.-J.; Sperfeld, F.; Beermann, J. (2020): Evaluation des Umweltinformationsgesetzes (UIG). Analyse der Anwendung der Regelungen des UIG und Erschließung von Optimierungspotentialen für einen ungehinderten und einfachen Zugang zu Umweltinformationen. UBA Texte. 235/2020, Umweltbundesamt, Dessau-Roßlau, 2020.
- [UBA18] UBA, Umweltbundesamt (Hrsg.): Corporate Design des Umweltbundesamtes. Umweltbundesamt, Dessau-Roßlau, 2018.

⁶ UBA-Umweltatlas: <https://www.umweltbundesamt.de/daten/umweltatlas/>

Umweltinformationssysteme für den Naturschutz

Potenziale, Risiken, Anwendungsfelder und Entwicklungsperspektiven

Christian Schneider¹, Klemens Mrogenda¹, Bettina Dibbern², Rene Höfer², Marlen Davis¹, Heike Sommer² und Michael Biló¹

Abstract: Umweltinformationssysteme (UIS) bestehen aus den Komponenten Erhebung, Datenhaltung, Analyse und Bewertung sowie Kommunikation von natur- und umweltbezogenen Informationen. Dem Beitrag liegt ein breites Verständnis von UIS zugrunde, das sowohl den organisatorischen Überbau als auch die technische Infrastruktur einschließt. Das Bundesamt für Naturschutz (BfN) stellt UIS für zahlreiche Informationsprozesse zur Verfügung und fördert zugleich den Aufbau externer Systeme. Es betreibt nicht notwendigerweise alle Komponenten selber, sondern nimmt jeweils unterschiedliche Rollen ein. Diese Rollen betreffen Koordination, Betrieb, Beauftragung, Förderung oder Forschung und können je nach Zweck des Informationsprozesses variieren. Für Vollzugsaufgaben ist das BfN unmittelbar zuständig. Informationsprozesse, die auf einer Bund-Länder-Zusammenarbeit beruhen, wie zum Beispiel die nationale Berichterstattung über den Erhaltungszustand von Arten und Lebensräumen nach der europäischen Fauna-Flora-Habitat-Richtlinie (FFH-Richtlinie), erfordern eine Aufteilung von Rollen für die verschiedenen Komponenten des UIS. Der Beitrag umreißt exemplarisch drei Fachprozesse am BfN und die zugehörigen Komponenten. Anhand der Beispiele werden Entwicklungsperspektiven und Herausforderungen skizziert.

Keywords: Naturschutzinformationssysteme, FFH-Bericht, Natura 2000, FloraWeb, CITES Online

1 Einleitung

Das Bundesamt für Naturschutz (BfN) ist eine wissenschaftliche Bundesoberbehörde und unterstützt das Bundesumweltministerium (BMU) in allen Fragen des Naturschutzes, der Landschaftspflege und zur internationalen Zusammenarbeit. Die Bereitstellung von Naturschutzinformationen für die Öffentlichkeit gehört ebenfalls zu den zentralen Aufgaben des Amtes. Darüber hinaus übernimmt es Vollzugsaufgaben zum Beispiel für den internationalen Artenschutz, den Meeresnaturschutz und zur Umsetzung des Gentechnikgesetzes.

Zur Erfüllung dieser Aufgaben fördert, berät und entwickelt das BfN umfangreiche Aktivitäten zur Beobachtung und zum Schutz der Biodiversität und des Naturhaushaltes. Das

¹ Bundesamt für Naturschutz, Fachgebiet I 1.1 Strategische Digitalisierung in Natur und Gesellschaft, Alte Messe 6, 04103 Leipzig, Christian.Schneider@bfn.de

² Bundesamt für Naturschutz, Fachgebiet I 1.2 Naturschutzinformation, Geoinformation, Open Data, Konstantinstr. 110, 53179 Bonn

Bundesamt ist also eine zentrale Schnittstelle für den Transfer von Informationen und wissenschaftlichen Erkenntnissen über Naturzustände an politische Entscheidungstragende, Aktive im praktischen Naturschutz und die Öffentlichkeit.

Aus diesem breiten Aufgabenspektrum resultieren unterschiedliche Fach-, Informations- und Berichtsprozesse. In Form von Umweltinformationssystemen (UIS) stellt das BfN dafür sowohl den organisatorischen Überbau als auch die technische Infrastruktur zur Verfügung. Dieses weite Verständnis von UIS spiegelt sich auch in den Definitionen nach [Ge21] und [EC19] wider.

Die Umweltinformationssysteme am BfN ermöglichen einerseits sehr komplexe Abstimmungs- und Berichtsprozesse zwischen zahlreichen Akteuren, von lokalen Naturschutzorganisationen und Fachgesellschaften bis hin zu Delegationen in internationalen Verhandlungen der UN. Andererseits sind sie die Grundlage für Vollzugsaufgaben wie die Durchsetzung des Artenschutzes im internationalen Wildtierhandel. Aber auch für die Zusammenführung und Bereitstellung großer Datenbestände zum Thema Naturzustand und -schutz in Deutschland setzt das BfN UIS ein.

Es ist das Ziel des vorliegenden Beitrags die Vielfältigkeit von UIS am BfN darzustellen und zugleich die enge Verschränkung von organisatorischen Prozessen und technischen Systemen anhand exemplarischer Fachprozesse zu dokumentieren. Dabei wird darauf geachtet die unterschiedlichen Rollen des BfN transparent zu machen. Anhand der Beispiele werden zudem die Potenziale und Herausforderungen für die Weiterentwicklung von UIS im Naturschutz diskutiert.

2 Umweltinformationssysteme am Bundesamt für Naturschutz

2.1 Fachprozesse und exemplarische UIS

Im Folgenden werden drei Fachprozesse und die zugrundeliegenden UIS vorgestellt, welche die große inhaltliche Spannweite des BfN widerspiegeln:

- Nationaler Bericht über den Erhaltungszustand von Lebensraumtypen und Arten nach Art. 17 der Fauna-Flora-Habitat-Richtlinie (FFH-Richtlinie),
- Florenkartierung Deutschlands,
- Artenschutzvollzug nach dem Washingtoner Artenschutzabkommen (CITES).

Der Fokus liegt dabei nicht auf den technischen Details, sondern auf dem Zusammenspiel von naturschutzfachlichen Aufgaben, organisatorischen Prozessen, den zugrundeliegenden technischen Komponenten von UIS und den Rollen des BfN. Dafür werden die Fachprozesse vorgestellt, die Grundrisse der technischen Systeme skizziert, sowie deren Verschränkung mit politischen und verwaltungstechnischen Abläufen dargestellt.

Zur besseren Vergleichbarkeit der UIS wird ein einheitliches Schema verwendet, welches auf der Unterteilung der technischen und organisatorischen Komponenten in Datenerhebung, Datenprozessierung, Datenanalyse und Kommunikation der Daten basiert (Abb.: 1-3) (Günther 1998).

2.2 UIS zur Erfüllung von Berichtspflichten am Beispiel der FFH-Richtlinie

Die EU-weite, in Deutschland seit 1998 rechtsverbindliche FFH-Richtlinie zielt auf den Schutz von Pflanzen, Tieren und Pilze sowie deren Lebensräume ab, die in den Anhängen der Richtlinie definiert sind [Bf21]. Alle sechs Jahre ist Deutschland verpflichtet, einen umfassenden Bericht über den Erhaltungszustand geschützter Lebensräume und Arten an die Europäische Kommission zu übermitteln.

Am Prozess der Erstellung des Nationalen Berichts gemäß FFH-Richtlinie Art. 17 sind zahlreiche Akteure aus Bund und Ländern beteiligt.

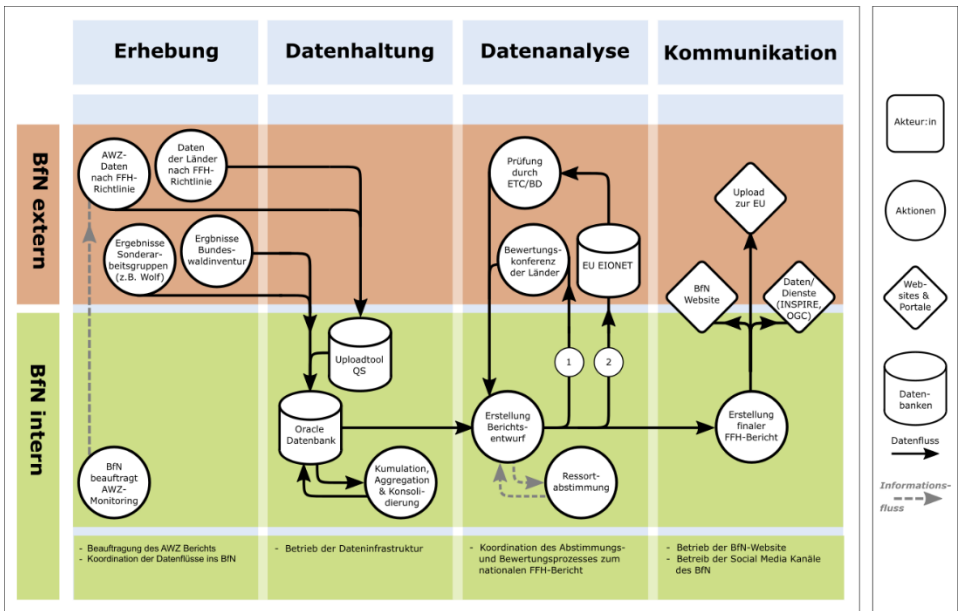


Abb. 1: UIS zur Erstellung des Nationalen Berichts gemäß FFH-Richtlinie

Das BfN nutzt ein UIS um diesen komplexen Abstimmungs- und Erstellungsprozess umzusetzen. Im Ergebnis umfasst ein solcher Bericht:

- die aktualisierten Ergebnisse des Fauna-Flora-Habitat-Monitorings, also der systematischen, regelmäßigen Erfassung von Erhaltungszuständen geschützter Arten und Lebensräume,
- die Bewertung der Erhaltungszustände der geschützten Arten und Lebensräume,
- einen Vergleich mit den Bewertungsergebnissen der vorherigen Berichte.

Der Bericht dient einerseits der Bilanzierung von Erfolgen und verweist andererseits auf bestehende Handlungsbedarfe. Darauf aufbauend wird eine Einschätzung der Effizienz des Natura 2000-Schutzgebietsnetzes möglich.

Die Bundesländer sind für die Erhebung der notwendigen Daten und die Bewertung des Zustands von Arten und Lebensräumen in ihren Gebieten verantwortlich. Parallel zu den Bundesländern verantwortet das BfN die Erhebung von Zustandsdaten der FFH-Gebiete in der jenseits der 12 Seemeilen vor der Küste liegenden ausschließlichen Wirtschaftszone (AWZ) in Nord- und Ostsee [Bf17] (Abb. 1: Erhebung).

Die von Bundesländern und Bund erhobenen Daten werden im BfN über ein Uploadtool plausibilitäts- und qualitätsgesichert gesammelt und zu Werten mit Aussagekraft für das Bundesgebiet konsolidiert (Abb. 1: Datenhaltung). Ziele dieses Bearbeitungsschrittes sind es, die Gesamtflächen der Verbreitung und der Zustand eines Schutzgutes, also spezifische Arten oder Lebensräume, in den biogeographischen Regionen zu erfassen.

Auf Basis dieser konsolidierten Daten und weiterer Datenquellen erarbeitet das BfN einen Berichtsentwurf. Dieser enthält bereits eine Aus- und Bewertung des Erhaltungszustands der geschützten Arten und Lebensraumtypen in Deutschland (Abb. 1: Datenanalyse).

In einem folgenden Schritt werden die Daten in Bewertungskonferenzen für jede biogeographische Region mit den Bundesländern abgestimmt. Am weiteren Verfahren sind in unterschiedlichen Schritten das Umweltministerium, weitere Bundesministerien und final das europäischen Themenzentrum für Biologische Vielfalt (ETC/BD) beteiligt. Nach der Überarbeitung des Berichts erfolgt ein Upload per XML an die Europäische Kommission (Abb. 1 Kommunikation).

Im Verlauf des gesamten Prozesses übernimmt das BfN neben der Koordination unterschiedliche Aufgaben, die in Abbildung 1 genannt sind.

Der FFH-Prozess steht exemplarisch für die Nutzung eines UIS zur Koordination und Erfüllung von Berichtspflichten durch das BfN. Perspektivisch wird eine Weiterentwicklung der organisatorischen Abläufe und der genutzten technischen Komponenten des UIS in Richtung eines teilautomatisierten Reportings angestrebt. Die Automatisierung kann und soll dabei die oben skizzierten Abstimmungsprozesse nicht ersetzen. Sie kann aber bei der

Analyse von Daten sowie bei der Aufbereitung und Bereitstellung von Daten für den Bewertungsprozess sinnvoll vorangetrieben werden.

Darüber hinaus bestehen große Potenziale auf dem Feld der satellitengestützten Fernerkundung, um neue Informationsquellen beispielsweise des europäischen Copernicus Programms einzubinden. Dabei scheint es neben der direkten Erfassung von Schutzgütern sinnvoll, aussagekräftige Proxy-Informationen zu identifizieren, welche Rückschlüsse auf den Zustand von Schutzgütern zulassen, oder die Triebkräfte hinter Naturzustandsveränderungen erfassbar machen. Die dafür zu entwickelnden Fernerkundungsprodukte können dann als zusätzlicher Input in das UIS eingespeist werden.

Große Potenziale liegen weiterhin in der Sicherstellung der Konnektivität und Nutzbarmachung der im Rahmen von Berichtsprozessen erhobenen Datenbestände für andere Prozesse. Ein Beispiel dafür ist die Bereitstellung von Informationen über Arten und Lebensraumtypen nach der INSPIRE-Richtlinie der EU [EP07]. Den Vorgaben der Richtlinie folgend, müssen Geo-Informationen zu den Arten und Lebensraumtypen in das INSPIRE Datenmodell transformiert und als interoperable Web-Services in Darstellungs- und Downloaddiensten zur Verfügung gestellt werden.

Um Informationen über Natur- und Umweltzustände wirkungsvoll einsetzen zu können, werden sie zukünftig noch stärker zielgruppenorientiert aufbereitet und über entsprechende Kanäle verfügbar gemacht.

Besondere Herausforderungen ergeben sich daraus, dass die Vorgaben und Rahmenbedingungen für die Erstellung des FFH-Berichts dynamisch weiterentwickelt werden und diese fortlaufend im UIS umgesetzt werden müssen.

2.3 UIS zur öffentlichen Bereitstellung von Naturinformationen am Beispiel von FloraWeb

FloraWeb.de ist ein Onlineangebot des BfN, das über die Verbreitung von wildwachsenden Farn- und Blütenpflanzen sowie Pflanzengesellschaften in Deutschland informiert [MS07]. Das System führt dafür Informationen aus unterschiedlichen Quellen zusammen. Neben Verbreitungskarten und Fotos macht es tausende Steckbriefe zu Pflanzenarten und Pflanzengesellschaften zugänglich, mit wissenschaftlichen Informationen zu Taxonomie, Systematik und Ökologie. Damit erfüllt es den im Umweltinformationsgesetz (UIG) formulierten Auftrag, wonach verfügbare Umweltinformationen systematisch für die Öffentlichkeit zur Verfügung gestellt werden müssen.

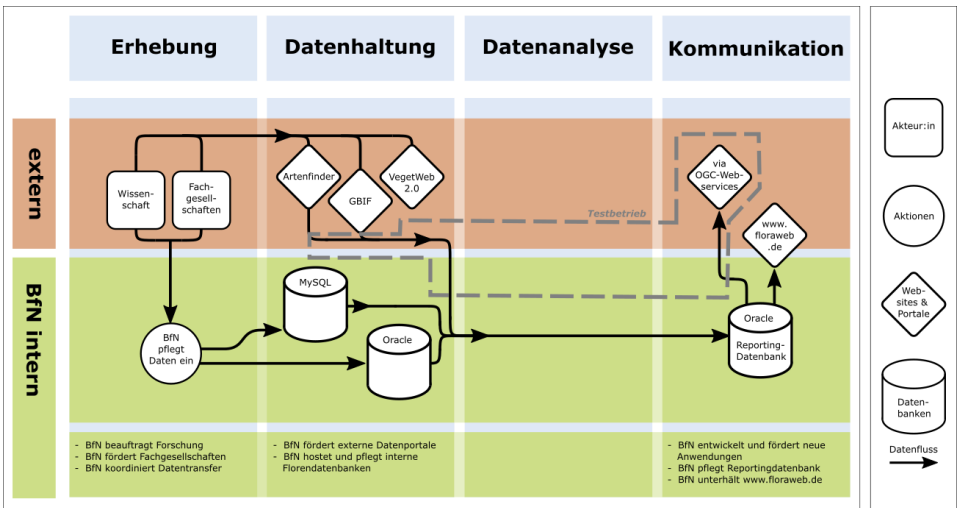


Abb. 2: UIS FloraWeb - Technische und organisatorische Komponenten

In vergangenen Jahrzehnten wurde mit FloraWeb ein komplexes UIS zur Erfassung, Speicherung, Pflege und Bereitstellung von Floreninformationen aufgebaut. Dieses System stellt einerseits die technischen Grundlagen zur Verfügung. Andererseits übernimmt das BfN auch eine koordinierende Rolle und stellt gemeinsam mit botanischen Fachgesellschaften und der Wissenschaft die Pflege und Weiterentwicklung der Inhalte sicher. Eine wesentliche fachliche Grundlage dafür ist die kontinuierliche wissenschaftliche Überarbeitung der zugrundeliegenden botanischen Taxonomie [Bu18]. Die enge und kontinuierliche Einbindung von Fachgesellschaften und Wissenschaftler:innen macht FloraWeb zu einer der wichtigsten Infrastrukturen zur Konsolidierung der Floren- und Pflanzenartenkartierung in Deutschland. Die Abbildung 2 zeigt die organisatorischen und technischen Komponenten von FloraWeb. Darüber hinaus werden die unterschiedlichen Rollen des BfN kenntlich gemacht.

Die mehr als 30 Mio. Einträge in den Datenbanken des UIS stammen vorrangig aus Kartierungsprojekten botanischer Fachgesellschaften und der Wissenschaft [FW21] (Abb. 2 Datenerhebung). Zur Unterstützung der ehrenamtlichen Kartierungen der Fachgesellschaften wurden eigene Erfassungsprogramme vom BfN gefördert und entwickelt.

Die so erhobenen Datenbestände zu Artenvorkommen wurden und werden in interne Datenbanken des BfN eingepflegt oder finden Eingang in Systeme von Fachgesellschaften, Citizen Science Projekten und internationalen Netzwerken (Abb. 2 Datenhaltung) [JS15]. Daten zur Verbreitung von Pflanzengesellschaften werden beispielsweise im System VegetWeb 2.0 zusammengeführt, welches vom BfN gefördert wurde.

Aus den internen verschiedenen Datenbanksystemen speist sich eine Reportingdatenbank. Daraus können über die Website www.FloraWeb.de wissenschaftliche Informationen und Verbreitungsangaben für tausende wildwachsende Pflanzenarten abgerufen werden. Aus

diesem Datenbestand werden zudem Verbreitungskarten kompiliert und können räumlichen Beobachtungsdaten per Webservices zur Verfügung gestellt werden. Zusätzlich können auch BfN externe Datenquellen in die Abfrage einbezogen werden (Abb. 2 Kommunikation).

Der Prozess zur Florenkartierung in Deutschland mit den zugrundeliegenden organisatorischen Prozessen und technischen Komponenten steht exemplarisch für die Nutzung eines UIS zur Zusammenführung und Bereitstellung von Informationen über Naturzustände in Deutschland. Das Besondere ist, dass auch Datenbestände aufgearbeitet wurden und werden, die von verschiedenen Akteuren außerhalb der Behördenstruktur erhoben worden sind. Das System wurde dafür laufend erweitert und wird organisatorisch, sowie technisch vom BfN als auch von externen Beteiligten getragen. Aktuell wird an der Erweiterung des Systems um einen bebilderten Bestimmungsschlüssel sowie um ein virtuelles Herbarium gearbeitet.

Große Potenziale liegen zudem in der Weiterentwicklung der Bereitstellung der Datenbestände zu wildwachsenden Farn- und Blütenpflanzen für externe Anwendungen via Webservices und Schnittstellen.

Mittel- und längerfristig kann der Ansatz Daten aus verschiedenen behördlichen und externen Quellen zusammenzuführen zum Vorbild für den Aufbau und die Integration von Infrastrukturen für ehrenamtliche und bürgerwissenschaftliche Datenerhebungen sowie der darauf aufbauenden Analyse und Kommunikation genommen werden. Zielführend wäre in diesem Zusammenhang eine Erweiterung des Systems z.B. um Verbreitungsdaten wildlebender Tiere.

Daraus ergeben sich allerdings nicht zu unterschätzende Herausforderungen. Diese betreffen neben der kontinuierlichen Finanzierung und Weiterentwicklung auch das Feld des Datennutzungsrechtmanagements, die Einbindung zahlreicher Akteure und nicht zuletzt langfristig gesicherte personelle Ressourcen.

Vielversprechende Schritte in Richtung Weiterentwicklung sind mit der Förderung von Projekten zu Entwicklung von digitalen, teils KI gestützten Arterfassungs- und -erkennungssystemen bereits vorgenommen worden.

Damit noch breitere Nutzer:innengruppen die Datenbestände abrufen können wird an neuen Lösungen für die visuelle und infrastrukturelle Weiterentwicklung gearbeitet.

2.4 UIS für Vollzugsaufgaben – Beispiel: CITES Artenschutzvollzug

Das Washingtoner Artenschutzabkommen (Convention on International Trade in Endangered Species of Wild Fauna and Flora CITES) regelt den Handel mit geschützten Tier- und Pflanzenarten. Der Begriff Handel umfasst in diesem Kontext den Handel mit lebenden oder toten Exemplaren geschützter Arten sowie mit Teilen und Erzeugnissen, wie z.B. Elfenbein, geschützten Holzprodukten oder Kaviar zwischen verschiedenen Staaten.

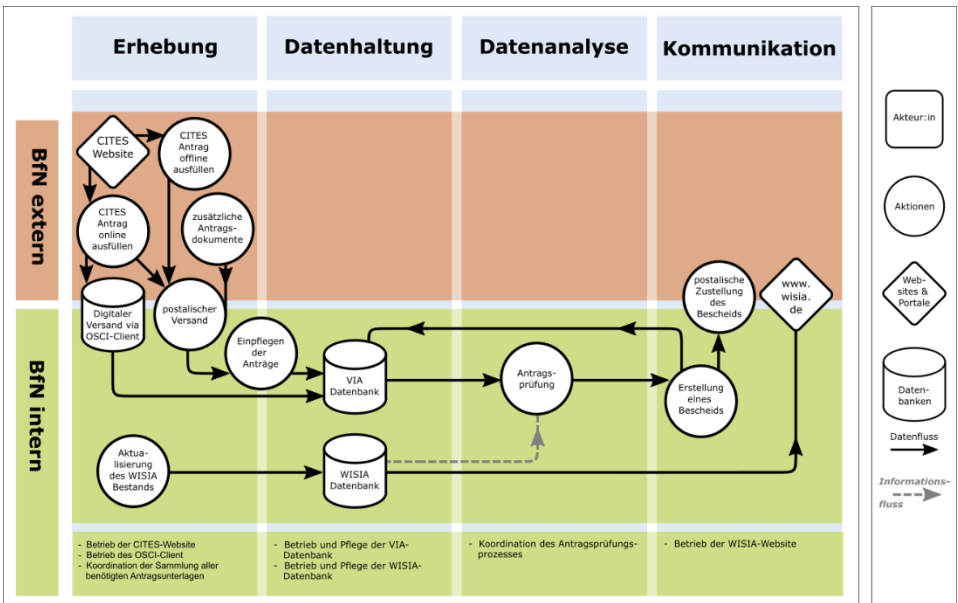


Abb. 3: UIS für CITES-Online - technische und organisatorische Komponenten

Dabei gilt der Grundsatz: je gefährdeter die geschützte Art, desto strenger sind die Handelsbestimmungen. CITES fokussiert also auf Im- und Exporte, einschließlich des Einführens zu privaten Zwecken, z.B. Touristenmitbringsel. Der Handel innerhalb eines Staates wird nicht durch CITES geregelt. In Deutschland wird CITES durch das Bundesnaturschutzgesetz und die Bundesartenschutzverordnung gesetzgeberisch umgesetzt.

Als nationale Vollzugsbehörde für dieses Abkommen erteilt das BfN Genehmigungen für die Ein- und Ausfuhr geschützter Arten und der Produkte, die aus diesen Arten hergestellt werden. Das BfN erarbeitet oder beauftragt auch die Überarbeitung wissenschaftlicher Grundlagen, die zugleich dazu dienen, das Artenschutzrecht fortlaufend an aktuelle Erfordernisse anzupassen und weiter zu entwickeln.

Mit dem CITES-Onlineportal bietet das BfN die Möglichkeit zur elektronischen Antragstellung für CITES-Genehmigungen zur Ein-, Aus- bzw. Wiederausfuhr von Exemplaren geschützter Arten. Neben der Online-Antragstellung stehen den Nutzenden zusätzliche Serviceleistungen zur Verfügung, wie die Übersendung von Nachweisdokumenten gemeinsam mit dem elektronischen Antrag und die lokale Speicherung des gestellten Antrages zur späteren Weiterbearbeitung und Nutzung als Vorlage für neue Anträge. Über eine Anlagenverwaltung können auch im Nachgang einer Antragstellung Anlagen zu bereits abgegebenen Anträgen online eingereicht bzw. zusätzliche Informationen per E-Mail übermittelt werden. Gesetzlich im Original vorzulegende Dokumente müssen per Post nachgereicht werden (Abb. 3 Erhebung).

Die Onlineanträge werden, wie auch die in Papierform eingereichten Anträge, in das getrennte Fachverfahren VIA überführt, dort bearbeitet und ein Genehmigungsbescheid erstellt. Derzeit werden rund 30.000 Genehmigungen im Jahr ausgestellt. Ungefähr 35 Prozent der Anträge werden auf elektronischem Weg via CITES Online gestellt. Die Genehmigungen müssen derzeit noch in Papierform ausgestellt werden, weil in Deutschland, so wie in den meisten der 183 Vertragsstaaten, die Voraussetzungen für elektronische Genehmigungen fehlen.

Im parallelen WISIA-System sind wissenschaftlichen Grundlagen wie die Taxonomien der geschützten Arten hinterlegt. Über die WISIA-Website werden diese Informationen der Öffentlichkeit zugänglich gemacht (Abb. 3 Kommunikation). Es ist somit die zentrale und rechtsverbindliche Informationsquelle, um zu erfahren welche Tier-, Pflanzen- und Pilzarten in Deutschland geschützt sind (<https://www.wisia.de>).

Der Prozess zur elektronischen Antragstellung für CITES-Genehmigungen steht exemplarisch für die Nutzung eines UIS zur Erfüllung von Vollzugsaufgaben durch das BfN. Aktuell wird an der technischen Überarbeitung der Beantragungplattform CITES Online und der VIA-Datenbank des nachgelagerten Prüfungsverfahrens gearbeitet. Darüber hinaus wird auch die technische Überholung der WISIA-Datenbank mit der dazugehörigen Website angestrebt. Die Einführung eines elektronischen Gebührenbescheids könnte darüber hinaus die Akzeptanz des Online-Antragsverfahrens erhöhen und einen Medienbruch vermeiden. Zukünftig wird dieser auch im Benutzerkonto bereitgestellt. Ein weiteres zentrales Vorhaben ist die Anbindung des BfN eigenen CITES-Antragsverfahrens an den e-CITES Prozess der Europäischen Kommission [Ci18]. Es ist das Ziel, eine weitestgehend papierlose Ein- und Ausfuhr geschützter Arten über alle involvierten Akteure auf Länder-, Bundes-, EU- und internationaler Ebene zu ermöglichen. Die Umsetzung in diesem Bereich ist eine große Herausforderung, da allein in Deutschland 238 Länderbehörden involviert sind. Dort existieren unterschiedliche Datenbestände und Datenhaltungsformen, die harmonisiert, zentral zusammengeführt und online bereitgestellt werden müssten.

Parallel zum Online-Antragsverfahren wurde und wird die Entwicklung teilweise automatisierter Erkennungs- und Bestimmungssysteme vom BfN gefördert. Das System CITESwoodID ermöglicht die Bestimmung geschützter Baumarten anhand von Holzanschnitten. Einen ähnlichen Zweck erfüllt auch das BfN-geförderte Bestimmungstool IvoryID. Auf Basis der Isotopenverhältnisse kann die räumliche Herkunft von Elfenbein nachgewiesen und mit den Genehmigungspapieren abgeglichen werden.

Ein wichtiges Feld des Artenschutzvollzugs ist auch die Bekämpfung des illegalen Wildtierhandels im Internet. Gestützt durch die umweltpolitische Digitalagenda des BMU soll dafür eine eigene Taskforce aufgebaut werden.

2.5 Ausblick

UIS sind von zentraler Bedeutung für die Arbeit des BfN. Die vorgestellten Beispiele stehen exemplarisch für eine Reihe weiterer Informationssysteme wie z.B. für das Monitoring nach Vogelschutzrichtlinie oder das marine Monitoring nach EU-Meeresstrategie-Rahmenrichtlinie. Der anhaltende Rückgang der Biodiversität wird in Zukunft den Bedarf nach amtlich gesicherten Informationen und wissenschaftlich fundierten Handlungsempfehlungen noch weiter steigen lassen. Aus diesem Grund ist das BfN auf die stete Weiter- und Neuentwicklung seiner UIS angewiesen. Hierfür spielen neue, naturschutzorientierte Digitalanwendungen und Technologien eine entscheidende Rolle.

3 Literaturverzeichnis

- [Bf17] BfN, Naturschutzgebiete in der ausschließlichen Wirtschaftszone (AWZ) der Nord- und Ostsee, <https://www.bfn.de/themen/meeresnaturschutz/nationale-meeresschutzgebiete.html>, Stand: 17.06.2017.
- [Bf21] BfN, Artenschutzbestimmungen der Fauna-Flora-Habitat-Richtlinie, <https://www.bfn.de/themen/artenschutz/regelungen/ffh-richtlinie.html>, Stand: 14.06.2021.
- [Bu18] Buttler, K. P.; May, R.; Metzger, D.: Liste der Gefäßpflanzen Deutschlands Florensynopse und Synonyme. BfN Skripten 519, 286 Seiten, Bonn, 2018.
- [Ci18] CITES Secretariat: Automation of CITES permit procedures and electronic information exchange for improved control of international trade in endangered species (eCITES). Genf, 2018.
- [EC19] European Commission: Promotion of good practices for national environmental information systems and tools for data harvesting at EU level. Brüssel, 2019.
- [EP07] Europäisches Parlament und Europäischer Rat: Richtlinie 2007/2/EG vom 14. März 2007 zur Schaffung einer Geodateninfrastruktur in der Europäischen Gemeinschaft (INSPIRE). ABl. L 108, Seiten 1-14, 25.4.2007.
- [FW21] FloraWeb, Datenquellen, <https://www.floraweb.de/ueberfloraweb/datenquellen.html>, Stand: 20.06.2021.
- [Ge21] GEMET, environmental information system, <http://www.eionet.europa.eu/gemet/concept/2853>, Stand: 08.06.2021.
- [JS15] Jacobs, C.; Schotthöfer, A.: Citizen-Science-Daten zur Biodiversität – Methoden zur Unterstützung der Qualitätssicherung. AGIT – Journal für Angewandte Geoinformatik. Seiten 470-479. Berlin, 2015.
- [MS07] May, R.; Schneider, T.: FloraWeb – the German web flora and the role of Recorder. Ferrantia 51, Seiten 115-121. Luxembourg, 2007.

envVisio: Universelle Bereitstellung von Umweltdaten

Management mit Umweltdaten, das alle Informationen themenübergreifend verknüpfbar und einheitlich strukturiert verarbeitet, über standardisierte Dienste mit einfach weiterverarbeitbaren Inhalten bereitstellt.

Janik Großmann¹, Werner Koch², Heino Rudolf³

Abstract: Bereits während der Arbeit am INSPIRE-Datenmodell „Umweltüberwachung“ wurde deutlich, dass die Datenbereitstellungen entsprechend der Geo-Normen der ISO 191xx für Umweltdaten ungeeignet sind. Die Datenstrukturen der INSPIRE-Anhänge sind phänomenal auf einige Anwendungsfälle zugeschnitten. Komplexe Ursache-Wirkbeziehungen, Zustandsbeschreibungen im Ökosystem und der Umgang mit dynamischen Daten, Strukturen und Anforderungen sind nur ungenügend (wenn überhaupt) abbildbar. Deshalb entwickelte RUDOLF in den letzten Jahren einen völlig neuartigen Ansatz zur Bereitstellung von Daten über die Umwelt: envVisio revolutioniert das Umweltdatenmanagement, denn alle Daten werden einheitlich weiterverarbeitbar bereitgestellt. envVisio wird in [1] ausgearbeitet, aktuell wird vom BMVI die Implementierung gefördert [2]. Die Methode ist in der Praxis mehrfach umgesetzt, am Beispiel der Datenförderierungsschicht DFS des LANUV Nordrhein-Westfalen kann diese präsentiert werden. – envVisio ist ein Meilenstein bei der nachhaltigen digitalen Transformation von Objekten und Prozessen unserer Umwelt.

Keywords: Umweltdatenmanagement, Datenbereitstellung, INSPIRE, Digitalisierung, digitale Transformation, Interoperabilität, Datenweiterverarbeitung, Ökosystemanalysen, Ursache-Wirkbeziehungen, Umweltzustand, envVisio

1 Digitalisierung und digitale Transformation

Auszug aus [1] Kapitel 9:

Die Digitalisierung bringt uns die vierte industrielle Revolution (auch ‚Industrie 4.0‘ bezeichnet). Es ist sehr auffällig, dass immer technische Innovationen die Triebkräfte dieser Revolutionen waren/sind und diese dann gewaltige gesellschaftliche, soziale, wissenschaftliche und wirtschaftliche Umwälzungen nach sich ziehen. Diese Kräfte waren bei den ersten drei industriellen Revolutionen für jedermann sicht- und anfassbar: Maschinen, Motoren, Computer...

Das ist jetzt ganz anders, denn Industrie 4.0 findet im Virtuellen statt. Die wahren technischen Hintergründe erschließen sich nicht auf den ersten Blick. Wenn aber die Potenziale der Digitalisierung erschlossen und die Auswirkungen beeinflusst werden

¹ Simplex4Data GmbH, Am Waldschlösschen 4, 01099 Dresden., janik.grossmann@simplex4data.de

² Simplex4Data GmbH, Am Waldschlösschen 4, 01099 Dresden., werner.koch@simplex4data.de

³ hrd.consulting, Am Waldschlösschen 4, 01099 Dresden, heino.rudolf@hrd-consulting.eu

sollen, müssen die (technischen) Ursachen und Entwicklungen gekannt und umgesetzt werden.

Folgende drei Aspekte bzgl. der Digitalisierung, ihrer Wirkmechanismen und der Historie seit Industrie 3.0 sind im Zusammenhang mit Umweltdaten von Bedeutung:

1. Umwandeln von analogen Werten in digitale Formate; Erstellen digitaler

Repräsentationen von physischen Objekten, Ereignissen, analogen Informationen:

Die ersten Digitalisierungen bezogen sich auf einfache Umwandlungen von analogen Werten in digitale Formate, zunächst einfache Zeichen und Zahlen in binären Code, später Texte, Musikstücke, Fotos, Karten, Filme ... Es entsteht ein binäres Abbild des analogen Gegenstands.

Wesentlich ist die Definition des Formats, wie dieser Gegenstand codiert wird. Damit ist ein entscheidender Aspekt der Digitalisierung gegeben: ein verständliches, leicht und effizient weiterverarbeitbares und auf zukünftige Anforderungen überführbares binäres Format.

Bezüglich der Umweltinformationen kann festgestellt werden, dass ein gewaltiger Anteil der Informationen heute digital vorliegt. Das betrifft Geodaten, Grenz- und Messwerte, Bewertungsergebnisse..., aber auch Daten zu Umweltprozessen und ihren Visualisierungen, zu Handlungen im Umweltrecht, -vollzug, in der Umweltüberwachung...

2. Wandel in Wirtschaft, Wissenschaft, Gesellschaft und Politik zu digitalen

Prozessen:

Zunächst waren Digitalisierungen auf Rechentechnik fokussiert. Durch die immer komplexeren Formate, die Bereitstellung großer Speicherkapazitäten und die gewaltige Erhöhung der digitalen Verarbeitungsgeschwindigkeiten wurden immer neue Bereiche digitalisiert und digitale Prozesse implementiert: Geodatendarstellung und -bereitstellung, Funk und Fernsehen, Musik-, Foto-, Kommunikationsindustrie, Aspekte des Autofahrens, Verkehrssteuerungen, Prozesssteuerungen in der Industrie...

Auch die Umweltdaten sind davon betroffen: digitale Gewinnung und Verarbeitung von zustands- und prozessbeschreibenden Daten sowie eGovernment-Prozesse der Verwaltung bei Analysen, Berichterstattungen, Überwachungen u. v. a. m.

3. Digitale Transformation: durch die Digitalisierung ausgelöster Umbruch, der einen

Wandel der Technik sowie (fast) aller Lebensbereiche bewirkt:

Technisch eingeleitet wurde die digitale Transformation durch die Vernetzung von Geräten, Maschinen, Unternehmen usw. Werden die Daten in standardisierten digitalen Formaten bereitgestellt, können sie jederzeit ausgewertet und weiterverarbeitet werden. So werden Prozesse angestoßen und gesteuert; es werden auch beschreibende Zustandsdaten zur Optimierung der Prozesse herangezogen u. v. m. Nach 2. sind immer mehr Maschinen und Geräte mit DV-Programmen ausgestattet, sodass sie digitale Daten auswerten, über Daten gelenkt werden und auch wieder Daten zur Weiterverarbeitung erzeugen. Der Trend ist heute längst erkennbar: Wenn so viele digitale Daten wie möglich gespeichert und digitalen Prozessen zur Verfügung gestellt werden, kann die Realität

virtuell umfassend erklärt; können immer mehr real ablaufende Prozesse vorhergesagt, gemanagt und beschrieben werden. Und das betrifft (noch mit unterschiedlicher Intensität) alle Bereiche der Gesellschaft: beginnend mit Speicherungen menschlicher Verhaltensweisen, über Gesundheitsdaten, Verkehrs- u. a. Prozessabläufe usw. - aber auch Erfassung der Umweltvorgänge und -zustände.

2 So wird es heute gemacht...

2.1 Beispiel INSPIRE

Weiterführende Informationen bzgl. der Digitalisierung und ihrer Umsetzung via INSPIRE können den Kapiteln 6...9 der Quelle [1] entnommen werden.

Da die datenhaltenden Stellen mittlerweile alle INSPIRE-Dienste kostenlos als Open Data zur Verfügung stellen und INSPIRE eine gesetzliche Aufgabe ist, sollte INSPIRE der Königsweg für den Beitrag der Umweltbehörden zur Digitalisierung sein. Nur scheint das so in der Praxis nicht zu funktionieren: wie viele Behörden versuchen herauszufinden, ob sie sich beteiligen müssen oder wie sie das Alles so minimal wie möglich umsetzen können. – Was für eine vergebene Chance... In [1] Kapitel 6 werden Gründe für diese Verhaltensweise zusammengetragen:

- INSPIRE hat keinerlei Bedeutung für die eigentliche Arbeit der Fachbehörden.
- Nicht einmal die Berichterstattungen können mit INSPIRE umgesetzt werden.
- Der Wunsch nach einem themenübergreifenden Umweltdatenmanagement wird durch INSPIRE nicht erfüllt.
- Die Art und Weise der Modellierung steht im Widerspruch zur Arbeitsweise der Fachbehörden, die „Dynamik in den Daten“ verlangt.
- Die Modelle sind so kompliziert notiert, dass kaum einer die Zeit und Mühe investieren möchte, sie zu lesen – geschweige denn, sie zu verstehen.
- INSPIRE ist etwas Zusätzliches – eine Pflicht, die keine Vorteile in der Arbeit bringt!

Und letztendlich ist auch eine Weiterverarbeitung der komplizierten und phänomenal an Anwendungsfällen orientierten INSPIRE-Datenstrukturen zur Schaffung neuer Anwendungen und Wertschöpfungen (wie das mit der Richtlinie angedacht ist) nur mit enormem Aufwand (wenn überhaupt) umsetzbar. INSPIRE hat fast alles Notwendige für die Digitalisierung bestens umgesetzt: Schaffung der technischen Infrastrukturen, der Organisationsstrukturen, Verwendung von Standarddiensten. ABER: es ist leider nicht geglückt, einfache, interoperable und weiterverarbeitbare Datenstrukturen zu kreieren. Doch damit stehen und fallen die Prozesse der digitalen Transformation, bzw. sie verschlingen riesige Kosten.

2.2 Warum wird es heute so gemacht?

Auf analoge Art und Weise zu INSPIRE werden auch Daten für andere Normen (z. B. XÖV) oder Berichterstattungen (z. B. zur WRRL) strukturell vorgegeben. Es werden einerseits immer mehr Datenpools für die Bereitstellung zumeist von Open Data angelegt, aber kaum einer davon kann vernünftig weiterverarbeitet werden. Andererseits werden für den Datenaustausch (höchst komplizierte) Datenübertragungen, also Schnittstellen, spezifiziert, die natürlich von der empfangenden Stelle verarbeitet werden können – hier liegt dann aber der Aufwand bei den Daten bereitstellenden Stellen, die ihre Daten in diese komplizierten Strukturen erst einmal überführen müssen.

In [4] hat RUDOLF die Ursachen für diese “Datenfriedhöfe“ bzw. Datenschnittstellen, die mit riesigen Kosten einhergehen, ausgearbeitet. Sie sind bereits in der Art und Weise der Datenmodellierung begründet.

1. Die Modellierung erfolgt in den Geo-Normen einstufig:

Diese Herangehensweise ist sehr speziell, denn Datenmanagementlösungen werden i. Allg. zweistufig kreiert: ein Modell der Datenhaltung – (darauf aufsetzende) Modelle der Datenbereitstellung. Wenn aber das ursprüngliche Ziel der Geo-Normen reflektiert wird, nämlich Geometrieobjekte beliebiger Geographischer Informationssysteme einheitlich anzubieten, dann ist eine Stufe für die Modellierung der bereitzustellenden Datenstrukturen völlig logisch. (Vgl. mit [4].)

2. Für jedes Fachthema gibt es ein eigenes, in sich geschlossenes Datenmodell:

Das ist eine zwangsläufige Konsequenz aus der einstufigen Modellierung, da die Einstufigkeit und parallel die Fokussierung auf ein Thema zu fachspezifischen Modellen führen muss. Die ursprünglich geometrisch geprägten Objekte werden um Sachdaten „angereichert“, wie es heißt. Und wie kann diese „Anreicherung“ unter den gegebenen Prinzipien anders umgesetzt werden als über spezielle Modelle für die verschiedenen Fachthemen? Bei INSPIRE werden in der Richtlinie [3] dafür Annexe definiert: 32 Fachthemen mit eigenen Datenmodellen. (Vgl. mit [4].)

3. Die Datenmodelle sind nicht dynamisch:

Die Geometrie ist zumeist die objektbildende Kategorie; der Fokus liegt auf geographisch geprägte Objekte mit ihren statisch zugeordneten Objektklassen und Verbindungen. Flexible Abfragen zu speziellen aktuellen Themen lösen (fast immer) zeitaufwendige Prozesse aus, um die Daten zusammenzustellen. Es ist kaum möglich, Objekte so miteinander zu verknüpfen, dass gewünschte und angefragte Ursache-Wirkbeziehungen abgebildet werden können.

2.3 Wie wird es heute umgesetzt?

Die oben diskutierte Art und Weise führt zu der für Umweltdatenmanager etwas überraschenden Praxis, bei Datenbereitstellungen nach einem anderen als dem eigenen Modell (z. B. für Reportings) die Daten von der eigenen Struktur in diese abweichende Struktur zu transferieren. (Die Geodatenbearbeiter nennen das „Harmonisieren“, was auch zum Schlagwort für die Bereitstellung von INSPIRE-Daten wurde.) Wie viele „Harmonisierungen“ sind umzusetzen, wenn beispielsweise die mehr als fünfzig Umweltberichterstattungen an die Europäische Kommission digital und automatisiert erfolgen sollen? Und jede „Harmonisierung“ ist kostenintensiv. (Vgl. mit [4].)

Die folgende Abbildung skizziert die heute übliche Bereitstellung von Umweltdaten:

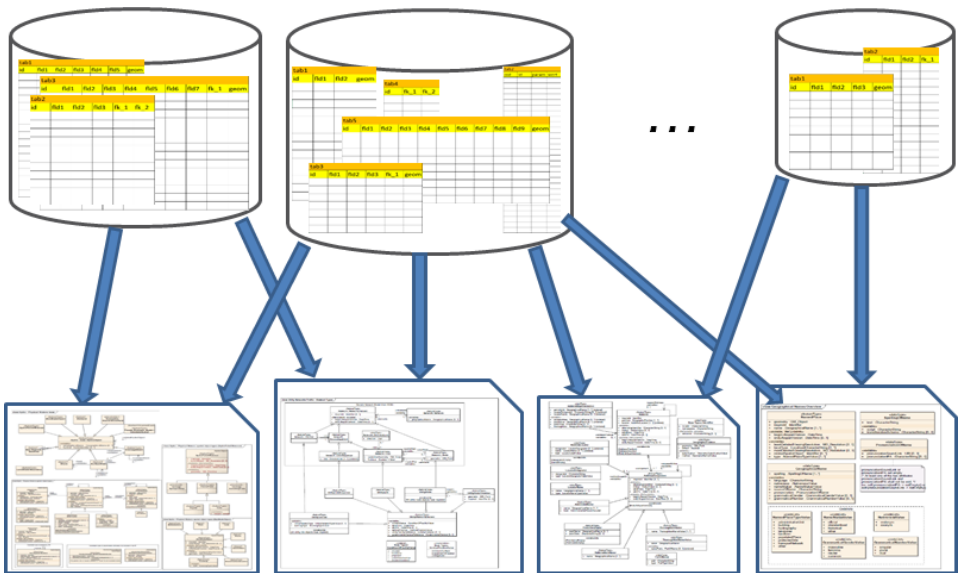


Abb. 1: Datenbereitstellungen via „Harmonisierung“

Die Töpfe in der oberen Reihe symbolisieren die unterschiedlichsten Fachsysteme mit immer wieder eigenen, der Fachaufgabe angepassten Datenstrukturen. Und die Praxis zeigt beispielsweise am Projekt DFS des LANUV, dass schon in einer Umweltverwaltung mehr als 30 Fachsysteme im Einsatz sind. Die Kästchen der unteren Reihe stehen für Modelle zur Datenbereitstellung, hier mit ausgewählten INSPIRE-Spezifikationen (im LANUV sind etwa 18 INSPIRE-Annexe zu bedienen, hinzukommen diverse Datenbereitstellungen für Open Data und Berichtspflichten). Und tatsächlich werden immer wieder Umsetzungsalgorithmen oder spezielle ETL-Prozesse programmiert. Ihre Anzahl ist schon gewaltig (bei m Fachsystemen und n Datenbereitstellungen können es theoretisch $m \cdot n$ Prozesse werden), aber auch ihre Implementierung ist nicht einfach. Es wird deutlich, welche Ressourcen hier gebunden werden.

3 So kann es anders gemacht werden...

3.1 Wie kann die Datenbereitstellung vereinfacht werden?

Das Prinzip ist die Datenbereitstellung über einen zentralen Datenpool:

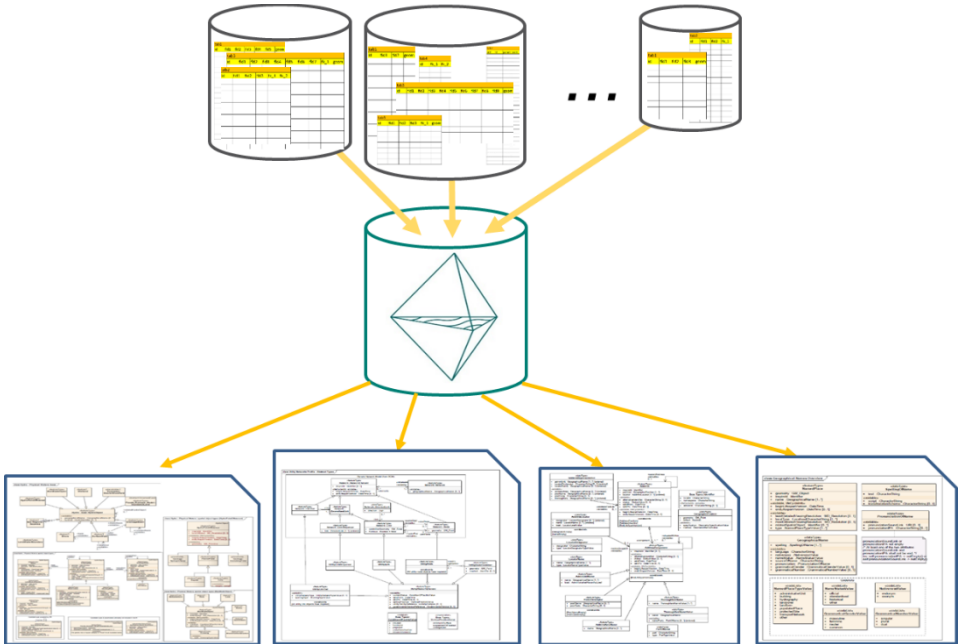


Abb. 2: Ein (wirklich) harmonisiertes Datenmanagement

Es werden die bereits vorhandenen Exportformate bzw. bereitgestellte Datenbanksichten der Fachsysteme (in der oberen Reihe) genutzt. Die exportierten Daten werden in einen zentralen Datenpool importiert. Niemand muss bei den Fachsystemen etwas anpassen oder verändern. Handelt es sich um standardisierte Datenbereitstellungen (z. B. nach INSPIRE oder für Reportings), so sind diese im zentralen Datenpool bereits hinterlegt. Neu eingespielte Daten stehen somit unmittelbar für die Datenbereitstellung bereit. Es wird deutlich, dass sich die Anzahl auf m Importe + n Datenbereitstellungen reduziert, wobei die Datenbereitstellungen für Standards einmalig zu erstellen und damit bereits im Funktionsumfang des zentralen Datenpools enthalten sind.

3.2 Wie sieht der zentrale Datenpool aus?

Der Schlüssel dieser Methodik steckt in der Struktur des zentralen Datenpools, der alle Daten über die Umwelt aufnehmen kann. Es wird sicherlich schnell klar, dass dieser nicht mit den oben charakterisierten Modellierungsmethoden erstellt werden sollte, sondern nach völlig anderen Grundsätzen:

1. Wir haben ein (!) Realitätsmodell „Ökosystem“, also genau ein einziges Datenmodell zur Verwaltung von Umweltdaten (z. B. statt der 34 INSPIRE-Annex-themen).

Nun ist dieses eine (!) Realitätsmodell, also ein (!) Datenmodell zur Beschreibung unserer Umwelt, nicht so einfach zu finden. Auf der Basis einer intensiven Ökosystemanalyse hat RUDOLF in [1] dieses Datenmodell ausgearbeitet und in [4] die praktische Umsetzung erläutert. Worauf fußt dieser Ansatz?

- Strukturen in der Natur: Vielfältige Erscheinungen bestehen aus wiederkehrenden Strukturen.

Viele in unserer Umwelt zu beobachtende Formen, Gestaltungen und Erscheinungen (Abbildung 3) sind mit herkömmlichen Geometrien nicht abbildbar, aber interessanter Weise sind Selbstähnlichkeiten zu finden. Mit der fraktalen Geometrie ist es möglich, immer tiefer in die Details einzudringen und unten dargestellte Umweltphänomene abzubilden. Deshalb kam der Gedanke, auch bei der Modellierung von Daten, die unsere Umwelt beschreiben, auf Methoden der fraktalen Geometrie zurückzugreifen.

(Auszug aus [4])

RUDOLF stellt ein Datenmodell auf, das immer wieder die Form eines Doppeltetraeders aufweist. Analog zu den Fraktalen können wir uns immer tiefer in die Details hineinmodellieren, es entstehen immer wieder analoge Grundstrukturen – im Datenmodell ein Doppeltetraeder.



Abb. 3: Natürliche Intelligenz: Fraktale Geometrie. Quelle [4]

Quellen: Kohl (Josef Schiller), Kristallwachstum (Thanks for your Like • donations welcome), Farn Spirale (leprechaun77), Margarite Pink (Mylene2401), Schnecke Hell (Dean Marston), Kaktus (MoneyforCoffee), Kaktuspflanze mit feinen, langen, gedrehten Blütenblättern (ralu_michael), Pusteblume die Blüte (Ich bin dann mal raus hier.), Gerissener Boden (Peter H) auf Pixabay [8]

Grafische Gestaltung: Sirid Fillinger Visuelle Kommunikation

- **Kreisläufe in der Natur:** Im Ökosystem sind die Prozesskreisläufe mit Stoffen und Energien die Grundlage allen Lebens.

Wesentlich ist das Verständnis der Stoff- und Energie-Kreisläufe, was zu einem ganzheitlichen Denkansatz bzgl. der Prozesse und Objekte der Umwelt führt und auch Ursache-Wirkbeziehungen als ein Kernelement aufnimmt.

Dieser Denkansatz muss selbstverständlich zu einem (!) Realitätsmodell für Umweltdaten führen (anstatt der 32 INSPIRE-Annex-Modelle). Im Realitätsmodell werden die Objektklassen immer entsprechend real auftretender Objekte, Prozesse, Handlungen und ihrer Wirkketten im Ökosystem definiert.

(Auszug aus [4])

2. Wir modellieren zweistufig und arbeiten für die Datenbereitstellungen Anwendungsmodelle aus.

Bei der zweistufigen Modellierung wird die Datenhaltung so gestaltet, dass sie die reale Welt abbildet. Im so genannten Realitätsmodell wird, von einem Verständnis ökosystemarer Zusammenhänge und der Strukturen in unserer Natur ausgehend, ein UML-Diagramm für das Umweltdatenmanagement kreiert.

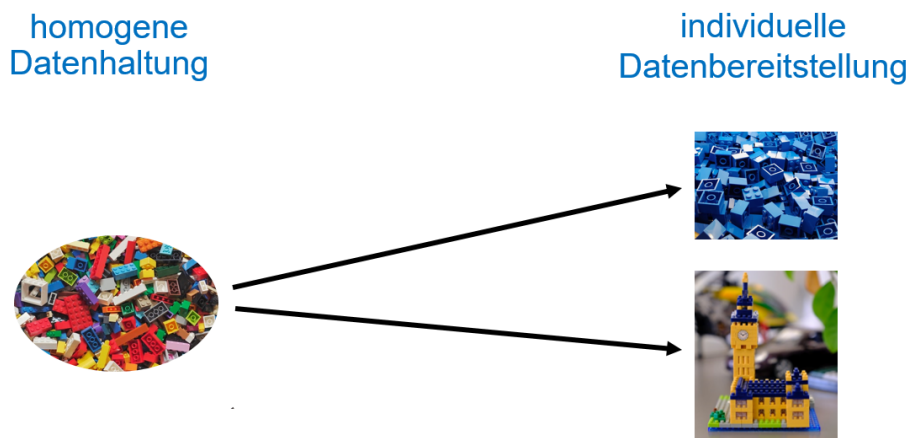


Abb. 4: Zweistufiges Datenmanagement, mit Lego symbolisiert

Die Datenhaltung können wir uns wie eine Kiste mit Legobausteinen vorstellen. Die Objekte in der Natur werden „atomar“ als einzelne Legobausteine in Kisten verwaltet und können beliebig miteinander zusammengesteckt/verbunden werden. Verrückter Weise ist für diese Art der Datenbereitstellung auch keine Datenmodellierung mehr notwendig. Das erschien den Autoren vor einiger Zeit noch undenkbar und mutet ein wenig wie Zauberei an... Es ist einfach nur noch notwendig, die Legosteine den Kästen, die nach den Kreisläufen im Ökosystem gebildet und im Doppeltetraeder modelliert sind, zuzuordnen.

Für die Datenbereitstellung werden dann individuelle Sichten kreiert, z. B. alle blauen Legosteine oder zusammengesetzte Legoobjekte. Damit können jederzeit beliebige Szenarien definiert werden. Die Sichten für Standard-Datenbereitstellungen werden einmal angelegt und sind dann für jede neue Datenbereitstellung bereits gegeben, was den oben beschriebenen Prozess der „Harmonisierung“ von Geodaten wirklich harmonisch umsetzt.

Es können aber auch ad-hoc Ansichten für konkrete Anfragen, Betrachtungen, Diskussionen, Entscheidungsfindungen u. v. m. schnell und einfach aus den Legosteinen mit immer wieder anzuwendenden Methoden und vorgefertigten Tools gebildet werden.

3.3 Was symbolisiert der Doppeltetraeder?

Das envVisio-Realitätsmodell basiert auf einer sehr simplen Grundstruktur, die im Doppeltetraeder (Abb. 5) symbolisiert wird. Das Dreieck in der Mitte beschreibt die Objekte und Prozesse im Ökosystem und bildet damit die Stoff- und Energie-Kreisläufe

sowie Ursache-Wirkbeziehungen ab. Die dafür definierten Objektklassen sind schematisch durch den Doppeltetraeder skizziert und werden in [1] Caput XIV definiert.

- Betrachtungsobjekte
- Umweltprozesse
- Handlungen zur Bewertung, Messung, Steuerung oder Regulierung
- Der Kopf des Doppeltetraeders enthält Metadaten, zentrale Register/Bezüge, Schlüssel Listen und Ontologie-Tabellen, die die Strukturen des unter ihm liegenden Dreiecks kennen.
- Der Fuß des Doppeltetraeders ist das „Bindeobjekt“ zur Realität: ihre Repräsentanz. Dort werden Ausprägungen wie Geometrien, sachliche Informationen, einzelne Werte (Mess-, Grenzwerte o. ä.), Visualisierungen, Dokumente, Filme u. v. a. m. verwaltet.

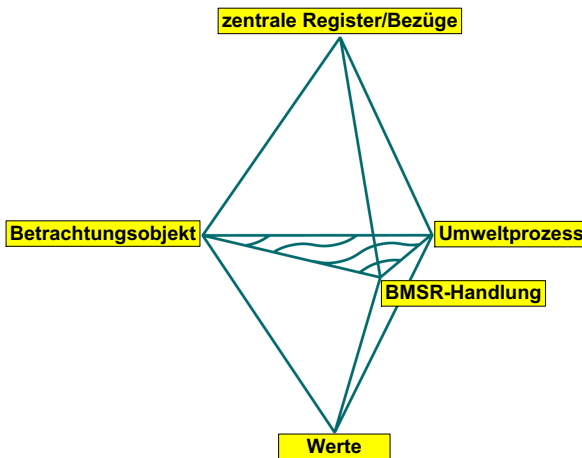


Abb. 5: Doppeltetraeder des envVisio Realitätsmodells. Quelle [1]

4 envVisio: So wird es anders gemacht...

Im Rahmen des Förderprojektes [2] wird aktuell die envVisio-Plattform implementiert. envVisio ist ein Datawarehouse, das nach den oben erläuterten Modellansätzen aufgebaut ist. Das Datawarehouse ist zweistufig gegliedert (Simplex zur Datenhaltung entsprechend dem Doppeltetraedermodell und Szenario zur Bildung von Anwendersichten). Um das Datawarehouse gruppieren sich notwendige Tools: der Importer zum Einspielen der Daten, eine Visualisierungskomponenten zum Anschauen der eingespielten Daten und Services zur Bereitstellung der Daten. Die folgende Abbildung skizziert envVisio.

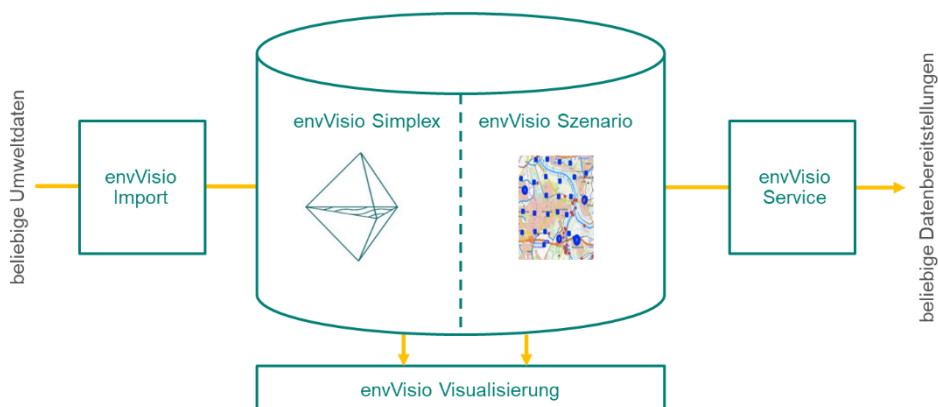


Abb. 6: envVisio – das Datawarehouse zur Umweltdatenbereitstellung

- **envVisio Import:** Über den Importer können Dokumente in einer Reihe von Formaten hochgeladen und in die Datenbank eingespielt werden. Aktuell werden folgende Formate unterstützt: (Geo)JSON, CSV, shapefiles und XML. Der Importer kann auch API-Features-Dienste (wie z. B. das Open Data Portal Dresden) ansteuern und deren Daten automatisiert abfragen. Der Konverter bietet ein Interface, um die Daten in die Doppeltetraederstruktur zu übertragen. Da das envVisio Datenmodell mit einfachen, wiederkehrenden Strukturen arbeitet, ist die Zuordnung (oft auch als Mapping bezeichnet) relativ überschaubar und wiederholt sich. Wesentlich ist beim Konvertieren aber auch die Definition der Verbindungen, die durch die Eingabe von Fremdschlüsseln umgesetzt wird.
- **envVisio Simplex:** Das ist die Kernkomponente von envVisio, mit der Datenhaltung nach dem Doppeltetraedermodell.
- **envVisio Szenario:** Da alle Objektklassen und Verbindungen eine einheitliche Grundstruktur in envVisio Simplex haben, können die Objekte und Verbindungen beliebig zusammengefasst, die Verbindungen kurzgeschlossen und damit neue Objektverknüpfungen erzeugt werden. Auf dieser Basis können sehr flexibel Sichten definiert und zu Szenarien vereint werden. Und dafür wird eine Bedienoberfläche angeboten, sodass auch relativ schnell und komfortabel Sichten für ad-hoc-Datenanfragen gebildet werden können.
- **envVisio Visualisierung:** Dieses Modul ist notwendig, um sich die einmal im Datawarehouse aufgenommenen und zu Szenarien zusammengestellten Daten ansehen zu können. Die Visualisierungskomponente arbeitet aber ganz anders als herkömmliche

Geodatenviewer, denn es ist z. B. keine Layerdefinition und -auswahl mehr vorzunehmen. Alle betroffenen Objektklassen können direkt angezeigt werden. Und man kann zwischen den Objektklassen entlang ihrer Verbindungen wechseln (z. B. von einem Gewässer zu den zugeordneten Messstellen, zur Adresse, weiter zur Straßenkante, von dort dann ggf. auf in der Straße gemessene Immissionen, auf Schutzgebiete ...). Außerdem kann über envVisio Visualisierung die Bildung von neuen Szenarien grafisch unterstützt werden.

- **envVisio Service:** Über den Service können alle Daten in envVisio selektiert werden und stehen zur Weiterverarbeitung zur Verfügung. Dieser Dienst ist als API-Feature-Service implementiert. API-Features-I ist der erste Dienststandard des OGC, der in wesentlichen Aspekten auf REST beruht. Es eröffnet Chancen, aus dem geoinformatischen Spezialbereich in ein breiteres Spektrum an Anwendungen und Akteuren zu gelangen. Das ist insbesondere für die Weiterverarbeitung von Umweltdaten höchst interessant, denn es macht die Daten für Amateure, Ehrenamtliche und professionelle Nutzer aus ganz anderen Fachrichtungen zugänglich. Die folgende Abbildung ist ein Auszug aus der Spezifikation des Dienstes und gibt eine Übersicht über die Abfragemöglichkeiten und die dahinter stehende Logik im Service. (Vgl. auch mit [5].)

```

/scenarios
/scenarios/{scenarioId}/
/scenarios/{scenarioId}/collections
/scenarios/{scenarioId}/collections/{collectionId}
/scenarios/{scenarioId}/collections/{collectionId}/neighbours
/scenarios/{scenarioId}/collections/{collectionId}/links
/scenarios/{scenarioId}/collections/{collectionId}/items
/scenarios/{scenarioId}/collections/{collectionId}/items/links
/scenarios/{scenarioId}/collections/{collectionId}/items/{featureId}
/scenarios/{scenarioId}/collections/{collectionId}/items/{featureId}/neighbours
/scenarios/{scenarioId}/collections/{collectionId}/items/{featureId}/links

```

Abb. 7: URL-Templates des envVisio Service

5 Fazit und Ausblick

envVisio macht es möglich, dass beliebige Daten über unsere Umwelt verfügbar gemacht werden. Aus den verschiedensten Fachverfahren und Anwendungen können die Daten in envVisio eingespielt werden. Damit sind alle Voraussetzungen gegeben, um die oft versteckten Datensätze zu heben.

Sind die Daten einmal in envVisio eingespielt, so können sie auch miteinander verknüpft werden. Wir können uns das wie ein Spinnennetz vorstellen, das ständig vergrößert und durch neue Maschen und neue Knoten weiter detailliert werden kann. Und alleine durch Verknüpfungen von Daten aus verschiedenen Quellen auf zentrale Register entstehen neue Zusammenstellungen (z. B. beim Einspielen von Daten zu Gewässern aus verschiedenen Anwendungen können diese alle gemeinsam vom Gewässer aus aufgerufen werden; oder es können alle Messergebnisse zu einem Parameter über alle Umweltmedien wie Wasser, Boden, Luft auf einmal selektiert werden).

Eine wesentliche Komponente ist der envVisio Service, über den die Daten selektiert und weiterverarbeitet werden können. Da das envVisio Datenmodell auf den immer gleichen, sich wiederholenden Strukturen beruht und der Dienst diese berücksichtigt, kann er beliebige Datenhaltungen entsprechend der envVisio Struktur abbilden. Die Menge und Heterogenität der Daten ist dabei unerheblich: das Modell kann sie bewältigen, der Dienst übernimmt sie nur. Und damit können die Daten in weiterverarbeitbaren Strukturen abgerufen werden, womit jetzt tatsächlich neuen Wertschöpfungen alle Türen geöffnet sind. (Vgl. mit [5].)

envVisio ist keine konkurrierende Applikation, sondern kann jederzeit ergänzend in die vorhandene Infrastruktur eingebunden werden. Mit envVisio wird im Gegenteil die IT-Welt bereichert, indem alle Daten zur Verfügung gestellt und über den envVisio Service überall auf einfache Art und Weise weiterverarbeitet werden können. Hinzu kommt sogar noch, dass die Quelldaten aus den Fachsystemen via deren Exportstrukturen für den envVisio Importer zur Verfügung gestellt werden können. D. h., dass keinerlei Programm-
anpassungen, Schnittstellen, ETL-Prozesse o. ä. bei der Bereitstellung der Daten notwendig werden. Die Fachsysteme werden nicht angefasst.

Sind die Daten in einem Standardformat bereitzustellen (z. B. für INSPIRE), so können sie lediglich in envVisio eingespielt werden. In envVisio Szenario wird die spezielle Sicht des Standards einmal implementiert und ist so im Funktionsumfang enthalten. Direkt nach dem Einspiel stehen die Daten im gewünschten Dienst bereit. Kein Fachexperte muss sich mit den Standardspezifikationen auseinandersetzen, er kann sich aber die Datensichten in envVisio Visualisierung anschauen.

Warum spielen wir die Daten dann überhaupt noch in den komplizierten Strukturen aus, wenn mit dem envVisio Service mindestens dieselben Dateninhalte viel einfacher und kostengünstiger bereitgestellt werden können? Im mFUND-Projekt [2] ist es vorgesehen, den envVisio Service zum Standard auszubauen. Wir können die INSPIRE-Infrastruktur nutzen, haben jetzt aber auch die Formate, die leicht zu verarbeiten sind. Es entsteht genau das, was INSPIRE anstoßen und erreichen wollte!

Zurück zu unserer Einleitung bzgl. der digitalen Transformation der Gesellschaft (siehe Kapitel 1): Mittels envVisio werden die Daten in standardisierten digitalen Formaten bereitgestellt und können jederzeit ausgewertet und weiterverarbeitet werden. So werden Prozesse angestoßen und gesteuert. Wenn so viele digitale Daten wie möglich gespeichert

und digitalen Prozessen zur Verfügung gestellt werden, kann die Realität virtuell umfassend erklärt; können immer mehr real ablaufende Prozesse vorhergesagt, gemanagt und beschrieben werden.

6 Quellenverzeichnis

- [1] Rudolf, H.: Umweltdatenmanagement: Eine Geo-Inspiration. Bernhard Harzer Verlag GmbH Karlsruhe (2018).
- [2] envVisio-GI - GeodatenIntelligenz für interoperables Datenmanagement mit der envVisio-Methode – envVisio-GI, mFUND-Projekt des Bundesministeriums für Verkehr und digitale Infrastruktur, <https://www.bmvi.de/SharedDocs/DE/Artikel/DG/mfund-projekte/envvisio-gi.html>
- [3] Richtlinie 2007/2/EG des Europäischen Parlaments und des Rates vom 14. März 2007 zur Schaffung einer Geodateninfrastruktur in der Europäischen Gemeinschaft (INSPIRE). In: ABl. der EU, L 108, S.1-14 (2007).
- [4] Rudolf, H.: Umweltdaten-Intelligenz. In: Freitag, U., Fuchs-Kittowski, F., Abecker, A., Hosenfeld, F. (Hrsg.): Umweltinformationssysteme – Wie verändert die Digitalisierung unsere Gesellschaft? Springer Vieweg, Wiesbaden 2020.
- [5] Großmann, J.: envVisio Service – Ein universeller Dienst für Umweltdaten. Digitale Tagung des Arbeitskreises Umweltinformationssysteme der Gesellschaft für Informatik, 2021.

Einschätzung des Lebensraumangebots für ausgestorbene Arten aus Fernerkundungs- und Forsteinrichtungsdaten am Beispiel des Haseluhns (*Tetrastes bonasia rupestris*) im Nationalpark Schwarzwald

Markus Handschuh¹, Sönke Birk², Christoph Dreiser³ und Marc Förschler⁴

Abstract: Das Haselhuhn (*Tetrastes bonasia rupestris*) ist in Baden-Württemberg vor allem aufgrund Lebensraumverlusts erloschen. Im Nationalpark Schwarzwald, Teil des EG-Vogelschutzgebiets Nordschwarzwald, entstehen derzeit durch unterschiedliche Prozesse wahrscheinlich wieder geeignete Haselhuhnhabitate in größerem Umfang. Dies wirft die Frage nach einer möglichen Wiederansiedlung der für das Vogelschutzgebiet als wertgebend aufgeführten Vogelart auf. Daher überprüfen wir in unserer Arbeit die saisonale Habitateignung von Waldbeständen im Nationalparkgebiet und quantifizierten das potenzielle Lebensraumangebot. Mangels Präsenzdaten der erloschenen Vogelart erfolgte dies durch eine wissens- und regelbasierte Ableitung des Habitatangebots aus Fernerkundungs- und Forsteinrichtungsdaten. Im Nationalparkgebiet umfassten geeignete Haselhuhn-Sommerhabitate im Jahr 2018 rund 1.550 ha, geeignete Winterhabitate rund 580 ha und Flächen mit ganzjähriger Habitateignung rund 275 ha, Tendenz jeweils zunehmend. Die Güte des Habitateignungsmodells ist exzellent. Da eine langfristig überlebensfähige Population des Haseluhns auf Landschaftsebene geeignete Lebensräume benötigt, ist die Prüfung der Übertragbarkeit des Modells auf Waldbestände außerhalb des Nationalparks geplant.

Keywords: Haselhuhn, *Tetrastes bonasia*, Lebensraum, Habitat, LiDAR, Forsteinrichtung, regelbasierte Habitatmodellierung, wissensbasierte Habitatmodellierung.

1 Einleitung

Das Haselhuhn (*Tetrastes bonasia rupestris*), eine Art des Anhangs I der EG-Vogelschutzrichtlinie, ist in Baden-Württemberg ausgestorben. Die letzten Brutnachweise gelangen im Schwarzwald in den 1990er-Jahren und die letzten gesicherten Artnachweise in den frühen 2000er Jahren [LR01, Ba16]. Als Hauptursache für das Erlöschen der Art gilt der großräumige Verlust geeigneter Lebensräume, insbesondere aufgrund der einseitigen Förderung von dunklen, einschichtigen Nadelbaumbeständen ohne Kraut- und

¹ Nationalpark Schwarzwald, Fachbereich 2 - Ökologisches Monitoring, Forschung & Artenschutz, Kniebisstraße 67, 72250 Freudenstadt, Markus.Handschuh@nlp.bwl.de

² Nationalpark Schwarzwald, Fachbereich 2 - Ökologisches Monitoring, Forschung & Artenschutz, Kniebisstraße 67, 72250 Freudenstadt, Soenke.Birk@nlp.bwl.de

³ Nationalpark Schwarzwald, Fachbereich 2 - Ökologisches Monitoring, Forschung & Artenschutz, Kniebisstraße 67, 72250 Freudenstadt, Christoph.Dreiser@nlp.bwl.de

⁴ Nationalpark Schwarzwald, Leitung Fachbereich 2 - Ökologisches Monitoring, Forschung & Artenschutz, Kniebisstraße 67, 72250 Freudenstadt, marc.foerschler@nlp.bwl.de

Strauchschicht und durch den systematischen Aushieb von Pioniergehölzen [AM89, Li94, LR01, Li03].

Das Haselhuhn ist ein sesshafter und wenig ausbreitungsfähiger, strikter Waldbewohner [z. B. Sw91, Mo05] und gilt als Indikatorart für struktur- und artenreiche Wälder von hoher ökologischer Wertigkeit [z. B. Mü78, Pa14]. Die Art ist das kleinste einheimische Rauhfußhuhn und spezialisiert auf ein Leben in der Kraut-, Strauch- und unteren Baumschicht deckungsreicher Wälder. Dabei ist das Haselhuhn nicht an bestimmte Pflanzengesellschaften, klimatische Bedingungen, Bodenverhältnisse oder Höhenlagen gebunden. Entscheidend für eine Habitatsignung ist vielmehr eine ganzjährige enge Verzahnung von Deckung und Nahrung [z. B. Py54, Li94, Be96].

Aufgrund seiner Lebensraumsprüche profitiert das Haselhuhn insbesondere von ausgedehnten Jungwäldern und kann dort hohe Bestandsdichten erreichen [GBB94, Be96]. Dies können Pionierwälder nach natürlichen Störungen wie Sturmwurf, Insektenkalamität oder Waldbrand sein [z. B. Ko18], gepflanzte Jungbestände nach Kahlhieb [z. B. AM89], Sukzessions- und Aufwuchswälder auf ehemaligen Freiflächen [z. B. Li94] oder durch Stockausschlag entstandene Niederwälder [z. B. Li86].

Durch die beiden „Jahrhundertstürme“ „Vivian & Wiebke“ im Februar 1990 und „Lothar“ im Dezember 1999 wurden im Nordschwarzwald ausgedehnte Freiflächen im Wald geschaffen und dabei insbesondere einförmige und für das Haselhuhn ungeeignete Fichtenbestände (*Picea abies*) beseitigt. Nach Daten der Nationalparkverwaltung entstanden allein im Gebiet des heutigen Nationalparks etwa 2.600 ha an Sturmflächen.

Auf vielen dieser Flächen wachsen derzeit ausgedehnte Jungwälder auf, welche allem Anschein nach die Lebensraumsprüche des Haselhuhns erfüllen. Auch im Wirtschaftswald scheint sich insbesondere durch die Schonung und Förderung von Mischbaumarten die Habitatsituation für das Haselhuhn in den letzten Jahren grundsätzlich verbessert zu haben. Zum Dritten entstehen vor allem in wenig oder nicht bewirtschafteten Waldflächen durch die Auflichtung von Fichtenbeständen aufgrund Trockenheit und Borkenkäferbefall und anschließende Sukzession vermutlich zunehmend geeignete Haselhuhnlebensräume.

Im Nationalpark Schwarzwald laufen alle drei Prozesse gleichzeitig ab und die Lebensraumsituation für das Haselhuhn scheint wieder günstiger zu sein. Allerdings lässt sich diese Vermutung nicht durch eine Wiederausbreitung des Haselhuhns oder durch zunehmende Bestandszahlen belegen. Denn eine natürliche Wiederbesiedelung des Gebiets ist mangels aktueller Haselhuhnvorkommen im weiteren Umfeld und aufgrund der sehr geringen Dispersionsfähigkeit der Art ausgeschlossen. Außerdem gibt es im Gebiet keine „Stellvertreterart“ mit ähnlichen Lebensraumsprüchen und aktuellen Vorkommen. Aus einem nachweislich günstigen Lebensraumangebot ergäbe sich die Frage einer möglichen Wiederansiedlung des Haselhuhns, zumal die Art für das Vogelschutzgebiet Nordschwarzwald, mit dem sich der Nationalpark wesentlich überschneidet, als Schutzgut von gemeinschaftlichem Interesse gelistet ist [RPK17].

Das Ziel der vorliegenden Untersuchung war daher, ohne Präsenzdaten des Haselhuhns das Lebensraumangebot für die Art im Nationalpark Schwarzwald einzuschätzen.

2 Material und Methoden

2.1 Untersuchungsgebiet

Der Nationalpark Schwarzwald wurde 2014 im nordwestlichen Schwarzwald im Westen Baden-Württembergs gegründet. Das Gebiet liegt zwischen 600 m und 1.100 m ü. NN und ist geprägt von zumeist fichtendominierten Fichten-Tannen-Buchen-Mischwäldern. Der Nationalpark umfasst eine Gesamtfläche von 10.076 ha, bestehend aus zwei etwa 3,5 Kilometer voneinander entfernten Einzelbereichen von ca. 7.600 ha und ca. 2.450 ha (vgl. Abb. 1). Fünf Bereiche mit einer Gesamtfläche von 5.115 ha sind als Kernzonen mit dem vorrangigen Ziel des Prozessschutzes und der freien, natürlichen Waldentwicklung ausgewiesen [FRB20]. Ein Großteil der Nationalparkfläche ist Teil des europäischen Schutzgebietsnetzes Natura 2000, darunter drei Fauna-Flora-Habitat (FFH) Gebiete und das EG-Vogelschutzgebiet Nordschwarzwald [FR19]. Insbesondere die Überschneidung mit letzterem ist erheblich: 77 % der Nationalparkfläche liegen im Vogelschutzgebiet und umgekehrt nimmt der Nationalpark 22 % des rund 36.000 ha umfassenden Vogelschutzgebiets ein.

2.2 Habitatmodellierung

Aufgrund artenschutzbedingter Betretungsbeschränkungen, der Flächengröße und des oft schwierig begehbaren Geländes im Nationalpark war eine flächige Habitatkartierung nicht möglich. Um dennoch das Gesamtgebiet ausreichend detailliert beurteilen zu können, wählten wir einen Modellierungsansatz. Hierzu verwendeten wir hochaufgelöste Laserscandaten (LiDAR = Light Detection and Ranging) aus einer Befliegung im Auftrag Nationalparkverwaltung im April und Mai 2015. Der Vorteil von LiDAR gegenüber passiven optischen Systemen wie Luftbild-Stereophotogrammetrie ist die Möglichkeit, die 3D-Vegetationsstruktur auch in dichten Waldbeständen zu charakterisieren. Da aus den Laserscandaten allerdings nicht die gesamte für Haselhuhnhabitate relevante Information abgeleitet werden konnte, wurde darüber hinaus auf Forsteinrichtungsdaten mit Referenzjahr 2014 zurückgegriffen.

Generell setzen Habitatmodelle bzw. Habitateignungsmodelle das Vorkommen von Arten mit Umweltvariablen in Beziehung [z. B. GZ00]. Für zuverlässige Ergebnisse ist daher eine Mindestanzahl an beobachteten Vorkommen erforderlich [z. B. Wi08]. Aus dem Untersuchungsgebiet lagen jedoch weder aktuelle Präsenzdaten von Haselhühnern noch detaillierte Untersuchungen vor dem Aussterben der Art vor. Deshalb wurden in mehreren Arbeitsschritten Präsenzdaten durch Literaturangaben und Expertenwissen zur Habitatwahl des Haselhuhns in anderen, gut untersuchten Gebieten mit ähnlichen

Lebensräumen ersetzt. Daraus wurden die für eine Habitatmodellierung im Nordschwarzwald erforderlichen Informationen abgeleitet.

Hierzu leiteten wir zunächst aus der Fachliteratur deskriptive Schlüsselparameter für die saisonale (Sommer / Winter) Eignung eines Waldes als Haselhuhnlebensraum her. Diese wurden anschließend in quantitative fernerkundungs- und forsteinrichtungsdatensatztaugliche Zustands- und Merkmalsvariablen umgewandelt. Die Variablen belegten wir mit ebenfalls aus der Literatur entnommenen Grenzwerten der Haselhuhn-Habitat eignung und fügten sie zu einem Regelwerk von Selektionskriterien zur Vorhersage der Haselhuhn-Habitat eignung von Rasterpixeln (s. u.) zusammen. Das Regelwerk wurde mit Hilfe von Python-Skripten im ArcGIS Softwarepaket Version 10.3 [Es15] auf die als Rasterdaten vorliegenden Fernerkundungs- und Forsteinrichtungsdaten angewendet. Es erfolgte somit eine wissensbasierte bzw. regelbasierte Habitat eignungsmodellierung.

Folgende Eingangsdatensätze wurden zur Modellbildung herangezogen: Klassifikation der Waldstruktur mittels LiDAR-Daten in sieben Waldstrukturklassen [La17], LiDAR-basiertes Vegetations-Oberflächenmodell mit einer Pixelauflösung von 40 cm x 40 cm, Pioniergehölzanteil auf Waldbestandsebene gemäß Forsteinrichtung, LiDAR- und Luftbild-basierte Waldbestandsklassifikation nach Laubholz- / Nadelholz-Anteil, Wege und Gewässerläufe als Sonderstandorte und Vernetzungselemente sowie Analysen von Nachbarschaftsbeziehungen und Flächensummen von Rasterpixeln.

Die Habitatmodellierung erfolgte auf Ebene eines Rasters mit einer Pixelauflösung von 20 m x 20 m. Dies entspricht einer Pixelfläche von 400 m² und damit etwa 0,1 % - 0,4 % eines typischen mitteleuropäischen Haselhuhnwohngebiets von 10 ha bis 40 ha Größe [vgl. Be96]. Diese Auflösung ist damit einerseits ausreichend fein, um räumlich differenzierte Aussagen treffen und der oft komplexen und feinkörnigen Mosaikstruktur von Haselhuhnlebensräumen gerecht werden zu können, andererseits ausreichend grob, um den Einfluss von Hintergrundrauschen, z. B. durch Einzelbäume, zu minimieren.

Die Vorhersage der Habitat eignung von Rasterpixeln erfolgte getrennt für Sommer und Winter, da das Haselhuhn deutliche saisonale Wechsel in der Nahrungs- und Habitatwahl zeigt, was je nach den lokalen Habitatbedingungen mit saisonalen kleinräumigen Ortswechseln einhergehen kann [z. B. Li94]. Ein 20 m x 20 m Rasterpixel konnte daher entweder als Sommerhabitat oder als Winterhabitat oder, bei gleichzeitiger Erfüllung der Kriterien für Sommer- und Wintereignung, als Ganzjahreshabitat geeignet sein. Die Ermittlung der saisonalen Habitat eignung erfolgte kategorial und ohne Zwischenwerte, das heißt jedes Rasterpixel war entweder geeignet oder nicht. Auf eine kontinuierliche Habitatbewertung oder Abstufungen wurde zur Minimierung der Komplexität des Habitatmodells und aus Gründen der Interpretierbarkeit der Ergebnisse verzichtet.

Die Ergebnisse wurden mittels Geographischer Informationssysteme visualisiert (ArcGIS Version 10.3: Es15 und QGIS Desktop Version 2.18.28: Qg19). Für weitere Details zur Methodik siehe [HB19].

Die Verifizierung (Überprüfung der zugrundeliegenden Algorithmen auf technische, logische und formelle Fehler), Justierung (Anpassung von Modellparametern, um die Übereinstimmung zwischen Modellvorhersagen und Wirklichkeit zu verbessern) und Validierung (Quantifizierung der Zuverlässigkeit von Modellvorhersagen durch deren Vergleich mit beobachteten Daten) sind wichtige Schritte in der ökologischen Modellbildung und in der Evaluierung von Modellen [z. B. Ry96].

Die Verifizierung und die Justierung des Modells erfolgten gutachtlich entlang von ca. 350 km ohnehin zurückgelegter Fahrstrecke im gesamten Nationalparkgebiet sowie an 200 im GIS festgelegten Zufallspunkten und an 83 weiteren Probepunkten, welche gezielt in hinsichtlich Modellvorhersage problematische bzw. fehlerbehaftete Flächen gelegt wurden. Zudem wurde im Rahmen eines Workshops von neun externen Haselhuhnexperten die Habitateignung an 27 weiteren, zum Teil zufällig verteilten, zum Teil gezielt in Problembereiche gelegten Referenzpunkten bewertet und auch diese Ergebnisse zur Justierung des Modells verwendet.

Die Endvalidierung und Güteermittlung des finalen Habitateignungsmodells erfolgte durch den Vergleich der Modellvorhersagen mit der im Feld festgestellten wirklichen Habitateignung an 100 „Gütepunkten“. Diese umfassten 18 durch die Experten weitgehend einhellig bewertete Referenzpunkte sowie 82 zufällig aus den oben genannten 200 Zufallspunkten ausgewählte Probepunkte. Eine Ermittlung der finalen Modellgüte an einem neuen Set an Zufallspunkten war aus logistischen Gründen nicht möglich. Aufgrund GPS-Ungenauigkeiten im teils steilen Gelände wurde die Güte des Modells auf einer Bezugsfläche von 3 x 3 Rasterpixeln bestimmt. Als Gütemaße dienten die Treffergenauigkeit und der Klassifikationsfehler sowie die Sensitivität und die Spezifität des Modells [vgl. Pe10].

3 Ergebnisse

In Abbildung 1 ist die durch unser Modell vorhergesagte saisonale Habitateignung für das Haselhuhn im Nationalpark Schwarzwald im Jahr 2018 dargestellt. Sowohl potenzielle Haselhuhn-Sommerhabitate als auch Winterhabitate kommen über das gesamte Nationalparkgebiet verteilt vor.

Auffällig ist das geklumpfte Auftreten von geeigneten Winterhabitaten. Dies ist Großteils methodisch bedingt, da die Wintereignung aus Forsteinrichtungsdaten und dadurch auf der räumlichen Ebene forstlicher Waldbestände mit Flächengrößen von 0,5 ha-15 ha abgeleitet wurde.

Neben den im Sommer und / oder im Winter für das Haselhuhn als geeignet klassifizierten Rasterpixeln sind in Abbildung 1 auch erklärende Elemente enthalten, welche keine eigenen Kategorien der Habitateignung bilden. Zum einen sind Bestandslücken als fein grau-schraffierte, dünn gelb umrandete Flächen dargestellt, denn Lücken sind bei geeigneter Größe und Struktur wichtige Element des Haselhuhnlebensraums. Zum

anderen sind Bereiche mit Sommereignung auf mindestens 50% der Fläche rosafarbig unterlegt und dünn schwarz umrandet dargestellt. Hierzu wurden alle Rasterpixel, bei welchen jeweils mindestens vier der acht unmittelbar angrenzenden Rasterpixel im Sommer geeignet sind, zu Komplexen zusammengefasst. Die so gebildeten Sommerereignungs-Cluster entsprechen in etwa der Aggregation von Rasterpixeln mit Winterereignung auf forstlicher Bestandesebene und tragen zur verbesserten Interpretierbarkeit der Karte bei.

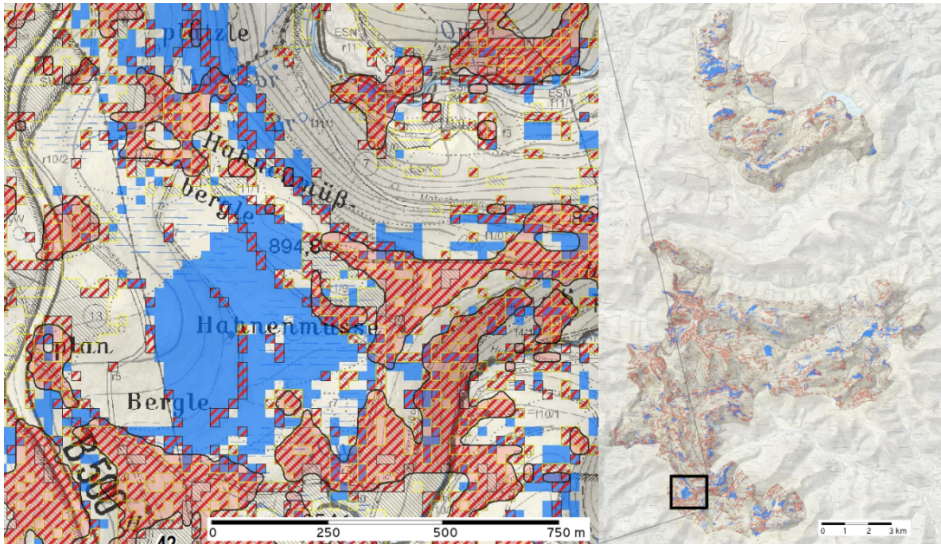


Abb. 1: Saisonale Habitateignung für das Haselhuhn im Nationalpark Schwarzwald im Jahr 2018 auf 20 m x 20 m Rasterbasis. Blau: Habitateignung im Winter; rot-schraffiert: Habitateignung im Sommer; blau-rot-schraffiert: Ganzjährige Habitateignung. Weitere Details siehe Text.

Tabelle 1 enthält Flächenangaben zu potenziellen Haselhuhnhabitaten im Nationalpark Schwarzwald als Summe der entsprechend klassifizierten 400 m²-Rasterpixel. Geeignete Sommerhabitate nahmen im Jahr 2018 eine Gesamtfläche von rund 1.550 ha ein, geeignete Winterhabitate von rund 580 ha und Flächen mit ganzjähriger Habitateignung 276 ha. Wird die Fläche mit ganzjähriger Habitateignung jeweils den beiden saisonalen Eignungen zugeschlagen und damit doppelt gezählt, betrug die Flächengröße von geeigneten Sommerhabitaten rund 1.830 ha und von geeigneten Winterhabitaten rund 850 ha. Somit waren im Jahr 2018 rund ein Viertel der Nationalparkfläche als Lebensraum für das Haselhuhn geeignet und rund drei Viertel der Fläche ungeeignet.

	Flächengröße in Hektar	Prozentanteil an der Nationalparkfläche
Habitateignung im Sommer	1.555	15

Habitateignung im Winter	578	6
Ganzjährige Habitateignung	276	3
Summe	2.409	24

Tab. 1: Flächengrößen und Flächenanteile von 20 m x 20 m Rasterpixeln mit saisonaler Haselhuhn-Habitateignung im Nationalpark Schwarzwald im Jahr 2018.

Die an 100 Referenzpunkten ermittelte Modellgüte und damit die Vorhersage der saisonalen Haselhuhn-Habitateignung durch unser Modell ist exzellent: Treffergenauigkeit > 90 %, Klassifikationsfehler < 10 %, Sensitivität = 89 % und Spezifität = 100 %.

4 Diskussion

Untersuchungen in verschiedenen Gebieten belegen, dass LiDAR-Daten generell geeignet sind zur Modellierung von Haselhuhnhabitaten [z. B. Ba14, Re17]. Ein Nachteil von LiDAR-Daten ist allerdings, dass derzeit noch keine Baumartenbestimmung, insbesondere von Winternahrungsgehölzen des Haselhuhns, möglich ist. Daher mussten wir in der vorliegenden Untersuchung auf die in mancher Hinsicht problematischen Forsteinrichtungsdaten zurückgreifen; auch eine Bestimmung der im Sommer für das Haselhuhn wichtigen Zusammensetzung der Krautschicht (z. B. Unterscheidung zwischen Naturverjüngung und Heidelbeere, Gräsern und Kräutern) ist anhand LiDAR-Daten derzeit noch nicht möglich.

Bei der Entwicklung des Habitateignungsmodells mussten mehrere Annahmen gemacht werden. Insbesondere die Frage, inwiefern Expertenaugen tatsächlich richtig sehen, kann mangels Präsenzdaten des Haselhuhns im Gebiet nicht abschließend geklärt werden. Zudem wurde die Habitateignung ausschließlich anhand der Struktur und Baumartenzusammensetzung von Waldbeständen ermittelt, während weitere mögliche Einflussfaktoren wie der Einfluss von Beutegreifern [vgl. Kä17] unberücksichtigt bleiben mussten. Andererseits wären trophische Interaktionen kaum modellierbar und die strukturelle und gehölzartenbedingte Habitatqualität sowie der Flächenanteil geeigneter Habitate gelten beim Haselhuhn als die entscheidenden Einflussgrößen für ein Artvorkommen [z. B. Li94, Sw95, Be96].

Die Haselhuhn-Habitateignung wurde durch unser Modell sehr gut vorhergesagt. Allerdings wurde die Modellgüte nicht anhand unabhängiger Zufallspunkte bestimmt. Andererseits fanden wir bei unseren Überprüfungen im Feld keine geeigneten Habitatflächen, wo nicht auch vom Modell als geeignet klassifizierte Rasterpixel vorhanden waren und umgekehrt war das Habitat nie völlig ungeeignet, wo durch das Modell Rasterpixel mit Habitateignung vorhergesagt wurden. Daher ist das Modell ausreichend für seinen angestrebten Zweck der Einschätzung des Lebensraumangebots.

Die Gültigkeitsdauer der Habitat-Flächenangaben beträgt aufgrund ständig ablaufender natürlicher Prozesse (v. a. Sukzession von Freiflächen, Aufwachsen von Jungbeständen, Borkenkäferbefall, Windwurf und Schneebruch) wahrscheinlich nicht mehr als 3-5 Jahre. Die natürliche Walddynamik wird im Nationalpark Schwarzwald in Zukunft mit Sicherheit zu weiterem und immer wieder neuem potenziellem Lebensraum für das Haselhuhn führen, möglicherweise sogar in größerem Umfang, ähnlich wie im Nationalpark Bayerischer Wald [Ko18; vgl. auch Bo18]. Dies bleibt durch eine neue Analyse der Lebensraumentwicklung in den kommenden Jahren zu überprüfen.

Des Weiteren ist geplant, die Übertragbarkeit des Habitateignungsmodells zu überprüfen, d. h. inwiefern das Modell auf Waldbestände außerhalb des Nationalparks anwendbar ist. Denn hier ist das Angebot an potenziell geeignetem Haselhuhnlebensraum derzeit offenbar mindestens ebenso gut wie innerhalb des Nationalparks. Eine Anhebung des Habitateignungsmodells auf die Landschaftsebene, beispielsweise auf die räumliche Ebene des EG-Vogelschutzgebiets Nordschwarzwald, wäre auch im Hinblick darauf wichtig, dass der maximal verfügbare Lebensraum im Nationalpark Schwarzwald allein wahrscheinlich nicht für die Etablierung einer dauerhaft überlebensfähigen Haselhuhnpopulation ausreicht.

Als Fazit und Ausblick lässt sich festhalten:

- Das Haselhuhn ist in Baden-Württemberg Anfang der 2000er Jahre vor allem aufgrund Lebensraumverlusts durch ungünstige forstliche Praktiken ausgestorben
- Mittlerweile hat sich die Lebensraumsituation aufgrund mehrerer Faktoren zumindest im Nationalpark Schwarzwald nachweislich wieder verbessert
- Es wäre sinnvoll, das Lebensraumangebot auch im Umfeld des Nationalparks zu überprüfen, insbesondere im EG-Vogelschutzgebiet Nordschwarzwald
- Da eine natürliche Wiederbesiedelung des Gebiets ausgeschlossen ist, sollte bei großflächig günstigem Lebensraumangebot eine Wiederansiedlung des Haselhuhns in Erwägung gezogen werden
- Möglicherweise ist die Situation in anderen Regionen Deutschlands ähnlich

5 Literaturverzeichnis

- [AM89] Asch, T.; Müller, G.: Haselwild in Baden-Württemberg. Schutzgemeinschaft Deutscher Wald e.V., Landesverband Baden-Württemberg, Stuttgart, 1989.
- [Ba16] Bauer, H.-G. et al.: Rote Liste und kommentiertes Verzeichnis der Brutvogelarten Baden-Württembergs. 6. Fassung. Stand 31. 12. 2013. Naturschutz-Praxis Artenschutz 11, 2016.

- [Ba14] Bae, S. et al.: Comparison of airborne lidar, aerial photography, and field surveys to model the habitat suitability of a cryptic forest species – the hazel grouse. *International Journal of Remote Sensing* 35:17, S. 6469-6489, 2014.
- [Be96] Bergmann, H.-H. et al.: *Die Haselhühner*. 4. Auflage. Westarp, Magdeburg, 1996.
- [Bo18] Bollmann, K.: Stürmische Zeiten für Raufusshühner im Bergwald. *Ornithologischer Beobachter* 115 (3), S. 239-254, 2018.
- [Es15] ESRI: ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems Research Institute, 2015.
- [FR19] Förschler, M.; Richter, C.: Nationalpark Schwarzwald – Arten und Biotopschutz durch Prozessschutz und Management. *Naturschutz-Info* 1/2019 + 2/2019, S. 63-66, 2019.
- [FRB20] Förschler, M.; Richter, C.; Birk, S.: Nationalpark Schwarzwald – Kernzonen-Erweiterung. *Naturschutz-Info* 1/2020 + 2/2020, S. 1-4, 2020.
- [Ge14] Gedeon, K. et al.: *Atlas Deutscher Brutvogelarten*. Münster (Stiftung Vogelmonitoring und DDA), 2014.
- [GGB94] Glutz von Blotzheim, U. N.; Bauer, K. M.; Bezzel, E.: *Handbuch der Vögel Mitteleuropas*. Band 5: Galliformes und Gruiformes (2., durchgesehene Auflage 1994). AULA-Verlag GmbH, Wiesbaden, 1994.
- [GZ00] Guisan, A.; Zimmermann, N. E.: Predictive habitat distribution models in Ecology. *Ecological Modelling* 135, S. 147-186, 2000.
- [HB19] Handschuh, M.; Birk, S.: Einschätzung des Lebensraumangebots für das Haselhuhn (*Tetrastes bonasia*) im Nationalpark Schwarzwald und fachliche Erörterungen zur Machbarkeit einer Wiederansiedlung der Art. https://www.researchgate.net/publication/351153946_Einschätzung_des_Lebensraumangebots_für_das_Haselhuhn_Tetrastes_bonasia_im_Nationalpark_Schwarzwald_und_fachliche_Erörterungen_zur_Machbarkeit_einer_Wiederansiedlung_der_Art, 2019.
- [Kä17] Kämmerle, J.-L. et al.: Range loss of a threatened grouse species is related to the relative abundance of a mesopredator. *Ecosphere* 8 (9), e01934, Doi: 10.1002/ecs2.1934, 2017.
- [Ko18] Kortmann, M. et al: Forest structure following natural disturbances and early succession provides habitat for two avian flagship species, capercaillie (*Tetrao urogallus*) and hazel grouse (*Tetrastes bonasia*). *Biological Conservation* 226, S. 81-91, 2018.
- [La17] Lang, F.: LiDAR based classification of forest structure of conifer dominated ecosystems in den Black Forest National Park. Master Thesis in Geoecology, Universität Bayreuth, 2017.
- [Li86] Lieser, M.: Untersuchungen zur Verbreitung und Ökologie des Haselhuhns (*Bonasa bonasia*) an der Mosel. *DBV-Mitteilungen Rh.-Pfalz* 1986/I, S. 1-98, 1986.
- [Li94] Lieser, M.: Untersuchungen der Lebensraumansprüche des Haselhuhns (*Bonasia L.* 1758) im Schwarzwald im Hinblick auf Maßnahmen zur Arterhaltung. *Ökologie der Vögel* 16, S. 1-117, 1994.
- [Li03] Lieser, M.: Probleme des Artenschutzes im Wirtschaftswald am Beispiel der Raufußhühner im Schwarzwald. *Natur und Landschaft* 78: S. 10-17, 2003.

- [LR01] Lieser, M.; Roth, K.: Haselhuhn. In (Hölzinger, J., Hrsg.): Die Vögel Baden-Württembergs (Avifauna Baden-Württemberg), Bd. 2. Nichtsingvögel. Ulmer, Stuttgart, 2001.
- [Mo05] Montadert, M.: Fonctionnement démographique et sélection de l'habitat d'une population en phase d'expansion géographique. Cas de la Gélinotte des bois dans les Alpes du Sud, France. PhD thesis, Université de Franche-Comté, Besançon, 2005.
- [Mü78] Müller, F.: Rauhfußhühner als Biotop-Indikatoren. Jagd und Hege 6, S. 57-68, 1978.
- [Pa14] Pakkala, T. et al.: Indicators of forest biodiversity: which bird species predict high breeding bird assemblage diversity in boreal forests at multiple spatial scales? *Annales Zoologici Fennici* 51, S. 457-476, 2014.
- [Pe10] Pearson, R. G.: Species' distribution modelling for conservation educators and practitioners. *Synthesis. Lessons in Conservation* 3, S. 54-89, 2010.
- [Py54] Pynnönen, A.: Beiträge zur Kenntnis der Lebensweise des Haselhuhns, *Tetrastes bonasia* (L.). *Pap. on Game Res.* 12, S. 1-90, 1954.
- [Qg19] QGIS Development Team: QGIS Geographic Information System. Open-Source Geospatial Foundation, 2019. URL: <https://qgis.org>
- [Re17] Rechsteiner, C. et al.: Remotely sensed forest habitat structures improve regional species conservation. *Remote Sens. Ecol. Conserv.* 3, S. 247-258, 2017.
- [RPK17] RPK, Regierungspräsidium Karlsruhe, Natura 2000-Managementplan Vogelschutzgebiet 7415-441 Nordschwarzwald, Gebietssteckbrief (Stand 19.01.2017), https://rp.baden-wuerttemberg.de/fileadmin/RP-Internet/Karlsruhe/Abteilung_5/Referat_56/Natura_2000/DocumentLibraries/Nordschwarzwald/Gebietssteckbrief.pdf, Abrufdatum 25.04.2021.
- [Ry96] Rykiel, E. J. Jr.: Testing ecological models: the meaning of validation. *Ecological Modelling* 90, S. 229-244, 1996.
- [Sw91] Swenson, J. E.: Is the hazel grouse a poor disperser? *Trans. IUGB-Congr.* 20, Gödöllő, S. 347-352, 1991.
- [Wi08] Wisz, M. et al.: Effects of sample size on the performance of species distribution models. *Diversity and Distributions* 14, S. 763-773, 2008.

6 Danksagung

Wir danken den neun Haselhuhnexperten für ihre Teilnahme am Expertenworkshop und für ihre aufschlussreichen Beiträge: Thomas Asch, Prof. Hans-Heiner Bergmann, Andreas Kämpfer-Lauenstein, Dr. Siegfried Klaus, Dr. Manfred Lieser, Dr. Jean-Jacques Pfeffer, Bernhard Scherer, Dr. Karl-Eugen Schroth und Dr. Ralf Siano. Wir bedanken uns außerdem bei Dr. Marc Montadert, Prof. Ilse Storch und Dr. Wolfgang Scherzinger für wertvolle Hinweise zu unterschiedlichen Aspekten der vorliegenden Arbeit.

lebensmittelwarnung.de - Rückrufe und öffentliche Warnungen in Deutschland

Ein Einblick in die Arbeit in der Lebensmittelüberwachung

Gregor Wengler¹, Karim Lindner², Theresa Gleich³, Silvia Raschke⁴

Abstract: In 2011 hat das Bundesamt für Verbraucherschutz und Lebensmittelsicherheit (BVL) in enger Zusammenarbeit mit den Bundesländern erfolgreich das Informationsportal lebensmittelwarnung.de ins Leben gerufen. Das Portal informiert Verbraucherinnen und Verbraucher über Warnungen und Rückrufe für die Produktkategorien Lebensmittel, kosmetische Mittel und Bedarfsgegenstände wie Spielzeug oder Kleidung. Die Entstehungsgeschichte von lebensmittelwarnung.de sowie die dazugehörigen Meldestrukturen zeigen auf, wie durch kommunikative Strukturen zwischen Bund und Ländern Verbraucherschutz im Alltag koordiniert wird.

Aktuell entwickelt das Bundesamt für Verbraucherschutz und Lebensmittelsicherheit (BVL) das Portal zusammen mit Vertreterinnen und Vertretern der Bundesländer weiter. lebensmittelwarnung.de soll gezielter an die Bedürfnisse der Verbraucherinnen und Verbraucher angepasst werden sowie weiterführende Informationen zur Lebensmittelsicherheit im Allgemeinen und aktuellen Ereignissen bieten.

Keywords: Lebensmittelsicherheit, Verbraucherschutz, Lebensmittelwarnungen, Lebensmittel, kosmetische Mittel, Bedarfsgegenstände, Spielzeug, Kleidung, Verbraucherwarnung, LFGB, Rückruf, Rücknahme, Verbraucherportal

1 Lebensmittelüberwachung in Deutschland

1.1 Föderale Strukturen in der deutschen Lebensmittelüberwachung

Die übergeordnete Zuständigkeit für das Lebensmittelrecht liegt beim Bundesministerium für Ernährung und Landwirtschaft (BMEL). In seiner Zusammenarbeit mit der Europäischen Union und den Bundesländern nimmt das BMEL diverse Rechtsetzungs-, Kontroll-, Überwachungs- und Informationsaufgaben wahr. Dem Bundesministerium ist das Bundesamt für Verbraucherschutz und Lebensmittelsicherheit (BVL) nachgeordnet. Als Zulassungs- und Managementbehörde ist das BVL zuständig für die Koordination

¹⁻⁴ Bundesamt für Verbraucherschutz und Lebensmittelsicherheit, Abteilung 1 – Referat 121, Mauerstr. 39-42, 10117 Berlin, silvia.raschke@bvl.bund.de

zwischen Bund und Ländern und übernimmt zahlreiche Aufgaben des Risiko- und Krisenmanagements sowie der Risikokommunikation im Bereich des gesundheitlichen Verbraucherschutzes.

Aufgrund der föderalen Verwaltungsstruktur in Deutschland sind die Bundesländer verantwortlich für die amtlichen Kontrollen zur Einhaltung der Gesetzgebung in den relevanten Bereichen des Verbraucherschutzes. In den meisten Bundesländern ist die Organisation der Lebensmittelüberwachung in drei Stufen gegliedert: Auf oberster Stufe koordiniert das zuständige Ministerium bzw. Senatsverwaltung die Überwachung. Darunter leisten Landesämter, Regierungspräsidien oder Bezirksregierungen die Fachaufsicht über die Überwachungsbehörden der über 400 Kreise und kreisfreien Städte. Diese kommunalen Lebensmittel- und Veterinärüberwachungsbehörden führen schließlich die amtlichen Kontrollen vor Ort aus.

1.2. Das Referat 121 (Warn- und Informationssysteme) im Bundesamt für Verbraucherschutz und Lebensmittelsicherheit (BVL)

Das in der Abteilung „Lebensmittelsicherheit“ in der Gruppe „Krisenmanagement“ angesiedelte Referat für Warn- und Informationssysteme (WIS) des Bundesamts für Verbraucherschutz und Lebensmittelsicherheit koordiniert die nationale und internationale Vernetzung und Zusammenarbeit der Bundesländer und der Europäischen Union im Bereich der Warn- und Informationssysteme. Über das Europäische Schnellwarnsystem für Lebens- und Futtermittel (RASFF⁵) koordiniert das Referat 121 als nationale Koordinationsstelle den Informationsaustausch über auffällige Lebens- und Futtermittel. Ebenso koordiniert das Referat den Kommunikationsaustausch der zuständigen Behörden für das Europäische Schnellwarnsystem für Verbraucherprodukte (RAPEX⁶), die in den Geltungsbereich des Lebensmittel-, Bedarfsgegenstände- und Futtermittelgesetzbuchs (LFGB) fallen (z. B. Kleidung, Spielzeuge, Kosmetik). Darüber hinaus stellt das Referat 121 die fachliche und administrative Betreuung des nationalen Portals lebensmittelwarnung.de sicher.

1.3. Aufgaben der Lebensmittelunternehmen

Die Verantwortung für Lebensmittelsicherheit liegt in erster Linie beim Lebensmittelunternehmen selbst. Die Risiken für die Lebensmittelsicherheit muss ein Unternehmen kennen und entsprechende Vorsorge treffen, dass nur sichere Lebensmittel die Verbraucherinnen und Verbraucher erreichen. Im Rahmen von Gefahrenanalysen und Eigenkontrollen entlang der gesamten Wertschöpfungskette eines jeden Lebensmittels erkennt ein Unternehmen mögliche risikobehaftete Produkte häufig bereits auf der Ebene

⁵ Rapid Alert System for Food and Feed, dt.: Europäisches Schnellwarnsystem für Lebensmittel und Futtermittel

⁶ Rapex Exchange of Information System

der Herstellung, der Verarbeitung oder des Zwischenhandels, bevor diese in den Einzelhandel gelangen. Betriebseigene Kontrollen und Laboruntersuchungen sorgen dafür, dass nicht verkaufsfähige Lebensmittel rechtzeitig erkannt und bereits beim Herstellungsprozess der Wertschöpfungskette entzogen werden. Zudem werden durch amtliche Kontrollen, bspw. durch Lebensmittelkontrolleure, regelmäßig Stichproben durchgeführt, um Produktrisiken frühzeitig zu erfassen. Lebensmittel, die die Verbraucherinnen oder Verbraucher noch nicht erreicht haben, werden innerhalb der Vertriebskette zurückgenommen (Rücknahme). Haben unsichere Lebensmittel bereits Verbraucherinnen oder Verbraucher erreicht, werden diese Produkte zurückgerufen (Rückruf).

2 Information der Öffentlichkeit über gefährliche Produkte

Die Informationspflichten und -regelungen für nicht sichere Lebensmittel werden in der so genannten Lebensmittelbasisverordnung (Verordnung (EG) Nr. 178/2002) sowie dem Lebensmittel-, Bedarfsgegenstände und Futtermittelgesetzbuch (LFGB) bestimmt. In der entsprechenden Verordnung ist u.a. vorgesehen, dass Lebensmittel, die nicht sicher sind, nicht in den Verkehr gebracht werden dürfen. Als nicht sichere Lebensmittel oder Produkte werden Waren bezeichnet, die beispielsweise auf Grund biologischer oder chemischer Verunreinigungen (Krankheitserreger, Pflanzenschutzmittelrückstände oder unerwünschte Zusatzstoffe), durch eine unzureichende Kennzeichnung oder durch Fremdkörper als gesundheitsgefährdend einzustufen sind. Haben Produkte, von denen ein Gesundheits- oder Täuschungsrisiko ausgeht, dennoch Verbraucherinnen und Verbraucher erreicht, informiert das verantwortliche Unternehmen unverzüglich und umfassend die Öffentlichkeit. Durch den öffentlichen Rückruf, zum Beispiel über Internetpräsenzen, Pressemitteilungen oder Ladenaushänge, werden Verbraucherinnen und Verbraucher aufgefordert, betroffene Produkte nicht zu verzehren oder zu verwenden, sondern an die Verkaufsstellen zurückzugeben.

Gleichzeitig hat das verantwortliche Unternehmen die zuständigen Überwachungsbehörden über den Mangel sowie die getroffenen Maßnahmen zu informieren. Auf dem Portal lebensmittelwarnung.de können die Landesministerien bzw. Senatsverwaltungen auf die Rückrufe von Unternehmen hinweisen. Wird eine neue Meldung erstellt, werden alle Bundesländer umgehend darüber informiert. Je nach Vertriebsweg bzw. Produktverteilung schließen sich die betroffenen Bundesländer dieser Warnung an. Erfolgt nach Bekanntwerden eines gesundheitsgefährdenden Mangels kein Rückruf durch das Unternehmen, so ist die Behörde gemäß § 40 LFGB ermächtigt, selbstständig eine öffentliche Warnung zu initiieren.

Eine Besonderheit bilden Meldungen über ein Produkt eines EU-Mitgliedstaates in einem der Europäischen Schnellwarnsysteme RASFF oder RAPEX, welches nicht in Deutschland hergestellt bzw. direkt in den Verkehr gebracht wurde. Zumeist handelt es sich hier um international verfügbare Produkte, die auf Onlineportalen auch von deutschen Verbraucherinnen und Verbrauchern erworben werden können. In diesem Fall darf das

Bundesamt für Verbraucherschutz und Lebensmittelsicherheit (BVL) eigenständig über das Verbraucherportal lebensmittelwarnung.de warnen.

3 Das Portal lebensmittelwarnung.de

3.1 Entstehungsgeschichte und Anwendung

Bis zur Inbetriebnahme des Portals lebensmittelwarnung.de mussten sich Verbraucherinnen und Verbraucher auf den einzelnen Internetseiten der 16 Landesministerien oder Senatsverwaltungen über Lebensmittellrückrufe informieren. Mit dem Ziel der Verbesserung des Informationsflusses wurde im Jahr 2011 in erfolgreicher Zusammenarbeit der Bundesländer mit dem BVL das Verbraucherportal lebensmittelwarnung.de etabliert. Hier können Verbraucherinnen und Verbraucher Informationen über Rückrufe und Warnungen aller Bundesländer zu der Produktkategorie Lebensmittel, und seit 2019 zu kosmetischen Mitteln und Bedarfsgegenständen (u.a. Spielzeug, Kleidung) finden. Das Referat 121 des BVL betreut im Auftrag der Bundesländer das Portal technisch und administrativ.

Über einen individuellen, passwortgeschützten Bearbeitungszugriff für registrierte Benutzer können die Mitarbeitenden der zuständigen Landesbehörden und des BVL Warnungen zu unsicheren Lebensmitteln in das Portal einstellen. Bei der Erstellung werden die wesentlichen Informationen der Meldung durch die Behörden erfasst. Registriert werden unter anderem das entsprechende Produkt, der Grund der Warnung, das verantwortliche Unternehmen sowie mögliche weiterführende Informationen wie beispielsweise Artikelnummer, Haltbarkeitsdatum, Kontaktinformationen der zuständigen Behörden sowie etwaige Bild- und Presseinformationen des Unternehmens.

Verbraucherinnen und Verbrauchern werden die Meldungen auf der Startseite des Portals in Kurzform präsentiert. Die weiterführenden Informationen können über eine Verlinkung angezeigt werden. Anhand eines RSS-Feeds können Warnungen nach Produktgruppen (Lebensmittel, Kosmetika, Bedarfsgegenstände) sowie nach Bundesländern kategorisiert werden. Zeitverzögert wird bei einer neuen Meldung automatisch ein Beitrag via Twitter (@LMWarnung) ausgelöst; zudem werden Medien- und Pressestellen über einen umfassenden Presseverteiler informiert.

3.2 Überblick: Zahlen und Fakten

Seit dem Start des Portals im Jahr 2011 bis zum 10.05.2021 wurden insgesamt 1469 Meldungen publiziert. Die Hälfte der Meldungen wurde in den letzten drei Jahren 2018 (186), 2019 (236) und 2020 (273) veröffentlicht (siehe Abbildung 1).

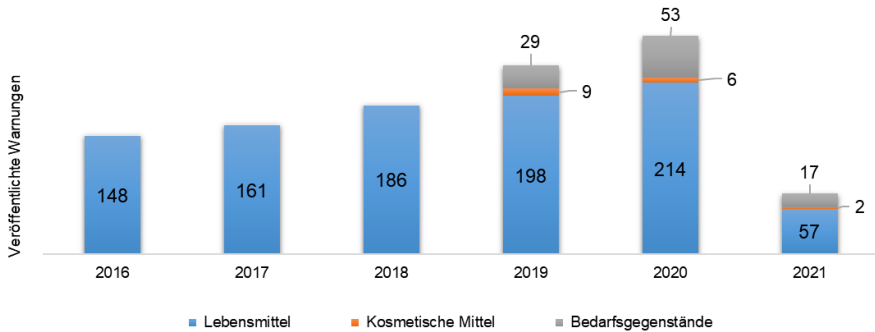


Abbildung 1 Warnmeldungen nach Produktkategorien seit 2016

Lebensmittel-Meldungen nach Produktkategorien (seit 2011)	
Fleisch, Wild, Geflügel und Erzeugnisse daraus	233
Milch und Milchprodukte	182
Getreide und Backwaren	158

Tabelle 1 Meldungen nach Produktkategorien (seit 2011)

Lebensmittel-Meldungen nach Warnungsgründen (seit 2011)	
Mikrobiologische Kontamination	485
Fremdkörper	363
Grenzwertüberschreitungen von Rückständen und Kontaminanten	232

Tabelle 2 Meldungen nach Warnungsgründen (seit 2011)

4 Verbraucherfreundliche Neuentwicklung

Das Portal lebensmittelwarnung.de wurde in den letzten Jahren für immer mehr Verbraucherinnen und Verbraucher zu einer wichtigen Informationsquelle in Bezug auf Produktgefahren. Insbesondere auch während Lebensmittelkrisen wird das Portal als Informationsquelle genutzt. Seit der Bereitstellung des Portals sind einerseits die Anforderungen an eine moderne und ansprechende Internetseite, andererseits die Bedürfnisse an behördliche Kommunikation und Transparenz stark gestiegen. Neue Formate wie mobile Apps und soziale Netzwerke bieten das Potential, Produktwarnungen gezielter und verbraucherfreundlicher zu verbreiten. Der Koalitionsvertrag der 19. Legislaturperiode sieht deshalb vor, dass die derzeitige Ausgestaltung der Webseite an die gestiegenen Anforderungen angepasst werden soll.

4.1 Planungsphase lebensmittelwarnung.de

Im Auftrag des Bundesministeriums für Ernährung und Landwirtschaft (BMEL) betreut und koordiniert das BVL aktuell die Neuentwicklung des Portals, die mit den Bundesländern durchgeführt wird. Aktuell ist die Planungsphase nahezu abgeschlossen. Die Produktivschaltung ist für Mai 2023 vorgesehen. In Zusammenarbeit mit dem BMEL und den Bundesländern sowie externen Stakeholdern konnten bereits spezifische Anforderungen und Ziele evaluiert werden. So konnte bereits ein so genannter Clickdummy für das neue Portal imaginiert werden.

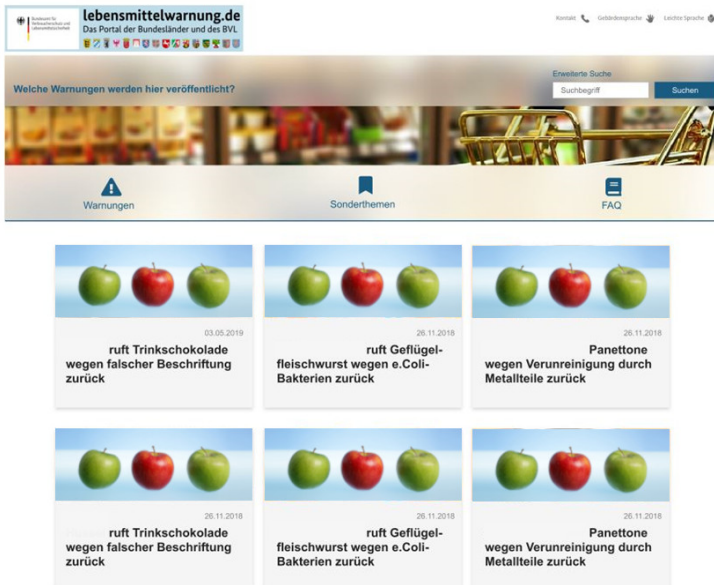


Abbildung 2 Beispiel für die Startseite der neuentwickelten Website

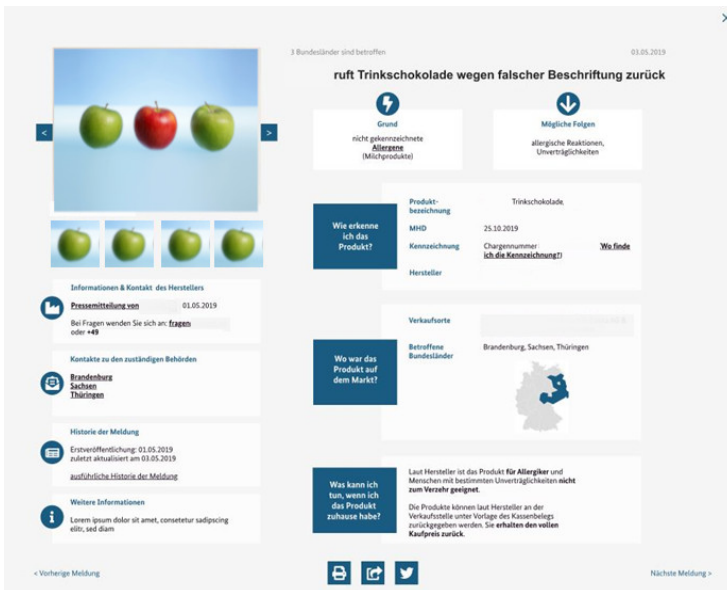


Abbildung 3 Beispiel für eine Lebensmittelwarnung der neuentwickelten Website

4.2 Planungsziele

Das übergeordnete Ziel ist die Entwicklung der Webseite lebensmittelwarnung.de zu einem benutzerfreundlichen, modernen und zukunftsfähigen Portal, das als zentrale Anlaufstelle („One-Stop-Shop“) für eine Mehrheit der Verbraucherinnen und Verbraucher dient, um sich über produktbezogene Gesundheitsrisiken sowie Täuschungsgefahren zu informieren.

4.2.1. Planungsziele – Technische Umsetzung

Aktuell wird das Portal als Zweikomponentensystem abgebildet. Beide Systeme wurden als Eigenentwicklung in JAVA IE angelegt. Die in der internen Bearbeitungsebene hinterlegte Datenbank wird durch Oracle dargestellt.

Zukünftig soll das Portal lebensmittelwarnung.de als Einkomponentensystem durch das Contentmanagementsystem DRUPAL umgesetzt werden, wodurch eine horizontale Veröffentlichung realisiert wird. Gleichzeitig werden Webschnittstellen über APIs implementiert. Aktuelle Meldungen sind somit über das Netzwerkprotokoll SOAP⁷ und über das Softwarearchitekturverfahren REST⁸ abrufbar.

4.2.2. Planungsziele – Inhaltliche Umsetzung

Neben aktuellen Veröffentlichungen nach § 40 Absatz 1, 2 und 5 LFGB soll die neue Webseite weiterführende Informationen zu Gesundheitsgefahren sowie Handlungsempfehlungen anbieten.

Zudem ist vorgesehen, das Portal als Kommunikationsmedium, z. B. in Krisensituationen, zu nutzen. Unter anderem ist die Installation flexibel gestaltbarer Sonderseiten geplant, um zeitnah über aktuelle Gefahren zu informieren.

Um die Reichweite der Informationen zu erhöhen bzw. Empfängerinnen und Empfänger schneller zu erreichen, wird eine mobile App entwickelt, die es Verbraucherinnen und Verbrauchern u.a. ermöglicht, personalisierte Push-Benachrichtigungen zu empfangen.

Die wesentlichen Inhalte des Portals sollen barrierefrei und mehrsprachig zugänglich sein. Eine verbesserte visuelle Darstellung der Portalkomponenten sowie ein prägnantes Logo sollen den Wiedererkennungswert sowie die Nutzerfreundlichkeit maßgeblich steigern. Verbraucherinnen und Verbraucher sind eingeladen, lebensmittelwarnung.de als informative und innovative Auskunftseite sowie als wichtige Quelle behördlicher Informationen zu nutzen.

⁷ Simple Object Access Protocol

⁸ Representational State Transfer

5. Weiterführende Links

Bundesministerium für Ernährung und Landwirtschaft:

https://www.bmel.de/DE/Home/home_node.html

Bundesamt für Verbraucherschutz und Lebensmittelsicherheit (BVL):

<https://www.bvl.bund.de/>

BVL Onlineauftritt auf Instagram: https://www.instagram.com/bvl_bund

BVL Onlineauftritt auf Youtube:

<https://www.youtube.com/channel/UCD9Yc3p7JgGbLfncc3oAvRA>

Lebensmittel-, Bedarfsgegenstände- und Futtermittelgesetzbuch: <https://www.gesetze-im-internet.de/lfgb/>

lebensmittelwarnung.de das Portal der Bundesländer und des BVL:

<https://www.lebensmittelwarnung.de/>

lebensmittelwarnung.de auf Twitter: <https://twitter.com/lmwarnung>

Bundesinstitut für Risikobewertung: <https://www.bfr.bund.de/>

Rapid Exchange of Information System (RAPEX), Verbraucherzugang:

<https://ec.europa.eu/safety-gate-alerts/screen/webReport>

Rapid Alert System for Food and Feed (RASFF), Verbraucherzugang:

<https://webgate.ec.europa.eu/rasff-window/screen/list?consumer=5007>

RASFF-Warmmeldungen auf der Seite des BVL gelistet:

https://www.bvl.bund.de/DE/Arbeitsbereiche/01_Lebensmittel/01_Aufgaben/04_Schnellwarnsystem/01_aktuelle_rasff_meldungen/aktuelle_meldungen_node.html

Verordnung (EG) Nr. 178/2002 des Europäischen Parlaments und des Rates vom 28.

Januar 2002 zur Festlegung der allgemeinen Grundsätze und Anforderungen des

Lebensmittelrechts, zur Errichtung der Europäischen Behörde für Lebensmittelsicherheit

und zur Festlegung von Verfahren zur Lebensmittelsicherheit: [https://eur-](https://eur-lex.europa.eu/legal-content/de/ALL/?uri=CELEX:32002R0178)

[lex.europa.eu/legal-content/de/ALL/?uri=CELEX:32002R0178](https://eur-lex.europa.eu/legal-content/de/ALL/?uri=CELEX:32002R0178)

Nutzerzentrierte Referenzmodellierung für Nachhaltigkeitskennzahlen

Datenstrukturierung für ein transparentes und Stakeholder-orientiertes Informationsangebot von Unternehmen an Investoren als Datennutzer

Raphaela Helbig ¹ und Sven von Höveling ²

Abstract: Investoren nutzen Nachhaltigkeitsinformationen zunehmend im Rahmen von Investitionsentscheidungen und der Bedarf nach soliden Nachhaltigkeitsbewertungen von Unternehmen steigt. Nachhaltigkeitskennzahlen sind dabei eine wichtige Teilmenge unter den Nachhaltigkeitsinformationen. Jedoch sind hochwertige Daten Mangelware – dies trifft aus dem Informationsangebot von Unternehmen und Drittanbietern zu. Dieser Beitrag stellt ein nutzerorientiertes Referenzmodell für die generische Reporting-Struktur von Nachhaltigkeitskennzahlen vor, das dazu beiträgt, dass für Investoren als Datennutzer Nachhaltigkeitsinformationen mit verbesserter Vergleichbarkeit und Qualität erhalten. Die Erarbeitung des Referenzmodells basiert auf einer Interviewstudie. Die Anwendbarkeit des Referenzmodells wird am Beispiel öffentlicher Daten der Volkswagen AG geprüft und kritisch hinsichtlich Informationslücken und Hindernissen diskutiert. Durch die Nutzerorientierung des Referenzmodells wird bezweckt, dass die so geschaffene Bereitstellung transparenter Fakten für Investoren als Nutzer von Nachhaltigkeitsdaten einerseits die Vergleichbarkeit und Qualität dieser Daten für die Weiterverwendung steigert und andererseits Greenwashing durch Unternehmen vorbeugt. Die Arbeit ergänzt die wissenschaftliche Fachliteratur der Wirtschaftsinformatik und des User Experience Designs mit einer nutzerzentrierten Methodik zur Entwicklung eines Daten-Referenzmodells, die explizit vom Mehrwert des Daten-konsumierenden Nutzers ausgeht. Das vorgestellte Referenzmodell für Nachhaltigkeitskennzahlen leistet einen Beitrag in der Fachliteratur zu Stakeholder-orientierter Unternehmensberichterstattung, indem es explizit die Anforderungen Daten-konsumierender Investoren berücksichtigt. Das Referenzmodell ermöglicht gesteigerten Informationsgehalt und eine standardisiert unternehmens-übergreifende Vergleichbarkeit der berichteten Daten.

Keywords: Referenzmodell, Nutzerzentrierung, Klassifizierung, Nachhaltigkeitsreporting, ESG, Stakeholderorientierung, ESG Disclosure, Nachhaltigkeit

¹ Volkswagen Aktiengesellschaft, Wolfsburg, Germany; Carl von Ossietzky Universität Oldenburg, Department of Computing Science, Business Information Systems / VLBA, Oldenburg, Germany, raphaela.helbig@uni-oldenburg.de; <https://orcid.org/0000-0003-3139-5562>

² Carl von Ossietzky Universität Oldenburg, Department of Computing Science, Business Information Systems / VLBA, Oldenburg, Germany, sven.von.hoeveling@uni-oldenburg.de; <https://orcid.org/0000-0003-3937-2429>

1 Motivation

Nachhaltigkeit gewinnt im Finanzsektor zunehmend an Bedeutung. Als Synonym für Nachhaltigkeit hat sich im Finanzsektor die Abkürzung ESG (Environment, Social, Governance) etabliert. Um Nachhaltigkeit im Rahmen von Investitionsentscheidungen zu berücksichtigen, greifen Investoren auf Nachhaltigkeitsinformationen zurück, die im Folgenden ESG-Informationen genannt werden. Aufgrund von Inkonsistenzen sind ESG-Informationen, die Unternehmen bereitstellen, schwer miteinander vergleichbar [KS19].³ Auch die Qualität der ESG-Informationen, die von Drittanbietern bereitgestellt werden, wird von Investoren als mangelhaft eingeschätzt, insbesondere aufgrund mangelnder Genauigkeit, Aktualität und der methodischen Rückwärtsgeandtheit⁴ [WP20]. In der Praxis folgen Investoren mit ihren Investitionsentscheidungen nicht blind den ESG-Bewertungen der Drittanbieter. Vielmehr haben Investoren eigene Herangehensweisen entwickelt, für die sie insbesondere ESG-Rohdaten der Drittanbieter nutzen [WP20].

Mit zunehmender Bedeutung der Nachhaltigkeit steigt der Bedarf nach aussagekräftigen und hochqualitativen ESG-Informationen über Unternehmen, die Investoren im Investitionsprozess nutzen können. Um den beschriebenen Herausforderungen zu begegnen und bisherige Kommunikationswege zu ergänzen, empfehlen Kotsantonis and Serafeim [KS20] Unternehmen eine proaktive Nachhaltigkeitsberichterstattung. Sie ermutigen Unternehmen dazu, die eigenen ESG-Kennzahlen für eine bessere Vergleichbarkeit mit den anderen Unternehmen anzupassen, zumindest auf Branchenebene und so Industriestandards zu etablieren [KS19].

Die Autoren dieses Beitrags arbeiten an einem Prototyp für die gezielte ESG-Kommunikation zur Nutzung für Finanzmarktakteure. Den Rahmen dafür bildet die Forschungskooperation *eco4fin – enabling ESG interaction* zwischen der Universität Oldenburg und der Volkswagen AG [Ca21]. Dabei wird auf Vorarbeiten zur Erforschung von Systemen und Datenstrukturen zur Nutzung einer digitalisierten Nachhaltigkeitsberichterstattung zurückgegriffen [IL01; Is08; IMS11; So11; So15].

In diesem Beitrag wird ein nutzerorientiertes Referenzmodell zur Bereitstellung nachhaltigkeitsbezogener Kennzahlen vorgestellt, das auf Grundlage einer Interviewstudie erarbeitet wurde. Die Bereitstellung von Kennzahlen nach dem vorgestellten Referenzmodell unterstützt Unternehmen bei einer nutzerorientierten Berichterstattung an Investoren und trägt auf diese Weise zu einer höheren Transparenz des Informationsangebots im Vergleich zum Status quo bei. Weiterhin erhalten Investoren, als Nutzer dieser Daten, durch die transparente Bereitstellung von Kennzahlen die Möglichkeit, die Vergleichbarkeit und Qualität besser als bisher einzuschätzen.

³ Zur Veranschaulichung listen Kotsantonis und Serafeim [KS19] für die sozialen Nachhaltigkeitsaspekte Mitarbeitergesundheit und -sicherheit verschiedene Kennzahlen aus der Unternehmensberichterstattung auf – beispielsweise Anzahl von Unfällen mit tödlichen Folgen [absolute Anzahl], Verletzungsrate pro 200 000 Arbeitsstunden [Index] und Unfallrate [Prozent]. Die mangelnde Vergleichbarkeit wird deutlich.

⁴ Rückwärtsgeandtheit: Rückschluss von vergangenen auf zukünftige Entwicklungen

2 Hintergrund – Spezifikationen für Metadaten in Datenportalen

Es gibt einen generellen Trend, Informationen zu digitalisieren, und diese maschinenlesbar in Form von Daten nutzbar zu machen. Dieses Kapitel beschäftigt sich mit Open Data und Datenportalen und verdeutlicht den Anwendungsbezug des generischen Datenmodells, das in diesem Beitrag vorgestellt wird. Ein großes und wachsendes Angebot an online zugänglichen Datenportalen ist zu beobachten, wie beispielsweise GovData⁵ und data.europa.eu⁶. Verwendete Plattformen sind beispielsweise Comprehensive Knowledge Archive Network (CKAN)⁷, Drupal-based Knowledge Archive Network (DKAN)⁸, OpenDataSoft⁹, Socrata¹⁰ und Junar¹¹ [Li16]. Die Nutzbarkeit der bereitgestellten Ressourcen ist unter anderem wesentlich von der Form der Bereitstellung abhängig. Die Qualität der Bereitstellung wird durch das Mitliefern von Metadaten erhöht. Vorteile von Metadaten und Anforderungen an Architekturen beschreiben Zuiderwijk et al. [Zi12]. Bei der Bereitstellung werden verschiedene Metadatenmodelle für die Metadaten eingesetzt, in der Regel abhängig von der Wahl der Plattform. CKAN verwendet beispielsweise ein eigenes Metadatenmodell, während viele andere Plattformen den Standard Data Catalog Vocabulary (DCAT) verwenden [Li16]. Dabei bauen die Modelle und Standards im Kern wiederum auf anderen etablierten Standards wie Dublin Core und Resource Description Framework (RDF) auf [Li16].

Die Verwendung von DCAT wird vom World Wide Web Consortium (W3C) empfohlen [KŠ18]. DCAT-AP¹² wurde auf Initiative der Europäischen Kommission für den Austausch zwischen den öffentlichen Verwaltungen der EU-Staaten aufbauend auf DCAT entwickelt [Li16]. DCAT-AP gilt als De-facto-Standard für Datenportale in Europa [KŠ18]. Eine Erweiterung von DCAT-AP ist DCAT-AP.de¹³, „das gemeinsame deutsche Metadatenmodell zum Austausch von offenen Verwaltungsdaten“ [Go21]. Die Umstellung von GovData auf dieses Voabular Ende des Jahres 2020 [Go20] verdeutlicht die Relevanz.

Ähnlich zu dem Vorhaben des vorliegenden Beitrags, aber in anderer Domäne, stellen Franke et al. [Fr20] ein Metadaten-Schema namens Plasma-MDS für die Domäne der Plasma-Technologie vor, das auf DCAT aufbaut. Paulet et al., [PFB19] haben anschließend für diese Domäne ein Konzept auf DKAN entwickelt, welches das Metadaten-Schema verwendet.

⁵ <https://www.govdata.de/>, Portal für Verwaltungsdaten von Bund, Ländern und Kommunen an

⁶ <https://data.europa.eu/de>, Portal für Daten zu Europa

⁷ <https://ckan.org/>

⁸ <https://getdkan.org/>

⁹ <https://www.opendatasoft.com/de/>

¹⁰ <https://www.tylertech.com/products/socrata>

¹¹ <https://www.junar.com/>

¹² <https://joinup.ec.europa.eu/collection/semantic-interoperability-community-semic/solution/dcat-application-profile-data-portals-europe/release/11>

¹³ <https://www.dcat-ap.de/>

3 Methode

Herleitung der Forschungsfrage

Institutionelle Investoren schätzen die verfügbare Vergleichbarkeit und Qualität der ESG-Informationen für die Nutzung im Investitionsprozess als ungenügend ein. Dies wird insbesondere am Beispiel von Nachhaltigkeitskennzahlen deutlich. Im Allgemeinen begegnen Datenportal-Anbieter dem Nutzerwunsch nach transparenten Angaben zu Vergleichbarkeit und Qualität von bereitgestellten Daten mit Spezifikationen der Datenmodellierung. Dieser Beitrag adressiert die Forschungsfrage: Welche Inhalte und Merkmale benötigt eine Reporting-Struktur für Nachhaltigkeitskennzahlen, damit Investoren diesen Daten für ihre Tätigkeit eine hohe Vergleichbarkeit und Qualität attestieren?

Zur Beantwortung der Forschungsfrage werden aus Sicht von Investoren als Datennutzer, wesentliche Inhalte und Merkmale für eine Reporting-Struktur für Nachhaltigkeitskennzahlen erarbeitet.

Interviewstudie zur Erarbeitung der nutzerzentrierten Sicht

Die nutzerzentrierte Perspektive wurde durch Erkenntnisse aus einer Interviewstudie ermittelt. Dazu diente das Material aus fünf Interviews mit institutionellen Investoren, die Teil einer weitergefassten Interviewstudie mit Finanzmarktakteuren von Mai bis Juli 2019 waren. Entwickelt wurde der halb-standardisierte Interviewleitfaden nach Helfferich [He11]. Alle Interviews wurden als Telefoninterviews durchgeführt, die Transkription erfolgte gemäß Dresing und Pehl [DP15]. Zur Auswertung des Interviewmaterials wurde die inhaltlich-strukturierende Qualitative Inhaltsanalyse angewendet [Ma14]. Eine deskriptive Beschreibung der Interviewpartner befindet sich im Anhang. Im Rahmen der Auswertung wurden Merkmale von Nachhaltigkeitsinformationen über Unternehmen mithilfe des Interviewmaterials identifiziert. Nach mehreren Konsolidierungsschleifen zeichneten sich Beschreibungen der Nutzer für Merkmale zur Qualität der Daten ab. In einem letzten Arbeitsschritt wurden die Kategorien den Dimensionen zugeordnet, die aus einem etablierten Modell zur Beschreibung der Datenqualität aus Nutzersicht [WS96] übernommen wurden.¹⁴ Für die anschließende Referenzmodellierung wurden schließlich aus den verdichteten Aussagen der Interviewpartner Anforderungen beschrieben. Daraus wurden Implikationen für das Referenzmodell abgeleitet und in Textform dokumentiert.

Deskriptive Beschreibung der Interviewpartner: Interviewpartner waren sieben Vertreter institutioneller Investoren, zu deren Tätigkeit der tägliche Umgang mit ESG-Informationen im Rahmen des Investitionsprozesses gehört. Drei Interviews wurden als

¹⁴ Das etablierte Modell zur Beschreibung der Datenqualität aus Nutzersicht geht auf mehrere Arbeiten am MIT zurück, wie beispielsweise die Studie von Wang und Strong [WS96]. Das Modell beinhaltet 16 Dimensionen, die vier Clustern zugeordnet werden.

Einzelinterviews durchgeführt, zwei weitere Interviews als Doppelinterviews¹⁵. Die Sprache der Interviews war Englisch (vier Interviews) und Deutsch (ein Interview). Die Interviewteilnehmer waren für Organisationen tätig, die im Bereich Vermögensverwaltung agieren. Die verwalteten Vermögenswerte der Organisationen lagen zum Zeitpunkt der Interviews zwischen 9 und 535 Milliarden Euro. Die Gesamtdauer der fünf Interviews betrug 172 Minuten, ein Interview dauerte durchschnittlich 34 Minuten.

Referenzmodellierung für ein Datenmodell

Die hier durchgeführte Referenzmodellierung¹⁶ diente der Spezifikation einer wiederverwendbaren Datenstruktur für Nachhaltigkeitskennzahlen, die eine nutzerzentrierte Bereitstellung dieser Daten speziell für institutionelle Investoren ermöglicht. Die Referenzmodellierung basierte auf den Kategoriensystemen, welche die Aussagen der Interviewpartner verdichten. Die Dokumentation des Referenzmodells erfolgte tabellarisch.

Evaluation der Anwendbarkeit als Fallbeispiel mit Inhalten aus einer verfügbaren Berichterstattung

Zur Evaluation des Referenzmodells wurde eine Prüfung auf Anwendbarkeit durchgeführt. Das nutzerzentriert erarbeitete Referenzmodell wurde für die Eingabe von Nachhaltigkeitskennzahlen eingesetzt. Dazu wurde ein generisches Template exemplarisch mit Beispieldaten von Volkswagen bestückt, wobei es sich um Nachhaltigkeitskennzahlen aus verschiedenen Themengebieten handelte. Als Quellen dienten die öffentlich verfügbaren Nachhaltigkeitsberichte und Geschäftsberichte der Volkswagen AG seit 2010.¹⁷

4 Ergebnisse und Interpretation

Während der Erarbeitungsphase des nutzerzentrierten Referenzmodell für die generische Reporting-Struktur für Nachhaltigkeitskennzahlen wurden Zwischenergebnisse dokumentiert. Kapitel 4.1 zeigt die Anforderungen der Nutzer, die die Erkenntnisse der Interviews verdichten. Das nutzerzentrierte Referenzmodell wird in Kapitel 4.2 vorgestellt und die Evaluation des Referenzmodells veranschaulicht die Anwendbarkeit.

¹⁵ Bei einem Doppelinterview wurden zwei Personen gleichzeitig interviewt.

¹⁶ Der Begriff Referenzmodell wird in der Wirtschaftsinformatik im deutschen Sprachraum mittlerweile mehrheitlich für wiederverwendungsorientierte Modelle für Daten und Prozesse verwendet, die für die Wiederverwendung entwickelt oder genutzt werden. Abweichend dazu wird der Begriff Referenzmodell in den Vereinigten Staaten generell für Referenz-Daten verwendet – das Begriffsverständnis dieses Beitrags orientiert sich jedoch am deutschsprachigen Verständnis. [FL07]

¹⁷ Die Nachhaltigkeitsberichte und Geschäftsberichte der Volkswagen AG erschienen jährlich. Genutzt wurden als Quellen öffentlich verfügbare PDF-Dateien, lediglich der Nachhaltigkeitsbericht für das Jahr 2015 war nicht öffentlich auffindbar.

4.1 Nutzerzentrierte Anforderungen

Die Aussagen der Interviewpartner werden im Folgenden anhand der gebildeten Kategorien verdichtet beschrieben. Die beschriebenen Merkmale zur Qualität der betreffen Nachhaltigkeitsdaten, die Investoren nutzen, beschränken sich nicht nur auf Daten, die von Unternehmen bereitgestellt werden. Es wurde deutlich, dass die Interviewpartner ähnliche Merkmale und deren Ausprägungen beschrieben. Auffällig war hierbei, dass je nach Anwendungsfall die Nutzer unterschiedliche Merkmale als besonders wichtig einstufen. Das Kategoriensystem ist in Tabelle 1 dargestellt.

Auf Grundlage der Erkenntnisse der Interviewauswertung wurden Implikationen für die Referenzmodellierung abgeleitet, z.B. Attribute in den Metadaten. Es wurden Anforderungen aus Sicht der Interviewpartner formuliert und Anknüpfungspunkte identifiziert, wie die Anforderungen in der Referenzmodellierung berücksichtigt werden. Jedoch lassen sich einige der Qualitätsmerkmale, die aus den Interviews hergeleitet wurden, nicht über ein Datenmodell abdecken, sondern betreffen vielmehr die Gestaltung von Prozessen und Anwendungsoberflächen sowie die Bereitstellung von Inhalten durch ein oder mehrere Datenproduzenten (Unternehmen) – beides ist nicht Teil dieses Beitrags.

Im Folgenden werden für ausgewählte Kategorien Ergebnisse beschrieben. Die Interviewpartner machen das Bedürfnis nach „Klarheit der Interpretation“ deutlich. Die Interviewpartner beschreiben, dass diese beispielsweise beeinflusst werden kann durch verbindliche Definitionen und Berechnungsarten sowie passende Formulierungen. Dies wird über die Attribute *berichtendes Unternehmen*, *Titel*, *Einheit*, *Region*, *Erhebungsmethode* (für Datenreihe) und *Erhebungsbeschreibung* (für Datenpunkt) berücksichtigt. Darüber hinaus beschreiben die Interviewteilnehmer die zeitliche Dimension des Informationsangebots auf unterschiedlichen Ebenen. Angaben zur „Gültigkeit der Information (Veröffentlichungszeitpunkt)“ und „Frequenz der Bereitstellung“ sind nicht immer transparent. Außerdem wird die zeitliche Verzögerung („Latenz“) der Bereitstellung von Daten kritisiert, deren Grund in den langen Reportingzyklen von Unternehmen gesehen wird. Im Referenzmodell soll mehr Klarheit geschaffen werden durch die Attribute *Häufigkeit der Eingabe* und *Erstellung*. Letzteres bezieht sowohl die Angabe zum Zeitpunkt als auch die Charakterisierung (Ersteintrag oder Update) ein.

4.2 Nutzerzentriertes Referenzmodell einer Reporting-Datenstruktur für Nachhaltigkeitskennzahlen und Evaluierung mit Beispieldaten

Bei der Referenzmodellierung wurde unterschieden in Datenpunkt und Datenreihe. Ein Datenpunkt umfasst neben der Zuordnung des Zeitraums zum Wert (Key-Value) zusätzlich umfangreiche Metadaten. Eine Datenreihe wiederum verweist auf beliebig viele

Dimensionen der Datenqualität erweitert nach Wang und Strong (1996)	Kategorien
Zugänglichkeit	- Sichtbarkeit
Bearbeitbarkeit	- Umrechnungsfähigkeit - Extrahierbarkeit und Integrierbarkeit *
Zugangssicherheit	- Zugangsberechtigung
Eindeutige Auslegbarkeit	- Klarheit der Interpretation
Einheitliche Darstellung	- Vergleichbarkeit
Übersichtlichkeit	- Umfang * - Formale Struktur und Systematik - Ästhetik * - Zentraler Ort der Bereitstellung *
Verständlichkeit	- Natürliche Sprache * - Klarheit und Nachvollziehbarkeit
Format	- Formen und Formate
Relevanz	- Zielgruppenorientierung - Materialität - Relevanz hinsichtlich der Branche - Individuelle Relevanz *
Informationsumfang	- Datendeckung * - Abdeckung mehrerer Unternehmen * #
Wertschöpfung	- Integration der Nachhaltigkeits- und Finanzinformationen
Aktualität	- Gültigkeit/Veröffentlichungszeitpunkt - Frequenz der Bereitstellung - Latenz
Vollständigkeit	- Repräsentativität - Vergangene Informationen u. zeitliche Verläufe - Abstraktionsgrad und Granularität
Glaubwürdigkeit	- Auskunftsfähigkeit * - Reaktions- und Dialogfähigkeit - Auskunftswillen * - Widerspruchsfreiheit
Fehlerfreiheit	- Einheitliche Prüfungsvorgänge - Absicherung/Prüfung durch Dritte
Objektivität	- Unabhängigkeit
Reputation	- Seriosität*

Tab. 1: Aussagen der Interviewpartner zu thematisch-inhaltlichen Schwerpunkten von ESG-Daten ohne speziellen Fokus auf Kennzahlen.

Anmerkungen: Der Stern * steht dafür, dass sich aus der Kategorie keine Implikationen für ein Datenmodell ergeben. Die Raute # kennzeichnet, dass diese Kategorie sich nicht auf das Reporting eines einzelnen Unternehmens bezieht.

Datenpunkte und besitzt ebenso umfangreiche Metadaten. Datenpunkte werden nicht nur als Key-Value paare innerhalb der Datenreihen betrachtet, da a) sowohl die Einträge für Datenpunkte als auch für Datenreihen historisierbar sein können und b) die Einträge zu einem Datenpunkt derart umfangreich sind, dass sie ein eigenes Modell begründen. Die Strukturen der beiden Modelle werden in den Tabellen 2 und 3 dargestellt. Die Bezeichnungen der Klassen, der Attribute und der Methoden sind für ein besseres Verständnis weniger technisch gewählt.

Die Evaluierung des Referenzmodells erfolgte mit Beispieldaten. Die Anwendbarkeit des nutzerzentriert entwickelten Referenzmodells wurde mit verschiedenen Nachhaltigkeitskennzahlen der Volkswagen AG geprüft. Die Eingabe der Daten gestaltete sich aufwendig, da die Daten aus PDF-Dateien übernommen wurden. Einige Inhalte konnten von den Autoren nicht direkt aus den Quellen des bestehenden Informationsangebots entnommen werden, in diesen Fällen wurden Annahmen mithilfe der eigenen Expertise getroffen. Die jeweils dritte Spalte in Tabelle 2 und Tabelle 3 veranschaulicht die Anwendung der nutzerzentriert erarbeiteten Struktur für die direkten Treibhausgas-Emissionen (Scope 1) der Volkswagen AG exemplarisch für eine Datenreihe und ausgewählte Datenpunkte.

Für die mit Stern markierten Attribute empfehlen die Autoren eine Historisierung, um das große Nutzerbedürfnisses hinsichtlich Transparenz zu adressieren. Technisch sind dabei zahlreiche Umsetzungsmöglichkeiten denkbar, beispielsweise mithilfe von Datenbanken wie immudb¹⁸, die persistente Historisierung ermöglichen, oder über eine Datenstruktur, wie im angedeuteten Klassendiagramm in Abbildung 1.

Im Forschungsprojekt eco4fin wurde ein UML-Klassendiagramm verwendet, wie in Abbildung 1 skizziert. In der Darstellung sind aus Gründen der Transparenz und Nachvollziehbarkeit die Klassen derart konzipiert, dass die meisten der (Meta-) Informationen der Datenreihe und des Datenpunktes eine Historie haben. Dazu wurden für die betroffenen Attribute eigene Klassen erstellt, deren Instanzen einen Attributzustand darstellen. Wird das Attribut verändert, wird eine neue Instanz der Klasse erstellt und die bisherige bleibt unverändert.

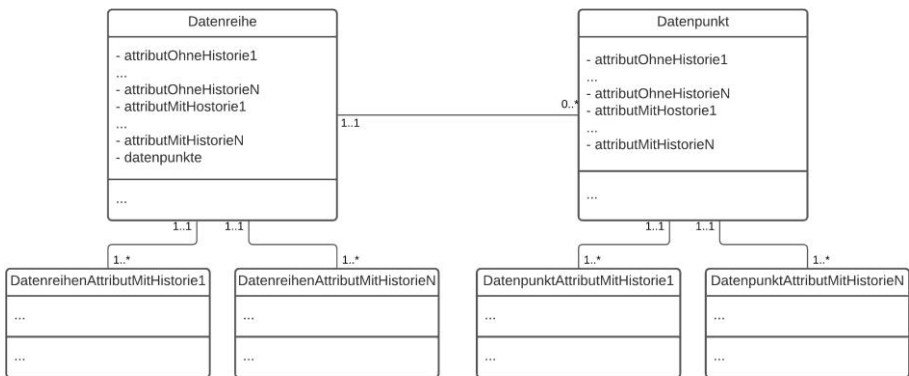


Abb. 1: Skizze für ein UML-Klassendiagramm zur Berücksichtigung einer Historisierung für einige der Attribute für eine mögliche Modellierung des Modells

¹⁸ <https://www.codenotary.com/technologies/immudb/>

Attribut (-Klasse)	Beschreibung	Beispieldaten
ID	Einzigartige Kennung der Datenreihe	2
Zeitpunkt der Erstellung	Zeitpunkt der Erstellung der Datenreihe	2010-05-01, no specific timestamp
Zeitpunkt der letzten Änderung	Zeitpunkt der letzten Änderung der Datenreihe	2011-05-01, no specific timestamp
Titel *	Kompakter Titel, der die Datenreihe charakterisiert und vom Nutzer leicht erfassbar ist.	Direct Group CO2 Emissions (Scope 1) per vehicle
Einheit *	Einheit der Punkte der Datenreihe	kg/vehicle
Region *	Lokale Bezugsgröße (z. B. Standorte, Regionen, weltweit)	Group worldwide, Group production sites
Editier-Status	Gibt Aufschluss darüber, ob die Datenreihe noch aktiv fortgeschrieben (mit aktuellen Werten), lediglich gepflegt oder lediglich noch vorgehalten (zur Ansicht wie ein Archiv) wird.	Active
Berichtendes Unternehmen	Benennt eindeutig das berichtende Unternehmen	Volkswagen AG, Group
Häufigkeit der Eingabe	Eintrag, der angibt, wie oft diese Datenreihe in der Regel aktualisiert wird.	Annually
Sichtbarkeit *	Grundlegende Sichtbarkeitseinstellung. Gib an, ob die Datenreihe öffentlich einsehbar oder nur für interne Zwecke gedacht ist. Nicht für vollständige Zugriffsbeschränkungen gedacht.	Public
Erhebung *: Beschreibung, Version, Erhebungsstandard	Details zur Erhebung. Version, Standard und Freitext (Beschreibung)	<i>Method Details: Green-house Gas Protocol Scope 1 Inventory. Version: 1. Description: [...]</i>
Themen-Tags (Liste) *	Ordnet fachliche Themenbereiche zu.	<i>[co2 emissions, greenhouse gas emissions]</i>
Verwandte Datenreihen (Liste) *: Datenreihe, Verwandtschafts-Beschreibung	Jeder Eintrag der Liste zeigt und erklärt die Beziehung zu einer anderen Datenreihe.	<i>[Data row: 1. Verwandtschaftsbeschreibung: Same Scope, but different unit "absolute value"]</i>
Beschreibung *	Ausführliche Beschreibung zum Inhalt dieser Datenreihe	Direct CO2 emission (Scope 1) per vehicle, includes emissions [...]
Verantwortliche Person *	Verantwortung für den Datenreihe und den Erhebungsprozess der Werte liegt bei einem Fachverantwortlichen. Hilfreich für Kontaktierung bei Rückfragen	<i>Person with ID 18</i>
Fachabteilung des Erstellers	Aufschluss zum Ursprung der Eingabe der Datenreihe	<i>Group Sustainability Management</i>
Datenpunkte (Liste) *	Datenpunkt der Datenreihe. Modell der Datenpunkte wird in Tabelle 3 beschrieben.	<i>[Data point with ID 13, Data point with ID 14]</i>

Tab. 2: Modell für Datenreihe mit Beschreibung und exemplarischen Inhalt für GHG Emissionen Scope 1 der Volkswagen Group

Bei den mit Stern * gekennzeichneten Attributen wird eine Historisierung empfohlen. *Kursiv* sind Inhalte gekennzeichnet, die von den Autoren nicht aus dem Volkswagen Nachhaltigkeitsberichten und Geschäftsberichten entnommen wurden, sondern mit Hintergrund der eigenen Expertise frei angenommen wurden. Eckige Klammern symbolisieren eine Liste.

Attribut (-Klasse)	Beschreibung	Beispieldaten
ID	Einzigartige Kennung des Datenpunkts	13
Zeitpunkt der Erstellung	Zeitpunkt der Erstellung des Datenpunkts	2019-05-01, no specific timestamp
Zeitpunkt der letzten Änderung	Zeitpunkt der letzten Änderung des Datenpunkts	2019-05-10, no specific timestamp
Zuweisung zu Reporting-Standards (Liste) *: Reporting-Standard, Beschreibung	Zuweisung und Beschreibung der Verknüpfung mit zugeteilten Reportingstandards und Unterkategorisierung	[Reporting standard: GRI 305-1. No further description provided]
Wert *	Zahlenwert des Datenpunktes	346
Erklärung *	Mitgelieferte Erklärung, z. B. bei starken Schwankungen des Werts oder gegenteiliger Zielsetzung	[Increased number of produced vehicles in comparison to previous reporting year]
Quellen (Liste) *	Publikationen, in denen der Wert berichtet wurde	[Volkswagen Sustainability Report 2019 page 70]
Erhebungsbeschreibung *	Konkrete Details zum Zustandekommen des Werts, wie sie z. B. in Fußnoten im Nachhaltigkeitsbericht vorkommen	Cars and light commercial vehicles
Sichtbarkeit *	Grundlegende SichtbarkeitsEinstellung. Gibt an, ob die Datenreihe öffentlich einsehbar oder nur für interne Zwecke gedacht ist. Nicht für vollständige Zugriffsbeschränkungen gedacht.	Public
Anmerkungen (Liste) *	Weiterführende Bemerkungen zum Datenpunkt	[Despite the inclusion of new production sites, direct CO2 emissions in the passenger ...]
Audits (Liste) *: Prüfungsgesellschaft, Prüfungslevel, Sonstiges	Prüfungsgesellschaft: Name der Prüfungsgesellschaft, die das Audit durchgeführt hat. Prüfungslevel: Angabe zum Level der Prüfung, z. B. Wirtschaftsprüfung als „Limited Assurance“ oder „Reasonable Assurance“. Sonstiges: Freitext zur weiterführenden Erklärung zum Audit.	[external Auditor: PricewaterhouseCoopers GmbH Wirtschaftsprüfungsgesellschaft. audit level: Limited assurance. others: Assurance from 2018]
Fachabteilung des Erstellers	Aufschluss zum Ursprung der Eingabe der Datenreihe	Group Sustainability Management
Zeitraum: Beginn, Ende, Beschreibung	Beginn: Konkreter Startpunkt des Zeitraums, auf den sich der Wert bezieht. Ende: Konkreter Endpunkt des Zeitraums, auf den sich der Wert bezieht. Beschreibung: Kurze Beschreibung des Zeitraums.	Start: 2018-01-01. End: 2018-12-31. Description: calendar years 2018.

Tab. 3: Modell für Datenpunkt mit Beschreibung und exemplarischen Inhalt für GHG Emissionen Scope 1 der Volkswagen Group

Anmerkungen: Bei den mit Stern * gekennzeichneten Attributen wird eine Historisierung mit Beschreibung empfohlen. *Kursiv* sind Inhalte gekennzeichnet, die von den Autoren nicht aus dem Volkswagen Nachhaltigkeitsberichten und Geschäftsberichten entnommen wurden, sondern mit Hintergrund der eigenen Expertise frei angenommen wurden. Eckige Klammern symbolisieren eine Liste.

5 Diskussion und Fazit

Dieser Beitrag präsentiert die nutzerzentrierte Erstellung eines Referenzmodells für Nachhaltigkeitskennzahlen im Kontext der Unternehmensberichterstattung –speziell zur Nutzung durch die Stakeholder-Gruppe der Investoren. Durch die Nutzerzentrierung des Referenzmodells wird bezweckt, dass Nutzer diese Daten möglichst zielgerichtet für ihre Tätigkeit weitenutzen können. Durch die elaborierte Datenmodellierung, die eine Vielzahl an Metainformationen vorsieht, wird Transparenz hergestellt, die auf eine bessere Vergleichbarkeit und Qualität der Daten einzahlt. Das Referenzmodell nutzt Metainformationen, mit denen zusätzlich zu den reinen Werten weitere Informationen den Kontext transparent machen (z. B. zum Zustandekommen der Daten).

Folgende Limitationen wurden identifiziert. Es konnte nicht allen Anforderungen an jene Qualitätsmerkmale mit dem Datenmodell begegnet werden, die im Rahmen der Interviewstudie identifiziert wurden. Einige identifizierte Anforderungen wirken sich vielmehr auf die Gestaltung von Dateninfrastruktur, von Prozessen und Anwendungsoberflächen aus – beides ist nicht Teil dieses Artikels. Weiterhin beschränkt sich die Nutzerzentrierung des Modells auf die Sicht der Daten-Konsumenten, die Produzenten-Sicht ist nicht Teil des präsentierten Modells. Folglich ist das entwickelte Modell möglicherweise noch nicht ausgereift, um als Grundlage für eine Implementierung zu dienen – es besteht Forschungsbedarf. Außerdem fehlt bislang der Abgleich des Datenmodells mit existierenden Spezifikationen, die ausführlich Metadaten einbeziehen.

Das vorgestellte Datenmodell wurde für die Verwendung im Prototyp einer ESG-Kommunikationsplattform für Finanzmarktakteure zur Speicherung und Präsentation von Kennzahlen inklusive Metadaten erarbeitet [Ca21]. Im Beitrag wurde die Anwendbarkeit des Referenzmodells veranschaulicht und die Struktur wurde mit Beispieldaten der Volkswagen AG ausgefüllt. Auf diese Weise wurden die identifizierten Nutzeranforderungen dem aktuell öffentlich verfügbaren Informationsstand für Nachhaltigkeitskennzahlen gegenübergestellt und bestehende Informationslücken deutlich. Darüber hinaus zeigte sich, dass die manuelle Erstellung der Datensätze gemäß der vorgeschlagenen Datenmodellierung folgende Risiken mit sich bringt: Fehleranfälligkeit (z. B. Vertippen) und Arbeitsaufwand des Prozesses. Dies könnte einen Hinderungsgrund für die Adaptation in der Praxis darstellen.

Weiterer Forschungsbedarf besteht zur Implementierung des Referenzmodells, die über den eco4fin-Prototyp mit Beispieldaten der Volkswagen AG hinausgeht. Im Rahmen einer Implementierung werden sich weitere Fragen stellen, die durch die Charakteristik eines Referenzmodells nicht ausreichend geklärt oder vorgegeben werden können. Beispielsweise ist es ratsam, dass Unternehmen, die Open Data ermöglichen wollen, sich an gängigen Standards wie RDF oder DCAT zur Modellierung der Metadaten hält.

Zur umfassenden Nutzerzentrierung ist zusätzlich die Perspektive der Unternehmen einzubeziehen, die die Daten bereitstellen (Producer-Seite). Denkbar ist eine weiterführende Evaluierung des vorgestellten Referenzmodells, indem überprüft wird, ob

Vorteile durch die Bereitstellung von Metadaten, wie von Zuiderwijk et al. [Zi12] beschrieben, realisiert werden. Als ein möglicher nächster Schritt ist die Prüfung einer Zuordnung des entwickelten Referenzmodells auf den DCAT-Standard oder auf verwandte Spezifikationen denkbar, ähnlich dem Vorgehen von Franke et al. [Fr20] in einer anderen Domäne. Weiterhin liegt die Erstellung eines konkreten, technischen Konzepts nahe, wie es Paulet et al. [PFB19] es für eine andere fachliche Domäne vorstellen.

Das präsentierte Referenzmodell leistet einen Beitrag für Stakeholder-orientiertes Nachhaltigkeitsreporting von Unternehmen, indem für Kennzahlen explizit die Anforderungen von Investoren berücksichtigt werden. Für Wissenschaftler der Wirtschaftsinformatik und des User-Experience-Design ergänzt die angewendete nutzerzentrierte Methodik, die explizit vom Mehrwert des Nutzers der berichteten Kennzahlen ausgeht, die Fachliteratur. Für Praktiker im Bereich Nachhaltigkeitsreporting werden mit Fokus auf Stakeholderorientierung neben dem Referenzmodell auch Hinweise zur Gestaltung der Datenarchitektur, des Datenmanagements und zur Berichterstattung gegeben. Das Referenzmodell liefert Praktikern eine Grundlage für eine generische Reporting-Struktur für Nachhaltigkeitskennzahlen.

Disclaimer

Ergebnisse, Meinungen und Schlüsse dieser Veröffentlichung sind nicht notwendigerweise die der Volkswagen Aktiengesellschaft.

Literaturverzeichnis

- [AT21] Abhayawansa, S.; Tyagi, S.: Sustainable Investing: The Black Box of Environmental, Social, and Governance (ESG) Ratings. *The Journal of Wealth Management*, 24(1), 49–54, <https://doi.org/10.3905/jwm.2021.1.130>, 2021.
- [Ca21] Carl von Ossietzky Universität Oldenburg: *eco4fin*. <https://eco4fin.de/>, accessed: 17/05/2021, 2021.
- [CSS21] Christensen, D.; M., Serafeim. G.; Sikochi, A.: Why is Corporate Virtue in the Eye of The Beholder? The Case of ESG Ratings. *The Accounting Review*, TAR-2019-0506. <https://doi.org/10.2308/TAR-2019-0506>, 2021.
- [DP15] Dresing, T.; Pehl, T.: *Manual (on) Transcription: Transcription Conventions, Software Guides and Practical Hints for Qualitative Researchers.*, 2015.
- [EK19] Eccles, R. G.; Klimenko, S.: The Investor Revolution. *Harvard Business Review*, 106-116, 2019.
- [FL07] Fettke, P.; Loos, P.: *Reference Modeling for Business Systems Analysis*: IGI Global. <https://doi.org/10.4018/978-1-59904-054-7>, 2007.

- [Fr20] Franke, S.; Paulet, L.; Schäfer, J.; O'Connell, D.; Becker, M. M.: Plasma-MDS, a metadata schema for plasma science with examples from plasma technology. *Scientific Data*, 7(1), 439, <https://doi.org/10.1038/s41597-020-00771-0>, 2020.
- [Gl19] Global Sustainable Investment Alliance: 2018 Global Sustainable Investment Review., http://www.gsi-alliance.org/wp-content/uploads/2019/03/GSIR_Review2018.3.28.pdf, accessed: 17/05/2021, 2019.
- [Go20] GovData: GovData auf DCAT-AP.de umgestellt, <https://www.govdata.de/neues/-/blogs/govdata-auf-dcat-ap-de-umgestellt>, accessed: 17/05/2021, 2020.
- [Go21] GovData: DCAT-AP.de-Start, <https://www.dcat-ap.de/>, accessed: 17/05/2021, 2021.
- [He11] Helfferich, C.: Die Qualität qualitativer Daten Manual für die Durchführung qualitativer Interviews. <https://doi.org/10.1007/978-3-531-92076-4>, 2011.
- [HM19] Helbig, R.; Marx Gómez, J.: Kommunikationsfäden im Nadelöhr – Fachliche Prozessmodellierung der Nachhaltigkeitskommunikation am Kapitalmarkt. In (Ludwig, T.; Pipek, V., eds.): Proceedings of the 14th International Conference on Wirtschaftsinformatik (WI2019): Human Practice. Digital Ecologies, 1359–1373, <https://aisel.aisnet.org/wi2019/track12/papers/1/>, 2019.
- [IL01] Isenmann, R.; Lenz, C.: Customized corporate environmental reporting by internet-based push and pull technologies. *Eco-Management and Auditing*, 8(2), 100–110, 2001.
- [IMS11] Isenmann, R.; Marx Gómez, J.; Süpke, D.: Making Stakeholder Dialogue for Sustainability Issues Happen – Benefits, Reference Architecture and Pilot Implementation for Automated Sustainability Reporting a La Carte. 44th Hawaii International Conference on System Sciences (HICSS), 1–11. 2011.
- [Is08] Isenmann, R.; Arndt, H.-K.; Christoph Bey; Marx Gómez, J.; Solsbach, A.: XBRL-Taxonomie für Nachhaltigkeitsberichte: Vorschlag einer harmonisierten Dokumentstruktur zur Online-Kommunikation von Organisationen. In (ed.): Proceedings of the Multikonferenz Wirtschaftsinformatik. MKWI 2008, München, 26.2.2008 - 28.2.2008, 602–613, 2008.
- [KŠ18] Klímeck, J.; Škoda, P.: Linked Pipes DCAT-AP Viewer: A Native DCAT-AP Data Catalog. International Semantic Web Conference, <https://www.semanticscholar.org/paper/LinkedPipes-DCAT-AP-Viewer%3A-A-Native-DCAT-AP-Data-Kl%C3%ADmek-%C5%A0koda/28ab7bcde1b5db660ac280e426556e96d599daed>, accessed: 17/05/2021, 2018.
- [KS19] Kotsantonis, S.; Serafeim, G.: Four Things No One Will Tell You About ESG Data. *Journal of Applied Corporate Finance*, 31(2), 50–58, <https://doi.org/10.1111/jacf.12346>, 2019.
- [Li16] Lisowska, B.: Metadata for the open data portals. <http://devinit.org/wp-content/uploads/2018/01/Metadata-for-open-data-portals.pdf>, accessed: 17/05/2021 2016.

- [ME14] Maali, F.; Erickson, J.: Data Catalog Vocabulary (DCAT). w3.org. <https://www.w3.org/TR/vocab-dcat-1/>, accessed: 17/05/2021, 2014.
- [Ma14] Mayring, P.: Qualitative content analysis: Theoretical foundation, basic procedures and software solution. 2014.
- [PFB19] Paulet, L.; Franke, S.; Becker, M. M.: Management interdisziplinärer Forschungsdaten mit DKAN. <https://doi.org/10.5281/ZENODO.3340938>, 2019.
- [Pr20] PRI Association: Principles for Responsible Investment., <https://www.unpri.org/download?ac=10948>, accessed: 17/05/21, 2020.
- [So11] Solsbach, A.; Süpke, D.; vom Berg, B. W.; Marx Gómez, J.: Sustainable Online Reporting Model: A Web Based Sustainability Reporting Software. In (Golinska, P.; Fertsch, M.; Marx Gómez, J., eds.): Information Technologies in Environmental Engineering: New Trends and Challenges. Springer, 165–177, 2011.
- [So15] Solsbach, A.: Document Engineering als Ansatz für eine überbetriebliche Nachhaltigkeitsberichterstattung. Shaker, <https://oops.uni-oldenburg.de/id/eprint/2481>, accessed: 17/05/2021, 2015.
- [TKS20] Threlfall, R.; King, A.; Shulman, J.; Bartels, W.: The time has come: The KPMG Survey of Sustainability Reporting 2020. <https://assets.kpmg/content/dam/kpmg/xx/pdf/2020/11/the-time-has-come.pdf>, accessed: 17/05.2021, 2020.
- [WS96] Wang, R. Y.; Strong, D. M.: Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems*, 12(4), 5–33. <https://doi.org/10.1080/07421222.1996.11518099>, 1996.
- [WP20] Wong, C.; Petroy, E.: Rate the Raters 2020: Investor Survey and Interview Results. SustainAbility (an ERM Group company). <https://sustainability.com/wp-content/uploads/2020/03/sustainability-ratetheraters2020-report.pdf>, accessed: 17/05/2021, 2020.
- [Zi12] Zuiderwijk, A; Jeffery, K., Janssen, M.: The Potential of Metadata for Linked Open Data and its Value for Users and Publishers. *JeDEM – eJournal of eDemocracy and Open Government*, 4(2), 222-244, 2012.

Die umweltgerechte Stadt – Entwicklung und Umsetzung einer GIS-gestützten, quartiersbezogenen Umweltbelastungsanalyse im Land Berlin

Dr.-Ing. H.-Josef Klimeczek¹

Zusammenfassung: Um einen Gesamtüberblick über die Umweltbelastungssituation im Land Berlin zu erhalten, hat die Senatsverwaltung für Stadtentwicklung und Umwelt in den Jahren 2010-2019 - bundesweit erstmalig - ein neues zweistufiges Umweltgerechtigkeitsmonitoring entwickelt und implementiert. Das neue Stadtbeobachtungssystem besteht aus einem umweltbezogenen „Kernindikatorenset“, der durch zusätzliche gesundheitsorientierte, soziale und städtebauliche/stadt-planerische „Ergänzungsindikatoren“ fachlich-inhaltlich untersetzt wird. Die hieraus entwickelte kleinräumige Mehrfachbelastungskarte (Berliner Umweltgerechtigkeitskarte 2016/17) zeigt, dass der größte Teil der drei-, vier- und fünffach belasteten Quartiere im hochverdichteten „Erweiterten Innenstadtbereich“ liegt.

Abstract:

In the years 2010 – 2019, the Senate Department for Urban Development and Environment of the Land of Berlin developed *and* implemented a two-stage monitoring of environmental justice. For the first time in Germany, this monitoring system aims to an overview of the condition of the environment in the Land of Berlin. The monitoring system consists of an environment associated core indicator set completed by indicators of public health, social state of the city dwellers, and indicators of urban planning/urban development. The resulting detailed small-scale map on multiple exposures (Berliner Umweltgerechtigkeitskarte 2016/17 – Berlin Map of Environmental Justice 2016/17) shows that the majority of the city districts (so-called planning areas) with high values of environmental burden for three, four or five parameters of exposure are located in the extended city centre area.

Keywords: Umweltgerechtigkeit, Post-Corona-Stadt, Umweltbelastungsmonitoring, urbane Gesundheit, Stadtentwicklung

1 Einleitung

Für eine erfolgreiche und umsetzungsorientierte Stadtentwicklung sowie Umweltpolitik ist eine ausreichende und aussagefähige Datenbasis eine unverzichtbare Grundlage. Dies

¹ Berliner Institut für Umweltgerechtigkeit – BIUg., Pfalzburger Str. 43/44, 10179 Berlin, E-Mail: E-Mail: dr.kli-

gilt vor allem für flächendeckende und kleinräumig vergleichbare Umweltdaten, denn inwieweit Umweltstandards festgelegt, auf welcher Ebene diese Festlegung erfolgt und wie sie umgesetzt werden, ist abhängig vom Wissen über Wirkungszusammenhänge und/oder Fragen der Zweckmäßigkeit und der politisch-administrativen Durchsetzbarkeit.

Um die Arbeits- und Entscheidungsgrundlagen in den Berliner Senats- und Bezirksverwaltungen zu verbessern, haben die Senatsverwaltung für Umwelt, Verkehr und Klimaschutz in enger Zusammenarbeit mit der Senatsverwaltung für Gesundheit und Soziales – bundesweit erstmalig – die Grundlagen für ein gestuftes Umweltgerechtigkeitsmonitoring entwickelt. Das neue Berliner Stadtbeobachtungssystem hat den Anspruch, durch die kleinräumige und handlungsorientierte Beschreibung der Umweltqualität einen Gesamtüberblick über die Umweltbelastungssituation in der Hauptstadt zu ermöglichen und gleichzeitig verbesserte Grundlagen für die Festlegung von Umweltzielen bereitzustellen. Der Ansatz ist integrativ und auf die gesamtstädtische Ebene bezogen. Das hierfür entwickelte praxistaugliche System des zweistufigen Umweltgerechtigkeitsmonitorings mit dem Kern- und Ergänzungsindikatorensatz gilt bundesweit und europaweit als richtungsweisend, da es erstmalig ermöglicht, „Räume mit Mehrfachbelastungen“ zu identifizieren. Das neue Monitoringsystem versteht sich als ein Frühwarnsystem, das die bereits etablierten Stadtbeobachtungssysteme im Land Berlin fachlich-inhaltlich weiter untersetzt und Grundlagen für die Definition von Handlungsfeldern bereitstellt. Fachspezifische Berichterstattungen werden durch das neue Monitoring nicht ersetzt, sondern durch eine zusammenfassende Betrachtung der Umweltsituation bzw. der sozialräumlichen Verteilung der gesundheitsrelevanten Umweltbelastungen ergänzt. Mit den vorliegenden Umweltgerechtigkeitsanalysen liegt für die Hauptstadt die erste kleinräumige Bestandsaufnahme zur gesundheitsrelevanten Umweltqualität in den Quartieren und Teilbereichen der Hauptstadt vor. Der Berliner Indikatorensatz wurde in enger Kooperation zwischen den Senatsressorts Stadtentwicklung, Gesundheit und Soziales, dem Umweltbundesamt, dem Amt für Statistik Berlin- Brandenburg sowie externen Forschungseinrichtungen in den Jahren 2010 – 2017 entwickelt und die Praxistauglichkeit im Kontext der Berliner Stadtentwicklung und Umweltplanung überprüft.

2 Methodik

Das Berliner Umweltgerechtigkeitsmonitoring beruht im Wesentlichen auf der Auswertung und Aggregation vorhandener Datengrundlagen aus verschiedenen Quellen, u. a. Umweltatlas, Luftreinhalte- und Lärminderungsplanung, Monitoring Soziale Stadtentwicklung (MSS), Einschulungsuntersuchungen - ESU. Es ist als zweistufiges Verfahren (Kernindikatoren und Ergänzungsindikatoren) angelegt. Für die Auswahl der Gesundheits-, Sozial-, Stadtplanungs- und Umweltdaten war entscheidend, dass die Daten in unterschiedlichen Abständen kleinräumig, auf der Ebene der Planungsräume erhoben werden, d. h. auch über längere Zeiträume hin verfügbar sind. Analog zum Monitoring Soziale Stadtentwicklung (MSS) wurde bei den Umweltgerechtigkeitsanalysen die räumliche Gliederung Berlins auf den drei Ebenen der Lebensweltlich orientierten Räume (LOR) zugrunde gelegt, die sich hierarchisch ineinander überführen lassen. Für das gestufte Indikatorensystem des Umweltgerechtigkeitsmonitorings wurde die kleinste Einheit – die

447 Planungsräume – mit einer Raumgröße von durchschnittlich 7.500 Einwohnern gewählt (Abb. 1).

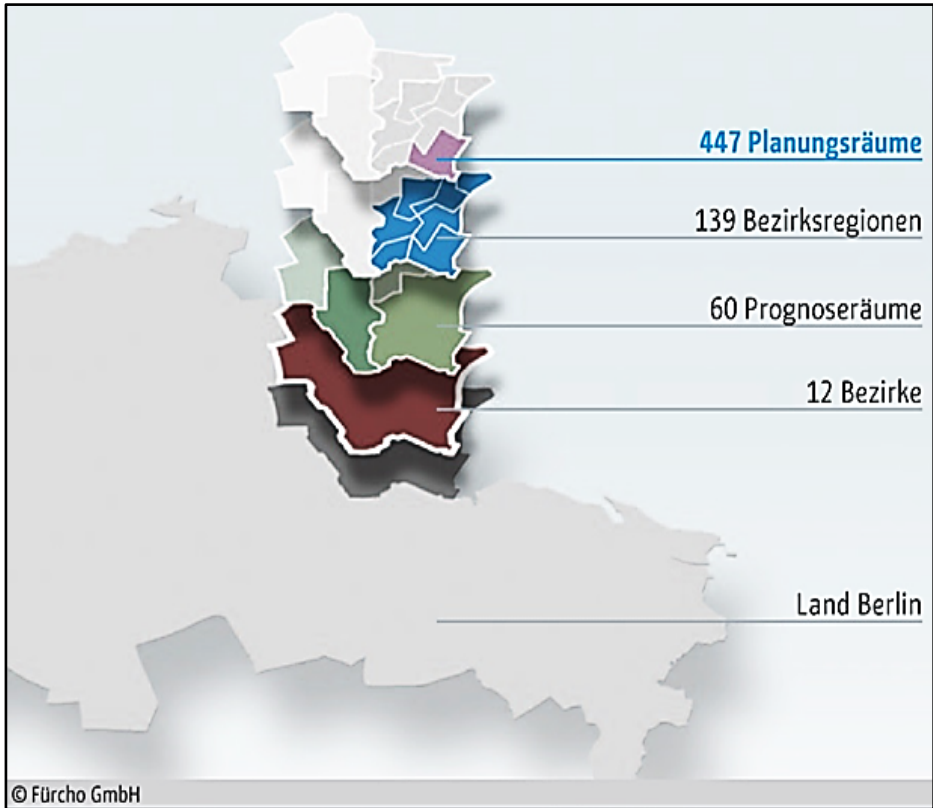


Abb. 1: Lebensweltlich orientierte Planungsräume – LOR

Quelle: Die umweltgerechte Stadt – Auf dem Weg zu einer sozialräumlichen Umweltpolitik, Senatsverwaltung für Umwelt, Verkehr und Klimaschutz 2019 (SenUVK 2019c)

Um die ungleiche Verteilung der kleinräumigen Umweltbelastungen für das Land Berlin messbar zu machen, wurden zunächst auf der Grundlage der Strategischen Umweltprüfung (SUP) vier wichtige umweltbezogene Themenfelder (Lärm, Luftbelastung, Grünflächenversorgung und bioklimatische Belastung) ausgewählt, deren Gesundheitsrelevanz wissenschaftlich belegt ist. Als weiterer gesundheitlich relevanter Bereich wurde die soziale Problemdichte (Statusindex aus dem Monitoring Soziale Stadtentwicklung) in das Grundindikatorenset der fünf Kernindikatoren integriert. Um die Aussagefähigkeit der Kernindikatoren zur Erfassung der Umweltqualität zu verbessern, wurden weitere Ergänzungsindikatoren aus den Bereichen Gesundheit, Städtebau, Stadtentwicklung oder

Risikomerkmale entwickelt. Diese ergänzenden Indikatoren untersetzen und präzisieren die Kernindikatoren. Aufgrund des gesundheitsorientierten Ansatzes der Berliner Umweltgerechtigkeitsanalysen haben vor allem Ergänzungsindikatoren aus dem Bereich Gesundheit Relevanz. Für Berlin liegen u. a. planungsraumbezogene Aussagen zur „vorzeitigen Sterblichkeit an Erkrankungen des Kreislauf- und Atmungssystems“, zu „durch Lärm verursachten Gesundheitskosten“, zu „Adipositas“ (Grundlage Einschulungsuntersuchungen - ESU), Daten zur Morbidität und zu den Krebs(neu)erkrankungen vor (Abb. 4 sowie Tabelle).

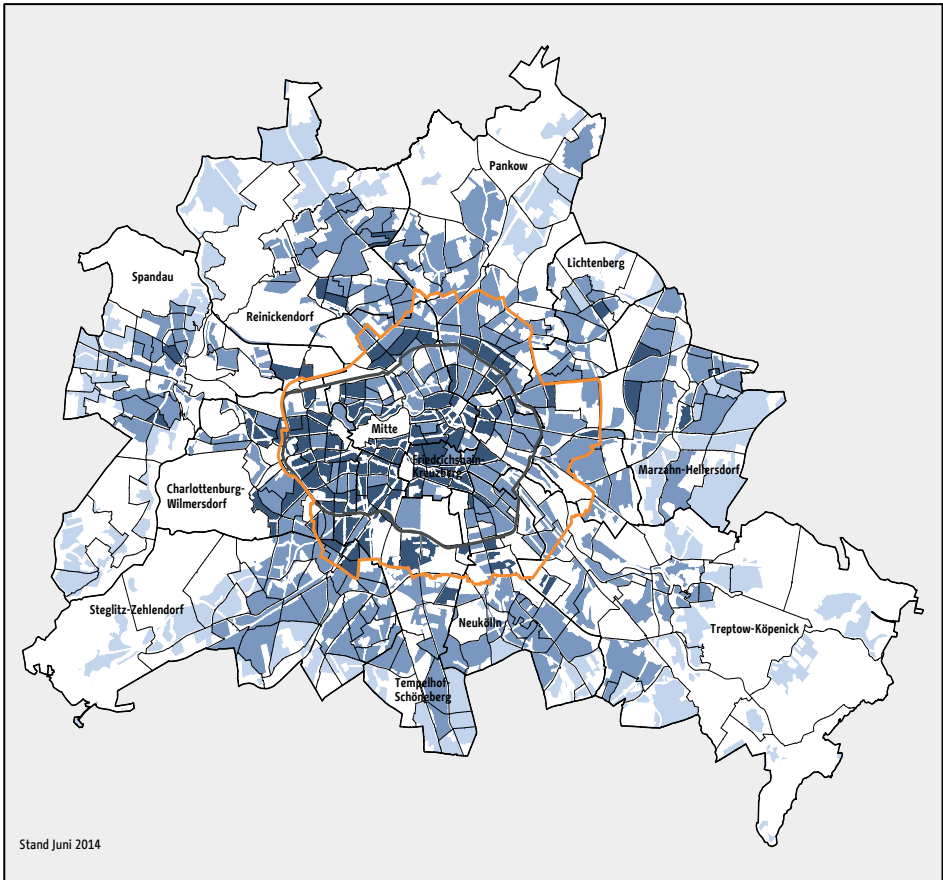


Abb. 2: Kernindikator „Luftbelastung“ ■ hoch ■ mittel ■ gering

Quelle: Basisbericht Umweltgerechtigkeit- Grundlagen für die sozialräumliche Umweltpolitik, Senatsverwaltung für Umwelt, Verkehr und Klimaschutz 2019 (SenUVK 2019b)

Im ersten Schritt des zweistufigen Analyseverfahrens wurden die Analysen zu den Kernindikatoren gesundheitlich gewichtet und entsprechend dem gesundheitlichen Risiko

(„gut“, „mittel“, „schlecht“) einer dieser drei Kategorien (Ampelsystem) zugeordnet. Danach wurde für jeden der 447 Planungsräume der Mehrfachbelastungsfaktor durch Summierung derjenigen Kernindikatoren ermittelt, die in dem v. g. „Ampelsystem“ der Kategorie 3 („schlecht“) zugeordnet wurden. Die entsprechende Kartierung bzw. Darstellung im Rahmen einer Mehrfachbelastungskarte identifiziert (nachvollziehbar und transparent) die mehrfach belasteten Räume auf gesamtstädtischer Ebene, wodurch gleichzeitig für das gesamte Stadtgebiet kleinräumige Umweltdaten auf der kleinräumigen Quartiersebene zur Verfügung gestellt werden (Abb. 2). Die Aussagen der Mehrfachbelastungskarte aufgrund der Kernindikatorenauswertung können in der zweiten Stufe bei Bedarf und mit Blick auf das Verwaltungshandeln durch Ergänzungsindikatoren aus den Bereichen Gesundheit, Soziales, Städtebau und Stadtentwicklung fachlich-inhaltlich weiter untersetzt werden (Abb. 3). Durch ein zusätzliches (gesundheitlich orientiertes) Risikomerkmäl können weitere Planungsräume gekennzeichnet werden, die aus umweltmedizinischer Sicht besonders stark belastet sind (Abb. 4).

Kernindikatoren (1. Stufe)	Ergänzungsindikatoren (2. Stufe)
Luftbelastung (PM10 / PM 2,5, NO2)	einfache Wohnlage (Mietspiegel)
Lärmbelastung (Betroffene/Kosten)	Flächennutzung (u. a. Gewerbe, Wohnen)
Grünflächenversorgung (EW/ha)	Stadtstruktur
Bioklimatische Belastung	Klimawandel (Hitzeinseln)/Betroffene
Soziale Problematik/Sozialstruktur (Monitoring Soziale Stadtentwicklung)	Einwohnerdichte/Betroffenheit in EW/ha
	Lärmkosten (Gesundheit) in Euro/EW
	vorzeitige Sterblichkeit (Erkrankungen des Kreislauf- und Atmungssystems)
	Adipositas (Einschulungsuntersuchungen - ESU)
	besonders hohes Gesundheitsrisiko
	Krebs(neu)erkrankungen / Morbidität
	Lichtverschmutzung

(Tabelle: Kern- und Ergänzungsindikatoren des zweistufigen Berliner Umweltgerechtigkeitsmonitorings auf der Ebene der 447 Planungsräume - LOR)

3 Ergebnisse

Von insgesamt 447 Planungsräumen – LOR (PLR) auf der gesamtstädtischen Ebene sind 3 PLR fünffach, 17 PLR vierfach, 71 PLR dreifach, 106 PLR zweifach und 105 PLR einfach belastet. 145 PLR sind unbelastet. Die Untersuchungen belegen, dass sich im Bereich des hochverdichteten Innenstadtbereichs der größte Teil der mehrfach belasteten Gebiete befindet. Im Bereich des „Erweiterten Innenstadtbereiches“ sind von 173 Planungsräumen 91 fünf-, vier- oder dreifach belastet. Hiervon sind 663.860 Einwohner betroffen. Die 145 Planungsräume ohne Mehrfachbelastung liegen vorwiegend in den Außenbezirken. Die Analysen zeigen weiter, dass ein Großteil der Gebiete mit einer hohen sozialen Problemdichte gleichzeitig auch von hohen gesundheitsrelevanten Umweltbelastungen betroffen ist. Dies betrifft insbesondere die Planungsräume im nördlichen Bereich des Bezirks Neukölln, im Bereich Mitte, im südlichen Bereich des Bezirks Reinickendorf sowie in Tempelhof-Schöneberg. Eine sehr geringe Umweltbelastung und ein sehr hoher bzw. hoher sozialer Status finden sich vor allem in den Außenbezirken.

4 Fazit:

Die kleinräumigen kartografischen Darstellungen bzw. die Kartierungen und der raumbezogenen Daten machen die quartiersbezogene Betroffenheit besonders deutlich. Sie zeigen den politischen Entscheidungsträgern, wo der Handlungsbedarf besonders groß ist. Die Form der Darstellung (Abb. 3 „Berliner Umweltgerechtigkeitskarte 2016/17“) kann als Frühwarnung für die Politik dienen, wenn sich umweltbezogene Risiken oder ungünstige Entwicklungen räumlich zu verfestigen scheinen. Durch die Darstellung werden direkte oder indirekte Zusammenhänge der dargestellten Sachverhalte, vor allem an der Schnittstelle der Ressorts Gesundheit/Soziales, Städtebau/Stadtentwicklung und Umwelt deutlich, da komplexe Sachverhalte auch für die Betroffenen in den belasteten Quartieren einfach und nachvollziehbar vermittelt werden.

Die Ergebnisse haben fach- und behördenübergreifende Bedeutung. Zur Verbesserung der Lebensqualität sind vor allem in den Gebieten mit Mehrfachbelastungen integrierte Strategien und Konzepte stärker als bisher ressortübergreifend auszurichten. Hierfür ist es notwendig, dass die Umweltbelastungsanalysen in die bereits etablierten kleinräumigen Stadtbeobachtungssysteme Berlins, wie beispielsweise die Gesundheits- oder Sozialberichterstattung des Landes, oder in die entsprechenden bezirklichen Berichterstattungen integriert werden. Wichtig ist, dass die Ergebnisse in die strategischen Konzepte und informellen bzw. formellen Instrumente der Stadtentwicklung und der Umweltplanung eingehen (Berliner Planungssystem). Dies gilt insbesondere im Hinblick auf die Bauleitplanung, die Bereiche Stadtentwicklung, Stadterneuerung, Stadumbau, das Bund-Länder-Programm Soziale Stadt, die Weiterentwicklung der Umweltprüfungen, Vorgaben für Planungswettbewerbe, Berücksichtigung im Rahmen Städtebaulicher Verträge und bei Aktionsprogrammen. Um die sozialräumlichen Veränderungen im Hinblick auf die Verteilung der Umweltbelastungen im Zeitablauf verfolgen zu können, ist es notwendig, das Umweltgerechtigkeitsmonitoring zu verstetigen. Hierbei ist zu beachten, dass die einzelnen

Bereiche bzw. Ressorts kleinräumige Daten in sehr unterschiedlichen Abständen erheben bzw. fachlich nutzbar aufbereiten und dass sich das Monitoring künftig den sich ständig verändernden Rahmenbedingungen anpassen muss. Dies gilt für alle Kern- und Ergänzungsindikatoren des zweistufigen Berliner Monitoringansatzes. Für eine erfolgreiche Strategie zu mehr Umweltgerechtigkeit ist die Entwicklung integrativer Planungsinstrumente und geeigneter Rahmenvorgaben und -bedingungen erforderlich. Mit Blick auf die Umsetzung nimmt das Land Berlin bundesweit weiterhin eine Vorreiterrolle ein. Die aktuellen Ergebnisse der Umweltgerechtigkeitsanalysen wurden bei der Erarbeitung des Stadtentwicklungskonzeptes 2030 (StEK 2030) berücksichtigt und im Bericht

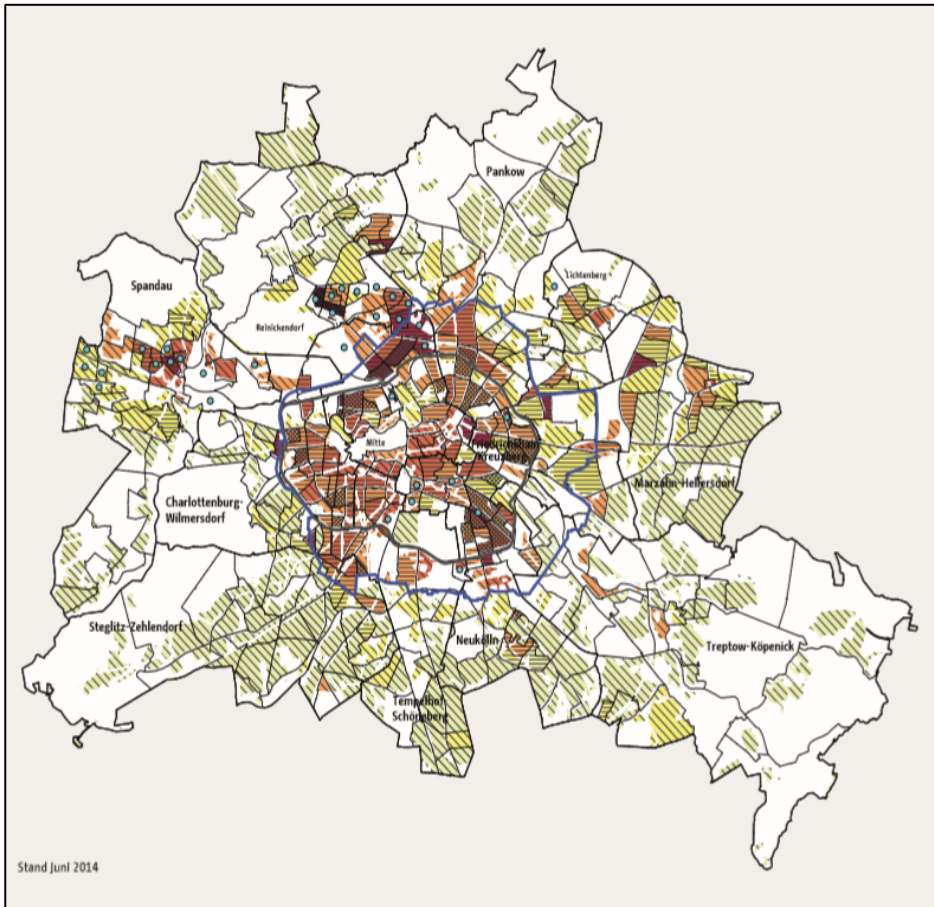


Abb. 3: Mehrfachbelastungskarte - „Berliner Umweltgerechtigkeitskarte“ („Berlin heute“)

Belastungen: 5fach 4fach 3fach 2fach 1fach nicht belastet

Quelle: Basisbericht Umweltgerechtigkeit- Grundlagen für die sozialräumliche Umweltpolitik, Senatsverwaltung für Umwelt, Verkehr und Klimaschutz 2019 (SenUVK 2019b)

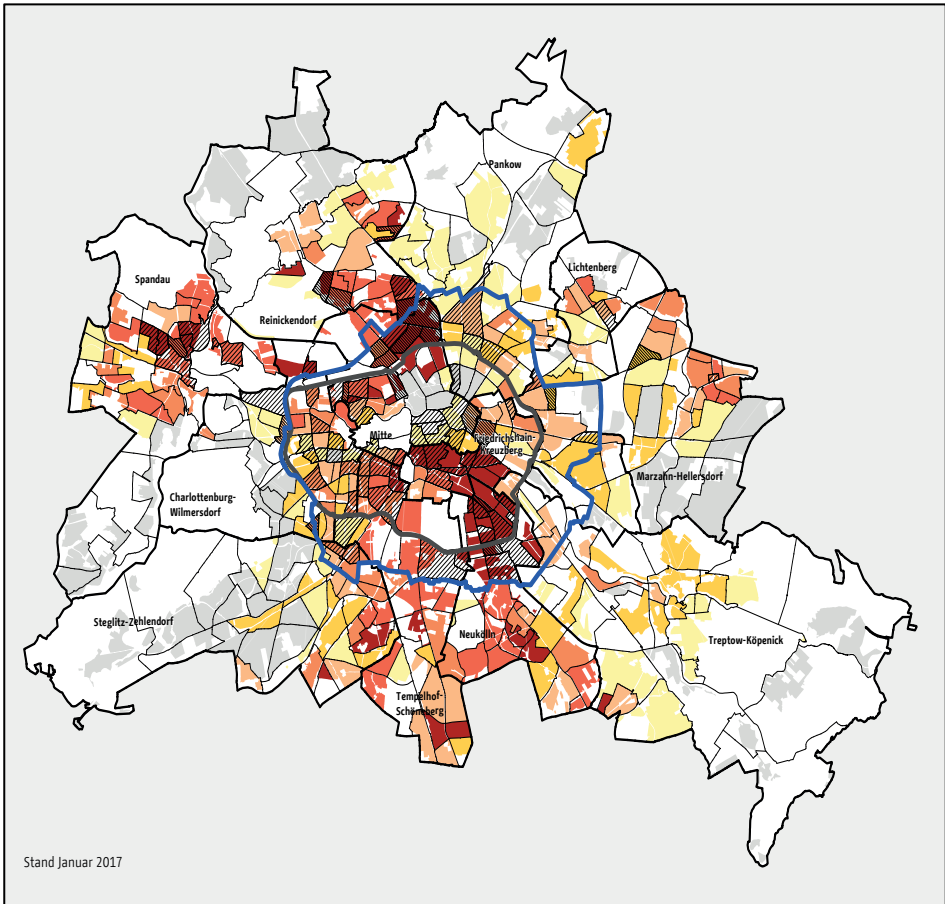


Abb. 4: Morbiditätsindex I (Gesundheitsberichterstattung 2010 / 2011) / Mehrfachbelastungen

Quelle: Basisbericht Umweltgerechtigkeit- Grundlagen für die sozialräumliche Umweltpolitik, Senatsverwaltung für Umwelt, Verkehr und Klimaschutz 2019 (SenUVK 2019b)

zum Flächennutzungsplan des Landes Berlin integriert. Die kleinräumige Umweltbelastungsanalyse könnte im Sinne eines Werteausgleichs in die Berechnungen der Finanzmittelzuweisungen des Senats an die Bezirke (vertikaler und horizontaler Finanzausgleich) mit herangezogen werden. Darüber hinaus haben die Umweltgerechtigkeitsergebnisse als zusätzliche Entscheidungshilfe bei der Umsetzung der Umweltförderprogramme Bedeutung. Die derzeitige Berliner Regierungskoalition hat 2020 einen Abgeordnetenhausantrag zur Entwicklung eines „Stadtentwicklungsplans Umweltgerechtigkeit 2030“ (StEP-Umweltgerechtigkeit) auf den Weg gebracht. Mit Blick auf das Ziel „Klimaneutralität Berlin 2050“ werden die Ergebnisse künftig Bedeutung haben, denn die Gebiete mit

„Mehrfachbelastungen“ sind gleichzeitig die vulnerablen Gebiete im Hinblick auf den Klimawandel. Aufgrund der interdisziplinären Vorgehensweise ist davon auszugehen, dass bei Umsetzung gesundheits- und klimapolitischer Strategien und Maßnahmen oder im Hinblick auf die Entwicklung smarter Infrastrukturen zur Verbesserung der Lebensqualität in den Quartieren die Umweltgerechtigkeitsanalysen eine wichtige Entscheidungsgrundlage werden.

5 Ausblick: Umweltgerechtigkeit in der Post-Corona-Stadt

Dies gilt auch im Hinblick auf die Corona-Pandemie. Die tagesaktuellen COVID-19 Fallzahlen (12.05.2021) sowie die weiterführenden Auswertungen der Berliner Gesundheitsverwaltung belegen, dass in den hochverdichteten Innenstadtlagen der Bezirke Mitte, Neukölln und Tempelhof-Schöneberg die Fallzahlen und Inzidenzen am höchsten sind (Landesamt für Gesundheit und Soziales Berlin). Erste Untersuchungsergebnisse legen nahe, dass neben der sozialen Lage, Siedlungsstrukturen und Mobilitätsfaktoren auch Alters- und Einkommensstrukturen zu berücksichtigen sind. Mit Blick auf die kleinräumigen Analysen und Ergebnisse der Berliner Umweltgerechtigkeitsanalysen kann vermutet werden, dass vor allem vier Kernindikatoren COVID-19-Relevanz: Soziale Problematik, Luftbelastung, Grün- und Freiflächenversorgung sowie

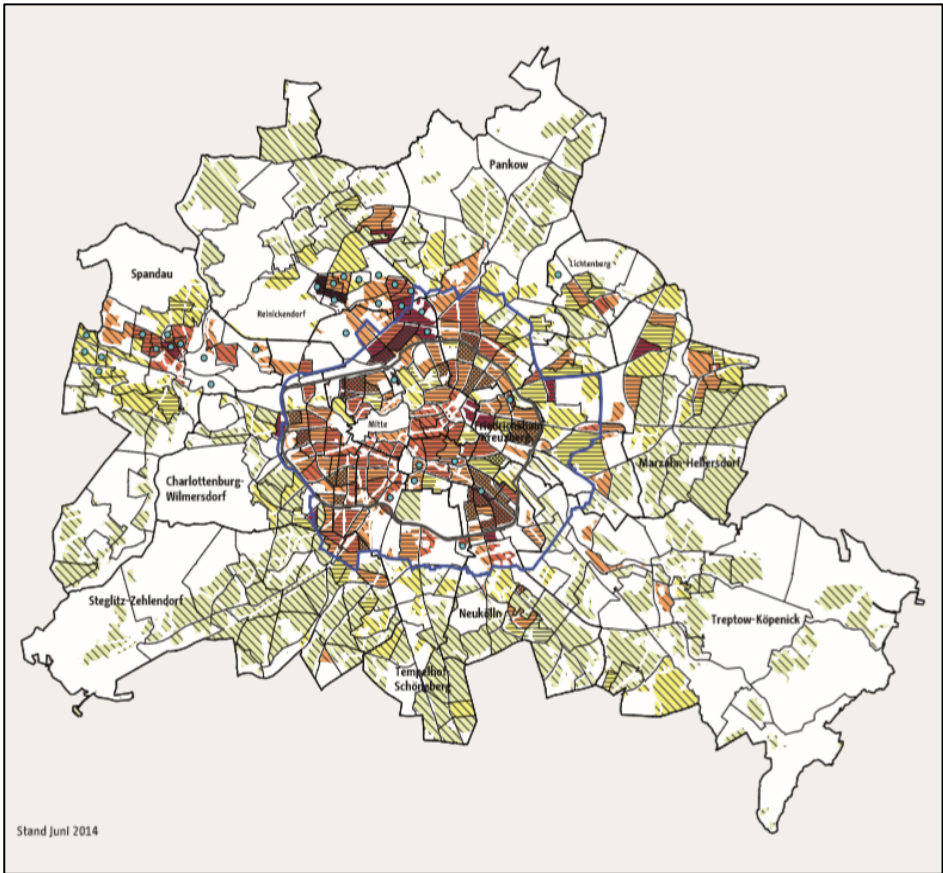


Abb. 5: Mehrfachbelastungskarte - „Berliner Umweltgerechtigkeitskarte“ („Berlin heute“)

Quelle: Basisbericht Umweltgerechtigkeit- Grundlagen für die sozialräumliche Umweltpolitik, Senatsverwaltung für Umwelt, Verkehr und Klimaschutz 2019 (SenUVK 2019b)

Wohnlage. Dies belegen die ersten wissenschaftlichen Analysen in den sozial benachteiligten Quartieren, insbesondere in Köln und Berlin, die den Zusammenhang von Wohn-dichte, Armut sowie individueller Vulnerabilität gegenüber gesundheitsrelevanten Um-weltfaktoren nahelegen. Umweltgerechtigkeit kann dazu beitragen, Aussage zu treffen, inwieweit es (sozial-)räumliche Erklärungsmuster für das Infektionsgeschehen in den hochverdichteten Metropolenregionen gibt und ob belastbare räumliche Zusammenhänge erkennbar sind. In der Post-Corona-Stadt wird Umweltgerechtigkeit für die Transfor-mation der Stadtquartiere, vor allem vor dem Hintergrund der Verteilungsgerechtigkeit und der bevölkerungsbezogenen Vulnerabilität Bedeutung gewinnen. Dies gilt insbesondere im Hinblick auf die Entwicklung von neuen Instrumenten des planerischen

Umweltschutzes, die – u. a. auch mit Blick auf die Pandemie - neue Sichtweisen und Prioritäten in den Vordergrund stellen.

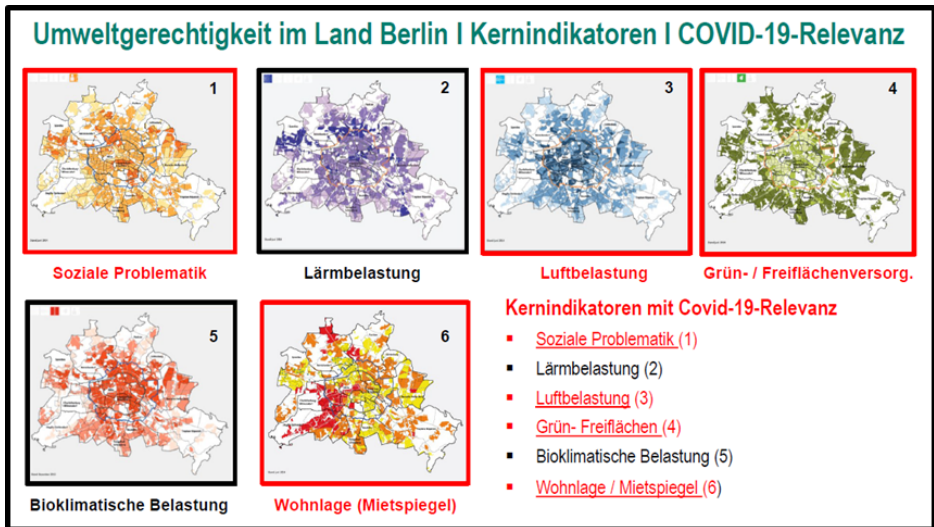


Abb. 6 Kernindikatoren mit COVID-19-Relevanz

Quelle: eigene Darstellung (2021) in Anlehnung an: Die umweltgerechte Stadt – Auf dem Weg zu einer sozialräumlichen Umweltpolitik, Senatsverwaltung für Umwelt, Verkehr und Klimaschutz, 2019 (SenUVK 2019c)

6 Literatur

Böhme, Christa; Franke, Thomas; Preuß, Thomas; Schwarze, Katrin; Winkler-Kühlken, Bärbel; Schipperges, Michael (2018): Möglichkeiten der verstärkten Nutzung von Synergien zwischen Umweltschutz und sozialer Gerechtigkeit in Programmen wie der „Sozialen Stadt“ – Abschlussbericht. Umweltbundesamt, Texte 74/2018, Dessau-Roßlau.

Böhme, Christa; Franke, Thomas; Preuß, Thomas (2019): Umsetzung einer integrierten Strategie zu Umweltgerechtigkeit – Pilotprojekt in deutschen Kommunen (Umwelt u. Gesundheit, 02/2019, Hrsg. Umweltbundesamt, Dessau-Roßlau.

Bolte, Gabriele; Bunge, Christiane; Hornberg, Claudia; Köckler, Heike; Mielck, Andreas (Hrsg.) (2012): Umweltgerechtigkeit. Chancengleichheit bei Umwelt und Gesundheit: Konzepte, Datenlage und Handlungsperspektiven, Bern.

Bolte, Gabriele; Bunge, Christiane; Hornberg, Claudia; Köckler, Heike (2018): Umweltgerechtigkeit als Ansatz zur Verringerung sozialer Ungleichheiten bei Umwelt und Gesundheit. Bundesgesetzblatt, 61. Jg (6): 674-683

Klimeczek, H.-Josef (2018): Umweltgerechtigkeit im Land Berlin – Der integrierte Umweltgerechtigkeitsansatz. In: Basisbericht Umweltgerechtigkeit. Grundlagen für die sozialräumliche Umweltpolitik. Hrsg. Senatsverwaltung für Umwelt, Verkehr und Klimaschutz, Berlin, S. 131 ff.

Klimeczek, H.-Josef (2018): Umweltgerechtigkeit im Land Berlin – Strategien, Maßnahmen und Projekte. In: Planung für gesundheitsfördernde Städte, Forschungsberichte der ARL 8, S. 350 – 352.

Klimeczek, H.-Josef (2019b): Die umweltgerechte Stadt – Auf dem Weg zu einer sozialräumlichen Umweltpolitik. Hrsg. Senatsverwaltung für Umwelt, Verkehr und Klimaschutz, Berlin. (S. 21 – 38)

Klimeczek, H.-Josef; Heiss, Christiane; Lohner, Herbert; Zipf, Markus (2020): Schöneberger Erklärung zur Umweltgerechtigkeit 2020, Bezirksamt Tempelhof-Schöneberg (Hrsg.) Deutsche Umwelthilfe (DUH)

Internetquellen:

DÄ–Deutsches Ärzteblatt 13/2021: Zugriff: <https://www.aerzteblatt.de/nachrichten/123448/Forderungen-nach-verstaerktem-impfen-in-sozialen-Brennpunkten> (abgerufen am 10.05.2021)

DÄ – Deutsches Ärzteblatt 15/2021: Zugriff: <https://www.aerzteblatt.de/archiv/218459> (abgerufen am 10.05.2021)

DIfU – Deutsches Institut für Urbanistik: Zugriff: <https://toolbox-umweltgerechtigkeit.de/> (abgerufen am 10.05.2021)

DUH – Deutsche Umwelthilfe: Zugriff: <https://www.duh.de/projekte/umweltgerechtigkeit/partizipation/partizipation/> (abgerufen am 10.05.2021)

LaGeSo: Zugriff: <https://www.berlin.de/corona/lagebericht/> (abgerufen am 10.05.2021)

SenUVK 2019b – Senatsverwaltung für Umwelt, Verkehr und Klimaschutz: Zugriff: file:///C:/Users/KLIMEC~1/AppData/Local/Temp/umweltgerechtigkeit_broschuere.pdf (abgerufen am 10.05.2021)

SenUVK 2019c – Senatsverwaltung für Umwelt, Verkehr und Klimaschutz: Zugriff: file:///C:/Users/KLIMEC~1/AppData/Local/Temp/umweltgerechtigkeit_broschuere_kurzfassung.pdf

Workshop: Workshop on Systems to
Support Renaturation Projects
(SyRePro 2021)

Report on the Correction of Erroneous Geometry Data in Land Reuse Projects

Yves Annanias,¹ Marc Wahsner,² Geric Scheuermann,³ Daniel Wiegrefe⁴

Abstract: Land Reuse processes are large planning and decision-making processes based on a large amount of geographic data. Therefore, it is essential that this data is as accurate as possible. However, errors can occur during the creation of the data and not all of them are directly noticeable. We report here what errors we have encountered while working with this geographic data, what problems they can cause, and how we have fixed them. Since the correction can be very time-consuming with the enormous amount of data, we have focused on an automatic correction. Not all of this data can be corrected this way, for the rest, we briefly indicate a procedure to support and simplify the manual correction.

Keywords: GIS; Data Preparation; Automatic Error Correction

1 Introduction

Opencast lignite mines are temporary, massive interventions in nature since after the coal has been mined, crater landscapes remain. To make these regions usable again and to pursue sustainable concepts, entire stretches of land must be renaturalized. Such concepts are, for example, the agricultural use, the use of these areas for recreation for people as well as nature reserves, and building areas for whole city districts. Whether an area is suitable for recultivation into a certain type of after-use depends on many factors, such as soil characteristics or the location itself. So it may also be that for some areas the subsequent use is not possible due to contaminated sites, while others can be used immediately.

Therefore, a large amount of information has to be considered in detail during planning to decide where and how this land reuse can take place so that it can be done safely and as quickly as possible. The data involved in these land reuse processes consists essentially of two types: geographical data and expert knowledge. The expert knowledge includes information about soil conditions, groundwater levels, forecasts of water level changes of lakes or restrictions on the use of areas (e.g., entering certain areas is prohibited because contaminated soils are present). Geographic data links this knowledge to specific locations on a map. Furthermore, this data also relates individual areas to each other, which allows

¹ Leipzig University, Image and Signal Processing Group annanias@informatik.uni-leipzig.de

² Leipzig University, Image and Signal Processing Group mw83waji@studserv.uni-leipzig.de

³ Leipzig University, Image and Signal Processing Group scheuermann@informatik.uni-leipzig.de

⁴ Leipzig University, Image and Signal Processing Group daniel@informatik.uni-leipzig.de

examining the distance between areas, gives information about how far they are from other important points on the map, and sometimes it shows that areas even overlap.

Tools that link these two types of data are called geographic information systems (GIS). A GIS can provide many ways to measure, interpret, visualize, generalize, and interpolate collected samples [GYC07]. Specifically, spatial interpolation techniques are used to complete those spatial locations where measurements are not available or missing. GIS has penetrated a variety of fields supporting experts in decision making and are successfully used to create continuous surfaces in, e.g., meteorology, climatology, and water management (see Kienberger and Steinbruch [KS05] with focus on cyclones, floods and droughts, Abbas et al. [Ab09] for the combination of disaster management with a GIS for a district disaster management system for floods, or Khan [Kh13] for flood analysis and prediction using GIS). In addition, Chen et al. [Ch18] use machine learning to create landslide susceptibility maps to reduce this hazard if possible or to integrate this information into land-use planning. In general, the aforementioned works show that especially in disaster risk management GIS can be useful, this concerns, e.g., the different phases of risk management activities, such as planning, mitigation, preparedness and response (see Bala et al. [BT17]).

Since these planning and decision-making processes depend heavily on geographic data, these data must be as accurate as possible. Otherwise, erroneous data would lead to incorrect decisions that unnecessarily delay the land reuse process, which can be costly and result in legal consequences or, especially in the case of disaster management, endanger human lives. Unfortunately, errors in the geographic data are not always directly detectable. For example, a GIS may visually represent an area supposedly correctly, but underlying algorithms generate incorrect values when calculating the area's content and other parameters. Furthermore, calculations of intersections may be wrong or are not possible at all, because the algorithms already abort with error messages before due to various reasons. As a result, important information would be missing from the decision-making process. Finding and correcting these errors is very time-consuming.

To facilitate error detection and correction, we provide a brief overview of the specific errors that can occur, what causes them, the problems they cause, and how to fix them. A large part of the errors are due to syntactic inaccuracies and can therefore be corrected automatically, which saves a lot of time. Some GIS tools offer support to detect syntactical errors [ESRa], some are able to fix them directly, like incorrect ring orientations. Even self-intersections are partially adjusted automatically by them [ESRb]. However, changes in the shape of a geographic area might occur here depending on the algorithm used. Cui et al. [Cu20], instead, specify an algorithm that can be used to split self-intersecting polygons into partial polygons without self-intersection. Due to insufficient modeling possibilities at the time of creation of the data, it is possible that semantic errors were explicitly used to create a certain visual result. An automatic correction can then lead to different results. Therefore, it is reasonable to check and correct this geographic data manually by an expert. For this, we provide a structured procedure, with the help of which an expert is guided during the manual correction.

2 Data

The geographical data follows the Simple Feature Access specification [Ope], and is called a feature, here. In general, a feature could be a point (e.g., a well), a line (e.g., a road), a polygon (e.g., an area or an area with a hole), or a multi-polygon (e.g., two areas without a connecting line). To store and process these geometry types the specification includes a definition for representation of the geometry as a well-known-text (WKT) format. The different types of geometry can thus be represented as follows:

Simple Point

Point(0,0)

Line

LineString(0 0, 100 0, 100 100)

Simple Polygon

Polygon((0 0, 0 100, 100 100, 100 0, 0 0))

Polygon with a Hole

Polygon((0 0, 0 100, 100 100, 100 0, 0 0), (20 20, 20 80, 80 80, 80 20, 20 20))

Two Polygons

MultiPolygon((((0 0, 0 100, 100 100, 0 0)), ((20 0, 120 100, 120 0, 20 0)))

The data set provided to us includes 36,623 features and was obtained from the Lausitzer und Mitteldeutsche Bergbauverwaltungsgesellschaft mbH (LMBV)⁵. The LMBV is a company which is responsible for the management and recultivation of abandoned opencast lignite mines, and one of its tasks is to provide the public with reliable information about the areas it manages. The provided data set was created continuously over several decades using various input methods and software systems, and large portions of the data were created manually. The large majority of features are areas described by polygons and multi-polygons (33,955 features, 92.7%). The remaining features are wells and other features described by point data (2,668 features, 7.3%). There are a few polygons with an area smaller than 10 m². However, there are also a few particularly large polygons with areas up to 5,000 km².

The polygons and multi-polygons describing areas were generated in different ways. One way was to use GIS stations that displayed a map and offered the possibility to create points on it in order to connect them afterwards via lines. The problem with this is that sometimes points were created by mistake, resulting in additional unnecessary lines. Another problem lies in the fact that especially older stations were only able to create simple polygons. For more complex polygons, workarounds had to be used, such as unnecessary connecting lines or self-intersecting shapes, since these visually showed the desired result but did not adhere to the correct format. In a second way, a person walked along the border of the area. Using

⁵ <https://de.wikipedia.org/wiki/LMBV>

a GPS tracker, points on that border were stored at regular time intervals. This also can result in problems that are not always visually recognizable (e.g. redundant points caused by standing longer at the same point). Redundant points and lines, as well as self-intersecting polygons, cause problems, e.g., when intersecting polygons with each other. This error handling is discussed in the following section.

3 Error Detection and Handling

Defects within the geometry of features are usually difficult or sometimes impossible to perceive visually and can impede the processing of those features. With the aforementioned format, it is possible to solve these errors by simple string substitutions. An example that occurs in the data is related to the creation of the polygon on a map. Here, two points were created on a map and connected to another one. By mistake, another point was then created on this line and connected to the others, creating unnecessary and redundant lines. The result of this is shown in Fig. 3 (top), visually there is only one line recognizable and the error is therefore difficult to detect later. The problem lies in the so created self-intersection, whereupon some algorithms cause problems because they deliver incorrect values when the size of an area is calculated or even stop working when the overlap with another polygon should be calculated.

In this section, we indicate how these and other problems within the geometry of features could be solved. At first, a definition of validity is required. Here, the validity of the geometry is determined using the *is_valid* function of the *shapely python package* [Gi], which in turn uses an *OpenGIS* specification [Ope] to define validity. For example, a polygon is defined by one exterior border and zero or more interior boundaries (see *Polygon* and *MultiPolygon* definitions in section 2). A valid polygon has to fulfil certain conditions, the most important ones here in terms of errors appearing in the data set are the following (for more conditions, see the specification [Ope]):

- a “Polygons are topologically closed.”
- b “No two rings in the boundary cross and the rings in the boundary of a polygon may intersect at a Point but only as a tangent [...]”
- d “A polygon may not have cut lines, spikes or punctures [...]”

While some errors are only tagged and have to be dealt with manually, the majority allow for automated processing.

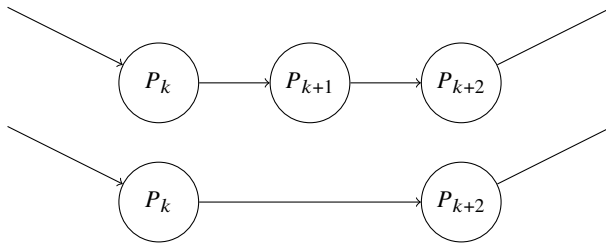


Fig. 1: Polygon with redundant point P_{k+1} (top) and fixed polygon (bottom)

3.1 Preprocessing

Prior to fixing any defects, the following redundancies are removed to allow for easier processing:

1. Duplicate Points: Whenever two adjacent points within the geometry of a feature are identical, one of those points is removed. Therefore, $Polygon(\dots, 0\ 0, 0\ 0, \dots)$ will be simplified to $Polygon(\dots, 0\ 0, \dots)$.
2. Redundant Points: Whenever a triplet of adjacent points (P_k, P_{k+1}, P_{k+2}) exists within the geometry of a feature, where P_{k+1} lies on the line segment $P_k P_{k+2}$, P_{k+1} is removed. This constellation can be seen in Fig. 1. Therefore, $Polygon(\dots, 0\ 0, 0\ 1, 0\ 2, \dots)$ will be simplified to $Polygon(\dots, 0\ 0, 0\ 2, \dots)$. However, it is assumed here that point P_{k+1} lies exactly on the line segment. Numerically, this may not always be exactly the case. For this an epsilon environment can be used as a threshold and it is assumed that the point belongs to the line segment if it lies within this threshold. However, such simplification can lead to problems, see section 4.

A redundancy that involves the first and last point of the polygon is currently not being fixed, as their can only exist one such defect within a polygon, which does neither affect the automated processing, nor does it prevent any other redundancies from being fixed. Hence, the $Polygon(0\ 2, 0\ 3, \dots, 0\ 1, 0\ 2)$ is not simplified to $Polygon(0\ 3, \dots, 0\ 1, 0\ 3)$. Two adjacent points are guaranteed to be non-identical because of the prior removal of duplicate points. Without this restriction, it would not be sufficient to merely check triplets of points to detect redundancies. In case of $P_k = P_{k+2}$, P_{k+1} can not be on the line segment $P_k P_{k+2}$. However, this case is considered to be a defect and a more general case will be automatically fixed in the next subsection.

3.2 Automated Processing

Although shapely is able to identify many invalidity reasons, the only invalidity that occurs and allows for automated processing is *self-intersection*. Here the following cases of *self-intersection* are automatically processed:

1. Within polygons and multi-polygons: Whenever a triplet of adjacent points (P_k, P_{k+1}, P_{k+2}) exists within a polygon, where the normalized vector $\hat{v}(P_k, P_{k+1}) = -\hat{v}(P_{k+1}, P_{k+2})$, P_{k+1} is removed. This operation can influence the shape of a feature, as can be seen in Fig. 2, but it will not change its area. This defect may also occur at the end of a polygon of size n , with $\hat{v}(P_n, P_1) = -\hat{v}(P_1, P_2)$, here P_1 has to be removed. It should be noted here, that such special lines can be meant as accesses to areas, e.g., paths and roads to the area. However, this should be created as extra line segments or as an extra polygon with a given width. Following our experts, however, this can actually be seen as an error in the present data set and corrected automatically. The removal of P_{k+1} may lead to P_{k+2} being made redundant, due to the following configuration: The points $P_k, P_{k+1}, P_{k+2}, P_{k+3}$ lie on a line in the order $P_k, P_{k+2}, P_{k+1}, P_{k+3}$, as is shown in Fig. 3. In this case P_{k+2} is also removed.
2. Within multi-polygons: If two polygons A, B of a multi-polygon have a polygon as an intersection, the first Polygon is reduced to $A = A - B$. This step is shown in Fig. 4.

3.3 Manual Processing

All invalid features that could not be processed automatically are tagged accordingly and can be fixed through user interaction. The hourglass serves as an example of such an invalid feature (see Fig. 5 left). In principle, it consists of two triangles that touch at one point. In the example shown, the direction in which the points are indicated also changes for each triangle. This has the consequence that the calculated area is incorrectly indicated as 0. In addition, the present self-intersection leads to further errors when intersecting the feature with other features.

An automatic correction would first correct the order of the points. Then it is possible to divide the intersection point into two points that are as close to each other as possible. The error handling described before is no problem here, even if a threshold is used when removing redundant points, since only points on a line segment are affected, which are thus directly connected by edges. The points I_1 and I_2 lie close together, but not on a line segment by connected points. This solution would solve the two problems mentioned above, which is shown in Fig. 5 (right). Nevertheless, this changes the shape of the polygon and an alternative is to create two single polygons from the two triangles, which then are combined into one multi-polygon.

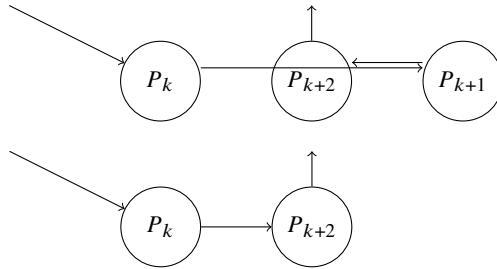


Fig. 2: Polygon with faulty point P_{k+1} (top) and fixed polygon (bottom)

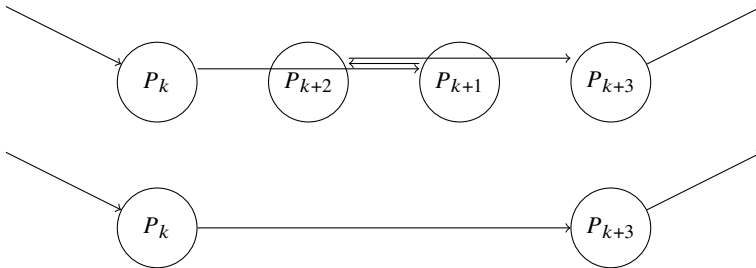


Fig. 3: Polygon with faulty point P_{k+1} , where the removal of P_{k+1} leads to P_{k+2} being made redundant (top) and fixed polygon (bottom)

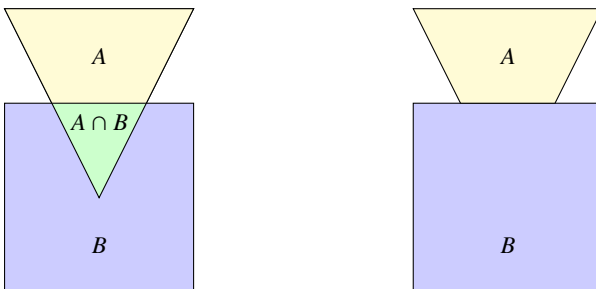


Fig. 4: Faulty multi-polygon (left) and fixed multi-polygon (right)

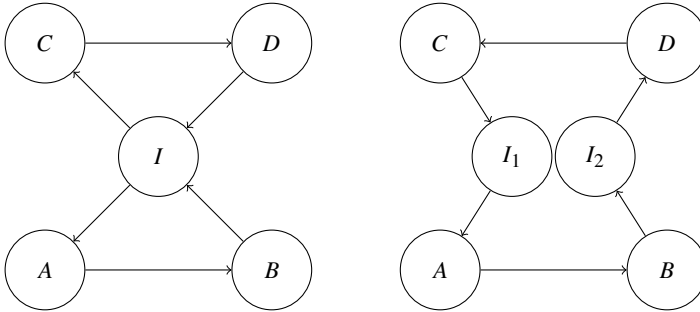


Fig. 5: **Left:** Hourglass polygon with self-intersection in point I and opposite point ordering in both triangles. The area of this polygon is calculated as 0. **Right:** The polygon with corrected point ordering and resolved self-intersection, where point I has been divided into the two points I_1 and I_2 , which are very close to each other. Therefore, the area is correctly calculated as ≈ 2 (if the area of each triangle is 1).

Not all of these problems are as simple as this one. Especially with polygons that have holes, changes could quickly lead to an altered representation and subsequently, the question arises as to what the polygon was originally intended to represent. In consultation with our experts, it is better to handle these cases manually. To simplify the process of manual correction, we use the following procedure:

1. Check feature for defects.
 - a) If there are no defects or they can be corrected automatically, correct them and calculate the required values.
 - b) If there are defects that need manual correction, tag the feature and go to the next step.
2. Mark the invalidity reason (a point which marks the start of the defect as determined by the *shapely python-package*).
3. Display invalid feature on a map and highlight marked position.
4. Allow the feature to be modified by an expert by repositioning, inserting, and deleting points.
5. After the feature was modified, go to 1. and check for other defects.

4 Large Polygons

Another problem we noticed, concerns polygons with a lot of points. As an example, there was a railroad track in the data set that was recorded using a GPS tracker. The tracker drove along the railroad track from start to end and back, recording data after each specific time interval. The result is a polygon that extends over a distance of approximately 400 *km* and contains over 66,000 points. Although this polygon is syntactically correct after error correction, it can lead to high memory and time consumption in other operations. In addition, the high number of points may degrade the response time of the GIS tool.

A possible solution would be, on the one hand, the simplification of the geometry. In this way, points can be removed that do not lie on a straight line, but where the angle between the lines and the neighbouring points is not too large. However, this changes the surface and is therefore not always an acceptable solution. In addition, such simplification can again lead to the previously mentioned problems, such as self-intersection. Therefore, rechecking for these defects after the adjustment is necessary. On the other hand, the geometry can be divided into several smaller parts. In this way, the entire polygon does not have to be considered when displaying it, but all parts must still be taken into account in the area calculation.

5 Conclusion

As a result, we were able to correct a significant amount of data from our data set, which consists of 36,623 features. Of these features, 30,278 (82.7%) were completely error-free. 5,802 features (15.8%) had errors that could be corrected automatically, although not all of these errors were problematic. For example, simple redundancies of points on a straight line do not cause problems. Only 543 features (1.5%) had to be adjusted manually, and the process of correction was supported by tagging the defects. The error handling ensured that all the data could be used and lead to correct results. Thus, the intersection of areas was also possible without any problems. This leads to further and reliable results of the data when used in land reuse processes. As a result, data created several decades ago can still be used.


ACKNOWLEDGMENT

This research was partially funded by the Development Bank of Saxony (SAB) under project number 100335729.

Bibliography

- [Ab09] Abbas, S.H.; Srivastava, R.K.; Tiwari, R.; Ramudu, P.: GIS-based disaster management: A case study for Allahabad Sadar sub-district (India). *Management of Environmental Quality: An International Journal*, 20:33–51, 01 2009.
- [BT17] Bala, Papiya; Tom, Santhi: GIS and Remote Sensing In Disaster Management. *Imperial journal of interdisciplinary research*, 3, 2017.
- [Ch18] Chen, Wei; Peng, Jianbing; Hong, Haoyuan; Shahabi, Himan; Pradhan, Biswajeet; Liu, Junzhi; Zhu, A-Xing; Pei, Xiangjun; Duan, Zhao: Landslide susceptibility modelling using GIS-based machine learning techniques for Chongren County, Jiangxi Province, China. *Science of The Total Environment*, 626, 06 2018.
- [Cu20] Cui, Yong; Liu, Qian; Chen, Guo; Zhang, Hujun: A general method for decomposing self-intersecting polygon to normal based on self-intersection points. *Theoretical Computer Science*, 842:118–129, 2020.
- [ESRa] ESRI Inc. ArcGIS Pro: Invalid Geometry. URL: <https://pro.arcgis.com/en/pro-app/latest/help/data/validating-data/invalid-geometry.htm> [accessed 2021-07-08].
- [ESRb] ESRI Inc. ArcGIS Pro: Repair Geometry (Data Management). URL: <https://pro.arcgis.com/en/pro-app/latest/tool-reference/data-management/repair-geometry.htm> [accessed 2021-07-08].
- [Gi] Gillies, Sean; Shapely: Manipulation and analysis of geometric objects in the Cartesian plane. URL: <https://pypi.org/project/Shapefile/> [accessed 2021-04-08].
- [GYC07] Goodchild, Michael F.; Yuan, May; Cova, Thomas J.: Towards a general theory of geographic representation in GIS. *International Journal of Geographical Information Science*, 21(3):239–260, 2007.
- [Kh13] Khan, Naveed: Flood Prediction and Disaster Risk Analysis using GIS based Wireless Sensor Networks, A Review. *Journal of Basic and Applied Scientific Research*, 09 2013.
- [KS05] Kienberger, Stefan; Steinbruch, Franziska: P-GIS and disaster risk management: Assessing vulnerability with P-GIS methods—Experiences from Búzi, Mozambique. 01 2005.
- [Ope] Open Geospatial Consortium. OpenGIS Implementation Standard for Geographic information - Simple feature access - Part 1: Common architecture. URL: <https://www.opengeospatial.org/standards/sfa> [accessed 2021-04-08].

TRIM4Post-Mining: Transition Information Modelling for attractive Post-Mining Landscapes – A Conceptual Framework


Jörg Benndorf ¹, Diego Alejandro Restrepo¹, Natalie Merkel¹, Andreas Knobloch², Martin Kressner³, Adriana Guatame-Garcia⁴, Harm Nolte⁵, Stefan Möllerherm⁶, Marinus Dalm⁷

Abstract: TRIM4Post-Mining brings together a consortium of European experts from industry and academia to develop an integrated information modelling system. This is designed to support decision-making and planning during the transition from coal exploitation to a revitalized post-mining landscape enabling infrastructure development for agricultural and industrial utilization and contributing to recover energy and materials from coal mining dumps. The smart system will be founded up on a high-resolution spatiotemporal database utilizing state-of-the-art multi-scale and multi-sensor monitoring technologies that characterize dynamical processes in coal waste dumps related to timely dependent deformation and geochemical processes. It will develop efficient methods for comprehensive spatiotemporal data analytics, feature extraction, and predictive modelling that allow for the identification of potential contamination areas and forecasting the waste dump dynamics. For the interactive exploration of alternative land-use planning scenarios in terms of residual risks, technical feasibility, environmental and social impact, and also affordability, up-to-date data and models will be embedded in an interactive planning system based on Virtual Reality and Augmented Reality technology forming a TRIM – Transition Information Modelling System.

Keywords: Geomonitoring, Data Analytics, VR/AR, Decision Support for Revitalization Planning

1 Introduction

With the currently implemented energy transition, most coal mines in Europe will transfer to the closing phase during the next two decades. The structural transformation of a region from an industrial mining area to an attractive landscape requires a good understanding of risks and opportunities associated with the post-mining landscape, in particular related to

¹ TU Bergakademie Freiberg, Germany, Joerg.Benndorf@mabb.tu-freiberg.de,  <https://orcid.org/0000-0002-2002-8639>

² Beak Consultants GmbH, Freiberg, Germany

³ Mitteldeutsche Braunkohlegesellschaft mbH, Zeitz, Germany

⁴ TU Delft, Delft, the Netherlands

⁵ Eijkkelkamp SonicSampDrill B.V., the Netherlands

⁶ Technische Hochschule Georgius Agricola Bochum, Germany

⁷ Spectral Industries B.V., Delft, the Netherlands

waste dumps, and ensuring open communication and engaging all stakeholders to explore fruitful options for future land use.

In the light of newly available monitoring-, data management-, modelling-, visualization- (VR/AR) and decision supporting technology, spatially and timely very dense data and information for both, geomechanical and geochemical attributes of coal waste dumps are available and can be effectively used. Examples include active and passive satellite sensors [Jo21], [Ha19], airborne and terrestrial LIDAR – data for areal information [Sc21], [Gun18], geochemical point sensor data including TIR and SWIR [Da17], [Gu19], classical surveying data or self-sufficient deformation sensors for local information. Many operations already continuously scan the dump during construction using laser-scanning technology [Kr21]. Currently, this information is mainly used for operations control. One opportunity but also a challenge is to make these data and information available in the future, especially during the post-mining phase after operations terminate. This leads to the following research questions of TRIM4Post-Mining:

- How can modern monitoring technology be best utilized, to characterize the geomechanical and geochemical behavior inside the heterogeneous waste dump body at an appropriate spatial and timely resolution? (**Revitalization Monitoring**)
- How can monitoring data be rapidly translated into a better understanding of local developing risks and opportunities? (**Data-Driven Predictive Modelling**)
- How can newly gained information be translated into a decision support for optimized revitalization planning at different stages? (**Revitalization Planning**)
- How can relevant data, information, and knowledge be conserved and made available to all stakeholders in the future? (**Long-term information access**)

2 Methodology

TRIM4Post-Mining is developed around a dense spatiotemporal geodatabase of coal waste dumps, from which the necessary information can be retrieved at any time and newly acquired data can be added (Figure 1). This database forms a grid model, which maps the spatial distribution geochemical and geomechanical properties. One particular feature to be captured is the spectral fingerprint of waste material within. Similar to a spectral library, this allows for the analysis with respect to particular elements also at later stages. The database will further integrate information from different sources, including:

- Operational data originating from the digital geological model and the material tracking system installed,
- Data from terrestrial sensors (face mapping, conveyor belt scanning), including RGB-imaging-, LIBS, Multi-spectral sensors, and terrestrial laser scanning,
- Data from airborne and spaceborne sensor, including radar and multi-spectral remote sensing, airborne laser-scanning and digital photogrammetry,
- In-situ data (drill-hole logging and CPT data).

The investigation of suitable sensor combinations, the development of related data analytics, and monitoring concepts are an intended outcome from TRIM4PostMining.

The two data analytic “engines” to be developed, “Geotechnical Application” and “Geochemical Application” will make use of the most current data to quickly predict risk- and opportunity related indicators, including slumping, settlement liquefaction potential [Di99],[Fö98],[We20], acidification potential or element concentrations etc. Newly evaluated information of system states will be fed back to the model, so it is “up to date”.

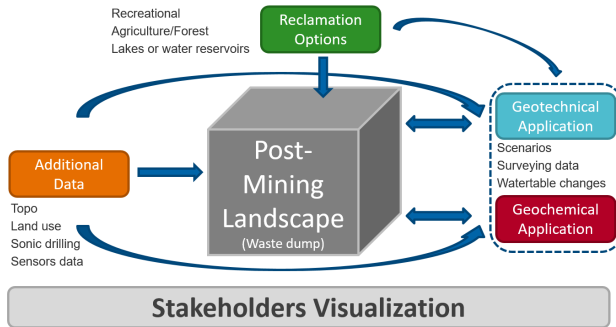


Fig. 1: TRIM4PostMining Approach

The revitalization decision support will apply the model and engines previously described to planning and re-cultivation decisions. Possible reclamation options include recreational areas, agricultural or forestry areas, and lakes or water areas for the remaining residual holes or any constructions related to renewable energy. Each of these options requires some range of geochemical and geotechnical parameters, which are evaluated as part of an objective function aiming for overall societal benefit.

Research activities related to “Stakeholder Visualization” aim to provide interpreted information to different stakeholders. This will include full access to the model to planners and reclamation executers, a web-GIS interface to authorities, and a VR/AR interface to communities and others. The key aspect is that information provided to stakeholders should be pre-interpreted at a level that is appropriate to the technical background of stakeholders and the involvement in the decision process.

3 Outlook

TRIM4Post-Mining aims to develop the framework for a Technology Readiness Level- 5 to 6 (framework tested in an industry-like environment). With this goal, the different research areas contributing and their integration will be developed and demonstrated based on the specific requirement for revitalization planning on two exemplary case sites and involves stakeholder engagement. The case study areas are a waste dump of a lignite

operation south of the city of Leipzig, Germany, and a hard coal waste dump close to the border of the Netherlands and Germany.

Acknowledgments

TRIM4PostMining receives funding from the Research Fund for Coal and Steel under grant agreement No. 899278

References

- [Da17] Dalm, M.; Buxton, M.W. N; van Ruitenbeek, F. J. A.: Discriminating ore and waste in a porphyry copper deposit using short-wavelength infrared (SWIR) hyperspectral imagery. *Minerals Engineering* 105 (2017), pp. 10–18.
- [Di99] Diener, U.: Sackungen von Kippen - Modellversuche und statisch gestützte Betrachtungen. PhD Thesis, TU Bergakademie Freiberg, 1999, Veröffentlichungen des Instituts für Geotechnik Heft 99-2.
- [Fö98] Förster, W.: *Bodenmechanik* Teubner Studienbücher Bauwesen. Teubner Studienbücher. Springer Verlag, 1998.
- [Gu19] Guatame-Garcia, A.: Detection of factors that determine the quality of industrial minerals: An infrared sensor-based approach for mining and process control. PhD Thesis, TU Delft 2019, 205p.
- [Gun17] Gunke, A.; Jany, S.: Ableitung eines Bergmännischen Risswerks, korreliert aus ALS-, TLS- und RGBDaten. In *Tagungsband 19. Geokinematischer Tag 2018*, TU Bergakademie Freiberg, Institut für Markscheidewesen und Geodäsie (Hrsg.), S 79-86.
- [Ha19] Haupt, T. (2019) *Kontaminationsmonitoring von Bergbaurückständen*. Unpublished Diploma Thesis TU Bergakademie Freiberg, 2019.
- [Jo21] John, A. Monitoring of Ground Movements Due to Mine Water Rise Using Satellite-Based Radar Interferometry—A Comprehensive Case Study for Low Movement Rates in the German Mining Area Lugau/Oelsnitz. *Mining* 2021, 1, 35-58. <https://doi.org/10.3390/mining1010004>
- [Kr21] Kressner, M.; Lohsträter, O.: Vom satellitengestützten Gerätefahrassistenzsystem zum markscheiderischen Aufmaß. In *Tagungsband Umwelt, Energie und Rohstoffe 2021*, TU Bergakademie Freiberg, Institut für Markscheidewesen und Geodäsie (Hrsg.), S 83.
- [Sc20] D Schröder et al.: i²MON: development of an integrated monitoring system for the detection of ground and surface displacements caused by coal mining. In *Proceedings Slope Stability 2020*, Australian Centre for Geomechanics, Perth, ISBN 978-0-9876389-7-7.
- [We20] Weißbach, J.: *Risikobeurteilung von Bodenverflüssigungsereignissen auf Innenkippen des Lausitzer Braunkohlereviere*. Unpublished PhD Thesis, TU Bergakademie Freiberg (2020), 347p.

Digitizing Drilling Logs - Challenges of typewritten forms

Kim Bürgl^{1,2}, Lea Reinhardt¹, Frank Binder², Lydia Müller^{1,2}, Andreas Niekler¹

Abstract: In this work, we show prospects of how mining and geological documentation in the form of drilling reports can be digitized and further processed. Processing these typed and handwritten forms poses challenges for document management in renaturation projects. We highlight the structural problems of drilling reports and present three approaches for recognizing and processing the information documented in them. We use optical character recognition and document layout analysis techniques to approach the problem. Layout analysis was performed using a heuristic approach and a neural network for layout recognition. In detail, we show the approaches Form Processing (A), Table detection by line counting (B) and processing with Mask-R-CNN (C). A case study is used to show initial results and challenges. B and C are more robust than A to small changes in the form. C can recognize columns better with more training data than B in cases where table boundaries are not respected. B and C also allow other language models to be used for OCR and can thus also recognize handwriting with appropriate training data.

Keywords: drilling logs; table recognition; information extraction; renaturation projects; OCR; forms processing

1 Introduction

The renaturation and recultivation of former mining areas poses major challenges to mankind worldwide. One of these challenges is the management of extensive geological and engineering documentation during the mining development of these areas. This documentation is significant for public safety and the definition of subsequent use concepts. In this scenario, old drilling logs documenting existing soil layers in the mining areas represent an important type of document. With this information, e.g. models of underground soil layers can be made. For this purpose, the positions of the boreholes and the measurement information need to be digitally available. In our case, a central German open pit mining region, the drilling logs are fixed paper forms in which the measuring engineers recorded their results by typewriter or handwriting. A geographic information system (GIS) is currently under development which gives experts access to all relevant information, including relief data, sensor data and important text passages in reports as shown by [Sc21]. Additionally, soil information from the drilling logs shall be integrated to model soil layers in the former mining areas. Therefore, the tables in the drilling log forms must be recognized, the meanings

¹ Natural Language Processing Group, Leipzig University, Augustusplatz 10, Leipzig, Germany {buergl,lydia, aniekler}@informatik.uni-leipzig.de,lr32coni@studserv.uni-leipzig.de

² Institute for Applied Informatics (InfAI), Goedelerring 9, 04109 Leipzig, Germany binder@infai.org

of the rows and columns must be semantically contextualized, and the values entered must be identified. Drilling logs contain the information of layers detected by drilling, metadata about the drilling process, drilling location, and the drilling time. We are aiming at extracting all relevant information from the form automatically, i.e. we are interested in some of the metadata but also in the information of the layers, their thickness and distance to ground level.

In this work, we discuss three methods how to approach the automatic extraction and contextualization of this information from scanned drilling reports. Thus, we report on the automatic processing of this information for further processes and information systems and show a contribution for the sustainable digitization of soil information documented in forms. While we have drilling logs filled out by hand and typewriter, we focus on forms filled in by typewriter. So far, we have not investigated handwritten forms further because they would require a different class of optical character recognition models. This is a separate task and is left out from this work in order to keep our focus on table contextualization.

In the following sections we present the related work, describe our dataset in more detail, explain the three approaches to extract the required information and examine the individual performances in a case study. We highlight the difficulties and challenges as this is work in progress which needs further investigation.

2 Related Work

Table detection and recognition has been researched for a long time [Em06]. Early works concentrated on the processing of scanned tables, where ruling lines and white space were used to detect table cells [HD95, TBB95]. More recent work mainly focused on digital-born PDFs, i.e. there were two ICDAR competitions on table recognition, 2013 [G3] and 2019 [Ga19]. The 2013 competition used a data set of digital-born PDFs for training and evaluation. Thus, problems arising from poor scan quality, varying lightning, skewing or warping were not considered. In the following the different approaches are summarized.

Fang et al. [Fa11] parses the PDFs using a commercial library. A page layout analysis is performed to filter out text paragraph regions. They use line separators and whitespace delimiters to detect tables. Line separators are detected based on PDFs parsing results, whitespace delimiters are detected with a whitespace detection analysis algorithm. Anssi Nurminen [Nu13] developed a table recognition algorithm that is based on the poppler library. It uses heuristics on page margins, page layout rows, text edges and text layout to detect table regions. Yildiz et al. [YKM05] converts the PDF into an XML file using pdf-to-html. Afterwards it detects tables based on the resulting XML structure. The TABFIND algorithm by Costa e Silva [CeS09] finds tables in text files using a Hidden Markov Model. Stoffel et al. [St10] use a decision-tree classifier and a conditional random field classifier to classify lines. Consecutive lines and table labels are then grouped together. KYTHE uses an optical character recognition (OCR) tool to preprocess PDFs and an interactive postprocessing of the results to detect tables. Liu et al. [Li07] uses a heuristics joining text

lines with uniform font size. Block containing tables are detected based on textual clues and whitespace.

The 2019 competition [Ga19] uses a new dataset containing archival and modern documents, handwritten documents, digital-born documents and images of documents. The Approaches use a variety of techniques described in the following. ABC Fintech uses segmentation model and rules to detect tables. It enhances the training data with simulation data. AIRL uses a bottom up approach detecting table cells with a neural network. Cells are then combined to rows and rows to tables. Applica robots, Table Radar, TJNU202-2, and Clova AI train and use a Faster R-CNN for table recognition and postprocessing steps. Cinnors table uses a U-Net model to detect tables. HCL IDORAN detects vertical and horizontal lines using Hough transform. Lines are then used to detect tables and table structure. Lenovo Ocean uses holistic nested edge detection to perform edge segmentation and table region segmentation. Results from both are combined to find the edge line of the table. NLPR PAL uses a fully convolutional network to classify pixels into table and background pixel. A connected component analysis is used to extract table regions. Table Fan uses YoloV3 model to detect tables.

Otha et al. [Oh19] detect table structure based on the result of the pdf2xml tools. Wei et al. [We21] provide a more fine-grained dataset for evaluation of table cell detection as well as an deep learning approach utilizing Faster R-CNN or Mask-R-CNN. Guerry et al. [GCL19] perform baseline detection. Tables are regarded since they are problematic for baseline detection. Using deep learning hypothetical baselines, begin- and end-indicators, and page boundaries are predicted. Table rulings are extracted using a line segment extractor. All information is combined with a grammar to recognize tables.

Based on this analysis, we decided to study three paradigms: Form Processing (A), Table detection by line counting (B) and Mask-R-CNN (C), since all three have made the best contributions so far in terms of different applications. Approach A leverages commercial OCR engines that performed well in the ICDAR 2013 competition. Approach B uses morphological properties. Such algorithms are still among the best performing approaches in table recognition competitions. Approach C employs Mask-R-CNN which is next to Faster R-CNN, the neural network most often applied to table recognition.

3 Data

Our data consists of about 3000 drilling logs of different quality, mostly from the 1970s. A drilling log is a form consisting typically of two pages. The first page contains information about the drilling process and result, i.e. the layers of soil. The second page contains information about back-filling of the well.

For this work we are only interested in the layers of soil, i.e. the first page of the form, which can be treated as nested tables. On the highest organization level, the form consists of a header, a metadata section about the drilling process, the layers of soil, and a footer. Each section has again a table like structure. While header and footer are not of interest for the drilling process itself, the metadata and the information about the layers of soil contain the

	1	2	3	4	5
Schluff grau		2,0		19,0	95,6
Feinsand grau		8,5		27,5	87,1
Schluff Braunkohle		2,5	13,0	39,8	84,6
Schluff grau tonig		2,0		45,0	69,6

Fig. 1: Section of a drilling log. (a) Detected lines are highlighted. Problems arising with overwritten lines can be seen in the text cells on the left and multiple lines per cell in the third row. (b) Detected Regions of interest (ROI) by Mask-R-CNN are highlighted. The Mask-R-CNN does not detect two ROIs for the overwritten words and numbers in row three. Columns: (1) soil type (2) diameter (3) diameter of lignite (4) total depth (5) height above sea level.

information of interest.

The most important information in the metadata is the geographic location given as coordinates. The documentation of the layers of soil is a table containing the material of the layer, the layer's diameter, the overall depth, and the depth with reference to sea level. Note that the layer's diameter are two columns. One column contains the diameters of layers which contain coal while the other contains all the other diameters (cf. figure 1). The height above sea level of the drilling well is given in the header of the table.

While our whole dataset consists of documents of almost always the same form type, it still has the following challenges: data is filled in by hand with a typewriter or a mix of both, information about a layer of soil may be found in one line in the table or as multi-line in the table, redundant information might or might not be recorded in the form, there might be extra data such as notes on the form, text in data cells may overflow leading to crossings of table borders and data, sometimes two lines are written into the same table line of the cell, and scan quality varies largely. For our case study we selected 10 drilling logs; 5 with good and 5 with lower quality. Some of the stated problems are observable in Figure 1 and 3.

4 Approaches

4.1 Forms Processing (A)

Extracting information from the scans of the drilling logs can be seen as a standard forms processing task, as the drilling logs appear to be in a consistent tabular layout. To derive a baseline for what extraction results can be achieved with relatively little technical effort,

we apply a commercial standard application for forms processing - FormPro³ - which is designed for end users in business settings and requires no specific technical knowledge. It does, however, require the user to initially scan or provide a blank form, and to manually define zones for data extraction on that blank form. To extract data from scanned tables, only the cells of the table's first data row need to be marked as zones.

With no blank form template available, we reconstruct one from the scans using Adobe Acrobat Pro, leveraging the observation made with pdfimages⁴, that the scanned pages in most of our PDF files contain several layered image objects including a pre-separated background image (showing the paper background, table borders and some text fragments). Data extraction from forms in batch mode proceeds as follows: FormPro identifies each page image with (one of) the predefined form template page(s), aligns the scanned image with the template image by automatically identifying anchor zones, performs image comparison against the blank form (or other filter methods), and extracts data from the zones by applying the integrated OmniPage OCR engines. The OCR can be parameterized per zone in terms of character set, including possible suppression of spaces or line breaks. Applying these parameters with care to each zone did considerably improve the extraction results.

However, the forms processing fails if the underlying form template differs from the initially defined (variety of) empty template(s), which causes FormPro's page-wide alignment to fail, even though the preceding identification step may have succeeded. The templates' differences can be subtle, such as slightly differing column widths or row heights, or missing footnotes or labels near the page margins - which are often used as anchor zones for alignment. In addition, after alignment, FormPro maps the predefined extraction zones onto the page with just a little local optimization. Slight changes in form layout and subsequent changes in zone positioning cause the extraction zones to be misplaced, cf. figure 2.

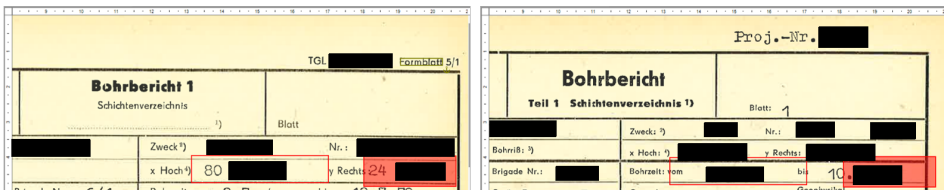


Fig. 2: Header parts of two drilling logs based on slightly differing form templates shown in the same scale. Left: Correctly mapped extraction zones (red) for geographic positions after successful page alignment using the anchor *Formblatt* in top-right corner. Right: Misaligned zones due to deviating positioning of header cells and lack of anchor. Note that the zones are placed in the same positions on both pages. Proprietary data is blackened.

³ Tested with FormPro 3.1.732.0, based on OmniPage Capture SDK v19.20.19161.100

<https://www.ocr-systeme.de/index/formpro/>

⁴ <https://poppler.freedesktop.org/>

4.2 Table detection by line counting (B)

As the drilling log forms are table-structured by vertical and horizontal lines, and most of the drilling logs are based on the same blank form or a slight variation of it, a simple approach is to detect those lines and select cells which are framed by the detected lines. Each selected cell is then processed independently by an OCR. The OCR output is then merged together to the table structure. Line detection parameters, cell count and OCR are adjusted/optimized on a single example document considered to be representative. Most of the functions needed are implemented in OpenCV [Br00]. The steps in detail are⁵: The scanned image is converted to grayscale. Edges are detected by a Canny-filter [Ca86], followed by Closing with a small elliptical kernel. Lines are detected with a Hough transformation [DH72] and are then grouped into vertical and horizontal ones (see figure 1a). Lines of the same orientation are integrated to a single line if close to one-another. Those lines very close to the image border are deleted as they mostly consist of scanning artefacts. Also very short lines and or not properly orientated lines are deleted. Both horizontal and vertical lines are then sorted by their occurrence from top to bottom and left-to-right respectively.

Each cell of interest is defined by its bordering lines. As the drilling logs do mostly have a table structure, one can iterate through vertical and horizontal lines. Each cell is cropped from a preprocessed binary image. This preprocessing consists of the removal of long lines and other standard preprocessing steps such as closing and binarization. The cropped image is then used as input for Tesseract [Sm07], which can be parametrized for each cell differently. E.g. in our case, there are certain columns where we expect only numbers to be present.

Some information however is written in cells which are not properly framed by lines (see figure 2), but are written in certain regions of a bigger cell. In that case one can define the cell of interest by visual judgement, e.g. only half of a cell width shall be used.

(a)	Schluff Braunkohle	2,5	13,0	49,8	84,6
(b)	Ton, braun, kohlig, schluff- fig, sehr fest	0,5		4,2	54,1
(c)	Schluff grau-grün	5,0		110	

Fig. 3: Typical challenges in drilling log table recognition, namely multiple lines per bordered cell (a and b), typewriting overwritten or very close to one-another (a) and forms filled with handwriting and low scan quality (c). Also note that column widths differ in (c).

4.3 Mask-R-CNN (C)

As a third approach, a Mask R-CNN [He17] was used to detect cells of interest. Mask-R-CNN is an extension of Faster-R-CNN. Faster-R-CNN has two steps. In the first step, candidate

⁵ inspired by this blogpost: www.levelup.gitconnected.com/text-extraction-from-a-table-image-using-pytesseract-and-opencv-3342870691ae

bounding boxes are predicted. The second step performs classification based on features of the bounding boxes and bounding box regression. Mask-R-CNN predicts a binary mask for each candidate in parallel to the second step. We annotated 20 drilling logs with Larex [RSP17], where every type of cell (i.e. each cell in the same column) was annotated with a different class. Empty cells were not annotated. There are typically between 35 and 100 cells/regions of interest (ROI) per drilling log. We use the standard training parameters from [Ab17] and their pretrained model trained on the COCO dataset [Li14]. Because of little annotated training data and better performance of this approach, we decided to train three separated models, one for the two coordinates, a second for the text cells describing layer of soil, and a third for the cells containing depth and layer thickness. We also observed that the model is more robust when all the numerical cells belong to the same class. This is probably due to the fact that we only use 20 samples as training data. Detected areas are then converted to a rectangle by using the outermost points (see figure 1b). Rectangles are then cropped from a preprocessed image (see section 4.2) OCR is then performed with Tesseract on these rectangles and the OCR-output is brought into table form by using the position of each detected cell and grouping similar placed cells in rows and columns.

5 Case studies

Due to no available ground truth, the incoherent form usage (e.g. soil layers written in multiple lines) and the ongoing work, we decided to evaluate the above mentioned three different approaches by comparing their behavior on interesting and complicated cases. From the more than 3000 drilling logs, we selected samples which represent different challenges to compare the behaviour of the three approaches on them.

5.1 Numerical cells

All three approaches perform comparably on numerical cells of high quality. However, FormPro fails on one document of this class in the test set, providing no useful results for the respective document at all. That particular document uses a slightly different form template, so that FormPro's page-wide alignment fails, and extraction zones are misplaced, see figure 2. While other available modes of alignment still need to be investigated, it becomes obvious that forms processing relies heavily on consistent form templates. Hence, FormPro may only be useful within classes of drilling logs based on identical templates. As approaches B and C do not rely on predefined zones they handle changes in the templates more robustly.

When examining the drilling logs, we often see characters written over bordering lines (e.g figure 3b: numbers touching the top line). As approach A depends heavily on the correct alignment with the blank form template and local positioning of the pre-defined data extraction zones, overwritten cell borders are problematic if any content crosses the borders of the extraction zones. Approach B is pretty vulnerable to overwritten lines, too.

Even touching a bordering line only slightly often leads to miss-classification. Widening the area of interest by a fixed offset is counterproductive. We guess it leads to different miss-classification from characters protruding into the ROI. Approach C is more robust on overwritten lines, as the Mask-R-CNN correctly adjusts the mask and therefore the region of interest in that direction which increases the performance in the cases we considered. There are issues in some cases, but they do not seem to come from misplaced masks, but from OCR, which is not fine-tuned on the typewriter-font.

As seen in figure 3a (both rightmost cells), there are cells where multiple lines were inserted into a single cell. All three approaches fail on the cases considered. We can only assume that the Mask-R-CNN could recognize multiple lines in a cell with more relevant training data. Overall Approach A, B and C extract 65%, 70% and 83% of the numerical cells correctly.

5.2 Text cells

The column describing the soil layers and their properties is filled with up to a dozen words. In the case of single lined texts, all three approaches produce a satisfactory result. Both approach B and C perform a little worse, but we assume this is due to differing OCR engines used in FormPro and will be fixable with post processing or a better suited OCR model. Additionally, in some cases the last and/or the first cell are not detected by the Mask-R-CNN as it takes surroundings into account. Again, this problem can be addressed with more training data. In general, overwritten text lines harbour the same problems as overwritten numerical lines (see section 5.1) where the Mask-R-CNN approach is the most robust.

5.3 Text fields not bordered by lines

As described in section 3, there is a single cell in the header of each drilling log which contains two coordinates. This is a special case as there are multiple pieces of information in a single cell, including the cell's label. We observe that both approach A and B solve this problem relatively well, disregarding smaller OCR errors and the known errors occurring with overwritten lines (see section 5.1 and 5.2). The Mask-R-CNN however sometimes fails to detect one of the two coordinates. We conclude that this problem is due to handwritten corrections and less training data, as these cell only occurs once per file.

6 Conclusion

We presented three different approaches to extract data from forms with a nested table structure. Approach A uses a combination of commercial tools and free PDF tools to generate a blank form and extract the data from predefined extraction zones after subtracting the blank form. The second approach (B) finds lines to detect cells. The contents of the

cells are then passed to an OCR engine for optical character recognition. Afterwards the table semantic is reconstructed using the cells and the OCR result. Approach C detects cells of different types using Mask-R-CNN. Similar to the second approach, cells are passed to the OCR engine and the semantic table structure is created afterwards based on the combination of the cells' positions and the OCR results. Both latter approaches perform comparable or even better than the commercial product and - albeit not being available out-of-the-box for non-technical users - are more easily adaptable for further investigation. Although quite a few of our samples of forms are filled by hand, we focused on text written with a typewriter in this work, as tools for handwritten text recognition (HTR) are more complicated to obtain. But as handwritten text overwrites bordering lines a lot more often than text written by typewriters (see figure 3c) we will further investigate the hypothesis that approach C would perform best with more training and using a handwritten language model for OCR.

With our case study and examination of three system designs that have been strongly supported in recent years, we have found strong evidence that the Mask-R-CNN method is probably the most complete solution approach in the end. As we only used 20 training samples for the Mask-R-CNN and still produced robust results in comparison, we conclude that this method can be strongly improved with a well annotated training set, especially in the challenging cases, which are overwritten bordering lines and multiple text lines per cell.

ACKNOWLEDGEMENT

This research was partially funded by the Development Bank of Saxony (SAB) under project number 100335729 and 100400221.

References

- [Ab17] Abdulla, W.: , Mask R-CNN for object detection and instance segmentation on Keras and TensorFlow. https://github.com/matterport/Mask_RCNN, 2017.
- [Br00] Bradski, G.: The OpenCV Library. Dr. Dobb's Journal of Software Tools, 2000.
- [Ca86] Canny, J.: A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence*, (6):679–698, 1986.
- [CeS09] Costa e Silva, A.: Learning rich hidden markov models in document analysis: Table location. In: 2009 10th International Conference on Document Analysis and Recognition. IEEE, pp. 843–847, 2009.
- [DH72] Duda, R. O.; Hart, P. E.: Use of the Hough transformation to detect lines and curves in pictures. *Communications of the ACM*, 15(1):11–15, 1972.
- [Em06] Embley, D. W.; Hurst, M.; Lopresti, D.; Nagy, G.: Table-processing paradigms: a research survey. *International Journal of Document Analysis and Recognition (IJ DAR)*, 8(2-3):66–86, 2006.
- [Fa11] Fang, J.; Gao, L.; Bai, K.; Qiu, R.; Tao, X.; Tang, Z.: A Table Detection Method for Multipage PDF Documents via Visual Separators and Tabular Structures. In: 2011 International Conference on Document Analysis and Recognition. pp. 779–783, 2011.

- [Ga19] Gao, L.; Huang, Y.; Déjean, H.; Meunier, J.; Yan, Q.; Fang, Y.; Kleber, F.; Lang, E.: ICDAR 2019 Competition on Table Detection and Recognition (cTDaR). In: 2019 International Conference on Document Analysis and Recognition (ICDAR). pp. 1510–1515, 2019.
- [GCL19] Guerry, C.; Couasnon, B.; Lemaitre, A.: Combination of Deep Learning and Syntactical Approaches for the Interpretation of Interactions Between Text-Lines and Tabular Structures in Handwritten Documents. In: 2019 International Conference on Document Analysis and Recognition (ICDAR). IEEE, sep 2019.
- [G3] Göbel, M.; Hassan, T.; Oro, E.; Orsi, G.: ICDAR 2013 Table Competition. In: 2013 12th International Conference on Document Analysis and Recognition. pp. 1449–1453, 2013.
- [HD95] Hori, O.; Doermann, D. S.: Robust table-form structure analysis based on box-driven reasoning. In: Proceedings of 3rd International Conference on Document Analysis and Recognition. volume 1. IEEE, pp. 218–221, 1995.
- [He17] He, K.; Gkioxari, G.; Dollár, P.; Girshick, R.: Mask R-CNN. In: Proceedings of the IEEE international conference on computer vision. pp. 2961–2969, 2017.
- [Li07] Liu, Y.; Bai, K.; Mitra, P.; Giles, C. L.: TableSeer: Automatic Table Metadata Extraction and Searching in Digital Libraries. Association for Computing Machinery, 2007.
- [Li14] Lin, T.; Maire, M.; Belongie, S.; Hays, J.; Perona, P.; Ramanan, D.; Dollár, P.; Zitnick, C. L.: Microsoft COCO: Common Objects in Context. In: European conference on computer vision. Springer, pp. 740–755, 2014.
- [Nu13] Nurminen, A.: Algorithmic extraction of data in tables in PDF documents. Master's thesis, 2013.
- [Oh19] Ohta, M.; Yamada, R.; Kanazawa, T.; Takasu, A.: A Cell-Detection-Based Table-Structure Recognition Method. Association for Computing Machinery, 2019.
- [RSP17] Reul, C.; Springmann, U.; Puppe, F.: Larex: A semi-automatic open-source tool for layout analysis and region extraction on early printed books. In: Proceedings of the 2nd International Conference on Digital Access to Textual Cultural Heritage. pp. 137–142, 2017.
- [Sc21] Schröder, C.; Bürgl, K.; Annanias, Y.; Niekler, A.; Müller, L.; Wiegrefe, D.; Bender, C.; Mengers, C.; Scheuermann, G.; Heyer, G.: Supporting Land Reuse of Former Open Pit Mining Sites using Text Classification and Active Learning. 2021.
- [Sm07] Smith, R.: An overview of the Tesseract OCR engine. In: Ninth international conference on document analysis and recognition (ICDAR 2007). volume 2. IEEE, pp. 629–633, 2007.
- [St10] Stoffel, A.; Spretke, D.; Kinnemann, H.; Keim, D. A.: Enhancing Document Structure Analysis Using Visual Analytics. Association for Computing Machinery, 2010.
- [TBB95] Turolla, E.; Belaïd, Y.; Belaïd, A.: Form item extraction based on line searching. In: International Workshop on Graphics Recognition. Springer, pp. 69–79, 1995.
- [We21] Wei, D.; Lu, H.; Zhou, Y.; Chen, K.: Image-based Table Cell Detection: a Novel Table Structure Decomposition Method with New Dataset. In: 2020 25th International Conference on Pattern Recognition (ICPR). pp. 1–7, 2021.
- [YKM05] Yildiz, B.; Kaiser, K.; Miksch, S.: pdf2table: A Method to Extract Table Information from PDF Files. pp. 1773–1785, 01 2005.

Object Detection and Classification in Digital Surface Models of the Lausitz Region in Germany

Benjamin Kalloch,¹ Toni Tontchev,¹ Mario Hlawitschka¹

Keywords: Digital Surface Models; Classification; Object Recognition; Image Processing

1 Introduction

Digital surface models (DSM) derived from light detection and ranging data comprise the height profile of entire landscapes with their natural and artificial structures. Detection and classification of these structures in time series of DSM data support the administration of these regions by assessing, e.g., the growth or shrinkage of forests, geomorphic changes, for instance, bank erosion of water bodies, or urban development. Employed techniques range from image processing, machine learning [Se19] to deep learning [LDZ20].

Typically, machine learning-based approaches provide high flexibility in terms of detectable objects but they require adequate training data, whose creation can be a tedious task. Classical image-processing approaches on the other hand may detect basic shapes and contours without prior information. Pre-classification based on features like object shape, size and structure can facilitate the training data generation process.

We present work in progress towards such a multi-modal approach. Its eligibility is assessed by independently classifying objects from DSM data of the Lausitz region in Germany.

2 Method description & interim results

The analyzed data comprised a squared region of 9 km × 10 km of the Lausitz region in Germany, with a spatial resolution of 1 m in the area, totaling 90,009,000 data points, and a height resolution of 1 cm. They were acquired between the years 2005 and 2018 with at least a year between consecutive time points. For comparability, we focused on data from 2016. No other data, such as multispectral satellite data, were available.

2.1 Image processing-based analysis

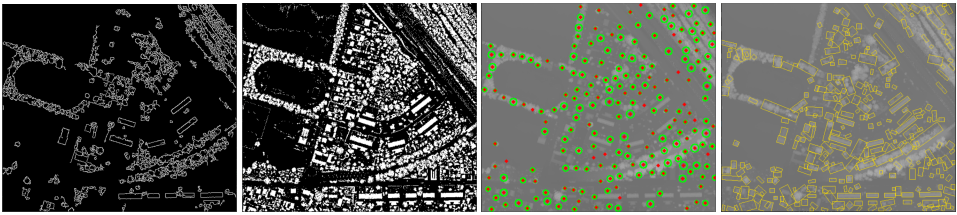
The DSM data were first transformed to grayscale images reflecting the height information as pixel intensities. This representation allows two strategies for object detection: a) analyses

¹ University of Applied Sciences Leipzig, Faculty of Computer Science and Media, Gustav-Freytag Straße 42a, 04107 Leipzig, Germany. *Corresponding author:* mario.hlawitschka@htwk-leipzig.de

of intensity changes using gradients for the pixel-wise identification of objects and their boundaries (fig. 1a), and b) analyses of height differences for the separation of objects from the ground through thresholding (fig. 1b). Here, objects are sets of pixels with homogenous intensities, i.e., uniform height. Changes in the height properties indicate separate objects.

Since neighboring objects are not strictly isolated, special methods are required to distinguish single objects from groups of objects. The shape and size of objects can predict their class. High vegetation (trees) shall have a rounded shape, buildings a mostly rectangular one. We tested Hough circle transformation with a maximum radius of 10 m over sections of the original image to detect single big trees. While functional for single trees, this method failed to recognize groves (fig 1c). Furthermore, nearby buildings impair the detection accuracy. Determining rectangular boxes or circles in which the object optimally fits might constitute another powerful solution for detecting buildings and trees, respectively (fig. 1d).

Further analyses focused on contours of objects encompassing several advantages. Contours can reduce the data volume by only storing segments of their boundaries instead of pixel clusters. They allow the application of powerful geometric algorithms. Finally, contours may have spatial hierarchies in form of common convex hulls enclosing multiple detected objects and facilitating the detection of areas of vegetation.



(a) Gradient edge detection. (b) Thresholding. (c) Hough transformation. (d) Box fitting.

Fig. 1: Image processing based analyses in a subset of the original DSM.

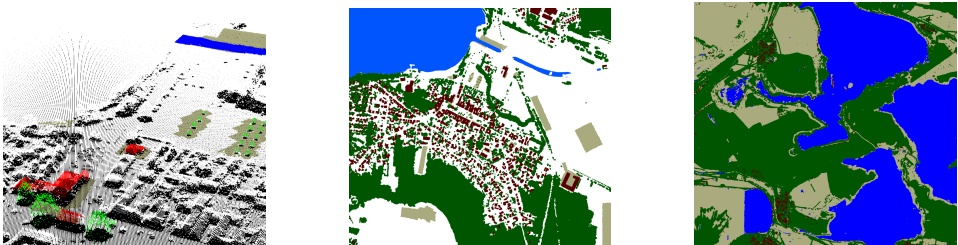
The output of the DSM analyses are closed polygons (for contours of irregular objects), rotated boxes (for rectangular objects), or circles (for round objects). Objects will be pre-classified accordingly as forests, buildings, or single trees.

2.2 Machine learning-based analysis

For the machine learning-based analyses, we employed the Computational Geometry Algorithms Library (CGAL) v.5.1.1. The DSM data were analyzed as point clouds using the CGAL Classification concept implementing a supervised random forest classifier [Wa14]. We distinguished between the classes ground, vegetation, buildings, and water.

Training data were created semi-automatically from few fully manually labeled objects (fig. 2a). Using these initial objects, a larger sub-region of 500×500 m was automatically classified (fig. 2b), manually corrected, and used as training data for the entire region.

The classification focused on the following features that were computed prior to classification: for each point the distance to 1) a locally estimated plane, 2) the global estimation of the ground, 3) the highest point of its surroundings, and 4) the lowest point of its surroundings was determined. For local point neighborhoods ($n = 12$), 5) the eigenvalues of their covariance matrices, 6) the dispersion along the z-axis, 7) the range between the highest and lowest points, and 8) the difference between the local normal vector and the vertical vector were computed. The classification was hierarchically executed in 6 levels of detail and regularized to ensure homogeneous results. While working reasonably in smaller sub-regions with accuracies > 0.9 , the accuracy of the classification of the whole dataset (fig. 2c) is limited between 0.4 and 0.5. Adjustments of the classifier parameters and further cleaning of the semi-automatically derived training data may be required.



(a) Manually labeled training data.

(b) Complete training data.

(c) Full classification result.

Fig. 2: Classification: red=building, green=vegetation, brown=ground, blue=water, white=not labeled

3 Outlook

Both methods, image processing- and machine learning-based, require further fine-tuning. We intend to facilitate and augment the creation of training data by applying the described image processing workflow. Finally, the classification results from the machine learning approach will be employed to detect changes in the vegetation, building development, or water banks across multiple years.

Bibliography

- [LDZ20] Liu, Bo; Du, Shihong; Zhang, Xiuyuan: Land Cover Classification Using Convolutional Neural Network with Remote Sensing Data and Digital Surface Model. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 3:39–43, 2020.
- [Se19] Sevgen, Sibel Canaz: Airborne lidar data classification in complex urban area using random forest: a case study of Bergama, Turkey. *International Journal of Engineering and Geosciences*, 4(1):45–51, 2019.
- [Wa14] Walk, Stefan: , ETH Zurich Random Forest Template Library. ETH Zurich, Department of Civil, Environmental and Geomatic Engineering, Institute of Geodesy and Photogrammetry, 2014.

Geodaten in der Öffentlichen Verwaltung

Modernisierung der öffentlichen Leistungserbringung durch Digitalisierung

Christoph Mengs¹, Christian Bender², Mario Hesse³

Abstract: Digitale Geodaten stellen eine wichtige Querschnittsressource für die öffentliche Leistungserbringung dar. Durch die INSPIRE-Richtlinie wurden hierzu europaweite Standards etabliert. In der Praxis zeigt sich jedoch, dass nicht nur anwendungsbezogene Probleme im Hinblick auf Datenqualität und Schnittstellen existieren, sondern dass auch die Datenerhebung und deren Pflege Kommunen, Länder und Bund auf eine harte Probe stellen. In diese komplexe Gemengelage ist das Onlinezugangsgesetz (OZG) einzubetten, das Bund, Länder und Kommunen dazu verpflichtet, 575 identifizierte Verwaltungsdienstleistungen auch digital anzubieten. Bei dessen Implementierung sind Geodaten ein entscheidender Baustein. In diesem Zusammenhang wurden im Rahmen des Forschungsprojektes Smart Development Infrastructure (SARDINE) Experteninterviews im Freistaat Sachsen mit Akteuren auf Ebene der Kommunen und des Landes durchgeführt. Diese zeigen, dass die Zielstellungen der Umsetzung des OZG und der Einbettung von Geodaten in das Verwaltungshandeln bekannt sind, jedoch Hemmnisse existieren, die dessen Implementierung erschweren. Der vorliegende Beitrag benennt und systematisiert diese Problemstellungen und zeigt mögliche Handlungsoptionen auf.

Keywords: Öffentliche Finanzen; Geodaten; E-Government; Geo-Government

1 Einleitung

Digitale Geodaten⁴ finden eine immer breitere Anwendung innerhalb der öffentlichen Verwaltung. Von den ersten Ansätzen von analogen Karten hin zu digitalen Geodaten (1970er), über die Entwicklung gemeinsamer einheitlicher Geodatenmodelle und interoperabler Geoinformationssysteme (1980er), hin zum Aufbau erster Geodateninfrastrukturen zum Datenaustausch (1990er) stellt die EU-Richtlinie

¹ Universität Leipzig, Institut für Finanzen und Public Management, Kompetenzzentrum für kommunale Infrastruktur Sachsen, Städtischen Kaufhaus, Universitätsstraße 16, 04109 Leipzig, mengs@wifa.uni-leipzig.de

² Universität Leipzig, Institut für öffentliche Finanzen und Public Management, Professur für Finanzwissenschaften, Augustusplatz 10, 04109 Leipzig, bender@wifa.uni-leipzig.de

³ Universität Leipzig, Institut für Finanzen und Public Management, Professur für Finanzwissenschaften, Augustusplatz 10, 04109 Leipzig, hesse@wifa.uni-leipzig.de

⁴ Vgl. Lakes (2019), S. 1348-1349.

Infrastructure for Spatial Information in Europe (INSPIRE) den aktuellen Umsetzungsstand bisheriger Aktivitäten zur Digitalisierung behördlicher Geodaten dar.⁵

Die Richtlinie 2007/2/EG ist seit 15. Mai 2007 in Kraft und verpflichtet die Mitgliedstaaten eine Geodateninfrastruktur (GDI) aufzubauen, die die interoperable Verwendung von Geodaten über Verwaltungsebenen hinweg ermöglicht. Primäres Ziel einer gemeinsamen europäischen Geodateninfrastruktur ist es, Grundlagen für eine gemeinschaftliche Umweltpolitik zu schaffen. Darüber hinaus soll eine europaweite Metadatenharmonisierung erfolgen und somit eine effizientere Verwendung von Geodaten in allen Verwaltungsleistungen ermöglicht werden.⁶ Hierbei gilt es, die Geodateninfrastruktur der Mitgliedstaaten so auszulegen, dass sie „auf optimal geeigneter Ebene gespeichert, zugänglich gemacht und verwaltet werden [...] [sowie] auf kohärente Art verknüpft und von verschiedenen Nutzern und für unterschiedliche Anwendungen genutzt werden können“⁷. Somit ist festzuhalten, dass die Umsetzung von INSPIRE zunächst eine interne Optimierung der Geodatenhaltung auf allen Ebenen der Verwaltung verfolgt.⁸ Die Implementierung sollte in den Mitgliedstaaten stufenweise nach Priorisierung erfolgen. Insgesamt sieht die Richtlinie 34 Geodaten-Themen vor, die die gemeinsamen Metadaten sowie Interoperabilität der Geodatenätze und -dienste sicherstellen sollen.⁹

Die Verknüpfung digitaler (Geo-)Daten hat einen hohen potenziellen gesellschaftlichen Mehrwert.¹⁰ Durch die Auswertung georeferenzierter Zeitreihen und Echtzeitdaten lassen sich relevante Entscheidungshilfen für die öffentliche Verwaltung bereitstellen. Die aktive Nutzung von Geodaten stellt hierbei eine nutzbringende Form der Digitalisierung in Staat und Verwaltung dar. Darauf gestützte deskriptive Analysemethoden umfassen (1) die kartographische Visualisierung von Informationen, (2) deskriptive Analysen zur Erkennung von Mustern, Anomalien und Ausreißern, (3) die Integration der räumlichen Lage von verschiedenen (Fach-)Datensätzen, (4) Diagrammvisualisierungen, (5) die Analysen nach topologischen Merkmalen und (6) nach geometrischen Merkmalen.¹¹ Auf Grundlage dieser Analysemethoden ist es möglich, bereits vorhandene Methoden zur Auswertung von Fachdaten durch die Betrachtung der räumlichen Lage um eine weitere Sichtweise zu erweitern, Fachdaten gemeinsam darzustellen und auf mögliche Zusammenhänge zu untersuchen. Je nach Verfügbarkeit weiterer Fachdaten und der individuellen Fragestellung können somit bisherige Monitoring- und Planungsprozesse in

⁵ Vgl. Bernhard und Mäs (2020), S. 105-109.

⁶ Zu den Zielen von INSPIRE siehe Richtlinie 2007/2/EG, Absatz 1 bis 35.

⁷ Richtlinie 2007/2/EG, Absatz 6.

⁸ Jedoch sollen öffentliche Daten, die im gesetzlichen Rahmen erhoben werden – so auch geographische Informationen – für Private der Zugang erleichtert werden. Folglich gilt es grundsätzlich die GDI-Daten, zur Verwendung für Dritte, bereitzustellen. Siehe dazu Richtlinie (EU) 2019/1024.

⁹ Siehe Richtlinie 2007/2/EG, Artikel 5 bis 10 sowie Anhang I, II und III.

¹⁰ Vgl. Huang und Ribers/Ullrich (2021), S. 240-246.

¹¹ Siehe dazu Bernhard und Mäs (2020), S. 101-110.

der öffentlichen Verwaltung verbessert werden. Auch Ad-hoc-Entscheidungen können durch die Einbindung dynamischer Geodaten schneller und effizienter gefällt werden.

Nach einer weiterführenden Beschreibung des föderalen Staatsaufbaus und der Einbettung der fortlaufenden Digitalisierung der öffentlichen Verwaltung in das Public Management wird die INSPIRE-Richtlinie als Treiberin der Einbeziehung von Geofachdaten in die Verwaltungsdigitalisierung qualitativ analysiert. Neben einer literaturbasierten Auswertung werden zudem Ergebnisse einer Expertenbefragung vorgestellt, die im Rahmen des SARDINE-Projektes im Freistaat Sachsen durchgeführt worden sind und als Anwendungsbeispiel dienen sollen. Mit der Verknüpfung beider Ansätze sollen neben einer Beschreibung des Status Quo vor allem Hindernisse und mögliche Lösungsansätze herausgearbeitet werden, die den Einbezug von Geodaten in das öffentliche Verwaltungshandeln betreffen.

2 Implementierung neuer Anwendungen in der öffentlichen Verwaltung

2.1 Staatsaufbau und Kompetenzverteilung

Die Bundesrepublik Deutschland verfügt über einen zweistufigen Staatsaufbau, bestehend aus dem Bund sowie den sechzehn Bundesländern.¹² Der deutsche Föderalstaat weist zudem einen dreistufigen Verwaltungsaufbau auf: Hierbei sind die Gemeinden Teil der Länder, verfügen jedoch gemäß Art 28 Abs. 2 GG über ein verfassungsrechtlich geschütztes Selbstverwaltungsrecht. Hieraus ergeben sich unterschiedliche Zuständigkeiten und Schnittstellen, die sich auf das Verwaltungshandeln und folglich auf die Möglichkeiten und Grenzen der Digitalisierung auswirken. Hinsichtlich der kommunalen Ebene ist die Ausgangslage bezüglich der Rahmensetzung zum Thema Digitalisierung heterogen, da diese durch die Länder mitgestaltet wird.

Wesentlich für den bundesrepublikanischen Staatsaufbau ist die ausdifferenzierte Kompetenzverteilung. So trennt das Grundgesetz nach Gesetzgebungsmöglichkeiten (Entscheidungskompetenz) und Aufgabenvollzug (Durchführungskompetenz). Die Finanzierungskompetenz obliegt nach Art. 104a Abs. 1 GG grundsätzlich der ausführenden Ebene. Hierbei liegt die Entscheidungskompetenz mehrheitlich auf Ebene des Bundes, wohingegen Durchführungs- sowie Finanzierungskompetenzen eher auf Ebene der Länder, respektive der Gemeinden, angesiedelt sind. Vor diesem Hintergrund ist das föderale System Deutschlands als ein ausgeprägtes Verbundsystem mit vielfältigen Verflechtungen zu charakterisieren.

¹² Siehe Art. 30, 70 ff. GG.

Der INSPIRE Richtlinie folgend liegt die Entscheidungskompetenz auf supranationaler europäischer Ebene. Durchführungs- und Finanzierungskompetenz obliegen hierbei in Deutschland den Bundesländern, die wiederum Aufgaben an die Gemeinden übertragen können. Falls auf dieser Ebene eine Aufgabenübertragung erfolgt, hat die Ausgabenlast, dem Bestellerprinzip folgend – Wer bestellt, muss auch zahlen –, durch die übergeordnete Ebene getragen zu werden.¹³

Somit unterliegt die Implementierung der INSPIRE-Richtlinie dem föderalen Wettbewerb zwischen den Ländern. Dieser enthält zwei Dimensionen: den allgemeinen Standortwettbewerb und die Möglichkeit einzelner Länder, als Innovationstreiber aktiv zu werden.¹⁴ Beide Aspekte können positive und negative Auswirkungen aufweisen. So wirkt sich der Standortwettbewerb erstens auf die direkte bedarfsgerechte Bereitstellung von Geobasisdaten sowie georeferenzierter Fachdaten für Unternehmen und Bürger aus und zweitens indirekt durch die allgemeine Optimierung der Erbringung öffentlicher Leistungen.¹⁵ Der Innovationstreiber kann positiv als Versuchsfeld für Neuerungen verstanden werden. Die bessere Lösung wird sich langfristig auch in anderen Ländern durchsetzen. Negativ kann hierbei der häufig als föderaler Flickenteppich bezeichnete Zustand der Gleichzeitigkeit einer Vielzahl verschiedener Lösungen verstanden werden. Insbesondere die überregionale Nutzung von Geodaten wird hierdurch erschwert – ein Aspekt, der durch den europäischen Impuls eigentlich erleichtert werden soll.

Um aus ökonomischer Perspektive zu entscheiden, inwieweit auf Ebene der Länder die Umsetzung der INSPIRE-Richtlinie auf die Gemeinden übertragen werden sollte, kann das Kriterium der allokativen Effizienz herangezogen werden.¹⁶ Dieses bemisst erstens die Beachtung regional unterschiedlicher Nutzerpräferenzen¹⁷ und zweitens die Kostenstrukturen. Falls die Nutzerpräferenzen räumlich stark variieren, gilt es, dieser Varianz durch eine Dezentralisierung der Leistungserbringung mit Abweichungsmöglichkeiten zu entsprechen. Falls dies nicht der Fall ist, erscheint nach dem Effizienzkriterium eine zentrale Lösung als zielführend. In Bezug auf die Kosten der öffentlichen Leistungserbringung gilt es, die Kostenentwicklung bei Änderung der Outputmenge zu berücksichtigen. Liegen sinkende Stückkosten vor (sog. Economies of scale), sind eher zentralisierte Lösungen als kosteneffizient zu betrachten.¹⁸

Mit Blick auf die Implementierung der INSPIRE-Richtlinie ist dem Allokationsziel folgend von einer zentralen Erbringung auszugehen: Die räumliche Präferenzverteilung

¹³ In dieser Hinsicht weichen die Regelungen der Landesverfassungen vom oben genannten Grundsatz des Art. 104a Abs.1 GG ab.

¹⁴ Vgl. Zimmermann (2019), S. 33-34.

¹⁵ Schätzungsweise 80 Prozent der Verwaltungsleistungen verfügen über einen Raumbezug.

¹⁶ Vgl. Zimmermann (2019), S. 34.

¹⁷ Nutzerpräferenzen umfassen nicht nur die Präferenzen der Verwaltung, sondern auch die der Bürger sowie der Unternehmen.

¹⁸ Ein Leistungsangebot mit sinkenden Durchschnittskosten ist dadurch charakterisiert, dass es hohe Fixkosten bzw. Kosten der Entwicklung und der ersten Inbetriebnahme gibt. Durch den Zuwachs an Nutzern jedoch kaum zusätzliche variable Kosten entstehen und die Fixkosten auf eine immer weiter steigende Zahl von Nutzern verteilt werden.

scheint in Bezug auf die Bereitstellung von Geobasisdaten homogen zu sein. In Bezug auf georeferenzierte Fachdaten kann dies allerdings nicht angenommen werden, da hier durchaus räumlich heterogene Präferenzen bestehen.¹⁹ Technische Lösungen – insbesondere bei digitalen Anwendungen – rufen in der Regel sinkende Stückkosten bei Erhöhung der Outputmenge hervor.²⁰ Die Entwicklung und Einführung entsprechender digitaler Verwaltungsprozesse haben zwar hohe initiale Kosten, die Ausweitung auf zusätzliche Nutzer dagegen verursachen nur geringe zusätzliche Kosten. In der mittleren und langen Frist sind somit potenzielle Effizienz- und Kostenvorteile zu heben. Die Zusammenführung der Aspekte zeigt, dass zentrale Lösungen nicht per se effizienter und damit gegenüber kleinteiligen dezentralen Lösungen überlegen sind. Beide haben eine Daseinsberechtigung. In Bezug auf den föderativen mehrteiligen Staatsaufbau ist es deshalb von hoher Relevanz, funktionierende technische Schnittstellen zu haben, um bei Bedarf dezentrale Module an zentralisierte Lösungen anbinden zu können. Dies kann als zentrale Bedingung angesehen werden, sofern Geodaten im Zuge der Digitalisierung der öffentlichen (Verwaltungs-)Ebenen übergreifend einen Mehrwert bieten sollen.

2.2 Public Management im Kontext evolutionärer Entwicklungslinien

Für die Implementierung neuer und digitaler Anwendungen in der öffentlichen Verwaltung bedarf es eines Instrumentes zur strategischen Umsetzung. Das E-Government ist ein „Instrument der organisatorischen Gestaltung von Interaktions- und Kommunikationsbeziehungen sowie Leistungsprozessen innerhalb des Staates wie auch zwischen dem Staat und seinen Anspruchsgruppen mittels Informations- und Kommunikationstechnologien“.²¹ Somit ist das E-Government als eine Umsetzungshilfe zu verstehen, mit welcher bestehende Prozesse digitalisiert und neue Prozesse in Bestehende integriert werden. Es ist kein Selbstzweck. Der Weg zum E-Government kann als evolutionärer Prozess gesehen werden, welcher durch verschiedene Konzepte beschrieben werden kann. MOON und WELCH (2015) haben hierzu ein Entwicklungsstufenkonzept entwickelt, das die unterschiedlichen Digitalisierungsfortschritte übersichtlich katalogisiert.²²

Stufe 1 Web-Präsenz mit passiver Informationsversorgung

¹⁹ Die im Rahmen des Projektes SARDINE durchgeführten Experteninterviews dienen als Grundlage für diese Einschätzung. Um einen allgemeinen Raumbezug zu erstellen sind homogene Basisdaten notwendig. Anders gestaltet sich dies für Geofachdaten. So sind erstens die Geofachdatenverfügbarkeit in Form von Quantität und Qualität über die Gebietskörperschaften hinweg sehr unterschiedlich. Zweitens sind das in der jeweiligen Verwaltung vorgehaltene Fachpersonal sowie die angewendete technische Ausstattung sehr verschieden. Zudem weichen die zu erbringenden Aufgaben je Verwaltungsebene voneinander ab. In der Konsequenz erscheint es als logisch, dass die Präferenzen heterogen ausgeprägt sind.

²⁰ Vgl. de Mello und Ter-Minassian (2020), S. 9-11.

²¹ Ritz und Thom (2019), S. 613.

²² Vgl. Moon und Welch (2015), S. 436-455.

Stufe 2	Zwei-Weg-Kommunikation zwischen Verwaltung und Bürger mit elektronischen Formularen, interaktiven öffentlichen Statistiken und Daten
Stufe 3	Reorganisierte Geschäftsprozesse mit online Service- und Finanztransaktionen
Stufe 4	Horizontal und vertikal integrierte Leistungsprozesse
Stufe 5	Übergangslose Integration unterschiedlicher öffentlicher Informationen und Leistungen und die Verfügbarkeit auf weiteren Medienträgern und Technologien

Beginnend mit der ersten Stufe wird eine reine Internetpräsenz zur einseitigen Informationsversorgung der Unternehmen und privaten Haushalte initiiert. Im fortschreitenden Prozess der zweiten Stufe öffnet sich der Informationskanal und eröffnet die Möglichkeit zur beidseitigen Kommunikation. Transmissionsmedien können hier elektronisch auszufüllende Formulare oder interaktive öffentliche Statistiken sein. Über reorganisierte Geschäftsprozesse der dritten Stufe werden online Service- und Finanztransaktionen möglich, sodass beispielsweise Zahlungen direkt online abgewickelt werden können. Im Zuge der vierten Stufe integrieren sich die Leistungsprozesse in horizontaler und vertikaler Weise. Ziel ist es, Ebenen übergreifende Aufgabenfelder und Geschäftsprozesse miteinander zu verknüpfen. Hier setzt auch das OZG nachfrageseitig an, welches zum Ziel hat, über einen einheitlichen föderativen Portalverbund zur Nutzbarmachung der OZG-Leistungen zu schaffen. Die fünfte Stufe kann mit dem Begriff des ‚Self-Governments‘ umschrieben werden, wobei der universelle und zeitlich unabhängige Zugriff auf Datenbestände und dessen Weiterverwertung die Bürger in die staatliche Leistungserbringung als Ko-Produzenten involviert.²³ Bürger und Staat interagieren somit auf Konsumenten- und Produzentenseite. Dabei ist die Evolution hin zur Stufe fünf der übergangslosen Integration kein Automatismus. Vielmehr muss eine fortschreitende Digitalisierung bekräftigt und forciert werden. Nach ANDERSEN und HENRIKSEN (2006) wird diese technologisch geprägte Sichtweise der Evolution von E-Government nicht vollständig gerecht. So wird durch die beiden Autoren herausgearbeitet, dass Evolutionsstufen durch den Mehrwert entstehen, der für die Nutzer generiert wird und welche organisatorischen Restrukturierungsimpulse für staatliche Stellen damit verbunden sind.²⁴

Folglich ist eine Evolutionsperspektive bei der Betrachtung des E-Government Ansatzes angebracht, welche jedoch um weitere Perspektiven erweitert werden muss. Dabei spielen neben der Frage, welche technologischen Möglichkeiten die Interaktion zwischen Bürger und Staat gegeben sind, insbesondere auch die organisationstechnische Ausgestaltung und die Dimension der Nutzbarmachung eine entscheidende Rolle. So muss der Grenznutzen

²³ Vgl. Ritz und Thom (2019), S. 615f.

²⁴ Vgl. Andersen und Henriksen (2006), S. 246.

der Implementierung von E-Government Anwendungen einen Mehrwert erkennen lassen, der einerseits eine Restrukturierung von Prozessen der staatlichen Leistungserbringung rechtfertigt und andererseits einen erkennbaren Zugewinn für dessen potenzielle Nutzer offeriert. Ist dies nicht der Fall, kommt der Evolutionsprozess trotz gegebener technischer Möglichkeiten zum Erliegen.

Neben der Integration von E-Government aus einer evolutionstechnischen Sichtweise sind dessen Antriebskräfte zu benennen, die eine permanente Umsetzung von staatlich-analogen hin zu staatlich-digitalen Prozessen befördern. Prozesse können exogen angestoßen und beschleunigt werden, wobei ein Schock von außen auf ein bestehendes (Verwaltungs-)System einwirkt. Dagegen beschreibt ein endogener Schock, dass eine Veränderung systemisch bedingt ist und demnach aus sich selbst heraus entsteht. Beide Perspektiven, die exogene, wie auch die endogene Sichtweise zeigen Ansätze auf, warum die Integration von E-Government, insbesondere in Bezug zu Geodaten, entlang der oben aufgezeigten Stufen erfolgt. So kann die INSPIRE-Richtlinie als exogener Schock für die kommunale Ebene gesehen werden, da diese im Rahmen ihrer Aufgabenzuteilung zur Mitwirkung verpflichtet sind. Endogene Prozesse hingegen, die aus sich selbst heraus entstehen, können abseits rechtlicher Vorgaben in der Verwaltungsdigitalisierung selbst gesehen werden. Das evolutionäre Stufenmodell nach MOON und WELCH (2015) verdeutlicht, wie ein endogener Prozess zu einer zunehmenden Verwaltungsdigitalisierung führen kann, wenn er Wirtschaftlichkeitsvorteile verspricht.

2.3 Geo-Government: Digitale Geodaten im Kontext des Public Managements

Digitale Geodaten können als Geoobjekte²⁵ umschrieben werden, die neben einem Raum- und Zeitbezug eine fachlich-inhaltliche Einordnung erfahren und eine Dynamik darstellen können.²⁶ Diese dienen als Abstraktionsmöglichkeit von beobachtbaren Objekten, wobei hier eine Zusammenfassung von mehreren Geoobjekten zu einem aggregierten Geoobjekt vorgenommen werden kann.²⁷ Digitale Geodaten bilden eine Querschnittsressource, da diese in unterschiedlichen fachlichen Bereichen eingesetzt werden und somit keine isolierte Verwendungsmöglichkeit aufweisen. Einerseits kann im Bereich der öffentlichen Daseinsvorsorge die Erhebung, Haltung und Nutzung von Geodaten selbst Inhalt der Verwaltungstätigkeit sein.²⁸ Andererseits stellt die Nutzung digitaler Geodaten auch eine Möglichkeit zur originären Erfüllung von Verwaltungsaufgaben dar, wobei diese als unterstützende Arbeitsgrundlage dienen.²⁹

²⁵ Das räumliche Referenzieren kann hierbei aus feldbasierten Daten bestehen (z.B. Luftbilder) oder amtlichen Geobasisdaten.

²⁶ Vgl. Bernard und Mäs (2020), S. 102.

²⁷ Vgl. ebd., S. 102. So kann eine Straße als Geoobjekt bezeichnet werden, während dessen Zusammenfassung zu einem Straßennetz ebenfalls ein Geoobjekt darstellt.

²⁸ Vgl. Neumann (2019), S. 537.

²⁹ Vgl. ebd., S. 537. Hauptsächliche Aufgabenfelder können indes beispielsweise das Vermessungswesen, den Umweltbereich, die Gefahrenabwehr und das Meldewesen umfassen.

Im Kontext des zuvor definierten Modells nach MOON und WELCH (2015) können digitale Geodaten nicht nur als Bestandteile einzelner Stufen gesehen, sondern auch als Treiber einer fortschreitenden Digitalisierung in der öffentlichen Verwaltung betrachtet werden – schlicht, weil deren analoge Verarbeitung zu aufwändig ist. Das daraus resultierende *Geo-Government* bietet demnach die Möglichkeit, durch die Verknüpfung von Basis- und Fachdaten Leistungsprozesse in der öffentlichen Verwaltung zu optimieren und sie den Nutzern zur Verfügung zu stellen. Die Visualisierung über Geoinformationssysteme (GIS) bietet zugleich die Möglichkeit, Daten zu suchen, darzustellen, zu transformieren und mit weiteren (Fach-)Daten zu verschneiden, wodurch das *Geo-Government* auch für nicht originäre Geo-Datensätze nutzbar gemacht werden kann.³⁰

Vor dem Hintergrund der angespannten Fachkräftesituation, welche auch die kommunale Ebene in den Blick zu nehmen hat, kann die Einbeziehung digitaler Geodaten und die damit verbundene Nutzung von GIS-Anwendungen zu einer Verschlinkung von Verwaltungsprozessen führen, die personellen Ressourcen schont. Insbesondere bei repetitiven Prozessen, die wiederkehrend personelle Belastungen erzeugen, kann *Geo-Government* Entlastung schaffen.³¹ Zusätzlich könnten Akteure durch den unmittelbaren Bezug von Geodaten und deren korrespondierenden Anwendungen Anfragen direkt über ein mögliches GIS einbinden, wodurch Bürger, Wirtschaft und Wissenschaft direkt auf Daten zugreifen können, ohne dass ein Mehraufwand in Form einer zusätzlichen Anfrage in der Verwaltung eingeht. So visualisiert beispielsweise die Stadt Leipzig ihre Investitionsvorhaben über einen Kartendienst und bietet Interessenten die Möglichkeit, geplante Investitionsvolumina transparent zu ermitteln.³² Korrespondierend zu dem evolutorischen Stufenmodell nach MOON und WELCH (2015) kann das *Geo-Government* zu einer erweiterten vertikalen und horizontalen Integration von Leistungsprozessen führen. Das *Geo-Government* dient somit selbst als Treiber einer fortschreitenden Verwaltungsdigitalisierung in Bezug auf die Nutzung von Geodaten und kann somit endogen auf die staatlichen Hierarchieebenen einwirken.

2.4 Die INSPIRE-Richtlinie als Treiberin des Geo-Governments

Zum Aufbau einer europäischen Geodateninfrastruktur ist die INSPIRE Richtlinie 2007 eingeführt und in den Folgejahren durch die Mitgliedsstaaten in nationales Recht umgesetzt worden – in Deutschland durch die Geoinformationsgesetze des Bundes und der Länder.³³ Die Koordinierungsstelle Geodateninfrastruktur Deutschland (GDI-DE) wird von Bund und Ländern betrieben, sowie finanziert und unterstützt die GDI-Aktivitäten der Kommunen.³⁴ Bund und Länder koordinieren demnach ihre Bestrebungen zum Aufbau einer einheitlichen Geodateninfrastruktur, was durch eine

³⁰ Vgl. ebd., S. 559 f.

³¹ Vgl. Jaenicke (2007), S. 66.

³² Vgl. Stadt Leipzig (2021).

³³ Vgl. Koordinierungsstelle Geodateninfrastruktur Deutschland (2019), S. 13.

³⁴ Vgl. ebd., S. 13.

Verwaltungsvereinbarung zwischen beiden Ebenen zudem rechtlich abgesichert ist.³⁵ Die Nationale E-Government Strategie (NEGS) beschreibt dabei die Ziele und Handlungsfelder für das E-Government in Deutschland und konzentriert sich auf die Anwendung und Gestaltung der Informationstechnik im Aufgabenbereich der öffentlichen Hand.³⁶ Über das Onlinezugangsgesetz (OZG) sollen bis Ende 2022 zudem 575 Verwaltungsdienstleistungen auch in digitaler Form angeboten werden. Überdies sind Bund und Länder verpflichtet, ihre Verwaltungsportale miteinander zu einem Portalverbund zu verknüpfen. Das korrespondierende OZG verpflichtet Bund, Länder und Kommunen hierzu und wird durch einen Umsetzungskatalog aller Verwaltungsdienstleistungen, welcher jedoch nicht Teil des originären Gesetzes ist, ergänzt.³⁷ Die Nutzung von Geodateninfrastrukturen, beispielsweise über GIS-Anwendungen, stellt dabei einen wesentlichen Baustein dar, sodass digitale Verwaltungsprozesse entstehen können, die zu einer Simplifizierung und Verkettung von Verwaltungshandeln führen.

Bevor auf die Ergebnisse der Experteninterviews im Rahmen des SARDINE-Projektes näher eingegangen wird, wird zunächst auf die Umsetzung der INSPIRE-Richtlinie im Freistaat Sachsen abgestellt. Hierbei werden zudem die Gebiets- und Funktionalreformen im Freistaat mit betrachtet, die beachtliche Auswirkungen auf die Aufgabenwahrnehmung der sächsischen Kommunen haben. Dies kann letztlich als Grundlage für die heutige Nutzung von Geodaten in der sächsischen öffentlichen Verwaltung gesehen werden und bildet somit die Ausgangslage für heutige endogene Entwicklungslinien.

3 Fallbeispiel Geodaten: Der Freistaat Sachsen

3.1 Geobasisdaten: Zwischen der Umsetzung der INSPIRE-Richtlinie und der Gebiets- und Funktionalreform 2008

Die Umsetzung der INSPIRE-Richtlinie in nationales Recht erfolgte im Freistaat Sachsen durch die Verabschiedung des sächsischen Geodateninfrastrukturgesetzes (SächsGDIG), wodurch die gehaltenden Stellen zur Durchführung bestimmter technischer und organisatorischer Maßnahmen verpflichtet sind.³⁸ Sachsen hat zur Umsetzung der INSPIRE Vorgaben ein System der zentralen Komponenten eingeführt, welches sich von der Systematik anderer Bundesländer unterscheidet. Nach INSPIRE sind die gehaltenden Stellen dafür zuständig, Informationen bereitzustellen. Dies schließt die Pflicht mit ein, Daten mit Metadaten zu beschreiben, zu standardisieren und darzustellen. Dies trifft im weiteren Sinne auf alle Kommunen und jede Landesbehörde zu. Selbst bei einer weitestgehenden Standardisierung wäre nicht ausgeschlossen, dass hierdurch

³⁵ Vgl. Verwaltungsvereinbarung GDI-DE.

³⁶ Vgl. Koordinierungsstelle Geodateninfrastrukturen Deutschland (2019), S. 80.

³⁷ Vgl. IT-Planungsrat (2021).

³⁸ Staatsbetrieb Geobasisinformation und Vermessung (2017), S. 3.

Doppelstrukturen aufgebaut und Ineffizienzen geschaffen werden würden. Deshalb hat der Freistaat eine zentrale Lösung favorisiert.

Während das Staatsministerium für Regionalentwicklung (SMR) organisatorisch zuständig ist, nimmt der Staatsbetrieb Geobasisinformation und Vermessung (GeoSN) unter anderem die Aufgaben wahr, die technische Funktionsfähigkeit der GDI Sachsen zu gewährleisten, ein landesweites Metadateninformationssystem und dezentrale Netzdienste für die gehaltenden Stellen bereitzustellen.³⁹ Dabei stellt der GeoSN zentral die Geobasisdaten bereit (GeoBAK) und bietet somit eine zentrale E-Government Applikation zur Nutzung von sächsischen Geobasisdaten.

Seit 1990 erfolgten im Freistaat Sachsen drei Gebietsreformen: die erste Kreisgebietsreform (1994/1996), die Gemeindegebietsreform (1998) sowie die zweigliedrige Verwaltungsreform (2008), bestehend aus der zweiten Kreisgebietsreform und einer Funktionalreform. Im Ergebnis (und begleitet von fortlaufenden Gemeindezusammenschlüssen und Eingemeindungen) sank die Zahl von 48 Landkreisen, sechs Kreisfreien Städten und 1 626 Gemeinden im Jahr 1990 auf zehn Landkreise, drei Kreisfreie Städte und 419 Gemeinden zum 1. Januar 2021.

Die im Jahr 2008 durchgeführte Funktionalreform hatte zum Ziel, strukturelle Aspekte der Aufgabenwahrnehmung zu ordnen. Dies umfasste insbesondere eine stärkere Kommunalisierung der öffentlichen Aufgaben sowie die Bündelung und Konzentration im innerstaatlichen Bereich. Für die weitere Betrachtung der Geodaten in der öffentlichen Verwaltung ist hierbei relevant, dass das Sächsische Verwaltungsneordnungsgesetz (SächsVwNG) eine Neuordnung der Landesvermessung, des Liegenschaftskatasters sowie amtlicher Geobasisinformationen vorsah. Letztlich drückte sich dies im neu gefassten Sächsischen Vermessungs- und Geobasisinformationsgesetz (SächsVermGeoG) aus. Die Aufgaben des Vermessungswesens nehmen seither die oberste Vermessungsbehörde (Staatsministerium des Inneren), die obere Vermessungsbehörde (Staatsbetrieb Geobasisinformationen und Vermessung Sachsen) und die unteren Vermessungsbehörden (Landkreise und Kreisfreie Städte) wahr.⁴⁰ Die obere Vermessungsbehörde ist zuständig für die Haltung und Bereitstellung von Daten des Liegenschaftskatasters, die Haltung und Bereitstellung von Daten des Botenrichtwertinformationssystems sowie die Einrichtung und den Betrieb von Geodiensten. Die unteren Vermessungsbehörden sind für die Fortführung der Daten des Liegenschaftskatasters ihres (Kreis-)Gebiets und deren Bereitstellung zuständig. Somit folgt die im Rahmen der Verwaltungsreform 2008 gefasste SächsVermGeoG mehrheitlich dem Allokationsziel einer möglichst hohen Zentralität der Aufgabenerbringung im Rahmen der Haltung und Bereitstellung von Geobasisdaten. Lediglich die Erfassung verbleibt in der dezentralen Zuständigkeit.

³⁹ Vgl. GDI Sachsen (2021).

⁴⁰ Siehe §2 Abs. 1 SächsVermGeoG.

3.2 Georeferenzierte Fachdaten: Digitalisierung der Verwaltung und die Verwendung georeferenzierter Fachdaten

Neben der Erhebung und Bereitstellung von Geobasisdaten ist die Verwendung georeferenzierter Fachdaten im Rahmen der öffentlichen Leistungserbringung in den letzten Jahren stetig gewachsen. Im Rahmen des eigenen Forschungsprojektes Smart Regional Development Infrastructure (SARDINE) sind 15 teilstandardisierte Experteninterviews⁴¹ mit Vertretern aus der Wissenschaft, der Wirtschaft sowie mit ausgewählten Vertretern auf Landes- und Kommunalebene im Freistaat Sachsen durchgeführt worden. Erste Ergebnisse deuten darauf hin, dass Gebietskörperschaften, welche Geobasisdaten fortschreiben (Landkreise und kreisfreie Städte), auch über umfangreiche georeferenzierte Fachdaten in ihrer Verwaltung verfügen. Zudem werden diese kontinuierlich um weitere Fachdaten ergänzt. Im Gegensatz dazu weisen kleinere kreisangehörige Gemeinden eine geringere Durchdringung von georeferenzierten Fachdaten auf – in Teilen existiert diese Form der Datenanwendung in einzelnen (Fach-)Verwaltungen noch nicht. Dieses Defizit kann zu einer ineffizienteren Leistungserbringung führen, was sich laut den Befragten in höheren Aufwendungen oder geringerer Qualität der erbrachten Verwaltungsleistungen ausdrückt. Ein entsprechender Lösungsansatz liegt in der zielgerichteten Forcierung des *Geo-Governments*, welches im Zuge der horizontalen und vertikalen Leistungsverknüpfung auch der kommunalen Ebene als Nutzer zugänglich gemacht werden muss.

Eine wesentliche Grundlage für die Etablierung eines *Geo-Governments* ist es, dass Daten harmonisiert, Schnittstellen geschaffen, die Benutzerfreundlichkeit von Daten und Anwendungen hergestellt, sowie Systeme weitestgehend interoperabel aufgesetzt werden. Hierin liegen jedoch zugleich Hindernisse, die einen weiterführenden Einbezug digitaler Geodaten erschweren können. Zwar existieren Standardisierungsbemühungen, wie etwa OGC-Standards, die ebenfalls laufend weiterentwickelt werden. Jedoch zeigt sich nach den Angaben der Befragten in der Praxis, dass diese oftmals den real existierenden Bedarfen mit einer gewissen Zeitverzögerung begegnen. Auch wenn Standards in einem Top-down-Prozess verordnet werden, ist deren Umsetzung zeitlich verzögert. Ursächlich hierfür sind limitierende Faktoren, wie etwa die personellen und budgetären Kapazitäten auf kommunaler Ebene. Dies hat nicht zuletzt einen negativen Einfluss auf die Datenaktualität, da der zugrunde liegende Standard unter Umständen veraltet ist. Auch wenn zudem durch das OZG vorgegeben ist, dass Verwaltungsdienstleistungen zu digitalisieren seien, ist dennoch fraglich, in welcher Intensität dies vonstattengeht. So ist beispielsweise aufgrund des kurzen Umsetzungszeitraumes nach überwiegender Einschätzung der Befragten davon auszugehen, dass eine große Anzahl von Kommunen zwar Dienstleistungen im Bereich des ‚Front-Office‘ digital anbieten werden, der zugrunde liegende interne Verwaltungsablauf des ‚Backoffice‘ jedoch nach wie vor zum Teil analog abläuft. Auch hier können GIS-Anwendungen bei der Systematisierung und Implementierung digitaler interner Leistungserstellungsprozesse nach Auffassung der

⁴¹ Der Befragungszeitraum umfasst das vierte Quartal 2020 und das erste Quartal 2021.

Experten hilfreich sein. Die Aufsetzung stellt Kommunen jedoch vor finanzielle, personelle und zeitliche Herausforderungen. Um eine Interoperabilität gewährleisten zu können, ist zudem von Interesse, in welcher Güte und Frequenz Daten erhoben werden und in welcher GIS-Anwendung diese bereitgestellt werden. So zeigt sich, dass durch die Nutzung unterschiedlicher GIS-Systeme eine geodatenbasierte interkommunale Kooperation erschwert wird. Deshalb ist als Lösungsansatz die Aufsetzung von standardisierten Schnittstellen von Relevanz, um über gebietskörperschaftliche Grenzen hinweg agieren zu können. Dies ist nicht nur vor dem Hintergrund einer gemeinsamen Leistungserbringung, sondern auch im Hinblick auf standortpolitische Entscheidungen, zu sehen, die unter Umständen nicht lokal, sondern regional getroffen werden müssen.

Die bisherigen Ausführungen zur Verwaltungsdigitalisierung sowie zur Verwendung georeferenzierter Fachdaten der Verwaltung lassen sich verallgemeinern und auf andere Bundesländer übertragen.⁴² Die aktuelle Verwaltungsdigitalisierung gilt es als Organisationsreform zu verstehen, die für den Verwaltungsaufbau eine wiederkehrende Herausforderung darstellt. Die Verwaltungsdigitalisierung erfolgt zudem in den Gebietskörperschaften in unterschiedlicher Geschwindigkeit, was sich auf allgemeine und spezifische Herausforderungen zurückführen lässt. Als Katalysator für den aktuell ablaufenden Prozess sind neben INSPIRE im Speziellen die Umsetzung des OZG und die Einführung der E-Akte im Allgemeinen zu nennen.

Die Handlungsfähigkeit der öffentlichen Verwaltung wird im Wesentlichen durch die Größe der jeweiligen Verwaltungseinheit determiniert. Erst bei genügender dezentraler technischer Ausstattung und personellen Kapazitäten ist es möglich eine eigene professionelle Georeferenzierung von Fachdaten durchzuführen. Dem stehen Größen- und Verbundvorteile (economies of scale and scope) entgegen, die u. a. im Rahmen der Verwaltungsdigitalisierung zu heben sind. Erst ab einer kritischen Verwaltungsgröße bilden sich Kostenreduktionen aus (Stückkostendegression), sodass es zielführend ist, möglichst viele Leistungen zentral zu halten. Somit besteht ein klassischer Zielkonflikt zwischen (kosten-)effizienter Zentralisierung und (handlungs-)effizienter Dezentralität.

Im Rahmen des Forschungsprojekts SARDINE konnten folgende spezifische, geodatenbezogene Herausforderungen für den Freistaat Sachsen identifiziert werden, die es zu beachten gilt, um Interoperabilität zwischen den georeferenzierten Fachdaten herzustellen. Hier zu nennen sind die Standardisierung der Daten sowie die Datenqualität. In Bezug auf die Standardisierung gilt es technische Standards⁴³, wie die verwendete Software oder verwendete Koordinatensysteme⁴⁴, zu harmonisieren, um den

⁴² Vgl. Zern-Breuer et al. (2020).

⁴³ In einigen öffentlichen Aufgabenbereichen existieren bereits einheitliche Austauschstandards. Beispielfhaft zu nennen sind XPlanung, die einen verlustfreien Datenaustausch zwischen Bauleitplänen, Regionalplänen und Landschaftsplänen sicherstellen.

⁴⁴ Innerhalb der Kernverwaltung wird das von der GDI vorgegebene Koordinatensystem einheitlich genutzt. Allerdings gilt dies nicht für die Gesamtheit des öffentlichen Sektors. So weisen einige öffentliche Unternehmen andere Koordinatensysteme auf, wodurch der Datenaustausch regelmäßig mit Fehlern einhergeht.

Datenaustausch weniger fehleranfällig zu gestalten. Zudem ist es notwendig, die fachspezifischen Standards in Bezug auf die Georeferenzierung zu beachten. Geodaten sind als Teil der ressortübergreifenden und somit ressortunspezifischen Basisinfrastruktur der Verwaltung zu verwenden.

Die Datenqualität umfasst laut den Befragten insbesondere die Vermeidung von Messfehlern bei der Erhebung und beim Verschnitt sowie bei der Integration in ein gemeinsames Geoinformationssystem, wobei die Qualität der georeferenzierten Fachdaten den Ansprüchen aller potentiell nutzenden Fachressorts und nicht nur dem der erstellenden Einheit genügen muss. Insbesondere die Auflösung ist darunter zu fassen. Dies stellt wiederum hohe Anforderungen an die Standardisierungen der Datenerfassung, die jeweils deutlich über den konkreten Anwendungsfall hinausgehen dürften, um spätere, im Einzelnen noch unbekanntere Verknüpfungsmöglichkeiten mit weiteren Daten zu gewährleisten. In vielen Bereichen der öffentlichen Leistungserbringung ist es trotz bestehender Kompetenzverteilung zielführend, eine vertiefende Kooperation anzustreben, sei es, um die Leistung gemeinsam zu erbringen, oder um einen Daten- und Erfahrungsaustausch zu betreiben.

4 Fazit

Die vierte Phase der Verwendung von digitalen Geodaten in der Verwaltung wurde durch die EU-Richtlinie INSPIRE angestoßen. Im Rahmen der Umsetzung der Richtlinie konnten die Mitgliedstaaten im Rahmen des Aufbaus der GDI sowie der Umsetzung der 34 Geodaten-Themen eigene Schwerpunkte setzen – wodurch u. a. verschiedene Standards entstanden.

Die Georeferenzierung von Fachdaten ist ein weiterer Prozess, der durch die Einführung von INSPIRE einen deutlichen Schub erlebt hat. Hierbei ist die Anwendung von Geodaten im Rahmen der öffentlichen Leistungserbringung Teil der Verwaltungsdigitalisierung. Ziel ist es, ein *Geo-Government* zu implementieren, um Verwaltung und deren Leistungen effizient auszugestalten.

Die Umsetzung der Verwaltungsdigitalisierung im Bereich der Verwendung von Geodaten wird im Freistaat Sachsen durch allgemeine Herausforderungen, wie Handlungsfähigkeit und Zuständigkeit, erschwert, sowie durch spezifische Herausforderungen, wie Fragen zur Standardisierung und der Datenqualität. Dies wurde durch die Analyse von geführten Experteninterviews im Rahmen des Projektes SARDINE nochmals verdeutlicht. Eine zunehmende kommunale Kooperation in diesem Bereich, wie auch der vertiefte Erfahrungsaustausch der Kommunen untereinander kann demnach dabei helfen, die Herausforderungen, vor denen die kommunale Ebene insgesamt steht, zu bewältigen.


5 Literaturverzeichnis

- [AH06] Andersen, K. V.; Henriksen, H. Z.: E-Government maturity models: Extension of the Layne and Lee model. *Government Information Quarterly*, 23(2), 236–248, 2006.
- [BM20] Bernard, L.; Mäs, S: Digitale Geodaten, in: Klenk, T.; Nullmeier, F.; Wewer, G. (Hrsg.): *Handbuch Digitalisierung in Staat und Verwaltung*, 1. Auflage, Springer-Verlag Wiesbaden, Seiten: 101-110, 2020.
- [Ri07] Das Europäische Parlament und der Rat der Europäischen Union: Richtlinie 2007/2/EG des Europäischen Parlaments und des Rates vom 14. März 2007 zur Schaffung einer Geodateninfrastruktur in der Europäischen Gemeinschaft (INSPIRE), in: *Amtsblatt der Europäischen Union* (Hrsg.): L 108/1 vom 25. April 2007.
- [Ri19] Das Europäische Parlament und der Rat der Europäischen Union: Richtlinie (EU) 2019/1024 des Europäischen Parlaments und des Rates vom 20. Juni 2019 über offene Daten und die Weiterverwendung von Informationen des öffentlichen Sektors, in: *Amtsblatt der Europäischen Union* (Hrsg.): L 172/56 vom 26. Juni 2019.
- [MT20] de Mello, L.; Ter-Minassian, T.: *Digitalisation. Challenges and Opportunities for Subnational Governments*, OECD Working Papers on Fiscal Federalism, 2020.
- [GD21] GDI Sachsen, Geodateninfrastruktur Sachsen, <http://www.gdi.sachsen.de/inhalt/devgdi/gdisn/gdisn.html>, Stand: 15.04.2021.
- [HRU21] Huang, S.; Ribers, M. A.; Ullrich, H.: Der gesellschaftliche Mehrwert verknüpfter Daten: Algorithmen als Entscheidungshilfen bei Antibiotikaverschreibungen, in: *DIW Wochenbericht*, Nr.13 und 14 (88), S. 240-246, 2021.
- [IT21] IT-Planungsrat, OZG-Umsetzungskatalog, https://www.it-planungsrat.de/DE/ITPlanungsrat/OZG-Umsetzung/Digitalisierungsprogramm/06_DigPro_OZG_Katalog/DigPro_OZG_Katalog_node.html;jsessionid=F984272028317BDE7B7B007342B50EBF.2_cid322, Stand: 15.04.2021.
- [Jk07] Jaenicke, K.: *Verfahren zur Berechnung der Wirtschaftlichkeit GIS-unterstützter Prozesse*, Dissertation zur Erlangung des Doktors der Naturwissenschaften, Technische Universität Münster, Münster, 2007.
- [Ge19] Koordinierungsstelle Geodateninfrastruktur Deutschland: *Geodatendienste im Internet. Ein Leitfaden*, 4. Auflage, Frankfurt am Main, 2019.
- [Lt19] Lakes, T.: Geodaten, in: Baur, Nina/Blasius, Jörg (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung*, 2. Auflage, S. 1345-1352, 2019.
- [MW15] Moon, M. J.; Welch, E. W.: *Managing e-government*, J. L. Perry & R. K. Christensen (Hrsg.), *Handbook of public administration* (3. Aufl., S. 436–455), San Francisco: Wiley, 2015.
- [Nc19] Neumann, C.: Geodaten – Anwendungsfeld und Pionier des E-Government, in: *Digitale Verwaltung. Vernetztes E-Government*, 2. Auflage, Erlich Schmidt Verlag, Berlin, Seiten: 535-564, 2019.

- [RT19] Ritz, A.; Thom, N.: Public Management. Erfolgreiche Steuerung öffentlicher Organisationen, 6. Auflage, Springer-Verlag, Wiesbaden, 2019.
- [St21] Stadt Leipzig, Investitionen der Stadt Leipzig, https://haushalt.leipzig.de/de/uebersicht_investitionen.asp?mm=8, Stand: 29.04.2021.
- [Sä16] Sächsisches Geodateninfrastrukturgesetz vom 19. Mai 2010 (SächsGVBl. S. 134), das durch Artikel 2 des Gesetzes vom 26. Oktober 2016 (SächsGVBl. S. 507) geändert worden ist, 2016.
- [St17] Staatsbetrieb Geobasisinformation und Vermessung: Umsetzung INSPIRE im Freistaat Sachsen (SAX4Inspire), Dresden, 2017.
- [Ve17] Vereinbarung zwischen dem Bund und den Ländern zum gemeinsamen Ausbau und Betrieb der Geodateninfrastruktur Deutschland (Verwaltungsvereinbarung GDI-DE), 2017.
- [Zr20] Zern-Breuer, R.; Seckelmann, M.; Regös, N.; Lorei, H.; Kruse, K. A.; Brunzel, M.: Voruntersuchung zur Einführung eines einheitlichen Geodatenmanagements in Rheinland-Pfalz, Speyrer Arbeitshefte Nr. 245, Speyer, 2020.
- [Zh19] Zimmermann, H.: Fiskalföderalismus – Vor- und Nachteile dezentraler Entscheidungsstrukturen – eine finanzwissenschaftliche Ableitung, in: Schweisfurth, T.,; Wallmann, W. (Hrsg.): Haushalts- und Finanzwirtschaft der Kommunen in der Bundesrepublik Deutschland, Schriften zur öffentlichen Verwaltung und öffentlichen Wirtschaft, Band 242, Seiten: 29-47, 2019.

Utilization and analysis of digital mining dump models

A case study from German lignite mining

Natalie Merkel ¹, André John  ²

Abstract: This extended abstract describes the main features of the planned activities to support decision-makers in the transformation process of converting mining areas into attractive post-mining landscapes.

Keywords: Trim4Post-Mining, mining waste, lignite, spatio-temporal data analytics, cluster analysis, dynamical processes, risk assessment

1 Introduction

The decision to phase out fossil energy production in Europe will ensure that most of the coal mines in Europe will go into the closure phase over the next two decades. This means that large areas that have previously been used for coal mining will have to be renewed and restored in the future, to successfully master the structural transformation from industrial to attractive post-mining landscapes in the European coal regions. The multi-partner research project ‘Trim4Post-Mining’ funded by EU Research Fund for Coal and Steel (RCFS / funding code 899278) addresses the challenges and possibilities associated with the coal phase-out. In one of the work packages in the project, data-driven spatio-temporal analysis and prediction methods for inside-dump deformations and geochemical processes in lignite or hard coal waste dumps will be developed to support decision making and planning during the transition from coal exploitation to a re-vitalized post-mining land. The following research work can be classified in this context.

2 Research Objectives

The current research regarding the decommissioning and renaturation of mining areas is focused on the analysis and utilization of digital mining waste dump models. The main objective is thereby to characterize the dynamical processes in overburden dumps related to timely dependent deformations and geochemical processes because of the inhomogeneous material composition in the dump. This can be used, for example, to

¹ TU Bergakademie Freiberg, Institute of Mine Surveying and Geodesy, Fuchsmuehlenweg 9 B, 09599 Freiberg, natalie.merkel1@mabb.tu-freiberg.de

² TU Bergakademie Freiberg, Institute of Mine Surveying and Geodesy, Fuchsmuehlenweg 9 B, 09599 Freiberg, andre.john@mabb.tu-freiberg.de, <https://orcid.org/0000-0003-0187-5071>

identify potential contamination areas or to plan new usage concepts for former mining areas in terms of residual risks, technical feasibility and environmental impact.

To estimate the potential post-use or development possibilities of a dump site a geotechnical hazard assessment is necessary. Settlement, slumping and liquefaction are considered as possible deformation scenarios in the following work. The first two damage mechanisms differ essentially in their cause. Whereas slumping can always occur as a result of a change in groundwater level, settlement occurs as a result of an external stress input or just due to the material's weight. Both damage scenarios are characterized in its occurrence as a vertical displacement, i.e. uplift or subsidence, at the surface. The third damage mechanism, liquefaction, is a complex process in which the soil spontaneously or also due to an initial event takes on characteristics of a liquid and results in large vertical and horizontal displacements.

3 Data and Methods

The central German lignite mining area is one of the example regions within the Trim4Post-Mining project and also provides the data basis for the current research. The primary data basis is an already available, operational digital dump model based on a long-term monitoring. This theoretical voxel model of the dump defined by the bounding box consists of approximately 740.5 million voxel cells ($n_x=2057$, $n_y=2011$, $n_z=179$) with a size of 1 m^3 ($1\text{ m} \times 1\text{ m} \times 1\text{ m}$) each. However, not all cells actually contain information. A total of approximately 72.5 million voxels of the model contain information relating to the following attributes:

- Geometry (x, y, z - coordinates)
- Date and time of interstratification into the dump
- Allocation of the voxel volume to the 28 defined material classes (geological layers) of the model by specifying the percentage of the volume.

That means the present dump model of the Central German lignite company (Mibrag mbH) gives a statement about the locality and the percentage of the different materials in the dump, but not about their properties. In order to be able to make a soil mechanical estimation, further parameters are needed. A statement about the exact grain distribution of the materials as well as their coefficient of uniformity and grain shape can be determined in the laboratory. For other parameters such as compactness, water content and degree of saturation, field tests are required. With the help of these parameters, the materials can be then divided into hazard classes that represent the probability with which settlement, slumping or liquefaction can occur due to the position of the materials in the dump. For this purpose, methods from the area of machine learning, specifically, methods of cluster analysis, i.e. unsupervised learning, are to be used. Clustering is the task of grouping a set of objects in such a way, that objects in the same group are more similar to each other than to those in other groups. Whereby the groups are also referred to as clusters. Basically, clustering can be formulated as a multi-objective optimization problem. The selection of a specific algorithm and the associated parameters, such as the

used distance function or the number of expected clusters, depend always on the individual data set and the intended objectives of the clustering. There is no single best clustering algorithm for all use cases, hence it is an iterative process to find an appropriate algorithm and its parameter. Due to the size of the model to be processed, the efficiency of an algorithm is a decisive selection criterion. The aim of the current work is to develop a cluster model that can independently create a hazard analysis for the damage scenarios mentioned on the basis of a dump model extended by geotechnical soil parameters. The representation of the clusters or respectively hazard classes could then be done for the respective damage scenario for example with the help of a geographic information system (GIS). A kind of traffic light map would be conceivable, in which the endangered areas are recognizable in the three-dimensional space and thus also clarify conclusions about their damage effect on the surface. And in a further development, an extension to design recommendations for structures subdivided according to different types of use would be conceivable.

Figure 1 shows a visualization of the described dump model. Exemplary the clustering of material groups in the dump model based only on the geometry, the date of the interstratification and the percentage material composition in each voxel, is presented. As a common pre-processing step all attributes were standardized to a value range of $[0,1]$. The number of clusters was defined using elbow method and silhouette analysis. The k-means algorithm was only used as an example to visualize the initial dump model. The choice of the cluster algorithm to achieve the objectives described will be part of the further work.

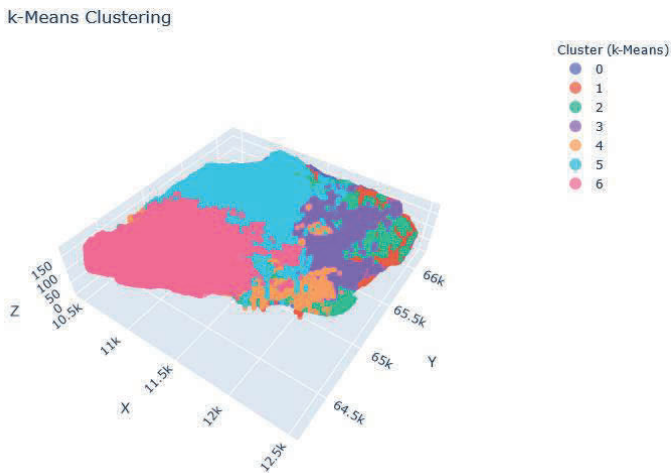


Figure 1: Example cluster result created by k-means clustering

Integrating External Data Sources into Internet of Things Architectures for Weather and Environmental Monitoring in Former Mining Areas

Jonah Windolph,¹ Robert Wehlitz,² Theo Zschörnig,³ Bogdan Franczyk⁴

Abstract: The management of former mining areas requires knowledge about local weather and environmental conditions in order to ensure the compliance with governmental regulations. Since mining areas are typically of large geographic scale, equipping them with sensors is a challenge at high cost. One approach to keep these low is the enrichment and in some cases substitution of sensor data utilizing external sources, such as weather services. Against this background, this paper describes requirements for a solution that allows the use of external data sources to enrich or substitute sensor data for historic and real-time analysis. An architecture that supports this approach is presented. The architecture is implemented as a proof of concept for the purpose of demonstrating its abilities by the use case of land use regulations.

Keywords: Internet of Things; System Architecture; Microservices; System Integration; External Data Sources

1 Introduction

Even with more and more countries deciding to phase out of coal-fired power generation in order to meet climate goals (cf. [En20]), the management of former mining areas remains an important task. It requires detailed knowledge of the environmental conditions at these sites in order to ensure the compliance with governmental regulations that are in place regularly due to the geographic conditions in these areas. Regulations often depend on the current environmental situation. For example, using an unpaved road with heavy equipment might be forbidden if the road is wet after it has rained. Internet of Things (IoT) approaches are able to support the management of former mining sites by providing real-time access to and analysis of sensor data. Past scientific research has mainly been focused on enabling access to devices or providing analytics or automation capabilities. Little attention has been

¹ Institute for Applied Informatics (InfAI), Competence Center Smart Energy Systems, Goerdelerring 9, 04109 Leipzig, Germany, windolph@infai.org

² Institute for Applied Informatics (InfAI), Competence Center Smart Energy Systems, Goerdelerring 9, 04109 Leipzig, Germany, wehlitz@infai.org

³ Institute for Applied Informatics (InfAI), Competence Center Smart Energy Systems, Goerdelerring 9, 04109 Leipzig, Germany, zschornig@infai.org

⁴ Leipzig University, Informations Systems Institute, Grimmaische Str. 12, 04109 Leipzig, Germany, franczyk@wifa.uni-leipzig.de; Wrocław University of Economics, Business Informatics Institute, ul. Komandorska 118-120, 53-345 Wrocław, Poland

given to integrating external data sources, such as public web services. The enrichment or substitution (if possible) of sensor data with data from external sources simplifies environmental monitoring of large areas immensely, since equipping them with sensors can be expensive (e. g. a single air quality sensor considered to be low cost still costs a few hundred euros (cf. [Al17])).

Therefore, the approach of integrating external data sources into IoT architectures is investigated in this paper. It is structured as follows: In the second section we describe related work. Requirements for a solution that enables the integration of external data sources are defined in the third section. In the fourth section we describe our proposed solution and demonstrate its abilities in section five. In the last section we summarize our findings and outline further research opportunities.

2 Related Work

In this section, preliminary work related to the topic of integrating web services into IoT architectures is presented.

Senožetnik et al. [Se18] describe an architecture they use to retrieve groundwater data from different sources for real-time queries and analysis. Their architecture consists of the components *retriever*, *watchdog*, *collector* and *API management*. *Retrievers* collect the data from their corresponding source and transform them into a unified format. The *watchdog* provides monitoring capabilities and ensures the restart of retrievers in case of errors. The *collector* collects the data from various retrievers and applies rule-based pre-processing before storing the data and making them available for querying. Finally, the *API management* allows for user and access management.

Kwon et al. [KOJ19] propose an IoT framework that is capable of handling sensor and externally available data. Measurements are collected by devices themselves or through data source-specific *external data collectors* that use the HTTP protocol to publish the data to a message broker. A *data collector manager* ensures error recovery by restarting failed data collectors. The collected data is saved in an OpenTSDB instance or further analyzed by *enrichers*, depending on a routing table in the message broker. All components run inside a Kubernetes cluster.

Heideker et al. [He20] introduce an IoT solution based on FIWARE to increase crop yield and decrease cost by optimizing water usage of agriculture in arid regions. In order to improve their computations they also include data from weather forecast services and historic records of agriculture yield by implementing custom mappers for the NGSI API, FIWARE uses. Terroso-Saenz et al. [Te19] follow a similar approach for energy consumption prediction of heating, ventilation and air conditioning systems based on sensor and weather data.

Luckner et al. [Lu20] propose an IoT architecture for a smart city IoT application. In order to integrate external data sources they implement custom data sources for Apache Flume. For less frequently updated data such as public transport timetables they use custom components that pull data from these sources regularly. Following the lambda architecture, their solution provides real-time analysis capabilities using a speed layer and analysis of historic data using a batch layer. Contrary to the lambda architecture they explicitly do not use a serving layer to make data available for downstream applications or users, but instead define a new layer they call *exposition layer* that provides application specific interfaces.

Reviewing the related literature shows that approaches for integrating external data sources into IoT architectures exist. However, existing approaches either require all sensors to send data via one specific protocol or rely on IT ecosystems like FIWARE that are very complex (cf. [We17]). Regarding analytics capabilities, previous approaches are based on the lambda architecture that requires two separate processing logics for historic and real-time analytics, resulting in higher complexity compared to the kappa architecture that only uses a single layer for analytics (cf. [St16]). Therefore, we present our own solution that eases integration of sensors (that may provide data via diverse protocols) and external data sources, does not rely on complex ecosystems and uses the kappa architecture for data analytics.

3 Solution Requirements

In this section, requirements for integrating external data sources into IoT architectures are defined as guidance for the implementation. These requirements arose in the ongoing research project Smart Regional Development Infrastructure (SARDINE), in which a novel data platform for renaturation projects in former mining areas is developed.⁵ Part of the project is the automated detection of weather conditions that affect governmental regulations. Since no weather sensors are deployed in the field at the time, integrating data from external sources is required. Our requirements are defined as follows:

- (R1): *Scalability*: With a lot of countries increasing their efforts to make data publicly available (cf. [Wo18]), the solution needs to be able to scale with both the amount of external data that could be integrated and an increasing number of potential users (i. e. when the data is meant to be publicly accessible).
- (R2): *Expandability*: Solutions should provide a way to integrate new data sources. This is necessary because external data sources are typically focused on a specific kind of data (e. g. weather data), but the management of former mining areas requires knowledge of diverse conditions. In order to be able to add support for new data sources, developers require assistance (e. g. in the form of a software development kit (SDK)).

⁵ <https://infai.org/sardine/>, accessed 12.04.2021

- (R3): *Error recovery*: Solutions should provide an ability to resume importing data after a failure. External data sources like weather services often provide huge data sets (e. g. historic records of weather data for whole countries). If the integration of external data sources fails (e. g. due to a hardware fault or an unforeseen programming error) reloading all data unnecessarily stresses the external data sources and own computation resources. Thus, a way to recover from errors is required.
- (R4): *Configurability*: The inclusion of external data sources should be reconfigurable in order to meet user requirements. Integrating external data sources in their entirety might not be required for the management of former mining areas. Therefore, it should be possible to configure integration processes in order to only integrate relevant data.
- (R5): *User friendly management*: Users without deep technical knowledge should be able to control integration processes and use configuration capabilities as described in (R4).
- (R6): *Sharing capabilities & access management*: After providing a way to integrate an external data source, users should be able to allow other users to use their implementation in order to avoid multiple implementations for the same data source. On the other hand, users might want to specifically use an integration for themselves (e. g. when private data is involved). Thus, fine-grained access management is required.
- (R7): *Interoperability*: Integrated data from external data sources should be usable in the same ways as device data. Both data types need to be used jointly. We conceive this to be a key requirement, because it allows to flexibly vary the level of substitution of sensors with data from external sources (e. g. an external data source needs to be replaced with a sensor for more detailed monitoring or a sensor needs to be substituted in order to decrease cost).

4 Solution Proposal

In this section, we describe an IoT architecture that supports the integration of external data sources. The architecture is shown in Fig. 1. It is based on the architecture previously described in Zschörnig et al. [Zs18] that has been evaluated in Zschörnig et al. [Zs20a] and Zschörnig et al. [Zs20b]. *Sensors* are connected to an *inbound data broker* that receives sensor measurements. If sensors are unable to connect to the inbound data broker themselves, a *device connector* is used to ensure connectivity of the device. This is typically required for sensors in wireless networks that are not directly connected to the internet, such as Zigbee or Z-Wave. *Platform connectors* use the stored metadata from the *device management* to consume the raw sensor data, transform them into a platform-wide unified format and produce the transformed data in a *streaming platform* that distributes the data to further system components, e.g. analytics or storage services. Besides holding metadata of

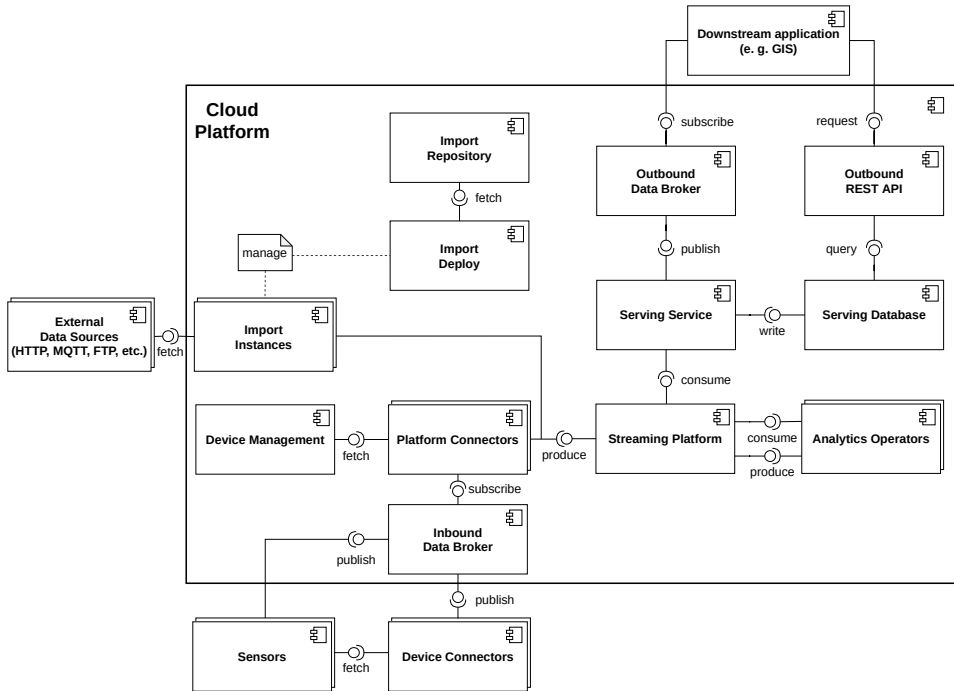


Fig. 1: Proposed architecture

different device types, the device management also provides information of which devices are currently connected to the platform and offers sharing capabilities among users (i. e. allowing other users to access sensor data). *Analytics operators* connect to the streaming platform as well in order to receive sensor data and perform analytics tasks. Analytics operators can be composed into complex analytics pipelines and are deployed automatically on user request. The results are then pushed back to the streaming platform again. Since the architecture is based on the kappa architecture, no additional analytics layer for batch processing is required. If necessary, historical data can be played back from the streaming platform. The *servicing service* consumes messages from the streaming platform in order to provide real-time access via the *outbound message broker* or to store the data in a *servicing database*. The *outbound REST API* is used to access historic data from the servicing database. *Downstream applications* (e. g. a geographic information system (GIS)) can consume data from the outbound data broker or the outbound REST API. The individual components offer REST endpoints that can be used comfortably via an overarching *web user interface*.

External data sources may provide the data in various different formats and make them available via a variety of different protocols. Therefore, integrating these data sources requires data source-specific implementations that we refer to as *imports*.

The *import repository* is responsible for storing metadata of implemented imports. It holds information on how an import can be instantiated, in which ways it can be configured and which data it provides. It also grants sharing capabilities among users with fine-grained access management mechanisms (R6). The *import deploy* component is responsible for managing instances of imports stored in the import repository. It provides an interface with operations to create, read, update and delete (CRUD) import instances.

Imports can be packaged using container technology in order to simplify deployment (R5) and scalability (R1). Container management software usually allows for automatic restarts of failed containers, thus fulfilling the requirement of error recovery (R3). Following the microservice principle, the new components are maintained as loosely-coupled applications. We opted for this principle because it eases scalability (cf. [Dr17]) and provides great flexibility in terms of development and deployment (cf. [Ch18]). Both, the import repository and import deploy component provide interfaces that can be integrated in the overarching web user interface for user-friendly import management (R5).

Developer support is supplied by a library that eases the implementation of new imports, simplifying expandability (R2). The library handles the publication of the data on the streaming platform in the same platform-wide unified format that is already in use for sensor data. It also parses user supplied configuration (R4) and improves error recovery capabilities (R3) by allowing access to already imported data, hence preventing possible duplication. After the data is published on the streaming platform it can be handled in the same way as data from sensors since it uses the same format. Thus, it is possible to use external data sources in analytics operators or to make it available for downstream applications (R7).

5 Demonstration

In former mining areas, land use is typically regulated due to the geographical circumstances in these areas. For example, driving on an unpaved road with heavy equipment (such as excavators) might only be permitted, if it has not rained in the last 24 hours. Since installing and maintaining weather sensors might be expensive, using data from a public weather service is a more practicable approach.

In order to check precipitation in the area of question, we implemented the necessary system components (cf. Fig. 1) as a proof of concept.⁶ Both, the import repository and import deployment component are implemented in Go⁷ and use a MongoDB⁸ instance as persistent data storage. The import deployment component provides the corresponding Docker image

⁶ All source code is available at <https://github.com/senergy-platform>

⁷ <https://golang.org/>, accessed 26.03.2021

⁸ <https://www.mongodb.com/>, accessed 26.03.2021

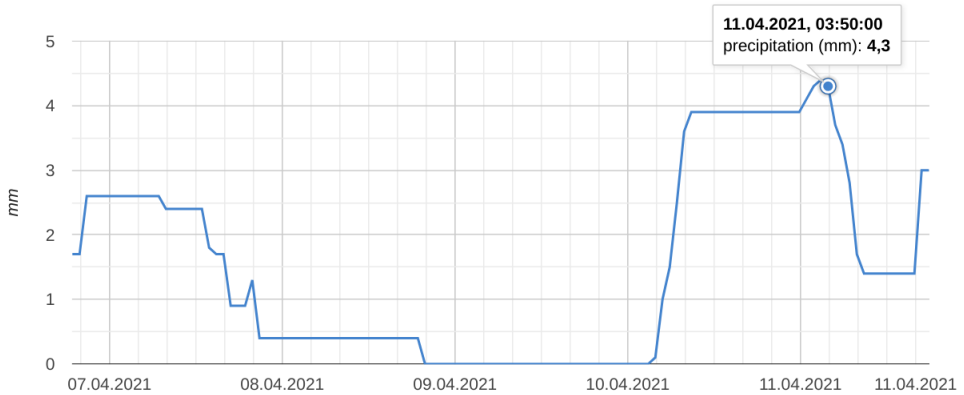


Fig. 2: Integrated precipitation data (24 hour sum) visualized in the web user interface

of an import either directly to a Docker⁹ host or to a Rancher¹⁰/Kubernetes¹¹ cluster. Using the import library (developed in python), an import for precipitation data from the German weather service was implemented. The German weather service provides precipitation data with a high geographic resolution of 1 km² by combining ombrometers with radar data (cf. [DW20]). Their service offers aggregated precipitation amounts over different time intervals. For this specific use case 24 hour precipitation sums were chosen that are updated hourly. The implemented import can be configured by supplying areas of interest. The import is registered at the import repository and deployed with the import deploy component using the area in question. Using the serving service, access to the imported data via the outbound REST API is enabled, the web user interface is used to visualize the imported data. As shown in Fig. 2, using the unpaved road with heavy equipment in a former mining area south of Leipzig (Germany) on 11th April 2021 is not permitted, since it rained less than 24 hours earlier.

Furthermore, the integrated data can also be used by other downstream applications. Fig. 3 shows how a GIS uses the precipitation data as an additional data layer. In this example, weather-dependent land use regulations are highlighted in green while the current precipitation levels are shown in shades of blue, hence allowing for real-time evaluation of regulations.

This demonstration shows that our solution is capable of providing current environmental data in order to evaluate the status of governmental regulations. Since the solution allows for expandability (R2), it could be used in other scenarios that require environmental data. The example of land use regulations based on weather conditions illustrates the relevance of the subject matter.

⁹ <https://www.docker.com/>

¹⁰ <https://rancher.com/>, accessed 26.03.2021

¹¹ <https://kubernetes.io/>, accessed 06.04.2021

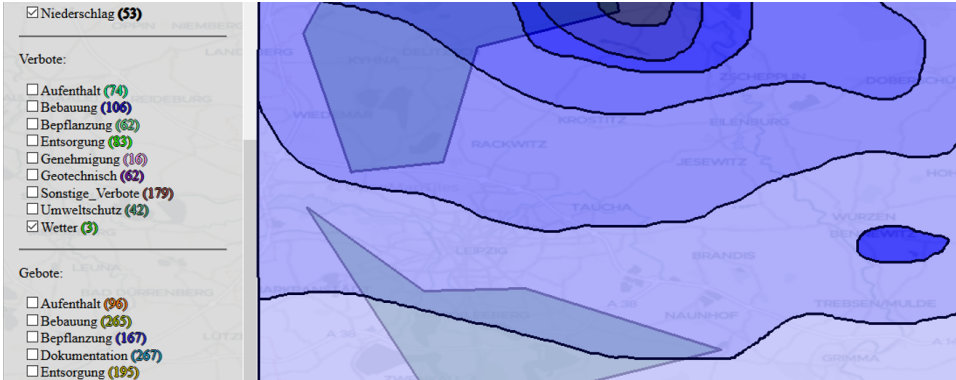


Fig. 3: A GIS displaying weather based regulations and current precipitation levels

6 Conclusion

In this paper, we described how integrating external data sources into IoT architectures supports weather and environmental monitoring in former mining areas. We defined a list of requirements a solution should be capable of in this context and proposed an architecture that tackles this goal. We implemented the necessary system components as open source software following the microservice approach and demonstrated the abilities of our solution in a real-world use case. Our solution is scalable, can easily be expanded to support new external data sources, supports error recovery, allows instance level configurability, provides end-users without deep technical knowledge with a web user interface for comfortable usage and offers sharing mechanisms among users.

While we were able to demonstrate the integration of external data sources into our architecture, some open questions remain. For example, unit conversions are not covered yet by our solution. Downstream applications depend on the unit external data sources provide. Support for unit conversion could be added using an analytics pipeline as described in section 4. Furthermore, data from external sources is usually not as geographically precise as data from own sensors. Enhancing geographic precision in the most relevant areas while maintaining a less precise overview of surrounding areas might be possible by combined analytics of both data sources. It is unclear however, how many sensors would be needed for an actual improvement and (depending on the geographic circumstances) where exactly sensors would need to be deployed. Future research might address these issues and develop generalized rules for effective sensor placement. Additionally, it would be interesting to integrate sensors, data from external sources and geographical models of former mining areas in order to create a digital twin. This model could be used to simulate scenarios (cf. [GV17]) such as flooding in order to gain new insights, e. g. where extend efforts in disaster prevention. Moreover, administration processes (e. g. requests for information on flooding risks based on current weather conditions) could be enhanced by combined analytics of sensor data and data from external sources in order to (partly) automate such processes.

Acknowledgements

The work presented in this paper is funded by the Free State of Saxony (SAB 100400223).

References

- [Al17] Alvarez-Campana, M.; López, G.; Vázquez, E.; Villagrà, V.; Berrocal, J.: Smart CEI Moncloa: An IoT-Based Platform for People Flow and Environmental Monitoring on a Smart University Campus. *Sensors* 17/12, p. 2856, Dec. 8, 2017, ISSN: 1424-8220.
- [Ch18] Chen, L.: Microservices: Architecting for Continuous Delivery and DevOps. In: 2018 IEEE International Conference on Software Architecture (ICSA). 2018 IEEE International Conference on Software Architecture (ICSA). IEEE, Seattle, WA, pp. 39–397, Apr. 2018, ISBN: 978-1-5386-6398-1.
- [Dr17] Dragoni, N.; Lanese, I.; Larsen, S. T.; Mazzara, M.; Mustafin, R.; Safina, L.: Microservices: How To Make Your Application Scale, Feb. 23, 2017, arXiv: 1702.07149 [cs], URL: <http://arxiv.org/abs/1702.07149>, visited on: 04/08/2021.
- [DW20] DWD: Radargestützte Analysen Stündlicher Niederschlagshöhen Im Echtzeitbetrieb Für Deutschland (RADOLAN) Und Mitteleuropa (RADOLAN-ME), Oct. 27, 2020, URL: https://www.dwd.de/DE/leistungen/radolan/radarniederschlagsprodukte/radolankurzbeschreibung_pdf.pdf?__blob=publicationFile&v=7, visited on: 03/29/2021.
- [En20] Enerdata: Coal Phase Out - What Are the Major Global Coal Trends?, Mar. 20, 2020, URL: <https://d1owejb4br3112.cloudfront.net/publications/executive-briefing/global-coal-phase-out.pdf>, visited on: 04/07/2021.
- [GV17] Grieves, M.; Vickers, J.: Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems. In (Kahlen, F.-J.; Flumerfelt, S.; Alves, A., eds.): *Transdisciplinary Perspectives on Complex Systems*. Springer International Publishing, Cham, pp. 85–113, 2017, ISBN: 978-3-319-38754-3 978-3-319-38756-7.
- [He20] Heideker, A.; Ottolini, D.; Zyrianoff, I.; Neto, A. T.; Salmon Cinotti, T.; Kamienski, C.: IoT-Based Measurement for Smart Agriculture. In: 2020 IEEE International Workshop on Metrology for Agriculture and Forestry (MetroAgriFor). 2020 IEEE International Workshop on Metrology for Agriculture and Forestry (MetroAgriFor). IEEE, Trento, Italy, pp. 68–72, Nov. 4, 2020, ISBN: 978-1-72818-783-9.

- [KOJ19] Kwon, D.; Ok, K.; Ji, Y.: IBFRAME: IoT Data Processing Framework for Intelligent Building Management. In: 2019 IEEE International Conference on Big Data (Big Data). 2019 IEEE International Conference on Big Data (Big Data). IEEE, Los Angeles, CA, USA, pp. 5233–5238, Dec. 2019, ISBN: 978-1-72810-858-2.
- [Lu20] Luckner, M.; Grzenda, M.; Kunicki, R.; Legierski, J.: IoT Architecture for Urban Data-Centric Services and Applications. *ACM Transactions on Internet Technology* 20/3, pp. 1–30, Oct. 2020, ISSN: 1533-5399, 1557-6051.
- [Se18] Senožetnik, M.; Herga, Z.; Šubic, T.; Bradeško, L.; Kenda, K.; Klemen, K.; Pergar, P.; Mladenić, D.: IoT Middleware for Water Management. *Proceedings* 2/11, p. 696, July 31, 2018, ISSN: 2504-3900.
- [St16] Stolpe, M.: The Internet of Things: Opportunities and Challenges for Distributed Data Analysis. *ACM SIGKDD Explorations Newsletter* 18/1, pp. 15–34, Aug. 2016, ISSN: 1931-0145, 1931-0153.
- [Te19] Terroso-Saenz, F.; González-Vidal, A.; Ramallo-González, A. P.; Skarmeta, A. F.: An Open IoT Platform for the Management and Analysis of Energy Data. *Future Generation Computer Systems* 92/, pp. 1066–1079, Mar. 2019, ISSN: 0167739X.
- [We17] Wehlitz, R.; Häberlein, D.; Zschörnig, T.; Franczyk, B.: A Smart Energy Platform for the Internet of Things – Motivation, Challenges, and Solution Proposal. In (Abramowicz, W., ed.): *Business Information Systems*. Vol. 288, *Lecture Notes in Business Information Processing*, Springer International Publishing, Cham, pp. 271–282, 2017, ISBN: 978-3-319-59335-7 978-3-319-59336-4.
- [Wo18] World Wide Web Foundation: *OpenData Barometer*, Sept. 20, 2018, URL: <https://opendatabarometer.org/doc/leadersEdition/ODB-leadersEdition-Report.pdf>, visited on: 04/12/2021.
- [Zs18] Zschörnig, T.; Wehlitz, R.; Rößner, I.; Franczyk, B.: SEPL: An IoT Platform for Value-Added Services in the Energy Domain - Architectural Concept and Software Prototype: In: *Proceedings of the 20th International Conference on Enterprise Information Systems*. 20th International Conference on Enterprise Information Systems. SCITEPRESS - Science and Technology Publications, Funchal, Madeira, Portugal, pp. 593–600, 2018, ISBN: 978-989-758-298-1.
- [Zs20a] Zschörnig, T.; Windolph, J.; Wehlitz, R.; Franczyk, B.: A Cloud-Based Analytics Architecture for the Application of Online Machine Learning Algorithms on Data Streams in Consumer-Centric Internet of Things Domains. In: *IoTBDS*. Pp. 189–196, 2020.
- [Zs20b] Zschörnig, T.; Windolph, J.; Wehlitz, R.; Franczyk, B.: A Cloud-Based Analytics-Platform for User-Centric Internet of Things Domains – Prototype and Performance Evaluation. In: *Hawaii International Conference on System Sciences*. 2020.

Workshop: Security, Datenschutz und Anonymisierung

Security, Datenschutz und Anonymisierung

Ralf Kneuper¹, Annika Selzer², Marion Steiner³

Sicherheit und Datenschutz rücken immer mehr in den Blickpunkt vieler Akteure, die vor die Herausforderung gestellt werden, einerseits die Chancen, die durch neue Technologien entstehen, nicht gänzlich durch normative Vorhaben „auszubremsen“ sowie andererseits für eine angemessene Umsetzung sicherheitstechnischer und datenschutzrechtlicher Vorgaben zu sorgen, u.a. um natürliche Personen vor Persönlichkeitseingriffen durch die exzessive Verarbeitung ihrer personenbezogenen Daten zu schützen. Vor diesem Hintergrund ist u.a. auch zu bewerten, ob die Verarbeitung personenbezogener Daten zwangsläufig erforderlich ist, oder ob die geplanten Ziele der Datenverarbeitung auch mit anonymen Daten erreicht werden können.

Der interdisziplinär angelegte Workshop „Security, Datenschutz und Anonymisierung“ widmet sich diesem Themengebiet. Neben technischen Aspekten werden auch rechtliche und wirtschaftliche Aspekte des Themengebiets diskutiert. Daher umfasst er die folgenden Tracks:

- Security & Privacy at Large
- Anonymität und Anonymisierung in der digitalen Welt (AnonDig)
- Nachhaltiges Informationssicherheitsmanagement (SEGMGT)

Für die Mitarbeit bei der Organisation des Workshops und die Begutachtung der eingereichten Beiträge danken wir den weiteren Organisatoren Agnes Koschmider (Christian-Albrechts-Universität zu Kiel), Fabian Prasser (Berlin Institute of Health und Charité - Universitätsmedizin Berlin) (beide Track AnonDig) und Christopher Rupprich (CR-IT.eu | it. network. security) (Track SECMGT) sowie den Mitgliedern der Programmkomitees Martin Bialke (Universität Greifswald), Rainer Böhme (Universität Innsbruck), Wolfgang Böhmer (Dr. Böhmer Consulting), Benedikt Buchner (Universität Bremen, IGMR), Stephan Fahrenkrog-Petersen (HU Berlin), Ulrich Heun (Carmao GmbH), Christoph Krauß (Fraunhofer SIT, Hochschule Darmstadt), Alexander Lawall (IU Internationale Hochschule – Fernstudium), Martin Macke (Würzburg), Felix Mannhardt (TU Eindhoven), Judith Michael (RWTH Aachen), Isabel Münch (Bundesamt für Sicherheit in der Informationstechnik BSI), Georgis Raptis (OTH Regensburg), Peer Reymann (ITQS Gesellschaft für Qualitätssicherung in der Informationstechnologie mbH), Murat Sariyar (Berner Fachhochschule), Martin Steinebach (Fraunhofer SIT, TU Darmstadt), Patrick Theobald (IT-Security@Work GmbH), Michael Waidner

¹ IU Internationale Hochschule - Fernstudium, ralf.kneuper@iu.org.

² Fraunhofer-Institut für Sichere Informationstechnologie SIT, annika.selzer@sit.fraunhofer.de.

³ IT-Security@Work GmbH (ISW), Marion.Steiner@isw-online.de.

(Fraunhofer SIT, TU Darmstadt), Bernhard C. Witt (it.sec GmbH) und Daniel Woods (Universität Innsbruck).

Teile des Workshops wurden vom Bundesministerium für Bildung und Forschung (BMBF) und vom Hessischen Ministerium für Wissenschaft und Kunst (HMWK) im Rahmen ihrer gemeinsamen Förderung für das Nationale Forschungszentrum für angewandte Cybersicherheit ATHENE unterstützt.

Corporate Digital Responsibility: Evaluating Privacy and Data Security Activities on Company-level

K. Valerie Carl¹

Abstract: The digital economy holds new opportunities for value creation but also threats for both companies and customers. Within this setting, the concept of Corporate Digital Responsibility (CDR) gains traction. Building on the well-established concept of Corporate Social Responsibility, CDR entails a set of principles through which it seeks to ensure an ethical and responsible development, deployment, and use of digital technologies. To date, the scholarly conceptualization of CDR is still in its infancy. This study pursues two main objectives: Firstly, it seeks to contribute to CDR theory by providing a more in-depth conceptualization with focus on privacy and data security. Secondly, this study provides first guidance for the evaluation of CDR activities on company-level, a benchmark corpus. As work-in-progress, the focus lies on identifying a starting point for the evaluation of CDR activities concentrating on privacy and data security, and hence research opportunities related to this assessment in future.

Keywords: corporate digital responsibility, ethical guidelines, data privacy, data security, benchmark corpus

1 Introduction

Advancements in digital technologies allow for the development of more sophisticated digital products and services. Especially, a plethora of Big Data and Artificial Intelligence applications emerged, further referred to under the umbrella term digital products and services. Nevertheless, besides all the new opportunities for value creation, digitalization also holds a range of threats and challenges [He16], [Th17], that managers need to cope with. Especially the collection of user data are accompanied by possible hazard of privacy and security related issues, sometimes causing economic, ethical, or legal issues for customers and firms alike [Ba19]. In an attempt to harness the advantages of digitization by adequately addressing its challenges, the emergence of a more comprehensive concept called Corporate Digital Responsibility (CDR) can be observed.

At its core, CDR is closely related to and has similar goals like the concept of Corporate Social Responsibility (CSR). While CSR aims to minimize the negative impacts and maximize the positive outcomes of corporate practices on socially and environmentally relevant issues [MR02], CDR intends to minimize the adverse effects of digitization while maximizing the positive impacts of corporate digital activities. In this vein, CDR seeks to

¹ Goethe University Frankfurt, Chair of Information Systems and Information Management, Theodor-W.-Adorno-Platz 4, 60323, Frankfurt am Main, kcarl@wiwi.uni-frankfurt.de

ensure an ethical and responsible development, deployment, and use of digital technologies and data. CDR puts, *inter alia*, privacy and data security attempts in a broader context to provide a more holistic approach.

To date, the CDR debate is strongly driven by practitioners and other policy-related initiatives, as its scholarly conceptualization is still in its infancy [Lo21] albeit digitization already brought up unprecedented challenges [Na17]. Recently, numerous initiatives evolved around the concept of CDR or widened their focus concerning this subject. Supranational organizations like the European Union, the OECD, or the UN are developing guidelines, laws (e.g., 'European General Data Protection Regulation' (GDPR), 'European Business Network for Corporate Social Responsibility', 'OECD Guidelines for Multinational Enterprises', 'UN Sustainable Development Goals', 'UN Global Compact'), or working groups which try to establish CDR in the corporate mind-set. Also, national and industry-led initiatives (e.g., the German 'Corporate Digital Responsibility Initiative') form up, where industry leaders want to set a good example by committing to ethical business practices in the digital world. Within this debate, experts have formulated eight CDR norms as a basis for ethical and responsible digital business practices [Th17].

Aiming to advance the scholarly debate on CDR, this study pursues two main objectives: Firstly, as a theoretical contribution, this study seeks to contribute to the emerging knowledge about CDR. An approach covering a total of 8 dimensions has already been devised [Th17]. To this end, this work theoretically links the concept of privacy and data security to the broader context of corporate responsibility, motivating research on CDR, and privacy and data security attempts as a distinct topic. Secondly, there are currently no concrete recommendations on how to evaluate CDR activities on company-level. This work seeks to close this theoretical gap, while at the same time contributing to the implementation of CDR in practice by providing a benchmark corpus. The implementation of CDR norms and practices can be pursued in many ways and at various levels [MM08]. This allows the company to position in relation to competitors and thus gain competitive advantages. As resources to be spent on CSR and CDR are limited, a successful CDR implementation hinges on companies' ability to evaluate activities regarding CDR norms.

This paper theorizes exemplary measures to evaluate CDR activities of a firm on the example of privacy and data security serving as an evaluation guideline. As work-in-progress, the focus lies on identifying new research opportunities related to the field of CDR with particular emphasis on privacy and data security to evaluate corporate activities. This research is based on previous research, own CDR surveys, and observations in practice.

2 Corporate Digital Responsibility

As mentioned previously, the Corporate Digital Responsibility (CDR) concept is closely related to the concept of Corporate Social Responsibility (CSR). Nevertheless, CDR merits scholarly attention on its own, as it accommodates the digital world's peculiar challenges [Lo21]. By drawing on existing literature regarding the unique ethical and social

challenges, that the digital context presents, this work discusses CDR's core components with a focus on privacy and data security. This will lay the foundation for evaluating CDR activities on the company-level.

Although CSR and CDR share common values, norms, and an organization's commitment towards ecological and social challenges at large, CDR should be considered separately from CSR. CDR addresses challenges to organizations' ethical behavior that are unique to the digital world and go far beyond CSR. Especially, "exponential growth in technological development, malleability of technologies and data in use, and pervasiveness of technology and data" [Lo21, p. 876] manifest particularities of digital technologies. Thus, CDR and CSR represent complementary but also sometimes overlapping concepts of business ethics. Compared to CSR, CDR addresses the unprecedented challenges related to the digital world. An approach comprising of a total of 8 dimensions to cover possible CDR activities has already been devised [Th17]. Table 1 presents these CDR norms proposed by practice and indicates exemplary related literature.

CDR norm [Th17]	Norm description [based on Th17]	Related work
I. Access	Consumers should have access to basic digital goods and services.	[HRK08];
II. Education and awareness	Consumers should be educated including their awareness of ecological, social, societal, and economic aspects of their consumption.	[LMH17]; [VS13]
III. Information and transparency	Consumers should have access to appropriate information so that they can be informed according to their individual wishes and needs.	[AK06]; [GGK10]
IV. Economic interests	The economic interests of consumers should be protected and promoted.	[BKV15]; [HHS11]
V. Product safety and liability	Consumers should be protected from risks to their health and safety.	[DR95]; [Sm17]
VI. Privacy and data security	The protection of consumers' privacy, the free flow of information, and protected and secure payment mechanisms should be ensured.	[BC11]; [Gi18]; [Oe17]
VII. Dispute resolution and awareness	Consumers should have access to effective dispute settlement and appeal procedures.	[Eu16]; [Th17]
VIII. Governance and participation mechanisms	Legal organizations and regulators should ensure that there are appropriate governance and participation mechanisms in place.	[Lo21]; [Th17]

Tab. 1: CDR norms that can serve as a preliminary conceptualization and exemplary related work.

As resources to be spent on CSR and CDR are limited, a successful CDR implementation hinges on companies' ability to evaluate CDR norms and measures. Due to the complexity of building a benchmark corpus covering all dimensions of CDR, a first focus is placed on the applications that are most important for key stakeholders involved [KBM13]. According to previous research, the dimension of privacy and data security has been evaluated

most important from the customers' perspective [Mi21]. Hence, the development of a benchmark corpus to evaluate CDR activities on company-level firstly concentrates on the dimension of privacy and data security, while also taking into consideration possible overlaps with further CDR dimensions.

3 Evaluating CDR activities regarding privacy and data security

3.1 Evaluating activities within the privacy and data security norm

While data privacy focuses on consumers' capability to exert control over the storing, processing, or forwarding practices concerning their data, information security refers to the protection of stored data against various threats [BC11]. To date, privacy and data security remain two of the significant concerns related to the adoption and use of information technologies [Ma86], [MZH17]. Hence, the importance of data privacy and security is a widely discussed research topic in information systems (e.g., [BC11], [Ha02], [HH18]). Amongst one of the undisputed major consumer concerns in the digital economy, information privacy refers, inter alia, to the consumers' capability to control their information stored and the handling of their data, including the monetization [BC11], [Go91], [We67]. The topic is subject to regulations (e.g., GDPR), which define the minimum requirements of privacy and data security companies must comply with. While non-compliance with the minimum requirements can have negative legal and financial consequences [GS09], [Oe17], compliance often does not stand out positively. Against the background that companies can positively influence consumers' perceptions through strategic initiatives [Ha07], the CDR norm regarding privacy and data security encourages companies to outperform the current (legal) regulations.

The topic of privacy and data security is multifaceted and features various aspects. Smith et al. [SMB96], for instance, suggest that privacy has four main aspects: the first aspect of privacy relates to data collection. The second aspect adheres to unauthorized secondary use of information for both organization-specific internal and other external purposes. Further, the third privacy aspect is improper access, while information accuracy (i.e., errors) represents the fourth important privacy aspect. In practice, these aspects reappear in data privacy and security regulations such as the GDPR or OECD Privacy Framework. According to the OECD [Oe13] guidelines, for instance, privacy and data security should consider eight main principles: (1) data collection limitation; (2) data quality; (3) purpose specification; (4) use limitation; (5) security safeguards; (6) openness; (7) individual participation; (8) accountability principle. Akin, the GDPR provides eleven privacy and security-relevant principles [Te18], similar to the OECD Guidelines. The selected sub-dimensions cover seven of the eight main principles within the OECD Privacy Framework. The principle of accountability was excluded as it can be seen as a framework condition for the appropriate fulfillment of the other principles. Nevertheless, the use of such a data controller, as part of the accountability principle, is of crucial importance and should not be neglected by companies. Based on scholarly work on data privacy and security, and the

current state of legislation (i.e., GDPR, OECD Guidelines), a theory link to seven exemplary implementation measures of the data privacy and security CDR norm was drawn (see Table 2).

To evaluate privacy and data security activities on company-level a measurement scale has to be developed. Each sub-dimension of a CDR norm represents a broad field of application, thus, one subordinate possible measure was chosen to exemplarily discuss the evaluation of privacy and data security activities on company-level. Each measure features three different levels to apply an ample-based approach to evaluate CDR activities of companies. The first level is used to illustrate minimal activities related to the CDR implementation and partly coincides with the (national) statutory requirements for companies. Accordingly, CDR activities below the minimal activities would not be counted as corporate activities related to CDR. Still, CDR is a concept based on voluntariness, accordingly no company has to meet such levels. The further levels represent increasing, voluntary acceptance of more responsibility with regard to CDR. In that way, the evaluation of privacy and data security activities on company-level can be performed by applying this ample-based approach to differentiate between minimal activities to prominent takeover of additional responsibility (see Table 2).

Limited or restricted data collection must be with the consent of the user. Nevertheless, many consumers struggle to understand what data companies are really collecting, thus making uninformed decisions [Wi21]. Even though companies have to inform about this in an understandable way according to the GDPR [Fe19]. Thus the measure information regarding data protection has been included to cover the aforementioned sub-dimension exemplarily.

In line with this, the **clear purpose of data collection** is often communicated in a way that is difficult to understand [Di13]. Still, the purpose of the data collection should be clearly stated at the time of collection avoiding contradictions between the threefold communication to the customer and the actual declaration, for example, in the data protection declaration. Additionally, consumers might not be aware of the general agreement on the acceptance of the data protection declaration at the time of providing additional data to a firm. Following, the communication of the purpose at the time of data collection is essential for a responsible handling of user data and therefore exemplarily incorporated in the benchmark corpus.

In the same vein, **restricted data use** should be established to avoid unknown or unintended secondary usage, especially when it comes to third-party access [Di13]. On the one hand, firms might not disclose secondary usage of data, which is not legal in some countries. On the other hand, firms might obtain the customers' consent for such use, for example, in the form of a complicated data protection declaration, so that the consumer is

Sub-dimension	Description	Measure	Measure levels (ascending commitment)	Related work
Limited/restricted data collection	The collection of (personal) data must be limited, lawful, and fair, usually with the knowledge and/or consent of the user.	Data protection declaration	Detailed; One Pager; Tabular form	[Eu16]; [SMB96]; [Wi21]
Clear purpose of data collection	The purpose of the data collection must be clearly stated at the time of collection.	Communicated purpose of data collection	Contradictory communication; Data protection declaration; At the time of collection	[Di13]; [SMB96]; [Th17]
Restricted data use	The use or disclosure of data must be limited to the agreed purpose(s) or only for closely related purposes.	Secondary usage	Unknown secondary usage; Unintended secondary usage; No secondary usage	[Di13]; [Eu16]; [Wi21]
Openness about data processing practices	Businesses need to be transparent about their data processing practices.	Explanation	Technical/legal explanation; Explanation for layperson; Personalized explanation	[Eu16]; [Wi21]; [TF09]
Secure storage and processing of user data	The storage, processing, and transmission of user data must be subject to appropriate security.	Notification of incidents	On request; Affected users only; Public broadcast	[Eu16]; [Th17]
Data quality	User data collected and stored by companies must be relevant, accurate, and up-to-date.	Connection between data and purpose	No connection; On purpose/data classes; On specific data and purpose	[Ma15]; [Oe13]; [SMB96]
Access and correction	Users must be able to view and correct the user data stored by companies.	Access and correction of personal data	Information; Information and correction; Information, correction, deletion	[Eu16]; [Ma15]; [Th17]

Tab. 2: Overview of privacy and data security sub-dimensions, its exemplary measures, and exemplary related work.

not even aware of his consent and the use of data [Di13], [Wi21]. Accordingly, restricted secondary usage has been exemplarily included in the benchmark corpus

Openness about data processing practices is (partly) covered by the GDPR but also suggested by research to enable consumers to make informed choices and gain some control over their (personal) data [Wi21]. Still, openness can occur in various ways including different levels of explanation which are more or less easy to understand.

The sub-dimension of **secure storage and processing of user data** covers a plethora of possible implementations. Secure storage and processing applies not only to internal (local) utilization but also, for example, to secure data transmission, to the inclusion of other companies in the value creation process. Secure storage and processing goes beyond the mere process in the eyes of the customers. Still, it is mainly perceived in the form of incidents and the related notifications from the consumers' perspective [Th17]. Hence, the notification of incidents concerning stored personal data is included in the benchmark corpus as an example for this sub-dimension.

Data quality can be characterized, amongst other influencing factors, based on the connection between collected data and data collection purpose [Oe13]. This sub-dimension applies to the internal data management processes. Hence, data and initial collection purpose can be stored decoupled from each other, based on formed classes of purposes and data, or for each specific data set with indication of the initial purpose.

The GDPR (partly) covers **access and correction** of personal data. Therefore, this sub-dimension is also captured in the benchmark corpus by its own measure [Ma15], [Th17]. Levels can range from solely information to correction and deletion.

3.2 Evaluating privacy and data security activities within further CDR norms

A more comprehensive approach of privacy and data security exceeds mere activities on the system-level (covered by CDR norm VI “privacy and data security”) and rather also impacts other dimensions of CDR. Hence, privacy and data security activities on company-level can also comprise activities rooted in other CDR dimensions (see Table 1 for an overview of the CDR norms). One subordinate possible measure per dimension was developed based on previous research to exemplarily discuss the evaluation of privacy and data security activities on company-level in the context of further CDR norms. Each measure features three different levels to continue the ample-based approach to evaluate CDR activities of companies (see Table 3).

In the context of CDR, the dimension **access** covers the possibility to get in contact with basic digital products and services [HRK08], [LMH17]. Access to everyone is not just strengthening social groups that did not have access until now but also the user experience of all consumers [Ne06]. In this vein, it might be favorable for consumers to be able to access basic services (i.e., advisory services, or (insurance) premium calculators) without providing any personal data. Consumers usually want to protect their personal data and reduce risks from providing such, but often have to weigh this risk against the lack of functionality, known in research as the privacy calculus [AK06], [Ha02], [Su13].

Hence, access related to privacy and data security can be exemplarily included in the benchmark corpus.

Additionally, **education and awareness** should comprise consumers' awareness for ecological, social, and societal aspects as well as the economic impacts of their consumption decisions. This dimension covers a wide spectrum of education fields. Examples are seeking information and advice, or coping with problems [Th17], [Un19]. This demand can even be reinforced through the ongoing digitization and the development of new digital security and privacy technologies. In order to offer secure and privacy-preserving products and services, new technology applications such as blockchain technology are increasingly being used to protect privacy and data security [Ay18], [Wa21]. Thus, associated new concerns arise, for example, regarding energy consumption or sustainability [Tr18], [Wa21]. Furthermore, many consumers are still often unaware of digital products or services intruding upon their privacy, their rights regarding data security and privacy, and how they can make use of them [Lo17], [MDK18]. Thus, education on consumer rights regarding privacy and data security should be offered and is consequently, as an exemplary instance of this dimension, part of the benchmark corpus.

Information and transparency is a broad application field of CDR also covering measures that refer to privacy and data security. For instance, companies can pursue more transparency by adequately informing (potential) customers about the collection, storage, or handling of personal data [Eu16], [TF09] also using certifications [CM20]. As privacy can be regarded as the individual's right, *inter alia*, to determine whom to disclose personal information [We67], users should be informed about all data processing companies involved in the purchase transaction including process steps such as distribution, packaging, or shipping. Hence, the information on involved companies has been exemplarily incorporated in the benchmark corpus.

Information privacy exceeds the consumers' capability to control their own information stored and the handling of their data especially concerning the monetization of provided data [BC11], [Go91], thus affecting consumers' **economic interests**. In these times, many business models (i.e., freemium, or free distribution through personal data intelligence) rely on analyzing supplied consumer data for own purposes, providing free products and services for the consumer [Lo17]. Hence, consumer data are called the "new currency on the Internet" [Ca12, p. 3834], although consumers might want to protect these sensitive data. Accordingly, the business model used can be decisive for a consumer decision, so consumers should be informed about this. In line with this, this aspect was included as an example for consumers' economic interests in the benchmark corpus.

Motivated by the complications arising from digital products and services, **dispute resolution and awareness** also covers disagreements originating in privacy and data security incidents. In general, dispute resolution refers to the mechanisms aiming to provide consumers who have suffered (economic) harm from transactions, to solve their complaints

CDR norm	Measure	Description	Measure levels (ascending commitment)	Related work
Access	Access without data input	Access, for example, to advisory services, or premium calculators can be open without entering personal data and thus with reduced privacy concerns.	Read only web page content; Limited service without data; Services without data	[AK06]; [LMH17]; [Ne06]; [Su13]
Education and awareness	Consumer rights education	Consumer education regarding their rights related to privacy and data security and how to make use of them.	Not provided; Passive offer; Actively focused	[Lo17]; [MDK18]; [Un19]
Information and transparency	Transparency about business partners	Users should be informed about all data processing companies involved in the purchase transaction.	No information; Hidden information; Proactive information	[Eu16]; [TF09]; [We67]
Economic interests	Deployed business model	Users should know how the company generates revenue with a (free) product or service (e.g., by the usage of collected data).	No information; Hidden information; Proactive information	[BC11]; [Go91]; [MZH19]; [SSL16]
Dispute resolution and awareness	Point of contact	In the event of disagreements originating in privacy and data security, dispute resolution can be secured differently.	Manufacturer specific; Manufacturer network; Independent agency	[Eu16]; [Oe07]; [Th17]

Tab. 3: Overview of further touched CDR norms, its exemplary measures, and exemplary related work.

and receive redress [Oe07]. As digitalization enables companies to operate across borders, the CDR concept envisions an uncomplicated, unified, and efficient dispute resolution and redress mechanism for all consumers. More specifically, CDR suggests that consumers should have the option to place complaints easily and free of charge, while the processing of the complaints should be fast, fair, and transparent [Eu16], [Th17]. The point of contact can have a strong influence on dispute resolution and redress depending on its independence of interests, therefore exemplarily included in the benchmark corpus.

Due to the complex, highly dynamic, and nationally fragmented legal debate on **product safety and liability** of digital products and services [De14], [HTW17] stable safety and liability measures could not be identified yet that would form a solid basis for the evaluation of CDR activities. Nevertheless, future advancements of the benchmark corpus should incorporate this CDR dimension in their considerations.

Because our work focuses on companies and how CDR activities can be evaluated on company-level, the dimension **governance and participation** has been excluded as it lies predominantly in the hands of policymakers and other non-governmental regulatory organizations [Th17]. Since this dimension is out of the direct reach of companies it can be considered to be an exogenous force within a CDR framework and is therefore not part of a benchmark corpus on company-level.

4 Conclusion

The goal of this work-in-progress paper is to start a discussion on how to measure CDR activities on company-level. Especially for companies operating with digital products and services, the understanding of responsibility has changed. The provided benchmark corpus above (see Table 2 and 3) should serve as a starting point for further research, providing exemplary measures to evaluate CDR activities in general and activities related to privacy and data security in particular. Future research should complement, value, and validate this part of the benchmark corpus in order to develop a comprehensive evaluation system to assess CDR activities on company-level. In addition, the benchmark corpus covering all dimensions of CDR equally needs to be expanded analogically.

In practice, the provided benchmark corpus should serve as an orientation for firms on how to evaluate privacy and data security activities on company-level. In order to adapt this benchmark corpus to the specific company's environment the applicability of individual measures must be assessed and supplemented with own criteria if needed. Besides, weighting can be used to adapt the benchmark corpus to specific circumstances within the company similar to a utility analysis. In practice, it might be worth considering to align

the status-quo evaluation with internal visions and missions and to create an action plan for the further development based on this benchmark corpus.

A significant limitation is that an all-encompassing benchmark corpus is not feasible even after further developments. Rather, the goal is to find a common, expanded consensus of the most relevant measures through discussion so that a comprehensive application to a wide variety of industries, products, and services is possible. A further development of the benchmark corpus can reduce but not remove this limitation.

5 Acknowledgement

This work has been supported by the Hessian State Chancellery – Hessian Minister of Digital Strategy and Development under the promotional reference 6/493/71574093 (CDR-CAT).

6 References

- [AK06] Awad, N. F.; Krishnan, M. S.: The Personalization Privacy Paradox: An Empirical Evaluation of Information Transparency and the Willingness to be Profiled Online for Personalization. *MIS Quarterly* 30/1, pp. 13-28, 2006.
- [Ay18] Ayoade, G. et al.: Decentralized IoT Data Management Using Blockchain and Trusted Execution Environment. In: 2018 IEEE International Conference on Information Reuse and Integration (IRI). Salt Lake City, Utah, USA, pp. 15-22, 2018.
- [Ba19] Baumann, A. et al.: The Price of Privacy: An Evaluation of the Economic Value of Collecting Clickstream Data. *Business & Information Systems Engineering* 61/4, pp. 413-431, 2019.
- [BC11] Bélanger, F.; Crossler, R. E.: Privacy in the Digital Age: A Review of Information Privacy Research in Information Systems. *MIS Quarterly* 35/4, pp. 1017-1041, 2011.
- [BKV15] Bourreau, M.; Kourandi, F.; Valletti, T.: Net Neutrality with Competing Internet Platforms. *Journal of Industrial Economics* 63/1, pp. 30-73, 2015.
- [Ca12] Camenisch, J.: Information Privacy?!. *Computer Networks* 56/18, pp. 3834-3848, 2012.
- [CM20] Carl, K. V.; Mihale-Wilson, C.: Consumer Privacy Concerns and Preferences for Certification and Accreditation of Intelligent Assistants in the Internet of Things. In (Roßnagel, H.; Schunck, C. H.; Mödersheim, S.; Hühnlein, D. eds.): *Open Identity Summit 2020*. Gesellschaft für Informatik e.V., Copenhagen, Denmark, pp. 157-162, 2020.
- [De14] Desai, D. R.: The New Steam: On Digitization, Decentralization, and Disruption. *Hastings Law Journal* 65/6, pp. 1469-1482, 2014.

- [Di13] Dinev, T. et al.: Information Privacy and Correlates: An Empirical Attempt to Bridge and Distinguish Privacy-Related Concepts. *European Journal of Information Systems* 22/3, pp. 295-316, 2013.
- [DR95] Daughety, A. F.; Reinganum, J. F.: Product Safety: Liability, R&D, and Signaling. *American Economic Review* 85/5, pp. 1187-1206, 1995.
- [Eu16] European Parliament: Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance), <https://eur-lex.europa.eu/eli/reg/2016/679/oj>, access 9 July 2019, 2016.
- [Fe19] Felzmann, H. et al.: Transparency You Can Trust: Transparency Requirements for Artificial Intelligence between Legal Norms and Contextual Concerns. *Big Data & Society* 6/1, p. 2053951719860542, 2019.
- [GGK10] Granados, N.; Gupta, A.; Kauffman, R. J.: Research Commentary – Information Transparency in Business-to-Consumer Markets: Concepts, Framework, and Research Agenda. *Information Systems Research* 21/2, pp. 207-226, 2010.
- [Gi18] Gimpel, H. et al.: The Upside of Data Privacy – Delighting Customers by Implementing Data Privacy Measures. *Electronic Markets* 28/4, pp. 437-452, 2018.
- [Go91] Goodwin, C.: Privacy: Recognition of a Consumer Right. *Journal of Public Policy and Marketing* 10/1, pp. 149-166, 1991.
- [GS09] Goel, S.; Shawky, H. A.: Estimating the Market Impact of Security Breach Announcements on Firm Values. *Information & Management* 46/7, pp. 404-410, 2009.
- [Ha02] Hann, I.-H. et al.: Online Information Privacy: Measuring the Cost-Benefit Tradeoff. In (Miralles, F.; Valor, J. eds.): *Proceedings of the International Conference on Information Systems 2002*, Barcelona, Spain. 2002.
- [Ha07] Hann, I.-H. et al.: Overcoming Online Information Privacy Concerns: An Information-Processing Theory Approach. *Journal of Management Information Systems* 24/2, pp. 13-42, 2007.
- [He16] Hess, T. et al.: Options for Formulating a Digital Transformation Strategy. *MIS Quarterly Executive* 15/2, pp. 123-139, 2016.
- [HH18] Heimbach, I.; Hinz, O.: The Impact of Sharing Mechanism Design on Content Sharing in Online Social Networks. *Information Systems Research* 29/3, pp. 592-611, 2018.
- [HHS11] Hinz, O.; Hann, I.-H.; Spann, M.: Price Discrimination in E-Commerce? An Examination of Dynamic Pricing in Name-Your-Own Price Markets. *MIS Quarterly* 35/1, pp. 81-98, 2011.
- [HRK08] Hsieh, J. J. P.-A.; Rai, A.; Keil, M.: Understanding Digital Inequality: Comparing Continued Use Behavioral Models of the Socio-Economically Advantaged and Disadvantaged. *MIS Quarterly* 32/1, pp. 97-126, 2008.

- [HTW17] Howells, G.; Twigg-Flesner, C.; Willett, C.: Product Liability and Digital Products. In (Synodinou, T.-E.; Jougoux, P.; Markou, C.; Prastitou, T. eds.): *EU Internet Law: Regulation and Enforcement*. Springer International Publishing, Cham, pp. 183-195, 2017.
- [KBM13] Kesavan, R.; Bernacchi, M. D.; Mascarenhas, O. A.: Word of Mouse: CSR Communication and the Social Media. *International Management Review* 9/1, pp. 58-66, 2013.
- [LMH17] Lameijer, C. S.; Mueller, B.; Hage, E.: Towards Rethinking the Digital Divide: Recognizing Shades of Grey in Older Adults' Digital Inclusion. In (Kim, Y. J.; Agarwal, R.; Lee, J. K. eds.): *Proceedings of the International Conference on Information Systems 2017*, Seoul, South Korea, 2017.
- [Lo17] Lopez, J. et al.: Evolving Privacy: From Sensors to the Internet of Things. *Future Generation Computer Systems* 75, pp. 46-57, 2017.
- [Lo21] Lobschat, L. et al.: Corporate Digital Responsibility. *Journal of Business Research* 122, pp. 875-888, 2021.
- [Ma86] Mason, R. O.: Four Ethical Issues of the Information Age. *MIS Quarterly* 10/1, pp. 5-12, 1986.
- [Ma15] Martin, K. E.: Ethical Issues in the Big Data Industry. *MIS Quarterly Executive* 14/2, pp. 67-85, 2015.
- [MDK18] Manikonda, L.; Deotale, A.; Kambhampati, S.: What's up with Privacy? User Preferences and Privacy Concerns in Intelligent Personal Assistants. In (Furman, J.; Marchant, G.; Price, H.; Rossi, F. eds.): *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, New Orleans, LA, USA, ACM, pp. 229-235, 2018.
- [Mi21] Mihale-Wilson, C. et al.: Corporate Digital Responsibility – Extended Conceptualization and a Guide to Implementation. In: *Proceedings of the 25th European Conference on Information Systems*, Marrakech, Morocco, 2021.
- [MM08] Matten, D.; Moon, J.: "Implicit" and "Explicit" CSR: A Conceptual Framework for a Comparative Understanding of Corporate Social Responsibility. *Academy of Management Review* 33/2, pp. 404-424, 2008.
- [MR02] Maignan, I.; Ralston, D. A.: Corporate Social Responsibility in Europe and the U.S.: Insights from Businesses' Self-Presentations. *Journal of International Business Studies* 33/3, pp. 497-514, 2002.
- [MZH17] Mihale-Wilson, C.; Zibuschka, J.; Hinz, O.: About User Preferences and Willingness to Pay for a Secure and Privacy Protective Ubiquitous Personal Assistant. In: *Proceedings of the 25th European Conference on Information Systems*, Guimarães, Portugal, pp. 32-47, 2017.
- [MZH19] Mihale-Wilson, C.; Zibuschka, J.; Hinz, O.: User Preferences and Willingness to Pay for In-Vehicle Assistance. *Electronic Markets* 29/1, pp. 37-53, 2019.
- [Na17] Nambisan, S. et al.: Digital Innovation Management: Reinventing Innovation Management Research in a Digital World. *MIS Quarterly* 41/1, pp. 223-238, 2017.

- [Ne06] Newell, A. F. et al.: Designing a Portal for Older Users: A Case Study of an Industrial/Academic Collaboration. *ACM Transactions on Computer-Human Interaction* 13/3, pp. 347-375, 2006.
- [Oe07] OECD: Recommendation on Consumer Dispute Resolution and Redress, <http://www.oecd.org/internet/consumer/38960101.pdf>, access 9 July 2019, 2007.
- [Oe13] OECD: The OECD Privacy Framework, http://www.oecd.org/sti/ieconomy/oecd_privacy_framework.pdf, access 9 July 2019, 2013.
- [Oe17] OECD: Key Issues for Digital Transformation in the G20. Report Prepared for a Joint G20 German Presidency/OECD Conference, <https://www.oecd.org/g20/key-issues-for-digital-transformation-in-the-g20.pdf>, access 9 July 2019, 2017.
- [Sm17] Smith, B. W.: Automated Driving and Product Liability. *Michigan State Law Review* 2017/1, pp. 1-74, 2017.
- [SMB96] Smith, H. J.; Milberg, S. J.; Burke, S. J.: Information Privacy: Measuring Individuals' Concerns About Organizational Practices. *MIS Quarterly* 20/2, pp. 167-196, 1996.
- [SSL16] Stohl, C.; Stohl, M.; Leonardi, P. M.: Managing Opacity: Information Visibility and the Paradox of Transparency in the Digital Age. *International Journal of Communication* 10, pp. 123-137, 2016.
- [Su13] Sutanto, J. et al.: Addressing the Personalization-Privacy Paradox: An Empirical Assessment from a Field Experiment on Smartphone Users. *MIS Quarterly* 37/4, pp. 1141-1164, 2013.
- [Te18] Tesfay, W. B. et al.: PrivacyGuide: Towards an Implementation of the EU GDPR on Internet Privacy Policy Evaluation. In (Verma, R. M.; Kantarcioglu, M. eds.): *Proceedings of the Fourth ACM International Workshop on Security and Privacy Analytics*, New York, USA, pp. 15-21, 2018.
- [TF09] Turilli, M.; Floridi, L.: The Ethics of Information Transparency. *Ethics and Information Technology* 11/2, pp. 105-112, 2009.
- [Th17] Thorun, C. et al.: Indicators of Consumer Protection and Empowerment in the Digital World. Results and Recommendations of a Feasibility Study, https://www.bmjv.de/G20/DE/ConsumerSummit/_documents/Downloads/Studie.pdf?__blob=publication-File&v=1/, access 9 July 2019, 2017.
- [Tr18] Truby, J.: Decarbonizing Bitcoin: Law and Policy Choices for Reducing the Energy Consumption of Blockchain Technologies and Digital Currencies. *Energy Research & Social Science* 44, pp. 399-410, 2018.
- [Un19] United Nations: Manual on Consumer Protection, https://unctad.org/en/PublicationsLibrary/ditccplp2017d1_en.pdf, access 12 November 2020, 2019.
- [VS13] Venkatesh, V.; Sykes, T. A.: Digital Divide Initiative Success in Developing Countries: A Longitudinal Field Study in a Village in India. *Information Systems Research* 24/2, pp. 239-260, 2013.
- [Wa21] Waheed, N. et al.: Security and Privacy in IoT Using Machine Learning and Blockchain: Threats and Countermeasures. *ACM Computing Surveys* 53/6, pp. 1-37, 2021.

- [We67] Westin, A. F.: Privacy and Freedom. 1st ed., Athenum, New York, 1967.
- [Wi21] Wieringa, J. et al.: Data Analytics in a Privacy-Concerned World. Journal of Business Research 122, pp. 915-925, 2021.

Chances and Limitations of Personal and Anonymized Data Processing

Implementing Appropriate Technical and Organizational Measures and Creating Added Value in Smart Cities

Annika Selzer¹ und Ingo J. Timm²

Abstract: Article 32 GDPR regulates the obligation to implement appropriate technical and organizational measures whenever personal data is being processed. In this paper, we want to link questions arising from taking appropriate technical and organizational measures with considering the chances and limitations of both, personal and anonymized data processing and the potential added value of personal and anonymized data exchange within a smart city context.

We demonstrate the link through a legal analysis and 30 structured interviews with smart city participants.

Keywords: Anonymized Data, GDPR, Personal Data, Smart City.

1 Introduction

A Smart city uses “information and communication technology to sense, analyze and integrate the key information of core systems in running cities” [SLF11]. By using such information and communication technology – including different physical devices that are connected through the internet – and by collecting data through these systems and devices, a city can be optimized in its efficiency of services and operations.

Therefore, a smart city can be enabled to intelligently respond to various needs of its citizens, e.g., making a city healthier, environmentally friendlier, safer and – in general – more enjoyable to live in. Through digitization, a smart city can assist in solving urgent problems of most cities, such as strains on basic public utilities, e.g., water and gas, inadequate social inclusion, and overcrowded traffic [GBP19; SLF11].

To achieve these advantages for their citizens, smart city participants – such as civil service offices, hospitals, basic public utility providers, public transport providers, schools, and universities – usually process personal and anonymized data. The

¹ Fraunhofer-Institute for Secure Information Technology, Rheinstraße 75, 64295 Darmstadt, Germany
Annika.selzer@sit.fraunhofer.de, <http://orcid.org/0000-0001-8859-4808>.

² German Research Center for Artificial Intelligence (DFKI), SDS Branch Trier, Behringstr. 21, 54296 Trier, Germany, ingo.timm@dfki.de, <https://orcid.org/0000-0002-3369-813X>.

differentiation of these two types of data is crucial: Whoever processes *personal data*³ for other than pure private reasons (such as writing down the address or birthday of one's family members or friends) needs to comply to the relevant data protection statutes. In contrast, *anonymized data*⁴ does not fall under data protection law. Therefore, if one processes anonymized data, one does not need to comply to data protection statutes.

1.1 Personal Data Processing in Accordance with the GDPR

In Europe, the European General Data Protection Regulation (short: GDPR) lays down rules to protect natural persons, the so called "data subjects", with regard to the processing of their personal data. The GDPR directly links the level of risks to the rights and freedoms of natural persons to the nature and extent of technical and organizational measures taken to protect them. Generally speaking, if the processing of personal data involves a high risk for the data subjects, a high level of technical and organizational measures needs to be taken, whereas a moderate risk for the data subjects requires a lower level of measures. In short: The implementation of technical and organizational measures needs to be taken in an *appropriate* manner [VB17].

To make sure an organization that determines the purposes of data processing, the so called "controller", implements appropriate technical and organizational measures, the GDPR names several factors that need to be considered, including the risks to the rights and freedoms of natural persons, the nature, scope, context, and purposes of processing, as well as the state of the art and the costs of implementation of technical and organizational measures. These selection criteria are very divers, which makes balancing them a difficult task: According to Art. 32 GDPR, the controller needs to weigh the individual circumstances of data processing and the risks to fundamental rights and freedoms of natural persons against his own economic interest in (potentially) implementing the least expensive technical and organizational measures [PP18; KB20; SHS19; Te20; Se21].

1.2 Research Questions

In this paper, we want to link questions arising from taking appropriate technical and organizational measures with considering the chances and limitations of both, personal and anonymized data processing and the potential added value of personal and anonymized data exchange within a smart city context:

We at first want to answer the question how taking into account the nature, scope, context and purposes of processing (when implementing appropriate technical and organizational measures) links to the chances, limitations and potential added value within a smart city

³ Personal data means any information relating to an identified or identifiable natural person (Art. 4 (1) GDPR).

⁴ Anonymized data is information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable (Recital 26 GDPR).

context. We assume that answering the question of chances and limitations of personal and anonymized data processing as well of potential added value through personal and data exchange should be considered a chance to “pre-determine” whether at all and how long the processing of personal data is necessary and economically reasonable for the specific purpose of the data processing and data exchange. We furthermore assume that the amount, the sensitivity, and the extend of exchanging of personal data is directly linked to the amount of technical and organizational measures that is appropriate to the relevant data processing.

Based on answering the first research question we want to answer the question to what extend personal data processing is necessary in a smart city context and what its chances are in comparison to the chances and limitations of a purely anonymized data processing within a smart city context. We want to answer this question from the point of view of an isolated data processing of each smart city participant as well as from the point of view of adding value for a smart city by exchanging data in between smart city participants. We assume that – from an individual smart city participants point of view – the amount of personal data in general, as well as the amount of sensitive data in particular can be reduced and that most goals of smart cities can be reached by processing anonymized data. We furthermore assume that – from an entire smart cities point of view – most potential added value can be reached by exchanging anonymized data so that – as a result – the necessity to implement technical and organizational measures in accordance with the GDPR could be reduces drastically.

2 Research Method

To answer the research questions of this contribution, we carried out a legal analysis (Sec. 2.1) and 30 structured interviews (Sec. 2.2).

2.1 Legal Analysis

The GDPR regulates the need to implement technical and organizational measures. Furthermore, the GDPR specifies that the state of the art, the costs of implementation and the nature, scope, context, and purposes of processing as well as the risk of varying likelihood and severity for the rights and freedoms of natural persons need to be taken into consideration when choosing appropriate technical and organizational measures. However, the GDPR lacks a legal definition of the mentioned terms.

Since we want to focus on taking into account the nature, scope, context and purposes of processing, we at first needed to assess the meaning of these four terms in the context of the GDPR. Therefore, we carried out a legal analysis of GDPR commentaries, including [ES18; Go18; KB20; PP18; SHS19; VB17]. The results of the legal analysis are summarized in Sec. 4 of this contribution.

2.2 Structured Interviews

To find out which chances and limitations smart city participants see in both, processing personal and anonymized data, we then carried out structured interviews – based on an interview guideline that we validated prior to the interviews with potential interview partners – with three different branches/industry sectors, namely

- health (e.g., public hospitals, private practices, and public ambulance control centers),
- service (e.g., basic public utility providers such as electricity and water, car- and bikesharing providers, and civil service offices), and
- education/research (e.g., schools, universities, and research institutes).

We did so with the aim of gaining a broad perspective of chances and limitations of personal and anonymized data within a smart city, especially regarding data exchange in between different smart city participants. We carried out ten interviews in each of the three branches, resulting in a total of 30 interviews. In the following table we show the key information of the structured interviews that we carried out.

Validation of interview guideline	22 February 2021, 3 via an online video conferencing system, ca. 30 minutes on average.
Interviews	23 February – 19 March 2021, 1 via telephone, 29 via an online video conferencing system, ca. 30-60 minutes each.
Aim of interviews	Identifying chances and limitations of personal and anonymized data processing in smart cities
Interview partner	30 organizations that are smart city participants (either the managing director, the data protection officer, or the smart city manager).
Documentation	Summarized transcription.

Tab. 1: Summary of Structured Interviews

The results of the structured interviews are summarized in Sec. 5 of this contribution.

3 Related Work

Personal data collection and processing according to GDPR. The GDPR compliant collection and processing of personal data has especially been discussed in the legal

literature [ES18; Go18; PP18; KB20; SHS19], focusing on the principles relating to processing of personal data, the rights of data subjects, technical and organizational measures, contract data processing, international data transfer and especially on the lawfulness of data processing, e.g., the requirements for a compliant consent of a data subject. Furthermore, the relevance of pseudonymization – especially in the context of processing personal data for research – has been discussed in regard to whether or not the effort for pseudonymization is worth the outcome, e.g., if pseudonymization can protect from internal attacks and if alternative privacy measures, such as an authorization concept, can do the same [KLP19]. The chances and limitations of personal and anonymized data processing in smart city contexts in accordance with the GDPR has – to the best of our knowledge – not yet been discussed.

Personal data erasure and anonymization according to GDPR. As well as the GDPR compliant data collection and processing, the GDPR compliant erasure and anonymization of personal data has especially been discussed in the legal literature, mainly discussing the relevant GDPR requirements and strategies on compliance in different processing context, such as e.g., E-Mail or AI [JS18; DSP19; ESS19; GV20; Ca20]. To the best of our knowledge, the chances, and limitations of a GDPR compliant personal data erasure versus anonymization have not been discussed before.

Value of personal data and privacy. The value of privacy and personal data has been assessed by various researchers throughout the last two decades. Most of these publications focus on the “willingness to pay”, i.e., the amount a data subject would pay to use a privacy friendlier service and the “willingness to accept”, i.e., the amount data subjects would demand to give up their privacy (e.g. [DLA05; AJL13; MIR14; SC16; WS19]). Another one suggests establishing a marketplace that helps controlling the data subjects’ personal data and/or anonymity by a monthly statement in regard to their privacy or anonymity [Ro17]. Unlike our contribution that identifies the chances and limitations of both, personal and anonymized data processing in a smart city context as well as identifying potential for added value in a smart city context through personal and anonymized data processing from a data controllers’ point of view, they put a price tag on personal data from a data subjects’ point of view, by using various research methods (e.g. direct and indirect surveys, field and laboratory experiments) which are summarized and discussed by [Wa18; ATW16], who carried out a broad overview of the literature of the economics of privacy, starting with publications from 1967 and including the aforementioned aspects).

4 Nature, Scope, Context and Purpose of Processing as Factors of Appropriate Technical and Organizational Measures

As stated in Sec. 1, the GDPR states that the nature, scope, context, and purpose of processing need to be considered when implementing appropriate technical and organizational measures. However, the GDPR lacks a legal definition of the mentioned four terms “nature”, “scope”, “context”, and “purpose” of processing regarding

appropriate technical and organizational measures. Therefore, we carried out a legal analysis of GDPR commentaries to find out the meaning of these terms.

4.1 Nature of Processing

The nature of processing comprises several aspects. On the one hand, this refers to the consideration of the types of processing. Accordingly, it must be considered whether personal data are to be e.g., collected, organized, modified, disclosed, transmitted, or erased. Compared to the lawful erasure of personal data, e.g., the disclosure of personal data generally poses a greater risk [Go18].

Furthermore, the consideration of the nature of processing focuses the types of data. In particular, the GDPR places a high level of protection on the processing of so-called special categories of personal data (e.g., political opinions, health data). Their processing may indicate an increased risk and accordingly also entail special requirements for the technical and organizational measures to be taken [PP18; KB20].

Finally, the processing technology used must also be considered. A distinction must be made here between automated and manual, paper-based data processing. A manual form of processing is generally to be classified as less intensive than similar data processing that is automated [ES18].

4.2 Scope of Processing

The scope of processing concerns two aspects: The amount of data subjects and the amount of data processed per se. In both aspects, mass processing is generally to be assessed more critically than individual processing [KB20].

The quantity of data subjects is relevant because as the number of data subjects increases, cross-connections can be established between them, which could lead to new conclusions/knowledge for each individual data subject [PP18].

The GDPR does not specify when exactly a particularly risky amount of personal data is processed. At least if a processing of personal data creates profiles of the data subjects, it will usually be a high-risk set of personal data in terms of quantity. This is true even if the merged data individually do not have a high informative value [Go18].

If further personal data are continuously added to an existing data set, the processing time has a direct influence on the scope of processing and should therefore also be considered. Another aspect of the scope of processing is the number of potential recipients, because each added recipient naturally increases the scope of processing [PP18; ES18].

4.3 Context of Processing

The context of processing must also be taken into account when selecting technical and organizational measures. This includes considering the systems used for processing, the technical and organizational measures that are already implemented in these systems, as well as the controller's economic interests in the data processing [KB20; Go18].

Regarding the system used and already implemented measures, it is relevant whether strong measures such as encryption, pseudonymization or even anonymization can and are implemented within the processing system. While encryption and pseudonymization are technical measures under the GDPR, aiming to ensure the security of personal data without the data losing their personal references, anonymization is a technical measure that results in the loss of personal references. Therefore, the processing of anonymized data does not fall under the GDPR.

The circumstance of whether data is processed locally at the controller or on servers of a cloud computing provider is also relevant for consideration. The latter is particularly relevant if the data processing by the cloud computing provider – or any other processors for that manner – takes place in so called “third countries” [KB20; ES18].⁵

The fact that the controller has a high economic interest in the data processing will regularly argue for higher technical and organizational measures to protect the rights and freedoms of the data subjects [PP18].

4.4 Purpose of Processing

When considering the purposes of processing, two aspects are relevant: On the one hand, it must be assessed whether the data processing is intended to serve a particularly critical purpose; on the other hand, it must be taken into account whether the processing purposes may be (too) broad.

Particularly critical purposes exist, amongst others, if they are aimed at strengthening the degree of personal reference or the economic exploitation of the personal data. Thus, if a controller processes data for the purpose of de-pseudonymization or de-anonymization or if he “collects” personal data in order to resell them later, these aspects will regularly argue for higher technical and organizational measures [Go18; ES18].

With regard to purposes that are too broad, the purpose limitation – one of the data protection principles regulated in the GDPR – must be complied with [KB20].

4.5 Relevant Data Protection Principles of the GDPR

According to the GDPR, personal data shall be collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those

⁵ Third Countries are all countries outside the European Economic Area.

purposes. Personal data shall also be adequate, relevant, and limited to what is necessary in relation to the purposes for which they are processed and only kept in a form which permits identification of data subjects as long as it is necessary for the purposes for which the personal data are processed. After the purpose of processing is fulfilled, personal data needs to be either erased or anonymized. As explained within the previous four subsections, complying to these rules is relevant when taking the nature, scope, context, and purposes of processing into account while choosing appropriate technical and organizational measures. The fewer and the shorter personal data are being processed, the lower could be the costs of implementing technical and organizational measures – either in the sense of not needing to implement certain measures at all since e.g. the processing of sensitive data is being skipped and therefore the risk of data processing is lower, or in the sense that personal data is being erased or anonymized after a short period of time and therefore ongoing costs of technical and organizational measures shrink to a minimum. Thus, in this paper we want to address the value of personal data from a controller's point of view, by assessing the chances and limitations of personal and anonymized data processing as the base consideration, a controller should undergo in terms of appropriately implementing technical and organization measures. We do so in the context of data processing within Smart Cities.

As a result, taking the aspects nature, scope, context, and purpose of processing into account can be seen as the first aspect every controller should consider regarding the implementation of appropriate technical and organizational measures, because it gives him the chance to pre-determine whether at all and how long the processing of personal data is necessary and economically reasonable for the specific purpose of the data processing and data exchange.

5 Chances and Limitations of Data Processing in Smart Cities

As a direct link to the result from our legal analysis we want to answer the question to what extend personal data processing is necessary in a smart city context, what its chances are in comparison to the chances and limitations of a purely anonymized data processing, and what added value the exchange of (personal and anonymized) data can bring within a smart city context.

5.1 Chances and Limitations of Personal and Anonymized Data Processing in Smart Cities

The results of the 30 structured interviews with smart city participants (e.g., hospitals, public ambulance control centers, basic public utility providers, civil service offices, public transport providers and universities) we conducted can be summarized as follows:

1. Currently, 13/30 of the interviewees only process personal data in their specific smart city context, another 13/30 process personal and anonymized data and another 4/30 only processes anonymized data (with one interviewee stating that they would process personal data in rare cases, if absolutely necessary).
2. Amongst our interviewees, 1/30 process 1-1000 data sets⁶ (personal data and/or anonymized data) in their specific smart city context, 9/10 1001-10.000 data sets, 8/30 process 10.001-100.000 data sets, and 12/30 process over 100.000 data sets.
3. 9/30 of our interviewees estimates that the data processing carried out by them in their specific smart city context results into little to no risk for the rights and freedoms of data subject, another 9/30 estimates a medium risk, 11/30 estimates a high risk, and 1/30 could not categorize the risk, because they only process data as a contract data processor within a smart city context, so they do not need to assess the risk of data processing.⁷ The high number of high risk processing mainly results from having 10/30 interviewees from the health sector.
4. Whereas personal data processing is usually necessary in the direct contact with citizens (e.g., when billing their use of electricity, granting them a birth certificate, or treating their medical condition), smart city participants usually do not need personal data to identify the potential of improvements in their city and work towards reaching these improvements. In the latter cases, processing anonymized data would be sufficient for 23/30 of the smart city participants we interviewed. However, re-identification could be beneficial for a specific citizen in a given context, such as having identified a cure to a disease based on the anonymized data of a patient.
5. Even though a majority of 24/30 of the interviewed smart city participants already exchange data with other smart city participants, the current state of data exchange is – generally spoken – reduces to the bare minimum (e.g., a public ambulance control center lets a hospital know how many patients they will receive due to an accident, hospitals state about birth and death to the civil service offices).
6. The great potential for exchanging data between individual smart city participants is currently not being exploited. The reason for this is, among other things, great uncertainty among smart city participants with regard to where the limits of data protection law are reached in the context of data exchange. Great potential could be reached by e.g., the exchange of data between public agencies and social networking platforms to better respond to disaster alerts of any kind and find the most effective ways to alert concerned citizens to a danger or get them help quickly; data sharing between nurseries, schools, universities, public health departments, laboratories, primary care physicians' offices, and clinical research institutions to help mitigate the spread of contagious diseases and respond quickly to new findings; data

⁶ We define a “data set” as being every information that belongs or – in case of anonymization – formerly belonged to a data subject.

⁷ In case of contract data processing, the controller is responsible to assess the risk of data processing.

exchange between hospitals, medical practices, clinical research institutions and civil service offices to identify areas in communities where certain diseases are occurring more frequently, so that the cause can be researched and help can be offered.

7. A considerable amount of added value – only 2/30 of all interviewees stated that anonymized data would not have any value for them within a smart city context – for smart cities could be achieved through the exchange of anonymous data. This potential is currently not being exploited. One of the main factors for this is that stakeholders are often unable to assess when data is anonymous from a legal perspective. Concerns that were expressed by the interviewees were e.g., that with anonymization, no one really knows how exactly this is supposed to work; that when it comes to anonymization, they are still “poking around in the dark”; that most of the time, they don't know what exactly has to be anonymized and what not; that their systems do not offer data anonymization or they do not have an established process for a (GDPR-compliant) data anonymization.
8. The uncertainty among smart city participants about where the limits of data protection law are reached in the context of data exchange and about anonymization seem not to be resulting from a lack of communication with the organizations data protection officer (DPO). 17/30 of our interviewees state that the organizations DPO is always being consulted and that the recommendations of the DPO are usually being followed, some of them even stated that whatever the DPO is recommending is being implemented, that their DPO has the right to veto against a planned data processing if it would not be carried out in a compliant way, and that the organizations head implemented a process in which every new data processing activity needs to seek approval of the DPO before it can be carried out. 4/30 of the interviewees answered that their DPO is following a very pragmatic approach and that it therefore is always most sensible to include the DPO. 2/30 stated that all their data processing activities that are being carried out in a smart city context follow a process that has been predefined by their DPO, so they do not need to contact him at all.

In contrast, only 3/30 of our interviewees stated that their DPO is not always being consulted, and each 1/30 stated that they only want to implement minimal legal standard and therefore do not want to get “extraordinary” suggestions for privacy measures, and that they “can do what they want as a decision-maker”, but usually seek at least a certain level of advice from their DPO. Another 1/30 chose not to answer the question.

Another 1/30 - where the answer came from the DPO itself – stated that he sometimes needs to threaten to contact the supervisory authority if a planned data processing would not be compliant to data protection law.

9. The great potential for exchanging data is jeopardized due to the fact that there is no uniform way of exchanging data between smart city participants. Whereas 20/30

of smart city participants exchange data through various electronic channels e.g., Excel-Lists, E-Mails, or secure download portals, 10/30 of all interviewees stated that the electronic exchange of data is not possible with government institutions, such as the civil registry office, and health organizations, such as private practices.

5.2 Robustness

The answers of interviewees must be seen as estimates from the interviewees (individual) point of view, which makes it necessary to ask each interviewee how sure they are about their estimates and how generalizable they think their estimates are.

To find out how reliable the estimates of our interviewees were, we asked them how sure they are about the answers they gave in regard to the worth of personal data for their employer – namely the chances and limitations of personal and anonymized data processing and the chances of data exchange that we discussed in the interview. Among the 30 interviewees, 21/30 answered that they are sure about their estimates, 8/30 answered that they are relatively sure about their estimates and one was not able to answer the question.

We furthermore asked the 30 interviewees to what extend their answers/estimates are generalizable in relation to other smart city participants with the same data processing context (e.g., are the answers of one school generalizable to other schools or are the answers of one civil registry office generalizable to other civil registry offices). 23/30 interviewees answered that their answers are generalizable, 5/30 interviewees stated that their answers are rather generalizable, 1 stated that his answers are rather not generalizable and 3/30 (further) stated that this depends on the amount of personal data a smart city participants needs to process, if a smart city participant is working more B-2-B or more B-2-C, and what the aims of research in a smart city context are.

As a result, the amount of personal data processing can be reduced, and most goals of smart cities can be reached by processing anonymized data. Furthermore, most potential added value can be reached by (solely) exchanging anonymized data. 29/30 of our interviewees are sure or relatively sure about their estimates and 28/30 think that their estimates are generalizable or at least rather generalizable.

6 Conclusion

In this paper, we answered the question how taking into account the nature, scope, context and purposes of processing when implementing appropriate technical and organizational measures links to the chances, limitations and potential added value within a smart city context. We conclude with the results regarding the implementation of appropriate technical and organizational measures (Sec. 6.1) and regarding potential for added value in smart city contexts (Sec. 6.2).

6.1 Implementing Appropriate Technical and Organizational Measures

The GDPR regulates the obligation to implement appropriate technical and organizational measures, such as pseudonymization and encryption. Such obligations do only concern the processing of personal data. In contrast, the processing of anonymized data does not fall under the strict rules of the GDPR.

To implement measures that are appropriate, the nature, scope, context, and purposes of processing must – among other factors – be considered. Considering the nature, scope, context, and purposes of processing, means e.g., that

- the disclosure of personal data generally poses a great(er) risk than other processing activities and therefore (generally) implicates higher technical and organizational measures;
- the more personal data and the more data subjects are exposed to a data processing, and the more sensitive personal data are, the higher technical and organizational measures (generally) need to be;
- a broad purpose – such as defining the processing purpose very generically as “research” – generally poses a greater risk than a narrowly defined processing purpose and therefore implicates higher technical and organizational measures.

Therefore, answering the question of chances and limitations of personal and anonymized data processing as well of potential added value through personal and data exchange should be considered a chance to “pre-determine” whether at all and how long the processing of personal data is necessary and economically reasonable for the specific purpose of the data processing and data exchange. Furthermore, the amount, the sensitivity, and the extend of exchanging of personal data is directly linked to the amount of technical and organizational measures that is appropriate to the relevant data processing.

Structured interviews with 30 smart city participants showed that the processing of personal data within a smart city context could be drastically reduced so that – as a result – the necessity to implement (a large amount of) technical and organizational measures in accordance with the GDPR could be heavily reduces: Even though 26/30 of all interviewees stated to currently process either only personal, or personal and anonymized data within the smart city context, 23/30 of all interviewees stated that processing anonymized data would be sufficient. Furthermore, most potential value can be reached by exchanging anonymized data (only 2/30 of all interviewees stated that anonymized data would not have any value for them within a smart city context).

6.2 Creating Added Value in Smart Cities

Through the 30 structured interviews with smart city participants, we also showed that there is a great potential for exchanging data between individual smart city participants, such as the exchange of data between

- public agencies and social networking platforms could improve the respond to disaster alerts of any kind;
- nurseries, schools, universities, public health departments, laboratories, primary care physicians' offices, and clinical research institutions could help to mitigate the spread of contagious diseases; and
- hospitals, medical practices, clinical research institutions and civil service offices could improve the identification of areas in communities where certain diseases are occurring more frequently, so that the cause can be researched, and help can be offered.

These – and other – potentials to add value to data within a smart city context, however, is currently not being exploited. The reason for this is, among other things, great uncertainty among smart city participants regarding data protection law and anonymization. To enable smart city participants to extract value of existing data by exchanging them in between smart city participants, a close collaboration of jurisprudence and computer science is necessary.

7 Acknowledgements

This research work has been funded by the German Federal Ministry of Education and Research (BMBF) and the Hessian Ministry of Higher Education, Research, Science and the Arts within their joint support of the National Research Center for Applied Cybersecurity ATHENE as well as by BMBF within their support of AScore (FKZ 13N15663) as a consortium project within the framework of the special programme Civil Security – Corona Pandemics (www.sifo.de).

References

- [AJL13] Acquisti, A., John, L. K. and Loewenstein, G. (2013). “What is privacy worth?” *Journal of Legal Studies* 42 (2), 249-274.
- [ATW16] Acquisti, A., Taylor, C. and Wagman, L. (2016). “The Economics of Privacy.” *Journal of Economic Literature* 54 (2), 442-492.
- [Ca20] Cabral, T. (2020). “Forgetful AI: AI and the Right to Erasure under the GDPR.” *European Data Protection Law Review* (6), 378-389.
- [DLA05] Danezis, G., Lewis, S. and Anderson, R. (2005). “How much is location privacy worth?” *Proceedings of the Workshop on the Economics of Information Security Series*, 1-13.
- [DSP19] Durmus, E., Selzer, A. and Pordesch, U. (2019). „Das Löschen nach der DSGVO – Eine Diskussion der datenschutzkonformen Umsetzung bei E-Mails.“ *Datenschutz und Datensicherheit* (43), 786-791.

- [ES18] Ehmann, E. and Selmayr, M. (2018). *Datenschutz-Grundverordnung, Kommentar*. 2nd Edition. München: C.H.Beck.
- [ESS19] Enzmann, M., Selzer, A. and Spsychalski, D. (2019). „Data Erasure under the GDPR – Steps towards Compliance.” *European Data Protection Law Review* (5), 416-420.
- [GBP19] Gassmann, O., and Böhm, J., and Palmie, M. (2019). *Smart Cities: Introducing Digital Innovation to Cities*. 1st Edition. Bingley: Emerald Publishing.
- [Go18] Gola, P. (2018). *Datenschutz-Grundverordnung, Kommentar*. 2nd Edition. München: C.H.Beck.
- [GV20] Groos, D. and van Veen, E. (2020). “Anonymised Data and the Rule of Law.” *European Data Protection Law Review* (6), 498-508.
- [JS18] Jandt, D. and Steidle, R. (2018). *Datenschutz im Internet*. 1st Edition. Baden Baden: Nomos.
- [KLP19] Kohlmayer, F.; Lautenschläger, R. and Prasser, F. (2019). „Pseudonymization for research data collection: is the juice worth the squeeze?” *BMC Medical Informatics and Decision Making* (19), 1-7.
- [KB20] Kuehling, J. and Buchner, B. (2020). *DatenschutzGrundverordnung / Bundesdatenschutzgesetz, Kommentar*. 3rd Edition. München: C.H.Beck.
- [MIR14] Morando, F., Iemma, R. and Raiteri, E. (2014). “Privacy evaluation: What empirical research on user’s valuation of personal data tells us.” *Internet Policy Review* 3 (2), 1-11.
- [PP18] Paal, B. and Pauly, D. (2018). *Datenschutz-Grundverordnung und Bundesdatenschutzgesetz*. 2nd Edition. München: C.H.Beck.
- [Ro17] Robinson, S. C. (2017). “What’s your anonymity worth? Establishing a marketplace for the valuation and control of individuals’ anonymity and personal data.” *Digital Policy, Regulation and Governance* 19 (5), 353-366.
- [Se21] Selzer, A. (2021). “The Appropriateness of Technical and Organizational Measures under Article 32 GDPR.” *European Data Protection Law Review* 7 (1), 120-128.
- [SC16] Sidgman, J. and Crompton, M. (2016). “Valuing personal data to foster privacy: a thought experiment and opportunities for research.” *Journal of Information Systems* 30 (2), 169-181.
- [SHS19] Simitis, S., Hornung, G. and Spiecker, I. (2019). *Datenschutzrecht*. 1st Edition. München: C. H. Beck.
- [SLF11] Su, K. and Li, J., and Fu, H. (2011). “Smart City and the Applications.” 2011 International Conference on Electronics, Communications and Control (ICECC), 1028-1031.
- [Te20] TeleTrusT (2020). Guideline “State of the art” – technical and organisational measures.” URL: https://www.teletrust.de/fileadmin/docs/fachgruppen/ag-stand-der-technik/2020-02_TeleTrusT_Gu_deline_State_of_the_art_in_IT_security_EN.pdf (visited on 4 January 2021).

- [VB17] Voigt, P. and von dem Bussche, A. (2017). The EU General Data Protection Regulation (GDPR). 1st Edition. Heidelberg: Springer.
- [Wa18] Wagner, A., Wessels, N. Buxmann, P. and Krasnova, H. (2018). „Putting a Price Tag on Personal Information – A Literature Review.” Proceedings of the Hawaii International Conference on System Sciences, 3760-3769.
- [WS19] Winegar, A. G. and Sunstein, C. R. (2019). “How much is Data Privacy Worth? A Preliminary Investigation.” Journal of Consumer Policy (42), 425-440.

The use of web tracking technologies and its compliance with the consent-requirement

Sarah Diel^{1,2}, Wendy Heereman^{1,2}, Dominik Spychalski^{1,2}

Abstract: Every day, numerous web tracking tools are used for advertising or as a basis for developing or improving products. The use of these technologies has a significant impact on the privacy of individuals, as they provide service providers with personal data and profiles can be created. From a user's perspective, it is not always transparent which kind of data is collected and how it is used and transmitted through an interconnected ad ecosystem with a number of participants with primarily monetary interests. This paper analyzes the use of tracking technologies in current practice and its compliance to the consent-requirement of the GDPR.

Keywords: Online advertising, Web tracking, Consent mechanisms, Privacy, Privacy policies

1 Introduction

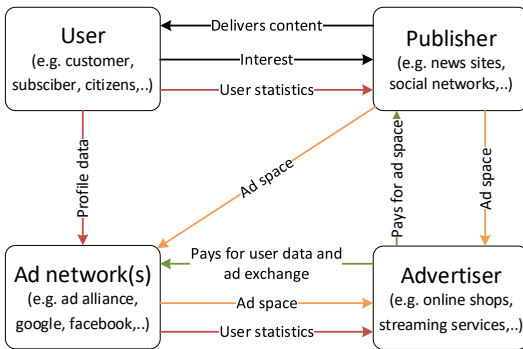
Advertising is the state of the art for increasing sales. In contrast to the offline business world, digital advertising is way more personalized. To paint a comprehensive and complete picture of a specific user or audience as well as to carry out personalized digital advertising the collection of user data as well as the re-identification of users while they move across the internet is required. Both can be performed based on the collection of data through link-able session identifiers, unique and distributed identifiers stored in e.g. cookies as well as measurements of the users' behavior and their used system (web tracking and advertising tools) and the potential sharing of the collected data with one or many global advertising networks and their participants as well as all connected publishers and advertisers (ad ecosystem).

In an ad ecosystem an *advertiser* buys ad space either from an *ad network* or from a *publisher* (green arrows in figure 1) to deliver and display its ads on one or many *publisher* sites. The *publisher* delivers content to a *user* (black arrows in figure 1) and sells ad space to an *advertiser* (orange arrows in figure 1). An *ad network* is an association of different tracking providers and their technologies that connects advertisers with publishers. Even

¹ Fraunhofer-Institute for Secure Information Technology, Cloud Computing, Identity & Privacy, Rheinstr. 75, 64295 Darmstadt, Germany, {firstname}.{lastname}@sit.fraunhofer.de

² This research work has been funded by the German Federal Ministry of Education and Research and the Hessen State Ministry for Higher Education, Research and the Arts within their joint support of the National Research Center for Applied Cybersecurity ATHENE.

though various ad networks exist, they are interconnected by so called *ad exchanges*. Once the user data has been collected through the tracking technologies, it is channeled into the ecosystem and flows independently over connected entities and in a manner that is neither transparent nor comprehensible for a user (red arrows in figure 1) [Ko21].



Since the processing of personal data in an ad ecosystem has a significant impact on the privacy of users and can lead to malicious ad serving, data breaches as well as to social risks, such as the influence of the users' purchasing decisions or political views through personalized ads and news based on their interests and values, it requires the processing to be legally compliant.

Figure 1: Abstract overview of an ad ecosystem, its participants and their economic relationships between each other.

2 Legal Analysis of consent tools

To ensure a legally compliant processing of personal data within the ad ecosystem certain privacy requirements must be met. These requirements arise in particular from Art. 5 (1) GDPR and require the processing of personal data to be *fair*, *transparent* and *lawful*.

2.1 Consent tools

Since the European Court of Justice (ECJ) has specified the requirement of lawfulness to the extent that data processing through tracking tools usually requires the data subject³ to actively consent [Eu19, margin number 44], a large number of websites have implemented a tool to obtain user consent (so-called consent banner).

A consent banner generally appears when a user accesses the website of a publisher for the first time. On level 1 of the consent banner the user is able to *accept* or sometimes *decline* all tracking elements with just one click. A differentiated configuration is possible via *settings*. Here, there are typically purpose-oriented groups of tracking technologies, each consisting of at least one tracker. After personalizing the settings, a user saves them with a click on *save settings*. The number of clicks required for consent personalization depends on the amount of levels the consent banner provides (click-depth) and the number of connected trackers

³ According to Art. 4 no. 1 GDPR the data subject is the identified or identifiable natural person to whom the personal data relate. In the terminology of the ad ecosystem, the user usually is the data subject.

per purpose-group. However, it should be noted that this is an abstract and generalized representation of consent banners, as they are used in most cases.

2.2 Lawfulness of the Processing

According to Art. 5 (1) (a) GDPR, the processing of personal data - and thus the use of tracking mechanisms - requires a legal basis. This usually requires the users' consent in terms of Art. 6 (1) (a) GDPR [Eu19, margin number 61-64].⁴

Pursuant to Art. 4 no. 11 GDPR a consent is *'any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her'*. The requirements for consents are specified in Art. 7 and Recital 42 of the GDPR.⁵ Accordingly, it must also be possible to change or withdraw consent in a way that is as easy as giving consent (Art. 7 (3) GDPR).

With regard to certain aspects identified as privacy critical in their implementation, the mentioned requirements are discussed below.

Freely given - Paying with money vs. paying with data Many sites give users the choice to pay or to consent to the tracking-related processing of their data, which is equivalent to paying indirectly with data. Since consent must be freely given, which is considered to be given if the user has the possibility to refuse or revoke consent without any disadvantages (Recital 43), it is questionable whether, in this case, consent can be considered freely given. Therefore, according to Art. 7 (4) GDPR, it shall be taken into account if the performance of a contract or a service is conditional on the consent.

Since not every website has the same purpose, this question should not be answered in general for every website. News sites, for instance, aim to trade information. For this, they usually demand money in return. Online, instead of paying money, they offer the possibility to pay with data. An online store, on the other hand, aims to sell goods. In this case, it is not the information provided on the website that is sold, but the goods users can buy. Thus, the user would have to pay directly or indirectly to access the website as well as to receive the good. Therefore, whether consent in this case can be considered freely given should depend on the website's purpose. The authors consider it only lawful when the website's content itself usually only is provided in consideration for money.

⁴ According to Art. 5 (3) of the ePrivacy directive consent is not required for cookies that either are used *'for the sole purpose of carrying out or facilitating the transmission of a communication over an electronic communications network'* or are *'strictly necessary in order to provide an information society service explicitly requested by the subscriber or user'*.

⁵ Even if the Recitals are not binding directly, they can be used for interpretation of the legislator's intent.

Active and specific consent For the consent to be valid [La20, margin number 37-54], it must also unambiguously indicate the user's wish to signify agreement to the processing. This requires a clear affirmative action and thus, an active rather than a passive behavior [Eu19, margin number 52-56].⁶ Hence, in context of consent banners, it requires the user to opt-in to the processing. Yet, some consent banners have cookies preselected - sometimes only technically necessary cookies, whereas it is not possible to remove the check mark or to turn the slider to *off*, sometimes more categories, e.g. preference cookies or functional cookies. According to Art. 5 (3) of the ePrivacy directive, strictly necessary cookies, which include technically necessary cookies as well as some preference cookies do not require consent but can be set without the users' consent, e.g. based on legitimate interests. Strictly necessary preference cookies include, for instance, the selection of a language that is actively set through the user [Ar12, p. 8]. However, according to the Irish Data Protection Authority, many publishers classify cookies as *necessary* or *strictly necessary* even though the stated function of the cookie is not [Da20b, p. 3].

Additionally, the consent and processing purpose must be *specific* and the request of consent must be requested for each purpose individually [Eu20, p. 14]. Therefore, the Irish Data Protection Authority suggests to outline on the banners first level that the consent for the use of tracking tools is requested for specific purposes and to provide more information about the categories of tracking tools with the option to accept or decline them on the second level [Da20a, p. 8]. Thus, the distinction between cookie categories and the request for consent to each of these categories separately is recommended. Additionally, it is recommended to allocate the cookies to these categories.

Informed consent According to Art. 4 no. 11 GDPR user consent must also be obtained in an informed manner. Except for the identity of the controller, the purpose of each of the processing operations (Recital 42) and the right to withdraw consent (Art. 7 (3) GDPR), the GDPR does not specify which information must be provided. However, according to the European Data Protection Board, the data subject must also be aware of the type of data that is processed, the intention of an exclusively automated decision and the intention of a transfer of data to third countries [Eu20, p. 15-16]. Since the use of tracking mechanisms and the potential channeling of data into the ad ecosystem can lead to the transfer of data outside the European Economic Area (EEA), pursuant to Art. 49 (1) (a) GDPR, additionally, it must be informed about the possible risks of transfers. In context of consent banners, the information must be provided on the first level. However, regarding the transfer to third parties the user must only be informed about the type of cookies used (first- or third-party) on the first level. The actual listing of third-party cookies is sufficient on the second level [Ag19, p. 22].⁷

⁶ This also follows from Recital 32, according to which '*silence, pre-ticked boxes or inactivity should not therefore constitute consent*'

⁷ In addition to the information provided when obtaining consent, according to Art. 5 (1) (a), 12 ff. GDPR for the processing to be transparent further information must be provided, inter alia, the legal ground of the processing and the duration of storage.

Whilst some websites provide all of the required information on the consent banners' first level, many websites provide the main information, such as the purposes of the processing, only on the second level (if they provide it at all). The information about possible risks of transfers of data outside the EEA sometimes even is found below the second level - even though consent to the transfer can already be given on the first level. Regarding the purposes, additionally, some websites inform about the processing in such a way that the users get the impression the processing through tracking mechanisms has only positive effects on them and that rejecting or withdrawing consent would lead to the impairment of their interests. This can lead to the users not seeing the need to reject consent or to enforce their rights.

Possibility to withdraw consent According to Art. 7 (3) GDPR withdrawal of the consent must be as easy as submitting it. Therefore some websites provide a link to the cookie settings either in their footer or hidden somewhere in the privacy notice. However, according to the Irish Data Protection Authority most websites lack on withdraw methods [Da20b, p. 7].

If withdraw methods are given within the settings, a user is able to withdraw consent by using toggle switches or removing checkmarks. Since the mechanisms use the same HTML-buttons, to withdraw is as easy as to give consent. However, if the withdraw mechanism is hidden in the privacy notice, it requires more effort (and clicks) for the user. It could even lead to the impression that there is no possibility to revoke consent. On the other hand, it is state of the art to address privacy related information, including the description of how to withdraw consent, within the privacy notice. Furthermore, since the consent banner usually informs about the location of the mechanisms, it can be considered as lawful. However, it is more privacy-friendly - and therefore recommended - to link to the cookie settings in the website's footer.

2.3 Design of banners: Fairness of the processing

In addition to the necessity of a legal basis for the processing, according to Art. 5 (1) (a) GDPR, personal data shall be processed fairly. The principle of *fairness* can be interpreted as a catchall element covering processing scenarios that cannot be qualified as lawful even though they do not violate a specific legal requirement [KB20, Art. 5 GDPR margin number 17]. Whilst some consent banners focus on being privacy-friendly, other consent banners use colors, fonts or terminology which lead to the user being induced to accept all cookies. The *accept all*-option, for instance, often is written bold and colored in such a way that the user instinctively is tempted to click on it. The *open the settings*-option as well as the *reject all*-option, on the other hand, usually is written in a very small font size or is colored in such a way that it can hardly be recognized. Sometimes these options even can only be found within the explanatory text. Hence, even though the design does not violate a specific legal requirement, it must still be considered unfair in terms of Art. 5 (1) (a) GDPR.

3 Conclusion

The use of tracking tools carries risks for users. Once a user visits a website and consents to the processing of its data through tracking tools, the data is collected, shared with multiple recipients and may end up in a complex ad ecosystem. For the average user the extent to which the data may be disclosed as well as the risks that result from the disclosure are unforeseeable. Current case law and legal provisions require that the user actively consents in the processing through tracking mechanisms. However, the legal requirements for a valid consent often are not met in practice.

References

- [Ag19] Agencia Española de Protección Datos: A Guide on the use of cookies, 2019, URL: <https://www.aepd.es/sites/default/files/2020-09/guia-cookies-en.pdf>.
- [Ar12] Article 29 Data protection working party: Opinion 04/2012 on Cookie Consent Exemption, 00879/12/EN WP 194, 2012.
- [Da20a] Data Protection Commission of Ireland: Guidance Note: Cookies and other tracking technologies, 2020, URL: <https://www.dataprotection.ie/sites/default/files/uploads/2020-04/Guidance%20note%20on%20cookies%20and%20other%20tracking%20technologies.pdf>.
- [Da20b] Data Protection Commission of Ireland: Report by the Data Protection Commission on the use of cookies and other tracking technologies, 2020, URL: <https://www.dataprotection.ie/sites/default/files/uploads/2020-04/Data%20Protection%20Commission%20cookies%20sweep%20REVISED%2015%20April%202020%20v.01.pdf>.
- [Eu19] European Court of Justice: Case C–673/17 Bundesverband der Verbraucherzentralen und Verbraucherverbände v Planet49 GmbH, ECLI:EU:C:2019:801, 2019.
- [Eu20] European Data Protection Board: Guidelines 05/2020 on consent under Regulation 2016/679, 2020, URL: https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines_202005_consent_en.pdf.
- [KB20] Kühling, J.; Buchner, B.: Datenschutz-Grundverordnung. Verlag C.H. Beck, 2020.
- [Ko21] Koop, M.: Preventing the Leakage of Privacy Sensitive User Data on the Web, Diss., Universität Passau, 2021.
- [La20] Landgericht Rostock (3. Zivilkammer): Urteil vom 11.08.2020 – 3 O 762/19, Aug. 2020.

Achieving Facial De-Identification by Taking Advantage of the Latent Space of Generative Adversarial Networks

Raphael Antonius Frick¹ Martin Steinebach²

Abstract: The General Data Protection Regulation (EU)2016/679 passed by the European Union prohibits any data collection and processing that was conducted without the consent of the individuals involved. Especially images showing faces are often subject to these regulations and therefore, either need to be removed or anonymized. Early approaches however were often troubled by strong visual artifacts. In this work, we propose a novel anonymization pipeline that generates a proxy face for a group of individuals by taking advantage of the semantics of the latent space of generative adversarial networks. Experiments have shown that by following a k -same approach and utilizing different clustering techniques, privacy for the individuals involved can be greatly enhanced, while preserving important facial characteristics.

Keywords: Facial De-Identification, k -Anonymity, Privacy, Generative Adversarial Networks, Facial Image Data

1 Introduction

The General Data Protection Regulation (EU)2016/679 passed by the European Union highly restricts the distribution and processing of data connected to personal information. Personal data can only be collected and processed in those cases, in which the data subject has given their consent or where it is justified in accordance with the regulation. Sharing datasets with third party organizations, such as research groups, either require additional statements of consent or the removal of identifiable data. As a result, numerical data and textual labels revealing personal information have been subject to pseudonymization for several years already. The data is modified in such a way, that they loose any reference that might lead to identifying the individual from whence the data originated.

Nowadays, image data holds great significance in datasets. They are used in multiple ways, e.g. to train AI-systems of self-driving cars that can distinguish between road signs and pedestrians. Images that are subject to privacy regulations usually show faces. For these, not many approaches exist that increase their privacy without changing the semantics of the image too much. For example, it is still common practice to blur or pixelize the entire face in an image, which results in loosing all its characteristics. Methods based on deep learning often anonymize a face to such an extend that even general characteristics are not recognizable.

¹ Fraunhofer SIT, Rheinstraße 75, 64295 Darmstadt, Germany raphael.frick@sit.fraunhofer.de

² Fraunhofer SIT, Rheinstraße 75, 64295 Darmstadt, Germany martin.steinebach@sit.fraunhofer.de

In this paper we propose an approach to achieving privacy of facial image data by trying to solve the facial de-identification problem while preserving important facial attributes. In other words: we provide privacy protection for the individual shown in an image without reducing the information value of the image more than necessary. Age, sex, color of skin or hair are typical attributes preserved by our approach.

This approach can be applied in various scenarios. The advantage of providing anonymity and as much information as possible at the same time is of interest whenever screening of groups is of relevance. In a criminal investigation scenario this can help to enable privacy-preserving forensics: groups of suspects can be provided to third party experts with anonymized data and portraits without masking e.g. the general background of the suspects. Male and female or young and old groups can quickly be recognized as anonymization will not mask these general attributes if all members of a group feature the same characteristics.

We deploy an anonymization pipeline, which takes a dataset containing facial image data as an input and returns anonymized versions of its data by utilizing several clustering strategies and by taking advantage of the latent space of generative adversarial networks.

The remainder of the paper is structured as follows: in section 2 introductions to privacy metrics and generative adversarial networks are given. The related works are reviewed in section 3, and then our main contributions are presented in section 4, followed by a showcase of the experimental results in section 5. The paper concludes in section 6 with a discussion on said results and with proposals for future work.

2 Background

In the following, an introduction to k -Anonymity, a common privacy concept, and an insight to generative adversarial networks and their latent space is given.

2.1 k -Anonymity

A dataset is said to fulfill k -Anonymity [Sw02b] when each data entry representing an individual cannot be distinguished from the data entries of at least $k - 1$ individuals. In order to assure that this property holds for a given dataset, two methods have been applied to datasets in the past [Sw02a]: *suppression* and *generalization*. Suppression removes the data yielding critical information, resulting however in overall information loss. With respect to facial images, suppression is achieved by heavily blurring the face or overlapping it with a black geometric shapes. Generalization tries to preserve most of the semantics by exchanging the data specific to one individual with data containing generalized information for the whole group of k -individuals.

In this paper, we propose an anonymization technique, which follows an approach based on generalization.

2.2 Latent Space of Generative Adversarial Networks

Proposed in 2014 by Ian Goodfellow et al. [Go14], *Generative Adversarial Networks* have gained a lot of popularity in the recent years. They are generative models, which allow inter alia the synthesis of realistic unseen facial images. These models consist of two networks: a generator and a discriminator. Both compete against each other in a *zero-sum game*. The objective of the generator is to generate synthetic images from random input noise, also known as the *latent vector*, so that the discriminator network is not able to tell apart generated data from authentic data. By receiving feedback from the discriminator, the generator continuously refines the input noise, whereby it learns to map latent vectors and its corresponding generated output to the n -dimensional *latent space*. In the latent space, common vector-arithmetic apply. For example, it is possible to add, subtract and interpolate between two or more latent vectors. Furthermore, points which are close in the latent space refer to outputs that are usually also similar to each other.

Well known implementations of these type of neural networks are *CircleGAN* [SC20], *BigGAN* [BDS19] and *StyleGAN* [Ka20]. However in this work, we utilize *AnyCost GAN* [Li21] a state-of-the-art generative adversarial network which allows for a fast and precise face image synthesis.

3 Related Works

In the past, several methods were proposed that try to enhance the privacy of the individuals inside facial image data.

Early works revolved around the application of *suppression-techniques* [BEG00], such as *blurring* and *pixelation*. Other approaches took advantage of *k-same selection* in order to ensure k -Anonymity. In these works, a proxy image representing k -individuals was created. Some works performed a *Principal Component Analysis* on a group of face images, while others calculated the average of the images pixel-values in order to form the proxy face image, here denoted as meta-face, from a set of face images [HS96; NSM05]. The images resulting from these methods are often troubled by strong visual artifacts. Thus, newer approaches take advantage of the advancements in machine-learning made in the last years. Recent papers [LL19; Me17] proposed approaches based on generative adversarial networks. Both proposed methods train specialized generative adversarial models. The GAN suggested in [LL19] takes a face image as input and only alters its attributes. Furthermore, their method deploys adversarial perturbation in order to decrease the likelihood of an image being identified by facial recognition systems. Since the attributes are only slightly changed and anonymization is based on the face image of a single individual, it is still possible for humans to identify the person. Moreover, the background stays unaffected which may lead to exposing the identity. The model in [Me17] applies a *k-same* approach, that first clusters a given dataset into partitions of size k based on their visual similarity in the pixel space. Then, it generates a proxy image using a generative model by incorporating information about the facial characteristics of the individuals in each cluster. In comparison

to the approach that is described in this paper, those methods do not take advantage of the semantics of the latent space of a pre-trained generative adversarial network. Training a generative model anew on a diverse dataset can be cost intensive and they are usually not able provide as good results in terms of image quality as pre-trained state-of-the-art models.

The privacy aspect of forensic investigations like addressed in our introduction has only rarely be mentioned in the literature. Srinivasan et al. [Sr06] introduce behavioral policies for preserving privacy during an investigation. Adams [Ad08] addresses the compliance of forensics tool for investigations in the United States, including the prevention of privacy breaches. Stahlberg et al. [SML07] explore the privacy risks of handling databases in investigation. Peter et al. [Pe12] discuss handling images in an investigation by partial encryption.

4 Proposed Approach

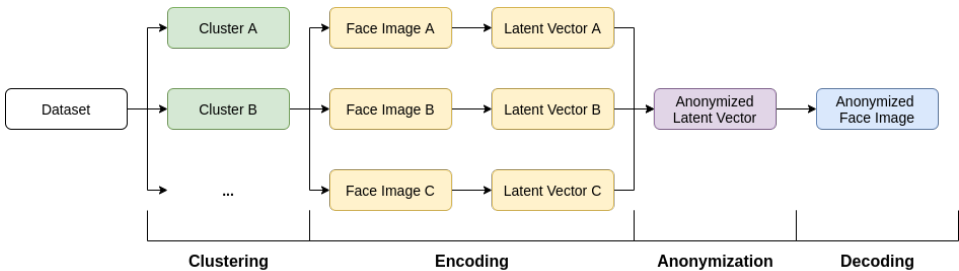


Fig. 1: Visualization of the Anonymization Pipeline

In this section, we present our anonymization pipeline, which is also visualized in Figure 1. The approach follows a *k-same* approach, meaning that a meta-face representing the face images of k -individuals is synthesized and exchanged with the face images that were used in the process. The pipeline consists of the following four stages: *data clustering*, *image encoding*, *anonymization* and *latent vector decoding*.

4.1 Clustering

At first, the dataset is divided into disjoint partitions of minimum size k . We propose three types of clustering strategies: *random clustering*, *similarity-based clustering* and *attribute-based clustering*. The results of clustering an example image using the three techniques is showcased in Figure 2.

Random Clustering: The dataset is randomly sorted and k samples are taken from it to build each cluster not taking any semantics of the given images into account.

Similarity-Based Clustering: In order to retrieve the similarity between a pair of facial images, the face images are transformed by a neural network into an abstract vector-representation. The similarity of each pair of sample images can then be estimated by calculating the *cosine similarity* or *euclidean distance* based on the newly computed representations.

Attribute-based Clustering: This clustering strategy tries to divide the dataset into groups sharing the same facial attributes. For this purpose, a classifier is used which identifies the presence of face attributes, resulting in a feature vector that is then fed into a model conducting the clustering.

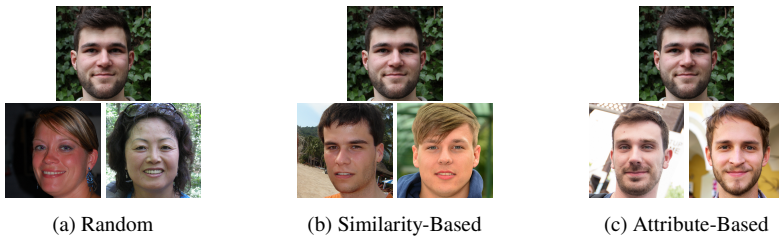


Fig. 2: Clustering Results of an Example Image using $k = 3$

4.2 Encoding

Once the splitting of the data into the individual groups has been completed, the face images in each cluster need to be transformed into latent vectors, as the anonymization is conducted on the latent vectors of the latent space. To achieve this, we take advantage of a *latent space encoder*. The encoder tries to solve the *GAN-inversion* problem, meaning that it tries to find the latent vector that corresponds to a given input image. However, this estimation is often prone to information loss. In the most cases, the found latent vector may not always be able to replicate all the facial details of the input face image. This is especially the case for the background of the image. Nonetheless, in regards to our task of anonymizing face images, this information loss can be neglected as it does not have any negative impact on the process. The encoding is then executed for each face image of a cluster.

4.3 Anonymization

In this phase of the pipeline the actual anonymization is performed based on the computed latent vectors. Common vector-arithmetic also apply for latent vectors of a latent space trained by a generative adversarial network. In this paper, we deploy two merging strategies: a *mean-based* and a *median-based* approach. The mean-based approach calculates the mean of k -latent vectors, whereas the median-based approach computes the element-wise median between each of the k -latent vectors. Both merging strategies result in a face

image, incorporating attributes of the faces which were used for the synthesis. Since several elements of the feature vector combined can be responsible for one single facial attribute, the median-based images will only differ slightly from the images constructed with the mean-based merging strategy. An example showcasing this behavior is displayed in Figure 3. Unlike previous methods, this approach also introduces changes to the background of the given image. This results in an overall better privacy, as single individuals could be identified if only the faces were to be modified.

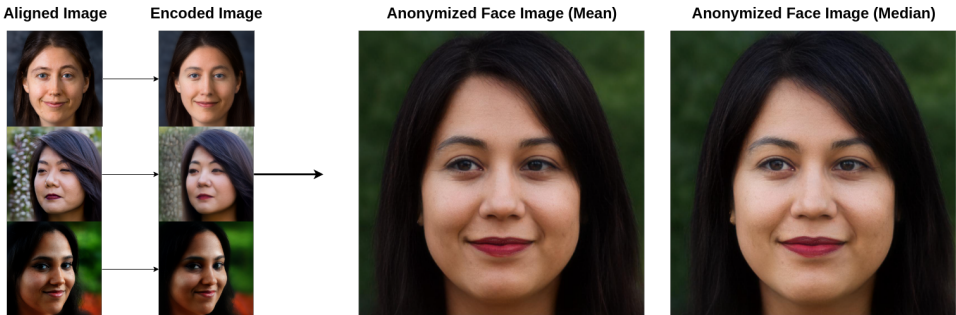


Fig. 3: Example of an anonymization conducted using $k = 3$ and showcasing both merging strategies

4.4 Decoding

As the last step, the latent vector representing the anonymized meta-face is transformed back into the color-space. The anonymized image is then replaced with each face image of the corresponding cluster, resulting in the finalized anonymized dataset.

5 Experiments

5.1 Setup

The anonymization framework was implemented in *Python*. The faces were synthesized using *AnyCost GAN* [Li21], a generative adversarial network created by Adobe Research. *Deepface*, a state-of-the-art face-recognition and identification framework created by Sefik Ilkin Serengil [SO20] was used for the re-identification tests during evaluation. Furthermore, it was used for transforming the face image into a vector representation, which was taken advantage of in order to cluster the face images by their similarity. Tencent’s implementation of the *FAN-Face Attribution Network* [He18] was utilized for the attribute-based clustering as well as for judging how many facial attributes were preserved in the meta-face after anonymization. It is able to extract 40 facial attributes from a given input image, ranging from the *sex* and *age* to the *shape of the face*. It also includes the detection of presence of *makeup* and *facial ornaments*. In order to force the size of each cluster to a fixed size of k ,

the python package *size-constrained-clustering* created by Jing Wang [Wa] was adopted. It contains a size-constrained *KMeans*-clustering technique, that tries to solve the *minimum cost flow* problem.

The application was executed on a computer having an *Intel i7 8700K @3.7 GHz* and a *NVIDIA RTX 2080Ti* installed, as well as access to *32GB RAM*.

5.2 Evaluation

The experiments were conducted on the *Flickr-Faces-HQ* [KLA19] dataset. The initial dataset contains around 210,014 image files, out of which only 1000 were considered for testing during evaluation. The dataset features peoples faces of different age, sex and color. As for the cluster size k , the values 3, 5, and 7 were considered during evaluation.

Since the *k-same* approach automatically fulfills the k -Anonymity requirement, we try to measure the privacy and facial feature preservation of multiple configurations in the following three test cases. Each configuration hereby consists of the cluster size k , the clustering strategy as well as the adopted merging strategy.

Test A: In the first two test cases we assume, that an adversarial has access to the anonymized dataset as well as to the FFHQ dataset. In this particular case they would like to identify, if a face of the FFHQ dataset was part of the initial non-anonymized dataset.

In order to measure the privacy preservation, a face recognition task is simulated by computing the similarity between an anonymized face and the face images in the FFHQ-dataset using the vector representation provided by *FaceNet*. For our analysis, we calculate the average euclidean distance between anonymized images and the images that were indeed used during the synthesis of the meta-face and define it as our privacy metric.

A low distance between the anonymized face and a sample image of the FFHQ dataset, suggests, that the image of the FFHQ dataset was part of the initial dataset. Thus, the configuration that maximizes the average *euclidean* distance bears the greatest privacy.

The test results are showcased in Table 1. Since random clustering does not take any similarities in regards to facial attributes into account during clustering, it results in a much better privacy than an anonymization conducted on clusters using an attribute-based or similarity-based approach. In addition, with increasing value k , the privacy properties of the dataset also increases. While the differences between face images created with a mean-based merging strategy approach only seems to slightly differ from the ones using a median-based approach, the median-based merging strategy happens to yield better results in achieving privacy.

Test B: The second test corresponds to the inverted case of *Test A*: this test case assumes it is known that the FFHQ dataset was the initial dataset. Thus, it is of interest, whether that

k	Random (Mean)	Random (Median)	Attribute-based (Mean)	Attribute-based (Median)	Similarity-based (Mean)	Similarity-based (Median)
3	3.54142245	3.54975472	3.27548274	3.28725609	2.93203361	2.94830903
5	6.15211511	6.15331176	5.70569627	5.71658666	5.0629013	5.07446723
7	8.79226881	8.78263694	8.27201167	8.27769541	7.29780857	7.27766072

Tab. 1: Results of *Test A* for different configurations. The values correspond to the computed average euclidean distance. A greater value indicates better privacy. Values displayed in bold indicate, which merging strategy resulted in the best privacy preservation for each cluster configuration (random, attribute-based and similarity-based).

knowledge can be exploited to identify which image of the FFHQ dataset contributed to the creation of the meta-face. Hence, this time the average euclidean distance between an image of the FFHQ dataset and the images of the anonymized dataset is computed and analyzed.

k	Random (Mean)	Random (Median)	Attribute-based (Mean)	Attribute-based (Median)	Similarity-based (Mean)	Similarity-based (Median)
3	1.16512798	1.1678693	1.07435834	1.07822	0.96463906	0.96999367
5	1.20581456	1.20604911	1.11261077	1.1147344	0.99232866	0.99459558
7	1.22212536	1.22078653	1.1332656	1.13404427	1.00709758	1.00431718

Tab. 2: Results of *Test B* for different configurations. The values correspond to the computed average euclidean distance. A greater value indicates better privacy. Values displayed in bold indicate, which merging strategy resulted in the best privacy preservation for each cluster configuration (random, attribute-based and similarity-based)

The results are displayed in Table 2. In comparison to the results of first test (Table 1), the average euclidean distances are drastically lower. This indicates that knowledge on the initial dataset can contribute to re-identifying the relationship between a sample image and anonymized face images. Those configurations, that preserved privacy the most during the first test had the same results in the second test. Hence, it is advisable to apply a median-based merging strategy in combination with either random clustering or an attribute-based approach.

Test C: In the last test, it was assessed, how well the facial attributes are retained after anonymization. We calculated the average *hamming*-distance between each feature-vector of a meta-face and its corresponding image samples from the FFHQ dataset. The feature vector was extracted using the *FAN-Face Attribution Network* which computes a one-dimensional vector of values 0 and 1 depending on the observed attributes.

According to the results showcased in Table 3, the attribute-based clustering technique is able to retain the characteristics of a face a lot better than the other clustering methods. As the amount of data points in a cluster increases, the preservation weakens. This is to be expected, as it may not always be possible to create a group of images of size 5 or 7 that

k	Random (Mean)	Random (Median)	Attribute-based (Mean)	Attribute-based (Median)	Similarity-based (Mean)	Similarity-based (Median)
3	0.1812	0.1816	0.1052	0.1055	0.1530	0.1520
5	0.1995	0.2007	0.1101	0.1090	0.1637	0.1647
7	0.2076	0.2084	0.1117	0.1138	0.1619	0.1632

Tab. 3: Results of *Test C* for different configurations. The values correspond to the computed average hamming distance. A lower value indicates better facial attribute preservation. Values displayed in bold indicate, which merging strategy resulted in the best attribute preservation for each cluster configuration (random, attribute-based and similarity-based)

contains solely faces sharing the same attributes. The results however also suggest, that a mean-based merging strategy is better at creating faces, that share the same attributes as faces used as reference.

6 Conclusions and Future Work

In this paper we propose a novel approach to achieve facial de-identification or anonymization in image datasets. It is based on a *k-same* method, which tries to create a face serving as a representative of the faces of *k*-individuals in a dataset. Unlike previous methods, we utilize the latent-space of a trained *generative adversarial network*. In that regard, we deploy two merging strategies: one calculates the mean values of all latent vectors from a designated group, while the other computes element-wise the median values of them.

Our experiments show how well each configuration, consisting of the cluster size *k*, the clustering strategy and the merging method, is able to achieve facial de-identification while preserving the most important characteristics of a face. By increasing the cluster size *k*, the privacy can be further enhanced at the cost of retaining the facial attributes. Although images created using the median merging strategy only differ slightly from the images created using the mean-averaging strategy, they have shown to yield higher privacy preservation. While random clustering techniques are able to achieve superior privacy, they do not preserve facial attributes very well. The worst results are produced using the similarity-based approach, as the vector-representation computed by the neural network conducting face recognition does not take facial attributes into account. As such, it is advised to take advantage of the attribute-based approach in order get a balance between privacy and information preservation.

Our initial scenario, where a police investigator or third party experts can quickly derived general information about a group of suspects without being shown their individual portraits can be realized with this approach. It enables privacy-preserving forensics with respect to suspect photos and therefore increases the state of the art in this domain of forensics research.

Deploying an image to latent space encoder, which tries to find the optimal face is subject to future work.

Acknowledgment

This work was supported by a framework agreement with “Hessisches Ministerium des Innern und für Sport”.

References

- [Ad08] Adams, C. W.: Legal Issues Pertaining to the Development of Digital Forensic Tools. In: 2008 Third International Workshop on Systematic Approaches to Digital Forensic Engineering. Pp. 123–132, May 2008.
- [BDS19] Brock, A.; Donahue, J.; Simonyan, K.: Large Scale GAN Training for High Fidelity Natural Image Synthesis. In: International Conference on Learning Representations. 2019, URL: <https://openreview.net/forum?id=B1xsqj09Fm>.
- [BEG00] Boyle, M.; Edwards, C.; Greenberg, S.: The Effects of Filtered Video on Awareness and Privacy. In: Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work. CSCW '00, Association for Computing Machinery, Philadelphia, Pennsylvania, USA, pp. 1–10, 2000, ISBN: 1581132220, URL: <https://doi.org/10.1145/358916.358935>.
- [Go14] Goodfellow, I. J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y.: Generative Adversarial Networks, 2014, arXiv: 1406.2661 [stat.ML].
- [He18] He, K.; Fu, Y.; Zhang, W.; Wang, C.; Jiang, Y.-G.; Huang, F.; Xue, X.: Harnessing Synthesized Abstraction Images to Improve Facial Attribute Recognition. Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI-18/, pp. 733–740, 2018.
- [HS96] Hudson, S. E.; Smith, I.: Techniques for Addressing Fundamental Privacy and Disruption Tradeoffs in Awareness Support Systems. In: Proceedings of the 1996 ACM Conference on Computer Supported Cooperative Work. CSCW '96, Association for Computing Machinery, Boston, Massachusetts, USA, pp. 248–257, 1996, ISBN: 0897917650, URL: <https://doi.org/10.1145/240080.240295>.
- [Ka20] Karras, T.; Aittala, M.; Hellsten, J.; Laine, S.; Lehtinen, J.; Aila, T.: Training Generative Adversarial Networks with Limited Data. In: Proc. NeurIPS. 2020.
- [KLA19] Karras, T.; Laine, S.; Aila, T.: A Style-Based Generator Architecture for Generative Adversarial Networks, 2019, arXiv: 1812.04948 [cs.NE].

- [Li21] Lin, J.; Zhang, R.; Ganz, F.; Han, S.; Zhu, J.-Y.: Anycost GANs for Interactive Image Synthesis and Editing. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2021.
- [LL19] Li, T.; Lin, L.: AnonymousNet: Natural Face De-Identification With Measurable Privacy. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Pp. 56–65, 2019.
- [Me17] Meden, B.; Emersic, Z.; Struc, V.; Peer, P.: K-Same-Net: Neural-Network-Based Face Deidentification. In: 2017 International Conference and Workshop on Bioinspired Intelligence (IWOBI). Pp. 1–7, 2017.
- [NSM05] Newton, E.; Sweeney, L.; Malin, B.: Preserving privacy by de-identifying face images. Knowledge and Data Engineering, IEEE Transactions on 17/, pp. 232–243, Mar. 2005.
- [Pe12] Peter, A.; Hartmann, T.; Muller, S.; Katzenbeisser, S.: Privacy-preserving architecture for forensic image recognition. In. Pp. 79–84, Dec. 2012, ISBN: 978-1-4673-2285-0.
- [SC20] Shim, W.; Cho, M.: CircleGAN: Generative Adversarial Learning across Spherical Circles. Conference on Neural Information Processing Systems (NeurIPS)/, 2020.
- [SML07] Stahlberg, P.; Miklau, G.; Levine, B. N.: Threats to Privacy in the Forensic Analysis of Database Systems. In: Proceedings of the 2007 ACM SIGMOD International Conference on Management of Data. SIGMOD '07, ACM, Beijing, China, pp. 91–102, 2007, ISBN: 978-1-59593-686-8, URL: <http://doi.acm.org/10.1145/1247480.1247492>.
- [SO20] Serengil, S. I.; Ozpinar, A.: LightFace: A Hybrid Deep Face Recognition Framework. In: 2020 Innovations in Intelligent Systems and Applications Conference (ASYU). IEEE, pp. 23–27, 2020.
- [Sr06] Srinivasan, S.: Security and Privacy in the Computer Forensics Context. In: 2006 International Conference on Communication Technology. Pp. 1–3, Nov. 2006.
- [Sw02a] Sweeney, L.: Achieving k-Anonymity Privacy Protection Using Generalization and Suppression. Int. J. Uncertain. Fuzziness Knowl. Based Syst. 10/, pp. 571–588, 2002.
- [Sw02b] Sweeney, L.: K-Anonymity: A Model for Protecting Privacy. Int. J. Uncertain. Fuzziness Knowl.-Based Syst. 10/5, pp. 557–570, Oct. 2002, ISSN: 0218-4885, URL: <https://doi.org/10.1142/S0218488502001648>.
- [Wa] Wang, J.: Size Constrained Clustering Solver, URL: https://github.com/jingw2/size_constrained_clustering.

Appendix

Further examples showcasing each configuration are displayed in Figure 4, Figure 5 and Figure 6. Although each of the people in Figure 6 wore eyeglasses, the synthesized image does not reflect this. Since data of people wearing spectacles are often sparse in the datasets the generative adversarial networks were trained on, the encoder often fails to preserve them when translating input images to the latent space.



Fig. 4: Experimental results on clusters of size 3,5 using a *random* clustering approach

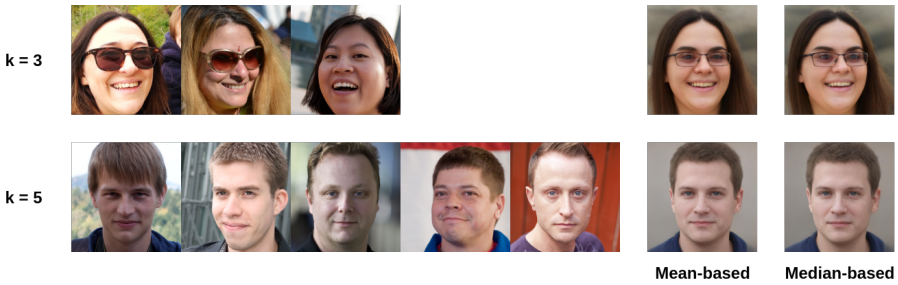


Fig. 5: Experimental results on clusters of size 3,5 using a *attribute-based* clustering approach



Fig. 6: Experimental results on clusters of size 3,5 using a *similarity-based* clustering approach

How to apply Database Anonymity Notions to Mix Networks

Marc Roßberger,¹ Alperen Aksoy,² Doğan Kesdoğan³

Abstract: Communication networks are an indispensable part of our society. By observing network traffic, one can acquire sensitive information about individuals, businesses, or governments. Thus, the protection of this traffic data is important and a well-researched topic. While encryption can protect the content of these messages, it can not hide the link between communication partners. Mix networks can be used to achieve this anonymity goal. However, these networks can not guarantee anonymity since they only apply technical protocols to improve anonymity. On the contrary, one can guarantee anonymity when releasing information from databases. Here, the principles of k -anonymity, l -diversity, and t -closeness can be applied to determine anonymity. In this work, we transfer these principles towards network anonymity, highlight problems that occur, and give an outlook on how a protocol could look like, with which users can build anonymity groups to guarantee anonymity.

Keywords: anonymity; mix networks; data privacy

1 Introduction

One of the biggest problems in our society is preventing the unauthorised creation of profiles on the Internet. The most basic defense against unauthorized profiling is the careful deployment of cryptographic techniques, which guarantees the privacy of exchanged messages. At the network level, however, a message's address information attributes it to both sender and receiver. Cryptography cannot hide this address information. Consequently, a network operator or intruder can read and collect a user's interactions with which to derive user-specific profiles. Thus the foundation of digital privacy is the network anonymity.

Although network anonymity has a few decades of experience, database anonymity has a solid foundation. In this paper, we ask ourselves at the definition and notions level how these two areas can be brought together. A fruitful discussion and exchange of notions and techniques could enrich both fields.

Contribution We are working on a new approach to measure and improve anonymity in mix networks based on metrics known from the field of database anonymity. For this purpose, we make a novel comparison between the two issues. In particular, we contribute with:

¹ University of Regensburg, Chair for Business Informatics IV, Universitätsstraße 31, 93053 Regensburg, Germany, marc.rossberger@ur.de

² Friedrich-Alexander-University Erlangen-Nürnberg, Computer Science Dept. Chair 1, Martensstrasse 3, 91058 Erlangen, Germany, alperen.aksoy@fau.de,
Thanks to the Turkish ministry of education for granting his Ph.D. studies.

³ University of Regensburg, Chair for Business Informatics IV, Universitätsstraße 31, 93053 Regensburg, Germany, dogan.kesdogan@ur.de

- an overview of anonymity goals in database anonymity and a suggestion on how to apply them to network anonymity,
- an analysis, why current mix networks can not fulfill these objectives, and
- an outlook of a possible protocol to meet these requirements.

Structure This paper is structured as follows. Section 2 gives an overview of related work. In section 3, we present the basics about mix networks and cover traffic. Afterward, section 4 introduces database anonymity and demonstrates how its principles can be applied to network anonymity. We conclude this work in section 5 by giving an outlook to future work.

2 Related Work

Anonymous communication is classified into two categories. These are high-latency and low-latency. Message-based systems are considered to be high-latency [EY09]. One of the well-known high-latency systems is the mix introduced by Chaum [Ch81]. Until 2004, several anonymous communication systems based on the mix were popular such as anon.penet.fi, cyberpunk remailers, and mixminion [DDM03; EY09]. After the release of TOR, which provides low-latency anonymization based on onion routing, mix-based systems went out of fashion because sending a message over mix-based systems can take longer times, even hours or days. In 2017, Pietrowska et al. revealed the Loopix system, which provides lower latency communication based on Mix networks by using cover traffic [Pi17]. In addition, the Loopix system is resistant to Sybil attacks, for which TOR has a vulnerability. Thus, Loopix shows that mix-based systems are able to compete with TOR-based systems.

By using database anonymization techniques, sensitive data can be protected from disclosure while still allowing statistical conclusions. Thus, these techniques hide the link between people and sensitive data. Anonymous communication networks have a similar purpose and hide the link between sender and recipient. To our knowledge, no study in the literature has investigated truly whether database anonymity techniques can be applied to anonymous communication networks in deep. In this work we will apply the principles of database anonymity to communication networks. This analysis also highlights how existing anonymization strategies can be used to try and achieve these goals and shows their flaws.

3 Mix Networks

Mix networks are systems that allow their participants to communicate anonymously by employing a special routing protocol. Most practical anonymous communication networks are inferred from the **Chaum Mix** [Ch81], also known as **Threshold Mix**. Its fundamental functionality is illustrated in figure 1. Here, 4 participants want to send their messages to 4 recipients. Both the sender set S and the recipient set R are part of a global set of users.

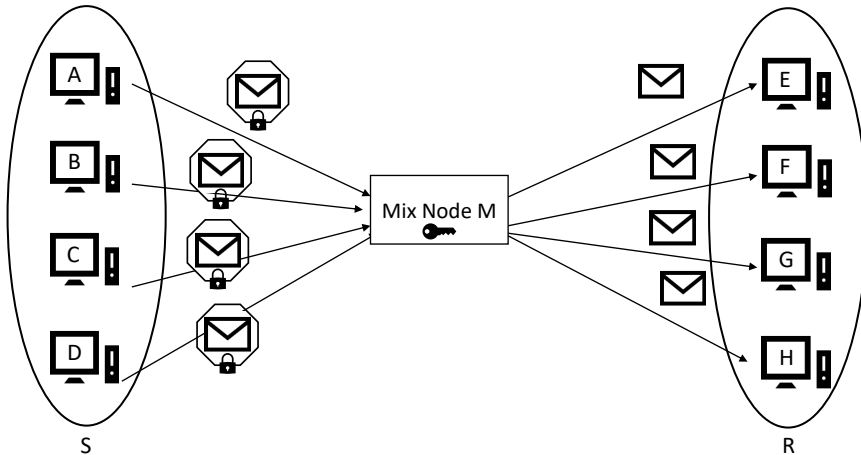


Fig. 1: One round of a Threshold Mix. The mix node decrypts and shuffles the messages. ($n = 4$)

To prevent possible attackers from observing the communication relationships between senders and recipients, they send their messages to a mix node M , encrypted with M 's public key. M then decrypts these messages, checks for duplicates to prevent replay attacks, and waits until it has received n messages, where n is its batch size. Once its threshold is reached, M shuffles the messages and relays them to their recipients. This protocol requires the messages to have a common size and appearance to prevent the linkage of messages. This is achieved, e.g., by the sphinx packet format [DG09].

A message can be routed through multiple mix nodes to improve anonymity. As long as at least one mix node on its path is honest, an attacker can not disclose the relationship between sender and receiver. Different types of mixes try to improve anonymity even further, e.g., by introducing indeterminism. One example is a **Pool Mix**, which collects $n + l$ messages but only sends l messages per batch. Another alternative is the **Stop-and-Go Mix**, which keeps messages for a randomly chosen time delay defined by the sender.

3.1 Anonymity Goals

Most mix network anonymity goals can be described using the terminology by Pfitzmann and Köhntopp [PK01]. The **sender-recipient unlinkability** describes that an attacker cannot determine which sender communicated with which recipient. If an attacker observes a message arriving at a receiver, his inability to determine the message's sender is described by **sender anonymity**. The list of all possible senders, which the attacker cannot differentiate, forms the **sender anonymity set**. Analogously, **recipient anonymity** and **recipient anonymity set** are defined. Further goals include **sender (recipient) unobservability**, which defines that an attacker cannot determine when a sender (recipient) is sending (receiving) a

message. How cover traffic can be employed to achieve these goals will be shown in the next section.

3.2 Cover traffic

One of the well-known ways to increase the anonymity set size is the usage of cover traffic. It protects messages from disclosure by covering them with dummy messages, making it difficult to track a particular message. Using cover traffic is a useful but costly application to prevent or delay successful disclosure attacks. According to Malleshwright et al. [MW07], cover traffic is beneficial under the conditions that (i) cover messages must be indistinguishable from real messages until they reach their final destination, (ii) dummy messages should be tolerable by receivers because the decryption phase does cost operations, and (iii) the cost of cover traffic should not exceed the limits of the mixes and receivers.

Cover traffic can be used for several purposes. Some of them are sender-recipient unlinkability, sender/recipient unobservability, detection of corrupt nodes, and reduced delays. Cover traffic can be handled in four ways. These are displayed in figure 2.

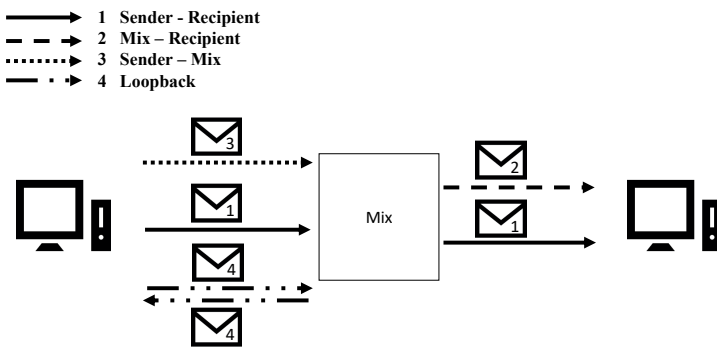


Fig. 2: Types of cover traffic

In the scope of sender-recipient cover traffic, dummy packets are created by the sender and reach the recipient via the mix network. The attacker cannot distinguish these packets from regular packets. Using this method makes a significant contribution to delay the success of disclosure attacks.

Another type of cover traffic is dummy packets created by mix nodes and sent to end-users. Although an attacker can notice a difference in count between packets arriving and leaving the mix, as long as the cover message's frame matches a real message's frame, he cannot distinguish cover packets from real packets. This method is convenient to increase the recipient anonymity set size. However, this method does not increase the sender anonymity set size.

Thirdly, dummy packets are created by a user and discarded by the mix, i.e. it does not increase the recipient anonymity set size. This method is only useful if every sender attends communication in every round. If senders were offline in some periods, an attacker might reveal the recipient set of a specific sender by comparing the recipient list to these rounds. If all users are online constantly, this method also provides sender unobservability for the network because an attacker cannot determine if a specific user is sending a message or not due to cover traffic.

Loop-back cover traffic is a dummy message sent by the user to himself. This type of cover message is used to detect corrupt nodes in the Loopix network anonymity system [Pi17]. If the message does not reach its sender, the user can detect the corruption in the path and use another route for future messages.

4 Comparison to Database Anonymity

We are trying to measure anonymity beyond the classical metrics from section 3.1, and for this purpose, we are looking at a related field: anonymity in databases. When releasing data from a confidential database to the public, e.g., for research purposes, one has to preserve the privacy of individuals listed in the database, including possible sensitive information, such as exact addresses, salaries, or diseases. For this purpose, the data is categorized into different classes, which, by the definition of Domingo-Ferrer et. al. [DSS16], include:

- **Identifiers** can be used to unambiguously identify an individual.
- **Quasi-Identifiers** are attributes, which can not identify a person alone, but a set of quasi-identifiers can. All entries sharing the same value for a quasi-identifier form an **equivalence class**.
- **Confidential attributes** have to be protected, and it should not be possible to determine an individual's confidential attribute (range).

	Identifier	Quasi-Identifiers		Confidential Attribute
	ID	ZIP Code	Gender	Salary
1	17394	93049	m	2794
2	52844	93049	m	3984
3	92961	93049	f	3104

Tab. 1: Example of a database table. Entries 1 and 2 form an equivalence class.

One example for these classes can be seen in table 1. The confidential attributes can be represented by the anonymity goals of section 3.1, and we will analyze these attributes from the perspective of an attacker trying to determine a message's sender. Except for the sender's signature, the only identifier is represented by the act of sending the message. A

global attacker, controlling all mix nodes, can observe this and thus identify the sender, while a local attacker can not follow a message beyond a mix node.

The link between an incoming and outgoing message at a mix node can represent quasi-identifiers in a mix network since the knowledge of one link does not reveal the message's full path, but knowledge of all links on the message's path would reveal the sender-recipient-relationship.

All messages arriving at a mix-node in the same batch form a message's local anonymity set. We will apply anonymity concepts developed for equivalence classes to these local anonymity sets. In contrast, we define the global anonymity set as the set of all sent messages, which could be the original message from an attacker's perspective. If every participant only uses one mix node, the local and global anonymity sets are identical. In addition to quasi-identifiers, the global anonymity set can be reduced even further by recognizing messages as cover traffic or knowledge of some messages' actual senders. Various privacy models from databases will be presented in the following sections.

4.1 k-Anonymity

The notion of k-anonymity indicates if the information for each person contained in the release cannot be distinguished from at least k-1 individuals whose information also appear in the release [Sw02]. This notion can be achieved in two ways. Some characters in quasi-identifier columns can be replaced with an asterisk (*), called suppression, or attributes are replaced with a broader category. This makes it difficult to identify an individual in the table.

When applying k-anonymity to mix networks, a release can be considered as messages that depart from the mix each round. In each round, if the mix flushes at least k messages, the system provides recipient k-anonymity. However, to provide sender k-anonymity, at least k messages must have arrived in the mix. The application of this technique in the Chaum mix, whose batch size is denoted as n, leaves three possibilities:

$$1) k < n \quad 2) k = n \quad 3) k > n$$

The Chaum mix, whose batch size n is equal to or greater than k , guarantees k-anonymity for both sender and recipient anonymity set. So, in the first and second conditions, k-anonymity can be provided by Chaum mixes easily. In the third scenario, this is not possible since it is impossible to send more than n packets per round. To achieve k-anonymity, indeterministic mixes may be used, such as a pool mix or a Stop-And-Go Mix.

When applying the k-anonymity technique to a pool mix, the mix has to wait until enough ($\geq k$) messages have been passed into the mix. From this point on, k-anonymity is guaranteed. With a Stop-And-Go mix, the situation is a little more complicated since it

is always possible for the Stop-And-Go mix to go idle. If one can accept this uncertainty, then the same statement as for pool mixes can also be applied to the Stop-And-Go mix. Otherwise, users must constantly send messages to the Stop-And-Go mix to guarantee it is not going empty.

4.2 l-Diversity

However, k -anonymity is not the only model used to evaluate database anonymity since it leaves weaknesses under certain conditions. In particular, two primary attacks can be employed against k -anonymity, even in the context of mix-networks:

- A **homogeneity attack** can occur when all confidential attributes share the same value. In a mix network, this can happen when all messages inside a batch go to the same recipient, thus revealing the relationship between all senders and that recipient.
- In a **background knowledge attack**, an attacker can reduce the (already small) set of possible values for the sensitive attribute with specific background knowledge. In a mix network, this can happen, for example, when the attacker knows that there is no communication between a sender and particular recipients of a batch.

Because of these problems, the concept of l -diversity was defined as follows: "An equivalence class is said to satisfy l -diversity if there are at least l well-represented values for the sensitive attribute." In this sentence, "well-represented" is not very clear, the following sections discuss more specific definitions and their possible implications for anonymity in mix networks.

4.2.1 Distinct l -diversity

This is the most straightforward definition and requires that l different values exist for the sensitive attribute. For a mix network, a message batch would have to include l different recipients, which could be end-users or other mix nodes. There are multiple possible strategies, how this could be achieved. A simple implementation would change a threshold mix to wait until it contains messages with l different recipients before relaying them. However, such a protocol would allow for new attacks. If an attacker observes messages arriving at the mix node and sees that the node is not sending its message batch even though l messages have arrived at it, he knows that at least two messages share the same recipient. He can also conclude that the last message arriving at the node before the batch is sent must have a new recipient. Thus, this protocol would require further adaptations, such as introducing indeterminism.

If the mix node is allowed to take active measures, known strategies from mix networks can be used to achieve l -diversity. One possible protocol would let the mix node generate cover

traffic and add it to its message pool until l distinct recipients exist. An alternative technique employed in Babel Mix [GT96] allows mix nodes to reroute messages. This could be used to spread messages, initially going to the same recipient, across different recipients.

In both previous suggestions, the end-user has to trust the mix node as only it can determine if l -diversity is fulfilled for its current message set. However, an attacker can abuse these options of withholding and rerouting packets to execute active attacks. If end-users do not trust the mix nodes, they can not guarantee l -diversity, but they can improve message diversity by sending cover traffic. This way, users can increase anonymity sets and thus diversity. However, known problems for end-to-end cover traffic apply. These include, e.g., how users should find their set of possible recipients for cover traffic. In the best case, every user is included in this set, but in real systems, knowing every end-user seems unrealistic.

4.2.2 Entropy l -diversity

A different interpretation is Entropy l -diversity, first suggested by Ohrn and Ohno-Machado [ØO99]. In this case, the entropy of a dataset is calculated with equation 1, where $p(s)$ describes the fraction of records in the dataset S with the same sensitive value as s . Entropy l -diversity is fulfilled if $H(S) \geq \log(l)$.

$$H(S) = - \sum_{s \in S} p_s(s) * \log(p_s(s)) \quad (1)$$

This definition of l -diversity does not alter how l -diversity can be applied to mix networks since only the calculation inside mix nodes has to be changed. However, unlike distinct l -diversity, it does not guarantee the latest message arriving at a mix node to have a new recipient. A new message going to a recipient underrepresented in the node's currently held messages could balance out the ratios of recipients and thus result in the fulfillment of entropy l -diversity, thus giving an attacker less information about the last messages' recipients.

4.2.3 Recursive (c, l) -diversity

In another definition of l -diversity, an attempt is made to limit the most frequent sensitive attribute's occurrence and provide a lower bound for the least common values. Equation 2 is used to determine if (c, l) -diversity is achieved, where r_1, r_2, \dots, r_n are the frequencies of the sensitive attribute's occurrence from most to least frequent.

$$r_1 < c * (r_l + r_{l+1} + \dots + r_m) \quad (2)$$

A special subcategory is defined as **Positive Disclosure-Recursive Diversity** and reduces the requirements towards diversity by allowing the disclosure or frequent appearance of a sensitive attribute if its value is not very sensitive, e.g., healthy. This concept could be transferred to mix networks since the knowledge of a message going to another mix node does not give a local attacker much information. However, this knowledge would allow an (n-1)-attacker (an attacker who controls all but one mix-node) to track a message completely, revealing the sender-recipient relationship.

4.3 t-Closeness

t-Closeness implies that the distance between the distribution of a sensitive attribute in each equivalence class and the distribution of the attribute in the whole table is no more than a threshold t [LLV07]. This technique is used to overcome some of the well-known attacks that cannot be prevented by l-diversity, such as the skewness attack and similarity attack. Let \mathbf{P} represent one of the distributions of an equivalence class, and \mathbf{Q} represents the whole table's distribution. To provide t-closeness, the distance between \mathbf{P} and \mathbf{Q} must be less than the threshold t .

This technique can also be applied to mix networks. The table can be considered as a union of recipient sets for all rounds (\mathbf{Q}), and classes can be considered as a recipient set of each round (\mathbf{P}). This distance can be calculated in several ways. The well-known techniques are variational distance and Kullback-Leibler distance. The formula of variational distance is given in Equation 3.

$$Dist[P, Q] = \sum_{i=1}^m \frac{1}{2} |p_i - q_i| \quad (3)$$

A global passive attacker, who is observing all routes in the network and wants to learn who a specific sender Alice is connecting with, can record the recipient sets of mixes over rounds in which Alice participates and does not participate. By comparing these recipient sets, he can intercept Alice's possible recipient set. These types of attacks are named intersection attacks and have some variants such as disclosure attack and statistical disclosure attack [Da03; KAP03]. In this scenario, if the distribution of recipient sets throughout all rounds are close to each other, creating an interception about Alice's recipient set takes more time.

In the scenario in figure 3, four users are communicating with four other users via a mix. P_i denotes the distribution of the recipients in the current round. The threshold value of t-closeness is set to 0.4, and the variational distance method is used to measure the distance between distributions. Q denotes the distribution of all rounds. In the first round, P_1 is equal to Q , and the distance is 0. It can be seen that all senders send a message and all recipients receive a message. In the second round, the distance between P_2 and Q is 0.1 and still less than the threshold. In round three, only two users send messages to the two recipients.

This information leakage can be reduced by applying t -closeness to this mix network. The mix can manipulate the third round because the distance between P_3 and Q is bigger than the threshold t . So, the mix can decrease the distance by producing cover packets. As seen in figure 4, these packets can be sent over two rounds, keeping a distance less than the threshold, preventing a decrease of the anonymity set size. In this example, recipient anonymity is considered. If sender anonymity is also considered, cover packets must be created by senders, requiring collaborative work.

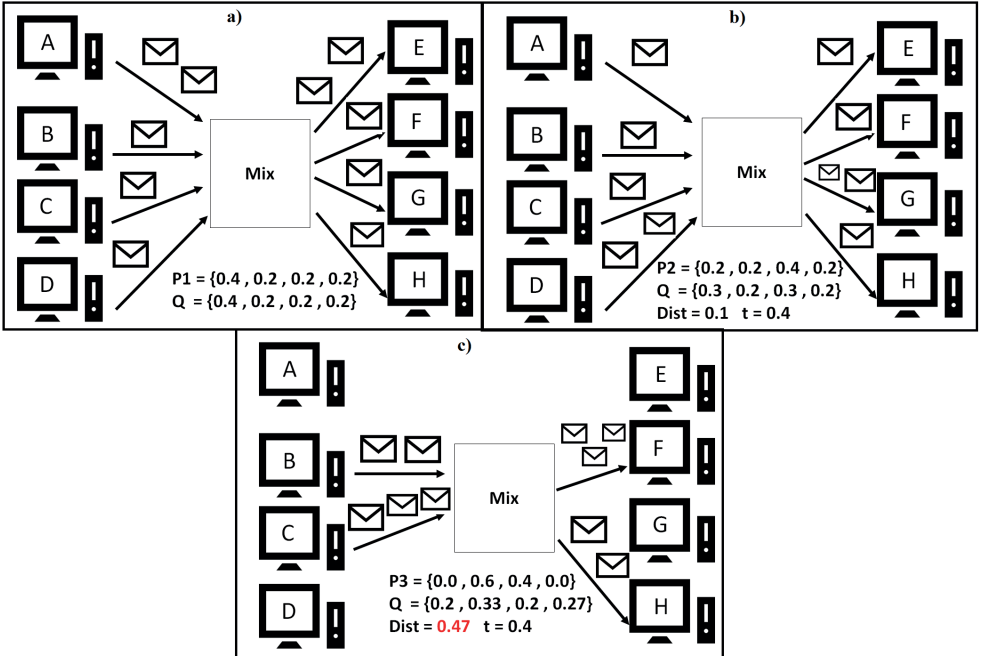


Fig. 3: Sample communication scenario via mix

5 Conclusion

This work has shown that the principles of anonymity used for databases can be applied to mix networks. However, achieving these strict privacy goals is a complex challenge. k -Anonymity, l -diversity, and t -closeness guarantee anonymity with specific parameters, while mix networks only apply techniques to improve anonymity, and we have shown that known attacks against database anonymity can also affect mix networks. The only way to guarantee anonymity under these conditions is by giving the mix nodes more control since only they can determine the diversity/closeness of their current message batch. However, it should not be necessary to trust the mix nodes. Instead, control over messages should be distributed among the end-users. If participants only work independently, they can improve

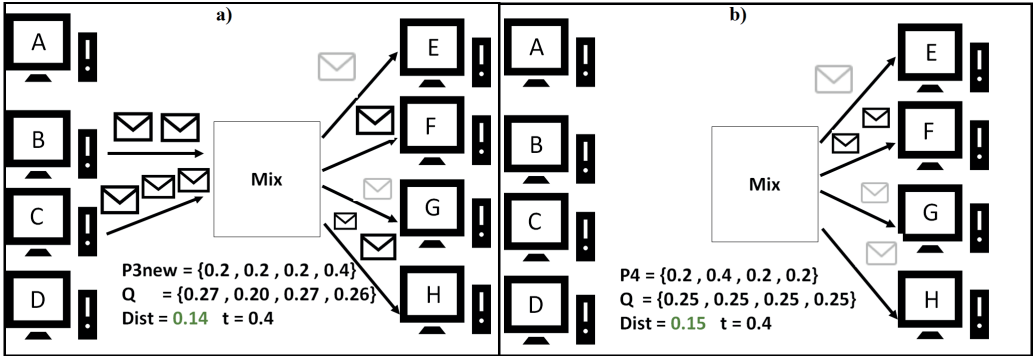


Fig. 4: a) third round after applying t-closeness. b) fourth round by applying t-closeness (cover messages are denoted by light grey color)

anonymity by introducing end-to-end cover traffic, but they can not guarantee it. Thus, multiple users have to work together to achieve privacy. For this purpose, they have to complete three objectives:

1. **Grouping:** Users have to build a group of size n of trusted users. This group needs to be big enough to achieve the anonymity goals, and it should not be possible for an attacker (group) to control most of its participants.
2. **Building:** The group needs to generate a message batch to insert into the network. Therefore, each user has to contribute a message, which can be real or cover traffic. The group has to ensure that this batch fulfills the anonymity requirements of diversity and closeness. If they are not fulfilled, the batch has to be changed, e.g., by randomizing the cover traffic or replacing some real messages with cover traffic.
3. **Embedding:** The group needs to deploy its messages into the mix network. To do this, one user has to transfer the batch into the mix network. The other users must be able to verify that their messages have been sent correctly without modification.

We are working on a protocol that allows mix users to build and send messages in anonymous groups under these conditions and want to present it in a future work.

References

- [Ch81] Chaum, D.L.: Untraceable electronic mail, return addresses, and digital pseudonyms. *Communications of the ACM* 24/2, pp. 84–90, 1981.
- [Da03] Danezis, G.: Statistical Disclosure Attacks. In: *Security and Privacy in the Age of Uncertainty*. Springer US, Boston, MA, pp. 421–426, 2003.

- [DDM03] Danezis, G.; Dingledine, R.; Mathewson, N.: Mixminion: design of a type III anonymous remailer protocol. In: 2003 Symposium on Security and Privacy, 2003. Pp. 2–15, 2003.
- [DG09] Danezis, G.; Goldberg, I.: Sphinx: A compact and provably secure mix format. In: 2009 30th IEEE Symposium on Security and Privacy. IEEE, pp. 269–282, 2009.
- [DSS16] Domingo-Ferrer, J.; Sánchez, D.; Soria-Comas, J.: Database anonymization: privacy models, data utility, and microaggregation-based inter-model connections. *Synthesis Lectures on Information Security, Privacy, & Trust* 8/1, pp. 1–136, 2016.
- [EY09] Edman, M.; Yener, B.: On Anonymity in an Electronic Society: A Survey of Anonymous Communication Systems. *ACM Comput. Surv.* 42/1, Dec. 2009.
- [GT96] Gülcü, C.; Tsudik, G.: Mixing E-mail With Babel. In: *Proceedings of the Network and Distributed Security Symposium - NDSS '96*. IEEE, pp. 2–16, Feb. 1996.
- [KAP03] Kedogan, D.; Agrawal, D.; Penz, S.: Limits of Anonymity in Open Environments. In (Petitcolas, F. A. P., ed.): *Information Hiding*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 53–69, 2003.
- [LLV07] Li, N.; Li, T.; Venkatasubramanian, S.: t-Closeness: Privacy Beyond k-Anonymity and l-Diversity. In: 2007 IEEE 23rd International Conference on Data Engineering. Pp. 106–115, 2007.
- [MW07] Malleš, N.; Wright, M.: Countering Statistical Disclosure with Receiver-Bound Cover Traffic. In (Biskup, J.; López, J., eds.): *Computer Security – ESORICS 2007*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 547–562, 2007.
- [ØO99] Øhrn, A.; Ohno-Machado, L.: Using Boolean reasoning to anonymize databases. *Artificial Intelligence in Medicine* 15/3, pp. 235–254, 1999.
- [Pi17] Piotrowska, A. M.; Hayes, J.; Elahi, T.; Meiser, S.; Danezis, G.: The Loopix Anonymity System. In: 26th USENIX Security Symposium (USENIX Security 17). USENIX Association, Vancouver, BC, pp. 1199–1216, Aug. 2017.
- [PK01] Pfitzmann, A.; Köhntopp, M.: Anonymity, unobservability, and pseudonymity—a proposal for terminology. In: *Designing privacy enhancing technologies*. Springer, pp. 1–9, 2001.
- [Sw02] Sweeney, L.: K-Anonymity: A Model for Protecting Privacy. *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.* 10/5, pp. 557–570, Oct. 2002.

Eine empirische Kontentanalyse zur Ermittlung von praxisorientierten Optimierungsfeldern zur Resilienz-Erhöhung der IT-Systeme im Sinne der ganzheitlichen Betrachtung der Informationssicherheit

Hemmnisse und Erfolgsfaktoren eines nachhaltigen und effizienten Informationssicherheitsmanagementsystems (ISMS)


Erfan Koza ¹

Abstract: Die Informationssicherheit ist in der Analogie zur Digitalisierung kein atomarer Zustand und muss infolge dieser Betrachtung als ein dynamischer, iterativer und inkrementeller Prozess verstanden werden, dessen primäre Zielsetzung in der nachhaltigen und effizienten Sicherung und Aufrechterhaltung der System- und Datensicherheit liegt. Vor dem Hintergrund der progressiven Digitalisierung der Kritischen Infrastrukturen befasst sich das Paper mit der Identifizierung sektoraler und intersektoraler Hemmnisse und kritischer Erfolgsfaktoren, welche für ein effizientes Informationssicherheitsmanagement eine essenzielle Rolle spielen. Ferner wird auch der Fokus auf das Herauskrystallisieren von differenzierten Sichtweisen hinsichtlich der Sinnhaftigkeit und Nützlichkeit der gesetzlichen Deklarationen, wie z. B. dem IT-Sicherheitsgesetz, gelegt. Zwecks empirischer Datenerhebung greift das Paper auf die quantitativen Methoden zu, indem eine heterogene Unternehmenslandschaft aus den Sektoren Energie, Wasserwirtschaft und Informations- und Kommunikationstechnik sowie dem Gesundheitswesen in die Forschungsstudie integriert wird. Als Resultat werden die empirisch validierten Faktoren sektorübergreifend in einem Kohärenzmodell zusammengetragen und im Sinne der Kausalität in Relation gesetzt.

Keywords: Informationssicherheit, KRITIS, Hemmnisse, Erfolgsfaktoren, IT-Sicherheitsgesetz, Kohärenzmodell, Informationssicherheitsmanagementsystem, ISMS

1 Einleitung

Der Begriff „Informationssicherheit“ setzt sich im Kontext der soziotechnischen Systeme aus den drei kohärenten Themengebieten: Technik, Mensch und Organisation zusammen. Im Zentrum dieser Auslegung stehen die komplementären Verhaltensweisen und die Interaktivität dieser drei Themengebiete, deren Zusammenspiel nach dem holistischen Lösungsansatz in einem Informationssicherheitsmanagementsystem (ISMS) konstatiert werden kann.

¹ Clavis Institut für Informationssicherheit der Hochschule Niederrhein, University of Applied Sciences, Webschulstraße 41 - 43, 41065 Mönchengladbach, Erfan.Koza@hs-niederrhein.de 

Die Entwicklung der nationalen Anstrengungen zur Steigerung der Resilienz der IT-Systeme sowie zur Initiierung der gesetzlichen Verbindlichkeiten an die Informationssicherheit kann grundsätzlich auf den Zeitraum zwischen 2005 bis 2019 zurückgeführt werden. Die Entwicklung der einzelnen Initiativen und gesetzlichen Grundlagen wurde in dem angegebenen Zeitraum inkrementell ausgeführt und wird gegenwärtig fortgesetzt. Nach dem Beschluss der ersten Cyber-Sicherheitsstrategie durch das Bundeskabinett im Jahre 2011 wurde in der zeitlichen Reihenfolge im Jahr 2015 das IT-Sicherheitsgesetz (IT-SiG) als Artikelgesetz verabschiedet. Initiiert durch die KRITIS-Strategie des Bundes mit dessen gesetzlichen Verankerung in Artikel 1 Nummer 2 und Nummer 7 werden durch das IT-SiG relevante Veränderungen innerhalb des Gesetzes über das Bundesamt für Sicherheit in der Informationstechnik (BSIG) für Betreiber Kritischer Infrastrukturen (KRITIS) getroffen [BM09], [BT09], [BT15], [BS17a].

Hierbei werden Betreiber Kritischer Infrastrukturen dazu verpflichtet, „(...) angemessene technische und organisatorische Vorkehrungen zur Vermeidung von Störungen der Verfügbarkeit, Integrität, Authentizität und Vertraulichkeit ihrer informationstechnischen Systeme, Komponenten oder Prozesse zu treffen (...)“ (§ 8a Abs. 1 Satz 1 BSIG), [BT09]. Zur Erfüllung der zuvor genannten gesetzlichen Anforderung haben die Branchenfachverbände: Deutscher Verein des Gas- und Wasserfaches e.V. (DVGW) und die Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall e.V. (DWA) für die Sparten der Trinkwasserversorgung und Abwasserbeseitigung, die Deutsche Krankenhaus Gesellschaft (DKG) für die stationäre Versorgung, der UP KRITIS – BAK Datacenter & Hosting mit CDN für das Rechenzentrums-Hosting und Rechenzentrums-Housing entsprechende Branchenspezifische Sicherheitsstandards (B3S) definiert. Dabei greifen die Fachverbände auf differenzierte methodische Ansätze zurück, um das B3S-Verfahren zu konzipieren. Während die beiden Fachverbände in der Wasserwirtschaft beispielsweise auf die BSI-IT-Grundschatz-Methodik (IT-Sicherheitsleitfaden WA Version 0.1 basierend auf den BSI-IT-Grundschatz und Version 0.2 basierend sowohl auf dem BSI-IT-Grundschatz als auch auf das BSI-IT-Grundschatz-Kompendium) zurückgreifen [DV17], greifen die anderen Fachverbände auf die Methodik der DIN ISO/IEC 27001 (B3S Krankenhaus greift auf die DIN ISO/IEC 27001, 27002, 27005 und 27799), (B3S zur IT-Sicherheit greift auf DIN ISO/IEC 27001 und ISO/IEC 22301) zurück [Up18], [DK19]. Als gemeinsame fachliche Schnittmenge deklarieren alle Verfahren jedoch die Implementierung eines ISMS. Das IT-SiG modifiziert auch weitere bestehende Gesetze, darunter auch das Energiewirtschaftsgesetz (EnWG), um hieraus gesetzliche Anforderungen für die Akteure in der Energiewirtschaft, Strom- und Gasnetzbetreiber sowie Anlagenbetreiber zu deklarieren. Die Verankerung der gesetzlichen Anforderungen befindet sich in § 11 Abs. 1 (a) EnWG für die Netzbetreiber und § 11 Abs. 1 (b) EnWG für die Anlagenbetreiber. In diesem Zusammenhang müssen die Netz- und Anlagenbetreiber die im IT-Sicherheitskatalog der Bundesnetzagentur definierten Anforderungen erfüllen und ein ISMS auf Basis der internationalen Norm DIN ISO/IEC 27001 und DIN ISO/IEC 27019 implementieren und erfolgreich zertifizieren lassen [BN15a], [BN15b].

Mittlerweile sind seit der gesetzlichen Deklaration über 6 Jahre vergangen. Infolgedessen mussten die ersten Nachweisverfahren (bspw. im Energiesektor) bereits bis zum 01.01.2018 erbracht werden [BS17b]. Somit müssten sich die Unternehmen mit einem KRITIS-Status seit ca. vier Jahren mit der Implementierung, der Umsetzung der kontinuierlichen Verbesserungsprozesse (PDCA-Zyklus) und der Zertifizierung und Auditierung von ISM-Systemen beschäftigen und daraus wertvolle empirische und praxisnahe Eindrücke generiert haben, die für die Forschung einen hohen Stellenwert darstellen [BM16]. Somit bildet diese Forschungsstudie eine empirische Datenerhebung und Datenanalyse zum Erkenntnisgewinn ab, sodass daraus primär ein Gesamtbild über die aktuellen Hemmnisse und kritischen Erfolgsfaktoren für ein nachhaltiges und effizientes ISMS abgebildet werden kann (siehe Forschungsfrage F1). Sekundär sollen die Eindrücke und Erfahrungswerte konsolidiert werden, d.h. inwiefern die gesetzlichen Grundlagen, wie bspw. das IT-SiG, zu einer nachhaltigen Sicherheit in der Informationstechnik beitragen (siehe Forschungsfrage F2).

F 1: Welche Haupthemmnisse und Erfolgsfaktoren spielen für ein nachhaltiges und effizientes Erreichen eines adäquaten Informationssicherheitsniveaus eine relevante Rolle?

F 2: Welche Rolle spielt der Gesetzgeber, um das Informationssicherheitsniveau flächendeckend, einheitlich und nachhaltig zu erhöhen?

2 Methodik

Zur Operationalisierung der Forschungsstudie werden im ersten Schritt die zu untersuchenden Felder (IKT: RZ-Betreiber, Wasser: Trinkwasserversorgungs- und Abwasserbeseitigungsunternehmen, Energie: Anlagen- und Netzbetreiber und Gesundheitswesen: Krankenhäuser und Hochschulkliniken) ausgewählt. In einem zweiten Schritt sind im Rahmen einer methodenpluralistischen Ausrichtung neben der Integration von quantitativen Methoden mit der „vollstrukturierten Online-Befragung,“ auch Methoden zur Struktur- und Inhaltsanalyse vorgesehen, um die Möglichkeit zu behalten, die gewonnenen Kenntnisse aus der Online-Befragung in einem strukturierten Format nach der jeweiligen Sektor-Zugehörigkeit (Aussagenverteilung in dem jeweiligen Sektor) und nach den jeweiligen sektorübergreifenden Gemeinsamkeiten validieren und diese ggfls. vertiefen zu können. Die konzipierten Themenblöcke und Fragestellungen adressieren insgesamt Hemmnisse für die Informationssicherheit (15 Fragen), Erfolgsfaktoren eines ISMS (18 Fragen) und Relevanz der Gesetzgebung (10 Fragen). Zur Initiierung der konzipierten schriftlichen Online-Befragung wurden in der zweiten Phase „Durchführung“ 200 Unternehmen bundesweit über eigene Online - Recherchen in die Befragung integriert. Additional hierzu wurden die Adressaten so ausgewählt, dass eine heterogene Unternehmenslandschaft involviert werden konnte, um sowohl die KMU als auch die Großunternehmen und Konzerne gleichermaßen in der Online-Befragung adressieren zu können.

3 Darstellung der Ergebnisse

Die Resonanz und somit die Aussagekraft der operationalisierten Online-Befragung wird anhand der absoluten Anzahl der bereinigten Datensätze gemessen. In der Summe existieren 117 Datensätze, wobei 31 Datensätze im Zuge der Datenbereinigungsprozesse eliminiert und aus der Auswertung entfernt wurden. Die übrig gebliebenen 86 Datensätze repräsentieren eine 43-prozentige Rücklaufquote $((86 / 200) * 100)$ und weisen einen einheitlichen Bearbeitungszustand auf, sodass alle Fragen von allen Unternehmen vollständig beantwortet wurden. Insgesamt lässt sich das n_{gesamt} auf 86, n_{Wasser} auf 25, $n_{Energie}$ auf 23, n_{IKT} auf 14, $n_{Gesundheitswesen}$ auf 14 und $n_{Energie\ u.\ Wasser}$ auf 10 sowie n_{KRITIS} auf 65, $n_{Nicht-KRITIS}$ auf 21 und $n_{Betriebsgröße}$ auf 86 summieren.

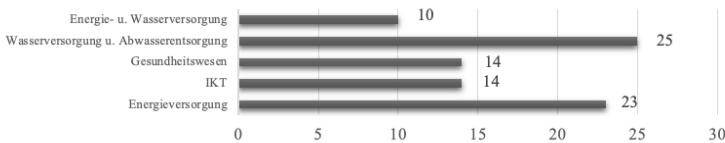


Abb. 1: Branchenzuordnung der involvierten Unternehmen

Die Diagrammkategorien der Hauptgeschäftsbereiche orientieren sich in der Abb. 1 an den vom BMI definierten Sektoren [BM09]. In der Summe werden 65 Unternehmen als KRITIS klassifiziert. Damit beträgt der prozentuale Anteil der KRITIS in der Gesamtheit 76 Prozent.

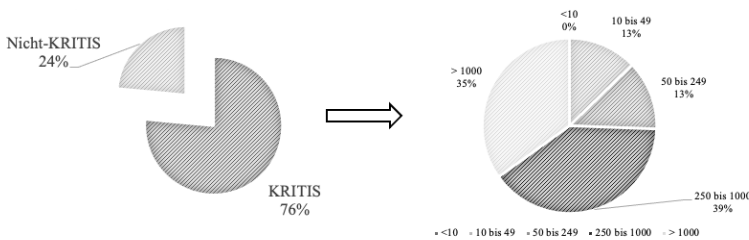


Abb. 2: KRITIS-Status und Betriebsgröße der involvierten Unternehmen

Die an der Studie beteiligte Unternehmenslandschaft zeigt ebenfalls eine gewisse Heterogenität auf, in der alle die von der Europäischen Union getroffenen Klassifizierungen, bis auf Kleinunternehmen (<10), hinsichtlich der definitorischen Auslegung der KMU vorkommen. In der kumulierten Betrachtung der einzelnen Klassifizierungen wird jedoch deutlich, dass die großen und mittleren Unternehmen einen höheren prozentualen Anteil an der Studie aufweisen. Diese Einschränkung bietet die Möglichkeit die jeweiligen Themenblöcke abhängig der Betriebsgröße auszuwerten und diese im Sinne eines korrelativen Zusammenhangs zu untersuchen. Diese Analyse wird jedoch im Rahmen dieser Publikation vernachlässigt, da der Fokus auf die kollektiven sektoralen und intersektoralen Faktoren gelegt wird.

3.1 Intersektorale Hemmnisse für die Informationssicherheit

Um die Hemmnisse für die Informationssicherheit zu identifizieren, wurden im Zuge der schriftlichen Online-Befragung insgesamt 15 dedizierte Aussagen mit unterschiedlichen Schwerpunkten integriert, die der Abb. 3 zu entnehmen sind. Das *größte intersektorale Hemmnis* für eine nachhaltige und ressourcenschonende Informationssicherheit in den betrieblichen Strukturen ist der erhöhte personelle und zeitliche Aufwand, sodass die im ISMS definierten Anforderungen nur mit einem hohen *internen* und/oder mit einem erhöhten externen Aufwand zu erfüllen sind (76 Unternehmen, 88 Prozent). Die in der DIN ISO/IEC 27001:2017 definierten Anforderungen (114 Maßnahmen im Anhang A) müssen bspw. in der Erklärung zur Anwendbarkeit auf den jeweiligen Anwendungsbereich projiziert und individuell spezifiziert werden. Der zu betrachtende Anwendungsbereich setzt sich i. d. R. aus mehreren zentralen und dezentralen IT-Systemen und Netzwerkkomponenten zusammen, die in separierten Netzwerksegmenten mit unterschiedlichen Kommunikationsprotokollen sowie Software- und Hardwaretechnologien zur Erfüllung differenzierter Aufgabenstellungen aufgesetzt werden. So wird die zunehmende Komplexität und Heterogenität von IT-Systemen und IT-Netzwerklandschaften von etwa 74 Unternehmen, also 86 Prozent der Befragten, als *zweitgrößtes intersektorales Hemmnis* für die Informationssicherheit angegeben.

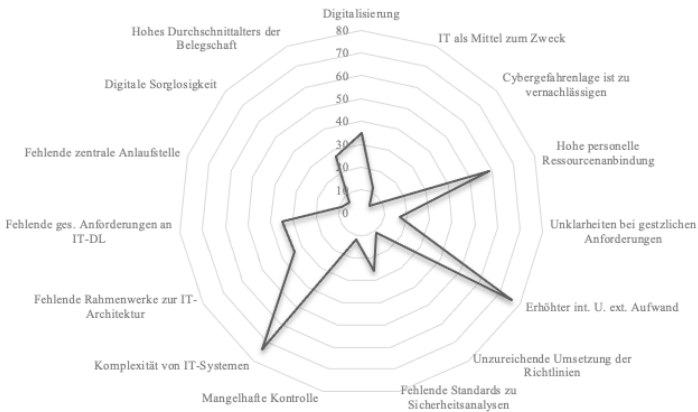


Abb. 3: Intersektorale Hemmnisse für die Informationssicherheit

In Anbetracht der zahlreichen gesetzlichen, aber auch selbstaufgelegten Anforderungen, müssen die Unternehmen mehrere Managementsysteme zeitgleich aufsetzen und kontinuierlich weiterentwickeln. Hieraus entsteht eine dauerhafte personelle Überlastung, die durch die starke personelle Ressourcenbindung verursacht wird. Beispielsweise werden die ISMS-dedizierten Aufgabenstellungen als On-Top-Aufgaben auf die regulären

Organisationseinheiten heruntergebrochen, wodurch diese mit der Aufgabenstellung faktisch überlastet sind.

So gaben insgesamt 59 Unternehmen mit einem prozentualen Anteil von 69 Prozent an, dass sie die starken personellen Ressourcenbindungen (=Überlastung) als *drittgrößtes Hemmnis* betrachten. Als *viertgrößtes Hemmnis* gaben insgesamt 35 Unternehmen (40 Prozent) an, dass die vorhandenen finanziellen Ressourcen bevorzugt in neue digitale Infrastruktur und Digitalisierungs- und Automatisierungsprozesse anstatt in die Informationssicherheit investiert werden. Zusätzlich hierzu spielen die IT-Dienstleister, die im Bereich der Entwicklung, Administration, Instandhaltung, und Installation der Software- und Hardware-Komponenten fungieren, eine wesentliche Rolle. So monieren 35 Unternehmen mit einem prozentualen Anteil von 40 Prozent das Fehlen von gesetzlichen Anforderungen an die IT-Dienstleister als Hemmnis. Durch die Kontentanalyse lässt sich das aufgeführte Hemmnis in Bezug auf das Fehlen von gesetzlichen Anforderungen an die IT-Dienstleister wie folgt verifizieren: 68 Unternehmen mit einem prozentualen Anteil von 79 Prozent sowie weitere 64 Unternehmen (74 Prozent) monieren die Defizitlage, dass das IT-SiG nur die KRITIS adressiert, ohne die Lieferkette in der Hardware- und Softwaretechnologie der Betreiber und Organisationen mit zu adressieren und hierbei verbindliche software- und hardwaretechnische Sicherheitsanforderungen zu deklarieren. Hierdurch entsteht eine fachliche Sicherheitslücke, da die KRITIS in einer hohen Kritikalität von den IT-Dienstleistern und deren Sicherheitsniveau abhängig sind. Ferner wird ersichtlich, dass das IT-SiG bspw. keine expliziten Anforderungen an die Entwicklungsprozesse in den Hard- und Softwaretechnologien der IT-Dienstleister stellt (bspw. Security by Design Ansätze oder obligatorische Produktzertifizierungen). Darüber hinaus gaben 34 Unternehmen mit einem prozentualen Anteil von 39 Prozent an, dass das Fehlen von geeigneten Standards, Methoden und Werkzeugen zur Sicherheitsanalyse von IT-Systemen ebenfalls als ein großes Hemmnis zu betrachten ist.

3.2 Intersektorale kritische Erfolgsfaktoren für die Informationssicherheit

Neben den identifizierten intersektoralen Hemmnissen für eine nachhaltige Informationssicherheit wird auch nach den kritischen Erfolgsfaktoren für den ressourcenschonenden und effizienten Betrieb eines ISMS geforscht. 83 Unternehmen mit einem prozentualen Anteil von 97 Prozent sehen das interne Vorhandensein von fachlichem, informationstechnischem Knowhow als den wichtigsten kritischen Erfolgsfaktor. Die thematische Ausprägung der Informationssicherheit ist vielfältig und adressiert ein breites Spektrum an mehreren differenzierten Themengebieten, deren Konzeption, Durchführung und laufende Überwachung ein tiefgehendes informationstechnisches Verständnis abverlangt. Allerdings ist das Vorhandensein von informationstechnischem Knowhow unmittelbar mit dem Vorhandensein von internen fachlichen Ressourcen verknüpft. So gaben insgesamt 77 Unternehmen mit einem prozentualen Anteil von 90 Prozent an, dass die personellen Kapazitäten mitunter durch die Integration eines Informationssicherheitsbeauftragten sowie durch die Integration weiterer

Sicherheitsfachkräfte den zweitwichtigsten kritischen Erfolgsfaktor für ein effizientes und nachhaltiges ISMS darstellt. Die Integration von internen Ressourcen bedarf jedoch der finanziellen und fachlichen Unterstützung der Geschäftsführung.

Vor diesem Hintergrund treffen 57 Unternehmen, die einen prozentualen Anteil von 66 Prozent darstellen, die Aussage, dass die fachliche und finanzielle Unterstützung der Geschäftsführung für den Erfolg eines ISMS unabdingbar ist (siehe Abb. 4).

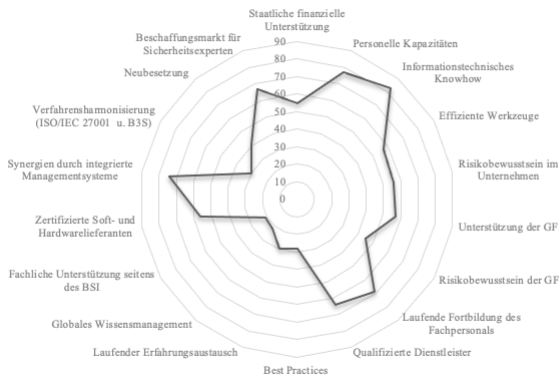


Abb. 4: Intersektorale Erfolgsfaktoren eines nachhaltigen ISMS

Insgesamt gaben 55 Unternehmen (64 Prozent) an, dass finanzielle staatliche Unterstützungen ebenfalls einen kritischen Erfolgsfaktor darstellen. In der Summe monieren 72 Unternehmen (83 Prozent) außerdem, dass das IT-SiG zwar die richtigen Themen deklariert, die KRITIS jedoch mit dem finanziellen Mehraufwand, der durch die Erfüllung der gesetzlichen Anforderungen entsteht, alleine lässt. Der dritte kritische Erfolgsfaktor lässt sich mit dem Themengebiet der integrierten Managementsysteme (integrated management systems) in Verbindung bringen. Durch den sogenannten „Common structure of ISO standards“ der Managementsysteme besteht die Möglichkeit, Managementsysteme mit unterschiedlichen Schwerpunkten mithilfe ihrer gemeinsamen Schnittmenge zu harmonisieren, sodass dadurch die anfallenden Doppel- bzw. Mehrfachbelastungen sowie dokumentarischen Redundanzen vermieden werden können. Damit kann dieser Erfolgsfaktor unmittelbar als ein effizienzsteigerndes Instrument definiert werden, um die starken personellen Ressourcenanbindungen (siehe drittgrößtes Hemmnis) aufgrund der vielen Regularien zu unterbinden und hieraus Synergien zu generieren. So treffen 74 Unternehmen mit einem prozentualen Anteil von 86 Prozent die Aussage, dass die Synergien, die durch die Synchronisation und Harmonisierung der Managementsysteme erreicht werden (Qualitätsmanagementsystem, Technisches Sicherheitsmanagementsystem, ISMS, Umweltmanagementsystem, Energiemanagementsystem etc.), den dritt wichtigsten kritischen Erfolgsfaktor für einen ressourcenschonenden und effizienten Betrieb eines ISMS darstellen. Sind jedoch die

personellen Ressourcen nicht vorhanden, so müssen diese zunächst auf dem Beschaffungsmarkt aufwendig erworben werden. So betrachten insgesamt 63 Unternehmen mit einem prozentualen Anteil von 73 Prozent das Vorhandensein von adäquaten Sicherheitsexperten als viertwichtigsten Erfolgsfaktor, um das Thema der Neubesetzung im Bereich der Informationssicherheit adäquat abbilden zu können.

Additional hierzu spielen das Vorhandensein von softwarebasierten Lösungsansätzen und Unterstützungstools eine relevante Rolle, um bspw. die dokumentarischen und analytischen Prozesse eines ISMS effizienter und ressourcenschonender abbilden zu können (52 Unternehmen mit einem prozentualen Anteil von 66 Prozent). Weiterhin gaben weitere 56 Unternehmen (65 Prozent) an, dass der Einsatz von zertifizierten Software- und Hardwareprodukten als weiterer intersektoraler kritischer Erfolgsfaktor zu definieren ist, um das definierte Sicherheitsniveau im Verbund erreichen zu können.

3.3 Sektorspezifische kritische Erfolgsfaktoren für Informationssicherheit

Die zuvor dargestellten Hemmnisse und kritischen Erfolgsfaktoren stellen eine intersektorale Auswertung dar. Der sektorspezifischen Auswertung nach spielt der Faktor „staatliche finanzielle Unterstützung“ primär bei den Akteuren in der Wasserwirtschaft und sekundär bei den Leistungserbringern im Gesundheitswesen eine relevante Rolle.

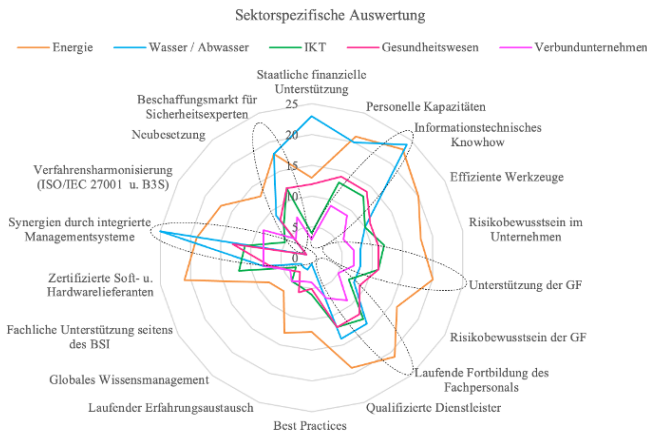


Abb. 5: Sektorspezifische Erfolgsfaktoren eines nachhaltigen ISMS

Ein weiterer sektorspezifischer Erfolgsfaktor, welcher lediglich von den Akteuren im Energiesektor und IKT-Sektor als wichtig erachtet, aber von den Akteuren in der Wasserwirtschaft und der stationären Versorgung überwiegend vernachlässigt wird, ist der Einsatz von zertifizierten Hardware- und Softwarelieferanten. Darüber hinaus zeigen die Verbundunternehmen eine ausgeprägte Konzentration auf den kritischen Erfolgsfaktor

„Verfahrensharmonisierung (DIN ISO/IEC 27001 und B3S)“, da diese möglicherweise als einzige Organisationen in zwei unterschiedlichen Sektoren tätig sind und somit zwei unterschiedliche Verfahren eingeführt haben, die mitunter dazu führen können, dass hierdurch gravierende finanzielle, fachliche und personelle Doppelbelastungen entstehen (siehe Abb. 5). Nachfolgend werden die Ergebnisse in einem Ursache-Wirkungsmodell zusammengetragen.

4 Konsolidierung der erreichten Ergebnisse in einem Kohärenzmodell

Das konzipierte Kohärenzmodell zur Informationssicherheit skizziert die korrelativen Zusammenhänge und fasst die fachlichen Einflüsse mit dem jeweiligen Wirkungsgrad zusammen, um die Hemmnisse (H-Faktor) und die kritischen Erfolgsfaktoren (E-Faktor) gegenüberzustellen und hieraus ein Gesamtbild zu schaffen. Die Digitalisierung und die von Unternehmen präferierten Investitionen in die Digitalisierungstendenzen (H-Faktor: 41 Prozent) werden in diesem Zusammenhang als Initialhemmnis verstanden. Das Dilemma hierbei ist, dass vorhandene finanzielle Ressourcen i.d.R. bevorzugt in die Digitalisierungs- und Automatisierungstendenzen anstatt in die Informationssicherheit investiert werden. Diese Investitionen werden unmittelbar zur Optimierung bzw. zur Effizienzsteigerung (Reduzierung der Prozesskosten und Personalkosten) der vorhandenen Geschäftsprozesse eingesetzt. Damit kann die Diskussion über den Mehrwert der Informationssicherheit entfacht werden, da die Vorteile der Informationssicherheit mit den herkömmlichen Methoden und Kennzahlen monetär nicht messbar sind und somit für eine Vielzahl an Entscheidungsträgern ein abstraktes Gebilde darstellen. Ein weiterer Umstand, der zur Verschlechterung dieser Sichtweise führen kann, ist die Tatsache, dass die konzipierten Redundanzen und Rückfallebenen als argumentative Kette zur Ablehnung von weiteren Investitionen in die Informationssicherheit missbraucht werden. Hierbei werden die Folgen einer Cybergefahr mit der Kompensationsfähigkeit der konzipierten Rückfallebenen, also mit den reaktiven und korrektiven Fähigkeiten denunziert. Hierbei wird die Sachdienlichkeit der Informationssicherheit im Sinne der Prävention nicht fachgerecht in Betracht gezogen. Im Kern bedeutet diese Sichtweise, dass man sich unabhängig der vorhandenen Rückfallebenen (reaktive bzw. korrektive Maßnahmen) mit den Gefahren und Auswirkungen auseinandersetzt und im Sinne der Prävention Maßnahmen konzipiert, die die Eintrittswahrscheinlichkeit von Gefahren reduziert. Die Ursache für ein solches nicht sachgerechtes Sicherheitsempfinden hat auch primär damit zu tun, dass die Anwender die IT-Risiken und IT-Gefahren üblicherweise anhand ihrer bereits gemachten praktischen Erfahrungen auf dem Gebiet der Versorgungssicherheit und dem Risikomanagement machen. Dies kann grundlegend mit der Tatsache in Verbindung gebracht werden, dass das vorhandene Sicherheitsgefühl dadurch ausgeprägt ist, dass bis jetzt keine flächendeckenden und gravierenden IT-Sicherheitsvorfälle und Informationssicherheitsvorfälle aufgetreten sind bzw. nicht erkannt wurden, die die

Betriebsfähigkeit gefährden konnten. Dieses trügerische Sicherheitsgefühl ist eine Implikation, da möglicherweise eine Vielzahl der Organisationen keine detektierenden und offensiv ausgerichteten Sicherheitssysteme integriert haben, die ihre IT-Systemschwachstellen aufzeigen bzw. Vorfälle vor dem Missbrauch entdecken, die sonst unentdeckt geblieben wären. Ein zweckdienliches Beispiel zur argumentativen Untermauerung dieser Ausführung bieten die Berliner Wasserbetriebe (BWB). Die BWB haben laut eigenen Angaben das Sicherheitsberatungsunternehmen Alpha Strike Labs im Frühjahr 2020 dazu beauftragt, in einer „Security Analyse“ einen Penetrationstest für das Abwasserentsorgungssystem durchzuführen [BW20].

Als Ergebnis des Penetrationstestes wurden mehr als 30 Schwachstellen trotz der erfolgreichen B3S WA Zertifizierung entdeckt, deren Reichweite sehr wahrscheinlich zu einem Zusammenbruch der Abwasserbeseitigung führen würde [JF20].

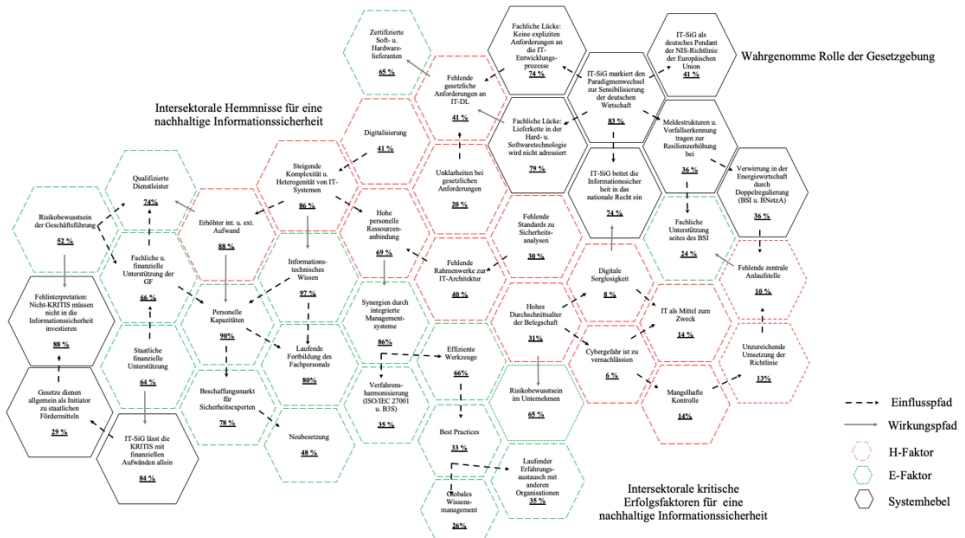


Abb. 6: Kohärenzmodell zur Relationssetzung der Faktoren

Das Vorhandensein eines sachdienlichen Sicherheitsverständnisses (E-Faktor: 52 Prozent) der Entscheidungsträger kann dazu beitragen, dass die notwendigen finanziellen und personellen Ressourcen zur Erfüllung dieser Aufgabenstellungen sogar unabhängig der gesetzlichen Anforderungen auf freiwilliger Basis freigegeben werden können. Hierzu dient das IT-SiG als Stimulanz, um den Beginn des Paradigmenwechsels zu markieren und die deutsche Wirtschaft in diesem Kontext durch die Ratifizierung des IT-SiG in Punkto Informationssicherheit zu sensibilisieren. Hierzu bettet das IT-SiG das Thema der „Informationssicherheit“ in das nationale Recht ein. Allerdings monieren die Unternehmen, dass trotz der Wichtigkeit der Deklaration des IT-SiG, dieses die Unternehmen mit der finanziellen Mehrbelastung, die durch die im IT-SiG definierten Anforderungen entsteht, alleine lässt. In der logischen Schlussfolgerung kann daher die

Interpretation erfolgen, dass die Möglichkeit finanzielle Unterstützung in Form von Fördermitteln zu erhalten, gleichzeitig einen positiven Impuls darstellt, der die Akzeptanz und die Entscheidungsbereitschaft der Geschäftsführung für Investitionen in die Informationssicherheit möglicherweise begünstigen kann. Derartige Impulse lassen sich derzeit im Gesundheitswesen beobachten, wo die Krankenhäuser sowie Hochschulkliniken, die nicht explizit unter das Regime des BSIG fallen, anhand der definierten Förderbestände nach dem KHZG mit dem § 19 KHZG in Verbindung mit dem § 19 KHSFV Abs. 1 Satz 1 Nr. 1 bis 10 KHSFV bis Ende 2021 dedizierte Fördermittel zur punktuellen Verbesserung der IT-Sicherheit bis zur ganzheitlichen Optimierung und Einführung eines ISMS beantragen können. Während das adäquate Risikobewusstsein der Geschäftsführung als elementare Voraussetzung für die Planung und Freigabe von Ressourcen sowie Überwachung und Steuerung von Maßnahmen verstanden wird, bildet das Risikobewusstsein der Mitarbeiter die grundlegende Voraussetzung zur Information Security Awareness, um hieraus eine zweckdienliche Sicherheitskultur nachhaltig in die Organisationsstruktur generieren zu können. Das Hemmnis, das hierbei als Gegenkraft wirkt, kann auf den demografischen Wandel zurückgeführt werden, in dem die durchschnittliche höhere Altersstruktur in der Belegschaft (digitale Sorglosigkeit) zur Vulnerabilität des Faktors Mensch führt, wodurch die Schwäche der menschlichen Firewall als Einfallstor durch Social Engineering ausgenutzt wird. Zudem werden die fachlichen Lücken innerhalb des IT-SiG moniert, da in diesem keine gesetzlichen Anforderungen an die IT-Dienstleister definiert werden. Diese fachliche Lücke ist von großer Bedeutung, da Organisationen unabhängig ihrer Sektorzugehörigkeit in der Regel mit externen IT-Dienstleistern im Bereich der Software- und Hardwaretechnologien arbeiten und somit auf Produkte und Dienstleistungen zurückgreifen müssen, die möglicherweise keine sicherheitstechnischen Merkmale aufweisen. Dabei muss konstatiert werden, dass mögliche Sicherheitslücken in den Hardware- und Software-Komponenten von den Unternehmen aufgrund fehlender Kapazitäten sowie fehlendem Wissen nicht autark analysiert und beseitigt werden können. Vielmehr müssen sich die Organisationen auf die Sicherheitsanalysen, Entwicklungsverfahren und Testverfahren der IT-Dienstleister verlassen, wobei deren Produkthaftungen in der Regel keine sicherheitstechnischen Defizite beinhalten. Innerhalb dieser Betrachtung fehlt die Übertragung der IT-Sicherheit auf Komponentenebene. Bereits bei der Entwicklungsphase sollte darauf geachtet werden, dass die Smartness der Software- und Hardware-Komponenten nicht isoliert für Produktivitätsvorteile eingesetzt wird, sondern auch als Früherkennungseinheiten und zur Detektion und Korrektur von Cyber-Angriffen genutzt werden. In der Summe muss die IT-Sicherheit in die Konzeption, Planung und Herstellung der IT-Komponenten und somit in die Unternehmensinfrastruktur einbezogen werden. Additional hierzu beeinflussen das Fehlen von geeigneten Methoden und Werkzeugen (H-Faktor: 30 Prozent) sowie das Fehlen von intelligenten Softwarelösungen zu Sicherheitsanalysen (H-Faktor: 40 Prozent) die personellen Ressourcenanbindungen negativ beeinflussen. Als Gegenkraft hierzu wirken das Vorhandensein von effizienten Werkzeugen und intelligentere Lösungen zur Sicherheitsanalyse (E-Faktor: 66 Prozent), Best Practices sowie Verfahrensharmonisierungen (E-Faktor: 35 Prozent), und -Synchronisationen (E-Faktor: 86 Prozent) zur Generierung von Synergien.

Allerdings stellen die erreichten empirischen Ergebnisse und das konzipierte Artefakt zunächst einmal eine Diagnose, also eine Momentaufnahme dar, die keinerlei Aussagen darüber macht, wie sich die einzelnen Aspekte im weiteren Verlauf (Prognostik), entwickeln werden. Ein denkbarer Ansatz zur Relaxion dieses Umstandes kann durch den Einsatz der Szenariotechnik erfolgen. Ein weiterer denkbarer Ansatz zur Verifikation und Vertiefung der empirischen Ergebnisse kann durch die Anwendung qualitativer Methoden (leitfadengestütztes Experteninterview) erreicht werden, aus denen ggfls. Handlungsoptionen, Chancen und Risiken definiert werden, die für eine nachhaltige und effiziente Informationssicherheit eine wesentliche Rolle spielen.

Literaturverzeichnis

- [BM09] BMI, Bundesministerium des Innern: Nationale Strategie zum Schutz Kritischer Infrastrukturen (KRITIS-Strategie), 2009.
- [BM16] BMI, Bundesministerium des Innern: Referententwurf des Bundesministeriums des Innern. Entwurf einer Verordnung zur Bestimmung Kritischer Infrastrukturen nach dem BSI-Gesetz (BSI-Kritisverordnung – BSI-KritisV), 2016.
- [BN15a] BNetzA, Bundesnetzagentur: IT-Sicherheitskatalog gemäß § 11 Absatz 1a Energiewirtschaftsgesetz, 2015.
- [BN15b] BNetzA, Bundesnetzagentur: IT-Sicherheitskatalog gemäß § 11 Absatz 1b Energiewirtschaftsgesetz, 2015.
- [BS17a] BSI, Bundesamt für Sicherheit in der Informationstechnik: Gesetz zur Umsetzung der Richtlinie (EU) 2016/1148 des Europäischen Parlaments und des Rates vom 06. Juli 2016 über Maßnahmen zur Gewährleistung eines hohen gemeinsamen Sicherheitsniveaus von Netz- und Informationssystemen in der Union, 2017.
- [BS17b] BSI, Bundesamt für Sicherheit in der Informationstechnik: Die Umsetzung des IT-Sicherheitsgesetzes aus Sicht des BSI, 2017.
- [BT09] BT, Deutscher Bundestag: Gesetz über das Bundesamt für Sicherheit in der Informationstechnik (BSI-Gesetz – BSIG), (BGBl. I S. 2821), Berlin, 2009.
- [BT15] BT, Deutscher Bundestag: Entwurf eines Gesetzes zur Erhöhung der Sicherheit informationstechnischer Systeme (IT-Sicherheitsgesetz). Gesetzentwurf der Bundesregierung, Drucksache 18/4096, Berlin, 2015.
- [BW20] BWB, Berliner Wasserbetriebe: Zur aktuellen Berichterstattung im Tagesspiegel zum von uns beauftragten Sicherheitsgutachten, <https://www.bwb.de/de/24762.php>, Stand: 11.06.2021.
- [DK19] DKGEV, Deutsche Krankenhausgesellschaft e. V.: Branchenspezifischer Sicherheitsstandard für die Gesundheitsversorgung im Krankenhaus, Berlin, 2019.
- [DV17] DVGW, Deutscher Verein des Gas- und Wasserfaches e.V.: Technischer Hinweis – Merkblatt DVGW W 1060 (M) IT-Sicherheit – Branchenstandard Wasser/Abwasser, Bonn, 2017.

- [JF20] Jansen, F.; Fiedler, M.: Wasserbetriebe gegen Hackerangriffe mangelhaft geschützt, <https://www.tagesspiegel.de/politik/gutachten-warnt-vor-zusammenbruch-wasserbetriebe-gegen-hackerangriffe-mangelhaft-geschuetzt/26045264.html>, Stand: 11.06.2021.
- [Up18] UP KRITIS: Branchenspezifischer Sicherheitsstandard zur IT-Sicherheit. UP KRITIS – BAK Datacenter & Hosting mit CDN, 2018.

Workshop: International Workshop on Digital Forensics (WDF)

Hypothesis-driven Case Analysis in 3D-Space as a Support for Forensic Casework

Sven Becker,¹ Tommy Bergmann,² Steffen Peldschus,³ Dirk Labudde⁴

Abstract: Interpreting the evidence found at a crime scene is essential in reconstructing the circumstances of a crime and, hence, solving it. In this paper a classical hypothesis-driven approach is combined with computer-aided modeling. Hereby, the paper focuses on the advantages of 3D models and their added value in the reconstruction of a case by visually supporting the analysis of evidence found at a crime scene. Using the example of blood patterns, it is explained what the static part of an evidence is and how the interpretation of this part, with the help of an interpretation function, results in additional information about the dynamic part of evidence. Further it is explained how this concept can assist the classical hypothesis-driven approach of reconstructing a crime. By using 3D models it is possible to improve the analysis of the formation of evidence and especially the time line. Thus, 3D-models can aid classical forensic casework with a new computer-based approach.

Keywords: Forensic Casework; 3D Modelling; Hypothesis; Evidence Highlighting; Bloodstain Pattern Analysis (BPA)

1 The concept/definition of evidence in forensic casework based on a real example

When working a case all evidence and witness reports need to be analysed objectively and conscientiously in order to reconstruct the circumstances of a crime and as a result form hypotheses. These then have to be defended in front of a court of law using the available evidence. In order to ensure the objectivity during this process, several models have been proposed of how to handle the information [AA12; Ag11; Ah02; Mo17; Sa17]. Evidence found at a crime scene is used as a foundation in all of the models.

However, it is not clear what should be collected as evidence when working at a crime scene. In 1789 Quistorp first mentions that all the evidence and other circumstances in connection with a crime should be examined by an expert [Qu89]. The definition used by German

¹ University of Applied Sciences Mittweida, Faculty of Applied Computer Sciences & Biosciences, Technikumplatz 17, 09648 Mittweida, Germany, becker1@hs-mittweida.de

² University of Applied Sciences Mittweida, Faculty of Applied Computer Sciences & Biosciences, Technikumplatz 17, 09648 Mittweida, Germany, bergmann@hs-mittweida.de

³ Ludwig Maximilian University of Munich (LMU), Medical Faculty, Institute of Forensic Medicine, Nußbaumstraße 26, 80336 München, Germany, steffen.peldschus@med.lmu.de

⁴ Fraunhofer Institute for Secure Information Technology, Rheinstraße 75, 64295 Darmstadt, Germany, dirk.labudde@sit.fraunhofer.de

law enforcement agencies states, that all visible or latent material changes that could be connected to a criminal event and can contribute to solving a crime should be considered evidence. This means, that all changes at a crime scene can be defined as evidence [Ge86].

Interpreting the meaning and chronological order of evidence found at a crime scene presents some limitations and risks, as it might be misinterpreted based on subjective assessments. Consequently, it can take longer until a case is solved and important details might be ignored. A study from 2017 has shown that prior knowledge of a crime scene can lead to evidence being treated significantly different [EPV16]. A logical evidence evaluation system is recommended for an objective assessment. Furthermore, even though misinterpretations of evidence may not be avoided it can be minimized by logically linking the evidence [SML18].

A basic representation of common analysis models is shown in Figure 1. In investigative procedures analogous conditions are attached to the process of hypothesis development. Scientific methods are used to generate information from the evidence (data, facts, numerals). Scientific methods are procedures, algorithms and systems that have been evaluated and meet a well-defined standard. Information are summarised into case-specific hypotheses and explanations and leads to the actual hypothesis, which is evaluated in the course of an iterative process. Fig. 1 shows a hypothesis cycle for the reconstruction of a crime scene. Information and the resulting knowledge provide the basis for the verification or falsification process of forensic hypotheses.[LS17] The hypotheses cycle can be continued until the circumstances of the crime are reconstructed to a satisfactory degree.

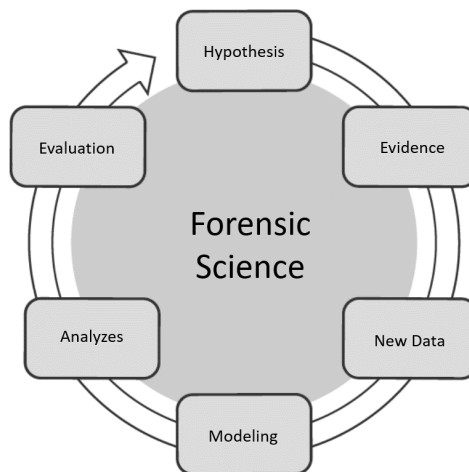


Fig. 1: Procedure of a classical analysis of a case. The hypotheses cycle represents an iterative process. The starting point is the formulation of hypotheses, followed by the analysis of the evidence in order to gain new data to model possible scenarios. The results are then evaluated and the original hypotheses can be accepted or rejected.[LS17]

Based on a real case, analyses and procedures of this concept will be explained in more detail in this publication. The case is a homicide from 2005 [Re18]. The victim was a middle-aged woman who was killed by several blows to the head in her apartment. In 2015, an extensive analysis of the case was performed. The data basis was formed by files, photographs and autopsy reports. During the investigation process from 2005 numerous evidence were found, especially blood traces. In order to gain a better understanding of the crime scene conditions at the time and to be able to analyse the relationship of individual evidence, a true-to-scale 3D model of the crime scene was created by an on-site visit. Due to the fact that the crime was committed in 2005, the actual crime scene was no longer in its original state and only the floor plan could be digitized. Based on the data mentioned above the victim, the evidence in form of the blood traces and pieces of furniture were enriched as digital models itself in the created 3D model of the floor plan. This made it possible to create a scale model of the entire crime scene. The case itself was chosen because the added value of the procedures presented here can be particularly emphasized and were explicitly applied to this case. Because of the usage of the hypothesis-driven approach in combination with the 3D model of the crime scene with the enriched blood traces, it was possible to formulate hypotheses and to chronologically order the sequence of events of the crime. On the basis of the case this mentioned combination of a hypothesis-driven approach of analysing a case and three-dimensional data is introduced. More detailed information on the procedure above, in particular on the significance of the static and dynamic part of an evidence (inspired by [Mu21]), can be found in the following chapters.

2 Static and dynamic part of evidence

In this paper we want to bring additional attention to an extension of the classic term of evidence: the static and the dynamic part of evidence.

In many cases this distinction, allows the development of new hypotheses without increasing the amount of the collected evidence. In this context, the static part refers to the appearance, condition and geometry of the evidence at the time of its collection. Examples include the shape and color of a blood pattern, the type and quantity of minutiae on a collected fingerprint or the material composition of a foreign synthetic fiber found on a carpet. In the presented model from Figure 1, the static part would form the basis for the first iteration of hypotheses. In our opinion, the dynamic part of evidence can provide information that should not be underestimated and refers to the formation of pieces of evidence. This means, that one has to consider why a piece of evidence was found at a certain location and how it was formed. The dynamics of how a piece of evidence was formed often provides a context about the relationship of different pieces of evidence to each other. Thus, it is also possible to reconstruct the circumstances of a crime. In the examples mentioned above the dynamic parts could be the calculated convergence area of the blood patterns and the order of their origin, the reconstruction of the position of the fingerprint-leaving person at the scene or the object, from which the synthetic fibre originated, and its current location.

For a better explanation of the formalisation of the presented hypothesis-driven approach a crime scene with crime events is given. These events have left evidence at a fixed point in time. This point in time can be defined as the true time of origin of the evidence, where each evidence has its individual true time of origin. In addition there is also the point in time when the evidence were secured at the crime scene. It should be noted that a secured evidence consists of the two parts, the static and the dynamic part like mentioned above. By analysing these two parts, it is possible to approximate the time of origin of an evidence through a hypothesis-driven approach with the aim of getting as best as possible to the true time of origin of the evidence.

In mathematical terms all evidence, defined as $E = \{e_k\}$ ($k \in \mathbb{N}$ and $k = 1 \dots n$ with $n =$ number of secured evidence e_k) (see 6) from a crime scene are secured at a fixed point in time T_S ($S:=$ secured). The true time of origin of the evidence is defined as $t_{o_true}^{e_k}$ which should be approximated as best as possible through the approximated time of origin $t_o^{e_k}$ of the evidence e_k , defined as $t_o^{e_k}$, by analysing the evidence e_k . On the description level these evidence e_k include a static part e_{k_s} and a dynamic part e_{k_d} (see 1). Based on the static part of evidence e_{k_s} the crime sequence should be analysed initially by a hypothesis-driven approach. This approach leads to an approximated time of origin $t_o^{e_k}$ (see 2), which results from the analysis of e_{k_s} . This approximation of $t_o^{e_k}$ is done via Δt_{e_k} an approximation step in relation to T_S . This results in an initial relation set S , which maps the relation between the evidence and their initially approximated time of origin defined as $S = \{e_k, t_o^{e_k}\}$, based on the analyses of the static part of evidence e_{k_s} (see 3). With the help of an interpretation function $I(e_{k_s})$ to interpret the static part, additional information of the dynamic part e_{k_d} can be generated in the form of connected time shares $t_o^{e_{k_s}}$ and $t_o^{e_{k_d}}$ for the best approximation of $t_{o_true}^{e_k}$ by $t_o^{*e_k}$. By analysing the dynamic part of evidence e_{k_d} , the approximated time of origin of the evidence $t_o^{e_k}$ can be adjusted to $t_o^{*e_k}$ based on the new information (see 4), where $t_o^{*e_k}$ are the best approximations of the true times of origin $t_{o_true}^{e_k}$. This results in an adapted relation set S^* with the evidence and the adapted time of origin $t_o^{*e_k}$ defined as $S^* = \{e_k, t_o^{*e_k}\}$ (see 5). A formalized representation of these processes is shown in equation (1)-(5) and visualised in fig. 2.

crime scene



$$e_k := e_{k_s} \cup e_{k_d} \tag{1}$$

$$e_{k_s} \Rightarrow T_S - \Delta t_{e_k} \Rightarrow t_o^{e_k} \tag{2}$$

$$S = \{e_k, t_o^{e_k}\} \tag{3}$$

$$e_{k_d} := I(e_{k_s}) \Rightarrow t_o^{e_{k_s}} \pm t_o^{e_{k_d}} \Rightarrow t_o^{*e_k} \tag{4}$$

$$S^* = \{e_k, t_o^{*e_k}\} \tag{5}$$

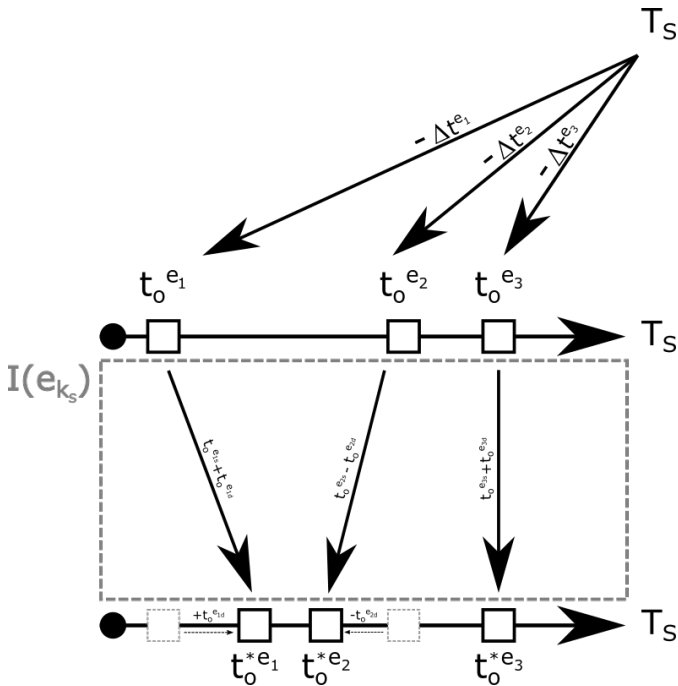


Fig. 2: Schematic representation of the formalisation presented in (see 1-5) based on three example evidence e_1 - e_3 . Through a hypothesis-driven approach, it is possible to approximate the time of origin for each individual evidence at the crime scene. This approximation is initially based on the static part of an evidence by an approximation step Δt_{e_k} in relation to the time of securing T_S of the evidence. By interpreting the static part of the evidence through an interpretation function $I(e_{k_s})$, information about the dynamic part of the evidence is obtained in the form of connected time shares $t_{e_{k_s}}^{e_{k_s}}$ and $t_{e_{k_d}}^{e_{k_d}}$. Based on this new information, approximated times of origin $t_0^{e_k}$ of evidence can be further adjusted to $t_0^{*e_k}$ to best approximate the true time of origin.

Only by analysing both parts of the evidence it is possible to get as close as possible to the true time of origin.

Nowadays, methods of 3D modelling are increasingly used to support the documentation of crime scenes because the technology allows investigators to make analyses at a later point in time [TSP19]. 3D modelling offers great added value in computer-assisted case analysis, especially in the possibilities of analysing a temporal sequence of events in 3D space. The following chapter first provides an introduction to the possibilities of digitising crime scenes, followed by a hypothesis-driven case analysis in 3D space based on blood traces of the presented case example.

3 Hypothesis-driven case analysis in 3D space based on blood traces

Modern visualization technology with software that is becoming increasingly user-friendly and constantly evolving hardware is finding its way into most areas of our everyday life. This is also the case in forensic casework, for which different tools can be used to digitally preserve crime scenes including complex evidence constellations [Be18; DLE16; Sp16; Sü19]. Many of these tools are already used routinely. However, despite their widespread use, they are often only used to create a simple model of a crime scene, which on itself offers only a limited added value. Currently, additional analysing options as well as the combination of these are rarely used. This means that the tools currently in use are mostly used to document a crime scene, at the best in 3D. Within such created 3D models, for example visual inspections and distance measurements between evidence are carried out. A possible limitation of the tools used may also be in the processing of cold cases. In these cases, there are usually no 3D data of the crime scene available, as 3D documentation may not have reached the necessary technical progress at that time. This means that a crime scene or the area where the crime took place in the past can theoretically be documented in 3D in the present, but a multitude of further process steps are necessary to reconstruct the actual crime scene in its entirety (e.g. including evidence) and thus to be able to use the added value of such models in combination with hypothesis-driven approach of case analysis like presented in this paper.

Within this publication, it will be shown how a combination of these two methods can improve forensic casework. Figure 3 shows a 3D model of a crime scene as it is already often created by crime scene investigators when processing the crime scene. The model shown here was created on the basis of on-site measurements at the crime scene. In open terrain, drones provide a means of creating aerial photographs. The refinement, texturing, and data enhancement of the 3D model presented here was done in the open source software Blender (<https://www.blender.org/>). Blender is a free open source 3D modeling suite that provides the entire 3D pipeline - modeling, rigging, animation, simulation, rendering, compositing, motion tracking and video editing. All secured evidences, possibly relevant to the crime, are included in the model. In order to keep the example as simple as possible, only blood patterns are of particular interest and used to illustrate the principle presented

in this paper. The blood traces secured at the crime scene were classified, if possible, according to the types defined in [BG08]. In order to optimally illustrate the principle of hypothesis-driven case analysis with the inclusion of 3D models, only those blood traces were used that could be clearly assigned to a type. This was crucial, as the dynamic part of the evidence is defined by the different processes of formation of different evidence, which can be used to extract information about the sequence of events. However, this is only possible if their type can be clearly determined. For example, a cast-off pattern found on the wall shows evidence of active violence and a pool of blood shows evidence of a blood spill that was influenced only by gravity. In this example all these secured and classified blood traces, corresponds to the defined $E = \{e_k\}$ with $k \in \mathbb{N}$ and $k = 1, 2, 3, 4, 5$ (see 6).

$$E = \{e_1, e_2, \dots, e_5\} \text{ with} \quad (6)$$

- e_1 := wipe pattern
- e_2 := cast-off pattern
- e_3 := drip pattern
- e_4 := contact pattern
- e_5 := pool of blood

Even though this is just a basic model and no additional information has been added, it can already offer some advantages for the criminal investigator. With a 3D model it is possible to virtually walk through the crime scene without the risk of changing anything. This makes it easier to examine individual pieces of evidence and their position (static part of the evidence). Furthermore, the relationship of the individual pieces of evidence to each other is easier to see in the model, as one is not limited to classical crime scene photographs as it is often not possible to capture all the relations between the pieces of evidence. Additionally, it is possible to discuss preliminary hypotheses about the time of origin of evidence directly in the model, which corresponds to the defined initial relation set $S = \{e_k, t_o^{e_k}\}$. It is also easy to share the findings as the model can be easily sent to other devices, where it can be viewed by others.

All these advantages thus allow better recognizability and assignability of individual evidence. The static part of evidence (appearance, condition and geometry) and especially the dynamic part of evidence (formation dynamic) can be captured and displayed more easily in the 3D model. Thus, 3D modeling becomes a helpful tool for investigators, supporting a hypothesis-driven case analysis.

Fig. 3A shows possibilities of using a 3D model to analyse the evidence of the crime scene of the presented real case example. Highlighting is used in order to quickly identify the blood traces and to form knowledge-based hypotheses about their origin. Here, it can be

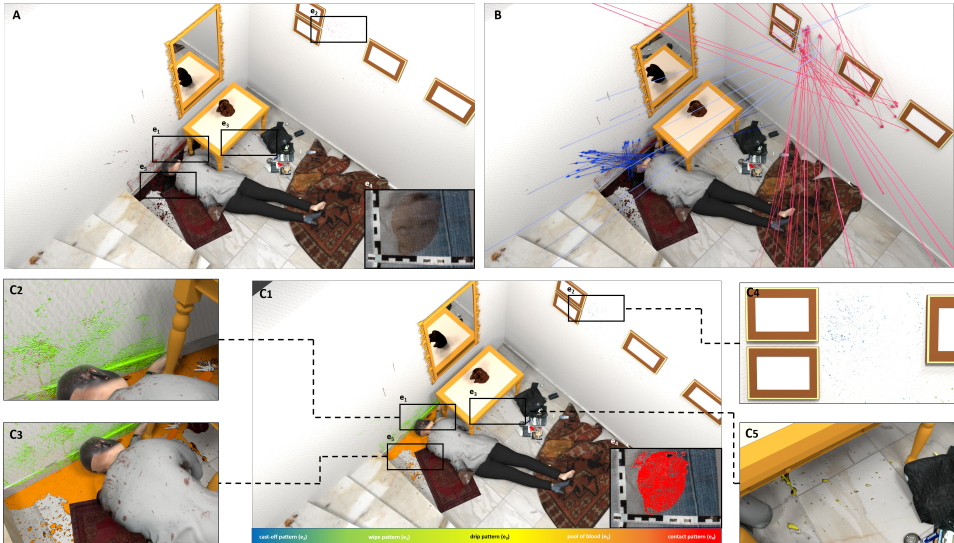


Fig. 3: A: 3D model of a crime scene, where the black squares show the different blood patterns on the ground and wall of the crime scene. The different blood patterns are: e_1 : wipe pattern e_2 : cast-off pattern e_3 : drip pattern e_4 : contact pattern and e_5 : pool of blood. The contact pattern shown is a part of trousers belonging to a witness who was allegedly at the crime scene. Even such blood traces, even if they were not found directly at the crime scene, can be classified in the chronology of the formation of blood traces. All these different blood patterns represent the static part of the evidence. B: 3D-Model after interpretation of the static part of evidence resulting in the area of origin (dynamic part). C: The different pieces of evidence were highlighted in different colors in order to visualize the chronological order. The coloring of the blood traces depends on the time of origin. Blue marked blood traces are younger than red marked ones. The smaller figures left- and right-hand-side of C₁ (C₂-C₅) represent the colored blood traces and serve as a better overview.

seen that the overview of the scene makes it easy to record and evaluate the position of the individual blood traces in relation to each other. Thereby, it is possible to create different categories in which individual blood traces can be grouped together. These categories can be easily identified, for example by using different colors for each category. This makes it easy to identify relationships between blood traces by using visual highlighting and to present these to a third party.

Analysing the static part of evidence (appearance, condition and geometry) in our example, it quickly becomes clear that the blood pool must have been created after the active blood traces on the walls, provided that all blood stains originate from the same person. The size of the blood pool says that the blood loss was fatal, so it is possibly the final position of the corpse. An origin of the evidence on the walls by active force impact on the victim walking around after the blood pool origin is therefore not plausible. Under what circumstances the contact pattern was created is not yet clear. After evaluation of the static part of evidence, a

first hypothesis about the time of origin of evidence arises, which can be summarized as the following relation set S in (see 7).

$$S = \{e_1, t_o^{e_1}; e_2, t_o^{e_2}; e_3, t_o^{e_3}; e_5, t_o^{e_5}\} \quad (7)$$

By using a model as shown in Fig. 3A it is easier to gain new knowledge from the given evidence. For example, by logically combining the location of the individual blood traces and their relationship to each other, it is possible to determine the dynamic part of the evidence. In the hypothesis-driven approach presented here, this corresponds to the interpretation of the static part with the help of an interpretation function. An example of this is the interpretation of the shapes of several blood traces and their relation to each other, which allows to determine the so-called area of origin [Gr21] (Fig. 3B), the location of the blood source at the time of trauma. Fig. 3C shows this new information with the help of color spectrum, where blue represents the piece of evidence that was formed first and red the one that was formed last. With the help of the color spectrum, it should be made clear how the blood traces were created temporal in relation to each other, independent of the time of securing.

This hypothesis of the sequence of events is based on the following knowledge-based analyses of the pieces of evidence in relation to each other and by including and by interpreting data (static part) recorded at the crime scene to gain new information (dynamic part). The static part of the blood traces corresponds to the position, shape and arrangement of the secured blood drops. These were interpreted with the help of an interpretation function, i.e. analyses were carried out with regard to the distance between individual blood drops and other blood patterns, as well as with regard to impact angles and direction. From this, calculations of the so-called area of origins were made. These area of origins correspond to the defined dynamic part of an evidence.

The cast-off pattern (evidence e_2) between the small picture on the wall indicates that the victim must have been standing near this position and was hit with great force. Therefore in comparison to the other secured blood traces this blood pattern was probably created in an initial phase of the crime act. In this case, interpreting the position of the blood traces to each other using the interpretation function $I(e_{k_s})$ results in the determined area of origin, which can be used to adjust the approximated $t_o^{e_k}$ to $t_o^{*e_k}$ of the evidence e_k . On this basis it was possible to derive that the cast-off pattern (evidence e_2) may have been originated first in relation to the other evidence (see 8). The approximated time of origin of the other evidence are not yet adjusted at this point.

$$t_o^{*e_2} < \forall e \in t_o^{*e_k} \text{ with } k = 1, 3, 4, 5 \quad (8)$$

Interpreting the location and geometry of e_1 and e_3 it comes to light that the wipe pattern is located on the wall just behind the drip pattern near the hand of the victim. Possibly, the wipe pattern was created when the victim was still alive but with blood adhesion on its hand. These adhesions could have occurred due to possible defensive postures and resulting injuries to the hands, thus leading to the secured blood drops which created the drip pattern. The blood drops had a circular shape due to a possible orthogonal angle of impact. Consequently, the use of the interpretation function $I(e_{k_s})$ to interpret the location and shape of the blood traces resulted in the knowledge that the drip pattern was created slightly staggered up to the same time as the wipe pattern was created (see 9).

$$t_o^{*e_1} \leq t_o^{*e_3} \quad (9)$$

The time of origin can also be adjusted for e_5 using an interpretation function ($I(e_{k_s})$). The size and shape of the contact pattern indicates that the trousers were touched by a large accumulation of blood. At the crime scene, there is only one such accumulation of blood, namely e_4 , the blood pool. This consideration leads to the assumption that the contact pattern on the trousers of the witness was possibly created after the victim was found in the floor (see 10).

$$t_o^{*e_5} < t_o^{*e_4} \quad (10)$$

After interpreting the static part of all evidence e_{k_s} with the help of the interpretation function $I(e_{k_s})$ and analysing the dynamic part of all evidence e_{k_d} the time of origin of the evidence were adjusted from $t_o^{e_k}$ to $t_o^{*e_k}$. In summary this adjustment results in the following order of time of origin of the evidence e_k (see 11).

$$t_o^{*e_2} < t_o^{*e_1} \leq t_o^{*e_3} < t_o^{*e_5} < t_o^{*e_4} \quad (11)$$

This results in the adapted relation set S^* (see 12).

$$S^* = \{e_2, t_o^{*e_2}; e_1, t_o^{*e_1}; e_3, t_o^{*e_3}; e_5, t_o^{*e_5}; e_4, t_o^{*e_4}\} \quad (12)$$

The deduced timeline can now be used to formulate new hypotheses about the people involved and their actions at the crime scene. These new hypotheses can again be visualized in order to get closer to what happened at the crime scene and the actual timeline. Of course, the described procedure of extracting information from the dynamic part of the evidence is

also possible at the actual crime scene or with the help of crime scene photographs and crime scene sketches. However, the intention of this paper is not to point out the novelty of the proposed method, but to demonstrate that a 3D model can assist in this procedure. The 3D model offers the possibility to quickly transfer all results and hypotheses and to present them clearly. Therefore, it should be recognized as an extension to the forensic toolbox.

4 Summary

In summary, it should be noted that by gathering and analysing the static and dynamic part of the evidence it is possible to group pieces of evidence and in result formulate hypotheses about the chronological order in which the evidence was formed to approximate as best as possible the true time of origin of evidence. These new hypotheses should be compared with the situation found at the crime scene in an iterative process until no more questions arise. 3D models are a valuable tool to reconstruct and analyse cases in this way, as they allow a non-invasive and clear analysis of the evidence. By using colors to make the evidence more prominent as demonstrated in this paper, relationships between pieces of evidence as well as different hypotheses can be visually highlighted. This simplifies the presentation of the reconstructed circumstances of the crime between those working the case as well as in front of a court of law. Therefore, including the dynamic part, in addition to the static part, of the evidence when working a case using 3D models is an efficient method and it should be considered to use it more often in the near future.

References

- [AA12] Al-Fedaghi, S.; Al-Babtain, B.: Modeling the forensics process. *International Journal of Security and Its Applications* 6/4, pp. 97–108, 2012.
- [Ag11] Agarwal, A.; Gupta, M.; Gupta, S.; Gupta, S. C.: Systematic digital forensic investigation model. *International Journal of Computer Science and Security (IJCSS)* 5/1, pp. 118–131, 2011.
- [Ah02] Ahmad, A.: The forensic chain of evidence model: Improving the process of evidence collection in incident handling procedures. In: *The 6th Pacific Asia Conference on Information Systems*. 2002.
- [Be18] Becker, S.; Spranger, M.; Heinke, F.; Grunert, S.; Labudde, D.: A Comprehensive Framework for High Resolution Image-based 3D Modeling and Documentation of Crime Scenes and Disaster Sites. *International Journal on Advances in Systems and Measurements*, 2018.
- [BG08] Bevel, T.; Gardner, R. M.: *Bloodstain pattern analysis with an introduction to crime scene reconstruction*. CRC press, 2008.

- [DLE16] Dustin, D.; Liscio, E.; Eng, P.: Accuracy and repeatability of the laser scanner and total station for crime and accident scene documentation. *J Assoc Crime Scene Reconstr* 20/, pp. 57–68, 2016.
- [EPV16] van den Eeden, C. A.; de Poot, C. J.; Van Koppen, P. J.: Forensic expectations: Investigating a crime scene with prior information. *Science & Justice* 56/6, pp. 475–481, 2016.
- [Ge86] Geerds, F.: Archive for criminology with special consideration of the judicial Physics. *Chemistry and Medicine* 177/, 1986.
- [Gr21] Griffiths, G.; Liscio, E.; Guryan, H.; Le, Q.; Northfield, D.; Williams, G. A.: Inter-observer error for area of origin analysis using FARO Zone 3D. *Science & Justice* 61/3, pp. 291–298, 2021.
- [LS17] Labudde, D.; Spranger, M.: *Forensik in der digitalen Welt: Moderne Methoden der forensischen Fallarbeit in der digitalen und digitalisierten realen Welt*. Springer, 2017.
- [Mo17] Morgan, R.: Conceptualising forensic science and forensic reconstruction. Part I: a conceptual model. *Science & Justice* 57/6, pp. 455–459, 2017.
- [Mu21] Mukhra, R.; Krishan, K.; Nirenberg, M. S.; Ansert, E.; Kanchan, T.: The contact area of static and dynamic footprints: Forensic implications. *Science & Justice* 61/2, pp. 187–192, 2021.
- [Qu89] Quistorp, J. C.: Versuch einer Anweisung fuer Richter bey dem Verfahren in Kriminal- und Strafsachen, wider solche, welche die Wahrheit nicht gestehen wollen. *Ländern, wo die Tortur abgeschafft wor den*. Leipzig/, 1789.
- [Re18] Reinhard, D.: *Der Münchner Parkhausmord*. Books on Demand, 2018.
- [Sa17] Sack, K.: *Selektion in der digitalen Forensik*, PhD thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), 2017.
- [SML18] Smit, N. M.; Morgan, R. M.; Lagnado, D. A.: A systematic analysis of misleading evidence in unsafe rulings in England and Wales. *Science & Justice* 58/2, pp. 128–137, 2018.
- [Sp16] Spranger, M.; Heinke, F.; Becker, S.; Labudde, D.: *Towards drone-assisted large-scale disaster response and recovery*, 2016.
- [Sü19] Sünksen, M.; Teistler, M.; Hamester, F.; Ebert, L. C.: Preparing and guiding forensic crime scene inspections in virtual reality. In: *Proceedings of Mensch und Computer 2019*. Pp. 755–758, 2019.
- [TSP19] Tredinnick, R.; Smith, S.; Ponto, K.: A cost-benefit analysis of 3D scanning technology for crime scene investigation. *Forensic Science International: Reports* 1/, p. 100025, 2019.

Automated Determination of Fingerprint Ridge Density and Fingerprint Size to Detect Sex Differences

Marleen Mohaupt¹, Sieke Stoeter and Dirk Labudde¹

Abstract. A fingerprint is probably the most important biometric feature when trying to link a suspect to a crime scene. So far, without a hit in a fingerprint database, it was impossible to use a collected fingerprint to narrow down the group of suspects. Moreover, in the existing studies about deriving phenotypic characteristics from fingerprints the analyses were done manually. In contrast, in this paper a procedure is presented to automatically determine the fingerprint ridge density and the fingerprint size, in order to derive information about the sex of the person the fingerprint belongs to. All 10 fingerprints of 140 individuals (70 males and 70 females) belonging to the German Caucasian population were secured and then analyzed. The best result was obtained for the ulnar area in combination with the fingerprint size of the left thumb with F_1 measures of 0.84 (k-nearest neighbors algorithm - KNN), 0.833 (Support Vector Machine) and 0.817 (logistic regression).

Keywords: Fingerprints, Ridge Density, Sex Classification.

1 Introduction

In forensics, fingerprints are the best-known pieces of evidence [1]. Overall, a fingerprint contains three levels of information [2]. The first level detail is the presence of one of the three basic patterns (loop, whorl or arch) [2]. Other anatomical features, the minutiae, e.g. interruptions and branches, are the second level details [2]. Furthermore, the third level details include special features, such as scars, creases or pores [2-4].

Without a hit in a fingerprint database such as the Automated Fingerprint Identification System (AFIS) from INTERPOL, it is impossible to use a collected fingerprint to narrow down the group of suspects, e.g. by sex, solely based on dactyloscopic evidence. In 1999 Mark A. Acree developed a method for determining the ridge density (RD - the number of ridges in a well-defined square) [5], which formed the basis for numerous subsequent studies [6-9]. It was found that the RD depends on several factors: age [5, 10, 11], population, sex [6] and the position of the well-defined area within the fingerprint. However, determining the RD manually is time-consuming and depends on the subjective view of an expert. In this paper, we present a novel approach, which performs this task automatically. As a result, a larger data set can be analyzed in a shorter amount of time.

¹ University of Applied Sciences Mittweida, Technikumpl. 17, 09648 Mittweida, marleen.mohaupt@hs-mittweida.de/ dirk.labudde@hs-mittweida.de

2 Related Works

As demonstrated in various studies [5-9] the ridges in fingerprints of females are closer together and, thus, show a higher RD compared to men. Depending on the location of the measurements (count areas) and the population of the suspects, the discriminant threshold to determine the sex differs. In most papers, these thresholds were determined using Bayes' theorem for both sexes. For a Spanish population, for the radial area, the fingerprint is more likely to be assigned to a male at a density of < 16 lines/ 25 mm^2 and more likely to be assigned to a female at a density of 16 lines/ 25 mm^2 or more. The average values vary due to different populations and the location of the count area.

For the Punjabi population in India Dhall and Kapoor chose a discriminant function and logistic regression to determine the sex. The results of five areas to measure the RD were analyzed and an accuracy of 97.4% using logistic regression was obtained. When the areas are considered individually, an accuracy of 87.2 % is calculated for the radial areas and 82.9 % for the ulnar areas. [6]

Verma and Agarwal, using only the RD without a clear definition of the selected area, achieved a classification rate of 53% using a SVM classifier. Their precision was improved by incorporating additional features. They additionally calculated the ridge thickness to valley thickness ratio and the ridge width and achieved an accuracy of 59.5 % and 68 %, respectively. Using all three features, an accuracy of 91 % was achieved. [12]

Badawi et al. used five features: ridge count, white lines count, ridge thickness to valley thickness ratio, ridge count asymmetry, and pattern type concordance. The ridge thickness to valley thickness ratio was calculated automatically for the scanned ink fingerprints by dividing the image into 30×30 pixel non-overlapping squares and determining the value for each block. Then, the average for all blocks was calculated. Blocks with low image quality were automatically excluded. All other features were determined manually. Three models were created with Fuzzy C- Means. With a linear discriminant analysis, they achieved an accuracy of 86.5 % and a slightly higher accuracy of 88.5 % using a multilayer backpropagation neural network. [13]

In another study (see [11]), fingerprint ridge count and fingertip size (FPS) of an Indian population were determined. The ridge count consists of three parameters: the number of lines located on the straight line between the core point and the delta, and the lines located on the diagonals at 45° and 135° beginning at the core point for the entire area of the fingerprint. The FPS was determined by the Fingkey Hamster II scanner. For the optimal score assignment method, the most frequent values were used as a reference for determining a score for fingerprint ridge count and FPS. Afterwards, a score for male and female was determined from the sum of both values. If the male score was higher than the female score, the fingerprint is most likely from a male, otherwise from a female. Using optimal score assignment, an accuracy of 88.41% was obtained for the right ring finger. [11]

For the analysis of large data sets, parameters for sex differentiation should be determined automatically. In this study, both RD and fingerprint size were automatically determined as possible parameters for sex differentiation and evaluated accordingly.

3 Materials and Methods

During the data acquisition, 1400 fingerprints of 70 females and 70 males were taken. The age of the subjects varied between 19 and 38 years (mean: 22.36, SD: 3.91). Since the RD depends on ethnicity, the study was performed exclusively for Caucasians residing in Germany. We have requested the biological sex and not the gender identity of the subjects (see [14]). Fingerprints were acquired by using a VERIFIER 300 LC 2.0 (CROSS MATCH) finger scanner. From each subject the prints of all 10 fingers were taken, starting with the little finger of the left hand and ending with the little finger of the right hand. Each image has a horizontal and vertical resolution of 500 dpi, matching the FBI standard for forensic applications in order to reliably extract all ridges and minutiae with a dimension of 640x480 pixels. Several studies indicate that arch patterns are observed in about 5 % of all fingerprints while loop patterns occur in about 60 – 70 % of all fingerprints and whorls about 30 % [15-17]. This distribution could also be observed in the data set.

3.1 Determination of the Ridge Density

In order to automatically determine the RDs, a MATLAB [18] program was implemented. Additional modules for image segmentation [19] and enhancement [20] were utilized. Eight steps were applied to each scanner image before the RD was determined (see Fig. 1). Since the determination of the RD requires closed ridges, some special features, such as scars, were closed by the image processing methods used. This represents a loss of information, that was, however, insignificant with respect to classification performance.

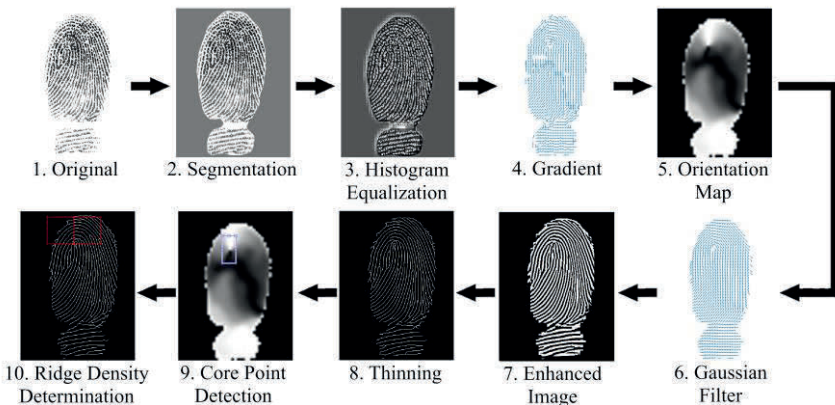


Fig. 1. Automated ridge density estimation, requires the reliable location of the fingerprint core. To that end, multiple image pre-processing steps are involved prior to core detection.

Segmentation, which separates the actual fingerprint from the background, was used to remove artefacts that occur during the scanning process. This method was developed and published by Fahmy and Thabet [19].

In the scanned fingerprint, the intensity of the grey values varies greatly, so that many papillary lines are only faintly visible. Image enhancement highlights, connects and evenly colors the lines so that the program can capture each line. Bifurcations and ending lines are not altered in the process. Islands are also preserved in the enhanced image. For the enhancement, the algorithm implemented by Houjun Jiang [20] was adopted and only the block size changed. Since local operations give a better result than the global processing of the whole image, the fingerprint was divided into several non-overlapping blocks and improved systematically. Fig. 1 above shows the effect of each step on a fingerprint scan. In the enhanced image, it can be seen that pores are no longer visible. In other scanner images, scars and wrinkles are no longer displayed and all lines are closed. Only the furrow of the first phalanx can be seen in many enhanced images.

The variations in grey levels are due to the varying pressure of the finger on the scanner. In order to simplify further image analysis, the contrast of non-high contrast areas had to be increased. This can be realized by using *histogram equalization* ("histeq" function). This involved stretching the intensity of the pixels analyzed in the histogram so that the differences between dark pixels (lines) and light pixels (valleys) are better emphasized. This also achieved an increase in contrast. A desirable side effect was the removal of visible sweat pores. However, scars and wrinkles were still visible after this step. Next, the gradient image was created. This step is necessary for the creation of the orientation map and the detection of the edges. The gradient indicates the strength of the change in direction of a mathematical function. With a small gradient value and, thus, a small change in direction, the papillary lines run parallel to the calculated line course. To obtain more uniform lines in the enhanced image, the gradient image is smoothed. For this purpose, the *Gaussian filter* with a block size of 5x5 pixels was used. However, this can cause the areas where the direction of the lines is not constant to be incorrect. This mainly affects the areas immediately around the core point and the delta. Therefore, some fingerprints may have slight artefacts that are more noticeable after thinning. This problem is addressed when counting the lines. To further highlight the papillary lines, a bandpass filter was applied. The Gabor filter is based on sinusoidal frequencies created using the Gaussian function ("imgaborfilt" function) which extracts and highlights the papillary lines with a block size of 32. This allows areas with a strong structural change to be captured and enhanced. The gradient image was also used to create an *Orientation Map*. The average change in direction was calculated for each block. The larger the value the darker is the area. Since the curvature of the ridges is strongest around the core point, a characteristic line was obtained in the orientation map, at which upper end the core point is located [21]. To simplify the counting of fingerprint lines, the lines in the enhanced image are reduced to a width of one pixel, using common *thinning* techniques and an inversion. To do so, the internal MATLAB [18] function *bwmorph* was used. The procedure described in [10] of setting squares to determine the RD directly above the core point was also applied here. Bahgat et al. used the orientation map and discovered that the image values around the core point follow a certain pattern: In the neighboring pixels, the change in direction increases from almost 0° to almost 180° anti-clockwise [21]. This *core point detection* was extended by using template matching to process the different types of orientation maps for the basic

patterns. Furthermore, a cross-correlation was carried out to detect the area of the orientation map that most closely matches one of four templates due to the different orientation maps of the ridge patterns. The center of the template corresponds to the core point of the fingerprint. For the left hand, the left square corresponds to the ulnar and the right square to the radial area. For the right hand, the assigned areas are reversed [10]. For the *RD determination*, the lines that intersect an imaginary diagonal of the respective square i.e. white pixels on this line should be counted (see Fig. 2). To detect all lines, neighbors of black pixels on the diagonal also had to be checked for white pixels. Possible additional pixels, which could result from the image enhancement, were only interpreted as one line. For this purpose, white pixels on the straight line with a distance less than or equal to three were counted as one line. Unfortunately, some fingerprint were not fully captured. Thus, the squares reach over the fingerprint, resulting in a lower RD. To use it anyway, the distance between the last found line and the upper square corner is determined, and additional lines are estimated using the median of the given ridge distances (see Fig. 2).

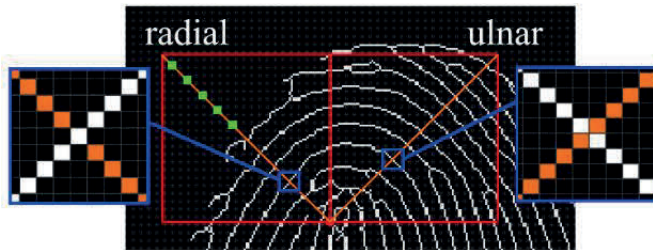


Fig. 2. Magnification of the intersections of the papillary lines (white) with the imaginary diagonal (orange) of the radial and ulnar regions (red) for fingers of the right hand and intersections with the estimated lines (green dots)

In the left magnification (blue), the white dot lies directly on the diagonal and is evaluated as a papillary line. In the right magnification (blue), the white pixels of the papillary line are not on the diagonal. Therefore, for black pixels, the top three neighboring pixels were checked for a white color value. After each hit, the following pixel on the diagonal was skipped to avoid double counting of lines.

3.2 Determination of the Fingerprint Size

As a second parameter to differentiate the sex, the fingerprint size (FPS) was determined to approximate the fingertip size. For its determination, the segmentation mask from the first step was used. The segmented area of the fingerprint was displayed in white pixels while the background was dyed black. Since there are significantly more artefacts in some images, which are not removed by the segmentation function in every case, the brightness of the image was increased. The FPS was calculated with the following formula:

$$fingerprint\ size = \frac{Number\ of\ white\ pixels * 25.4^2}{dpi^2}. \quad (1)$$

4 Results

1272 of 1400 core points were detected correctly, which corresponds to an overall accuracy of 90.85 %. The automatic core point detection shows an accuracy of 95.31 % for whorls, 91.29 % for loops and 87.61 % for arches. The orientation maps of plain arches do not show strong gradient changes and, therefore, the core point is difficult to find with the used templates. Hence, an adaptation of the core point detection for this specific problem is needed. The core point serves as a reference point for locating the squares with the chosen size of 75x75 for the radial and ulnar areas. Additionally, 27 fingerprints were incomplete and the squares covered an area outside the actual fingerprint, therefore the missing epidermal ridges had to be estimated.

For the male subjects a mean ridge density of 12.60 (1.82) was calculated in the radial and 12.56 (2.07) in the ulnar area. For the female subjects a mean value of 13.80 (2.00) in the radial and 13.71 (2.04) in the ulnar area were obtained. The exact distribution of the RD is visualized in Fig. 4. The values range from 7 to 19 with a mode of 12 (radial) and 7 to 20 with a mode of 13 (ulnar) in the male group. For the females the values are between 7 and 20 with a mode of 13 (radial) and 8 and 20 with a mode of 14 (ulnar). Overall, it can be seen that women have a higher ridge density than men for both radial and ulnar values and a statistically significant difference was observed (Mann-Whitney-U test, $p < 0.05$). However, the distribution of the RD for both sexes overlap. Instead of using a single threshold value to assign the correct sex, logistic regression, SVM and KNN were used to predict the sex. To increase the probabilities, a second feature was determined. The surface area of the segmentation mask was calculated in mm^2 for each fingerprint. For males, the values range from 102 mm^2 to 372 mm^2 with a mode of 203 mm^2 and a mean value of 201.45 mm^2 . For females, the values range from 94 mm^2 to 349 mm^2 with a mean value of 168.49 mm^2 and a mode of 145 mm^2 .

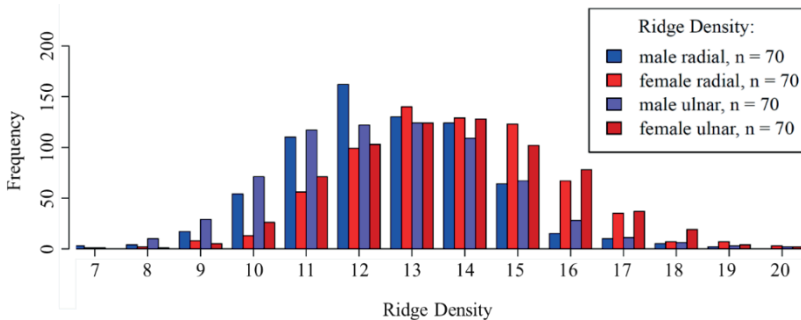


Fig. 3. Frequency distribution of the RD in the ulnar and radial area for the subjects

A logistic regression (LR) model and several support-vector machine (SVM) models were trained to classify and predict the sex. Additionally, the classic k-nearest neighbors algorithm (KNN) was used in order to evaluate the baseline classification performance. With $k = 11$ k-NN, LR and SVM models were trained and tested using a 10-fold cross validation. The experimental results show that with only one feature of the RD, the classification

rate is quite low. For the test data set, the best value for the radial RD was obtained for the right middle finger with an accuracy of 66.6 % and an F_1 measure of 0.666 using LR. For the ulnar RD the accuracy was 64.57% and the F_1 measure 0.717. For the left thumb, which provided the best results for the analyses with the FPS, the accuracy for the radial RD was 59.66 %. However, the F_1 measure was only 0.609. For the ulnar RD, both the accuracy with 63.98 % and the F_1 measure of 0.653 were slightly higher. The values calculated with the SVM and the KNN do not deviate much from the other values (see Table 1). Overall, the results improved when adding FPS as a feature.

Table 1. Classification Accuracy (A) and F_1 measure (F_1) for the Fingerprint Size (FPS), the radial RD (rRD) and ulnar RD (uRD) using LR, SVM and KNN algorithm

Finger	FPS		rRD		uRD		rRD+uRD+FPS		uRD+FPS	
	A [%]	F_1	A [%]	F_1	A [%]	F_1	A [%]	F_1	A [%]	F_1
1 (LR)	67.37	0.707	64.04	0.612	55.12	0.572	73.57	0.758	71.77	0.744
2 (LR)	73.68	0.767	67.35	0.627	57.99	0.589	73.37	0.766	73.39	0.753
3 (LR)	74.35	0.768	65.55	0.648	66.87	0.666	73.06	0.752	74.68	0.771
4 (LR)	67.17	0.704	60.43	0.599	58.85	0.588	69.12	0.711	68.61	0.715
5 (LR)	76.35	0.784	59.66	0.589	63.98	0.636	78.13	0.786	79.58	0.817
6 (LR)	76.88	0.775	61.86	0.604	56.25	0.577	78.23	0.794	74.52	0.768
7 (LR)	73.87	0.763	62.93	0.598	57.75	0.561	70.94	0.724	70.07	0.728
8 (LR)	68.53	0.734	66.60	0.650	64.57	0.674	66.88	0.695	65.20	0.685
9 (LR)	67.26	0.712	63.26	0.625	54.50	0.563	70.41	0.738	70.11	0.733
10 (LR)	68.42	0.719	61.31	0.627	59.67	0.556	72.68	0.750	74.35	0.762
1 (KNN)	67.80	0.701	57.13	0.560	55.24	0.611	68.14	0.687	71.76	0.735
2 (KNN)	71.28	0.735	61.65	0.673	58.98	0.514	73.58	0.751	68.09	0.711
3 (KNN)	72.44	0.671	59.83	0.633	63.97	0.661	73.97	0.742	66.47	0.658
4 (KNN)	64.43	0.655	64.03	0.652	65.15	0.661	69.30	0.715	65.92	0.700
5 (KNN)	72.94	0.754	57.13	0.552	69.01	0.624	79.03	0.802	82.37	0.840
6 (KNN)	73.32	0.754	53.84	0.552	51.67	0.523	74.70	0.762	76.36	0.749
7 (KNN)	74.75	0.757	62.42	0.640	45.00	0.497	71.49	0.739	64.48	0.655
8 (KNN)	68.12	0.692	65.38	0.703	63.59	0.658	66.90	0.691	60.12	0.629
9 (KNN)	68.23	0.727	62.53	0.608	62.16	0.634	69.54	0.722	69.71	0.725
10 (KNN)	66.79	0.696	54.79	0.596	47.89	0.551	69.54	0.725	69.15	0.729
1 (SVM)	72.67	0.749	60.73	0.613	54.79	0.507	73.69	0.767	70.08	0.729
2 (SVM)	73.26	0.750	64.81	0.681	61.73	0.641	74.75	0.779	73.26	0.754
3 (SVM)	74.05	0.770	63.00	0.632	60.96	0.605	73.04	0.734	69.48	0.705
4 (SVM)	66.93	0.649	60.48	0.6406	61.69	0.629	72.65	0.722	71.54	0.696
5 (SVM)	75.36	0.781	60.15	0.612	72.26	0.719	82.56	0.833	80.51	0.832
6 (SVM)	75.79	0.784	60.29	0.597	56.95	0.551	75.67	0.766	74.50	0.748
7 (SVM)	74.13	0.756	65.51	0.674	55.41	0.554	69.79	0.681	69.18	0.728
8 (SVM)	68.07	0.694	64.45	0.615	56.10	0.590	70.50	0.729	71.68	0.728
9 (SVM)	67.92	0.711	62.45	0.584	64.11	0.646	70.63	0.701	68.38	0.690
10 (SVM)	70.73	0.741	64.14	0.618	59.89	0.590	74.29	0.714	73.21	0.737

Figure 4 shows that the FPS given in mm^2 is higher for males than for females. The statistically significant difference in size resulted in a higher classification probability especially for the thumbs: for the left thumb, an accuracy of 76.35 % and an F_1 measure of 0.784 were calculated and for the right thumb, an accuracy of 76.88 % and an F_1 measure of 0.775 were estimated with LR (see Table 2). Thus, for the FPS a better prediction could be made than for the RD. However, the best results were obtained by combining both features. For the ulnar RD and FPS, an accuracy of 79.58 % and an F_1 measure of 0.817

were calculated for the left thumb using LR. The SVM increased the accuracy to 80.51 % and the F_1 measure to 0.832. The highest value was achieved with the KNN with an accuracy of 82.37 % and an F_1 measure of 0.84, which could not be exceeded even by the additional use of the radial RD.

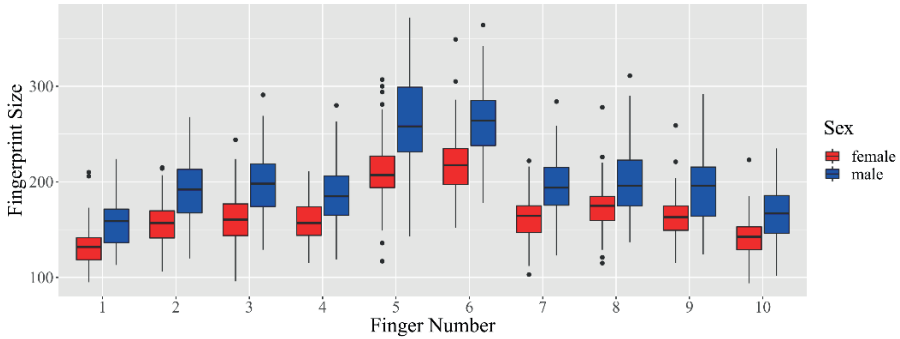


Fig. 4. Per-finger FPS distributions of all 140 subjects.

Table 2. Mean (meanFPS) and median fingerprint size (medianFPS) for each finger

Finger	meanFPS [mm ²] (SD) males	females	medianFPS [mm ²] (IQR) males	females
1	159.4 (26.6) 133.3 (21.8)		159.0 (35.0) 132.0 (23.0)	
2	190.3 (35.2) 156.1 (22.7)		192.0 (45.5) 157.0 (28.5)	
3	196.4 (37.1) 163.1 (28.2)		198.0 (44.5) 160.5 (33.3)	
4	187.4 (49.2) 159.0 (22.2)		185.0 (44.8) 157.0 (29.8)	
5	261.1 (41.8) 212.1 (35.8)		258.0 (67.5) 207.0 (33.0)	
6	263.7 (31.4) 219.8 (33.0)		264.0 (47.0) 217.5 (37.5)	
7	194.4 (34.3) 161.8 (24.2)		194.0 (39.5) 164.5 (28.0)	
8	201.3 (36.6) 172.9 (26.4)		196.0 (47.8) 175.0 (25.3)	
9	193.8 (29.3) 165.0 (23.5)		196.0 (51.3) 163.0 (25.5)	
10	166.8 (26.6) 141.2 (21.1)		167.0 (39.5) 142.5 (24.0)	

5 Discussion

Within this study, an automatic determination of the fingerprint ridge density and the fingerprint size, to derive information about the sex of the person the fingerprint belongs to, has been implemented. Dhall and Kapoor achieved a classification accuracy of 97.4%, by using five areas for the determination of the RD [6]. For our dataset, only the radial and ulnar regions could be used because the fingerprints were taken without rolls and, thus, the area to be analyzed is much smaller. Considering only the RD, Verma and Agarwal's value of 53% [12] could be raised to an average value of 63.29% (LR), 62.60 % (SVM) and 59.87% (KNN) for the radial region and 59.55% (LR), 60,38 % (SVM) and 58.26% (KNN) for the ulnar region. Their other features such as ridge thickness to valley thickness ratio (59.5 %) and ridge width (68 %) are also below the classification probability from

the FPS of 71.38 % (LR), 71.89 % (SVM) and 70.01 % (KNN) [12]. Combining the features, the result of Verma and Agarwal dominates. Ganasivam and Vijayarajan did not perform analyses for the individual features. Their top result of 88.41% includes three measurements for the manually determined ridge count and FPS [11].

In this study as expected, the FPS of males are on average larger than those of females and a statistically significant difference could be demonstrated, especially for the thumbs. However, many outliers could be observed. The broad distribution could indicate larger fingertips or even fingerprints with poor quality. Artifacts that were not removed during segmentation may extend the area beyond the actual fingerprint, resulting in a higher value. It is possible that the particularly large fingerprints are due to tall subjects with large hands and fingers. With only two measurements, our best result was an accuracy of 82.37% and an F_1 measure of 0.84 (KNN) for the right thumb. This result is better than that of Badawi et al. using Fuzzy C- Means, which achieved an accuracy of 80.39%. However, their best result using backpropagation multi-layer neural networks is an accuracy of 88.5% [13]. In the publications, it appears that a good result could not be obtained with only one feature [6, 11, 13]. To improve our predictive performance, another method like calculating the ridge count and ridge thickness should be implemented. The use of methods to determine the location of fingerprints on objects via photogrammetry or scanning techniques should also make it possible to determine features i.e. the hand geometry [22]. With this information, sex assignment can be improved and possibly further phenotypic characteristics can be derived.

Overall, connecting the fingerprint evidence to other evidence including biometric information found at a crime scene, such as footprints, can help improve the reconstruction of the crime scene depending on the location of the evidence. Furthermore, this method cannot be used to on its own to identify a person. However, in combination with other evidence it can be used to determine the sex of the suspect. With the help of sex differentiation, the group of persons can at least be narrowed down.

6 Conclusion

Within this study, a program for automatic fingerprint RD and FPS estimation was implemented. So far, this is only an estimate - an analysis and comparison of a manual analysis and the automated analysis is still required. A main part was the automatic detection of the core point (overall accuracy of 90.85 %) to define the regions for calculating the number of epidermal ridges. Furthermore, by estimating the RD, it is possible to analyze fingerprints resulting in an easier and faster analysis in comparison to the manual approaches. So far, studies using manual assessment have shown that the RDs of men and women differ. Within this study, similar results were achieved, using an automatic procedure. The best values were obtained using the RD of the ulnar region and the FPS for the left thumb with an F_1 measure of 0.817 (LG), 0.832 (SVM) and 0.840 (KNN). The results of the presented study support the hypothesis that woman have a higher RD than men and smaller fingertips as presented in former studies.

7 References

1. Datta, A. K.: *Advances in fingerprint technology*. 1st edn. CRC press, Boca Raton (2001).
2. U.s. Department of Justice: *The Fingerprint Sourcebook*. CreateSpace Independent Publishing Platform, USA (2014).
3. Jain, A., Chen, Y., Demirkus, M.: Pores and Ridges: Fingerprint Matching Using Level 3 Features. 18th International Conference on Pattern Recognition, (2006).
4. Khan, M. A. et al.: A spatial domain scar removal strategy for fingerprint image enhancement. *Pattern Recognition* (60), 258-274 (2016).
5. Acree, M. A.: Is there a gender difference in fingerprint ridge density? *Forensic science international* 102(1), 35-44 (1999).
6. Dhall, J. K., Kapoor, A. K.: Fingerprint Ridge Density as a Potential Forensic Anthropological Tool for Sex Identification. *Journal of Forensic Sciences* 61(2), 424-429 (2016).
7. Gutiérrez-Redomero, E., et al.: Variability of fingerprint ridge density in a sample of Spanish Caucasians and its application to sex determination. *Forensic Science International* 180(1), 17-22 (2008).
8. Gutiérrez-Redomero, E., et al.: A comparative study of topological and sex differences in fingerprint ridge density in Argentinian and Spanish population samples. *Journal of Forensic and Legal Medicine* 20(5), 419-429 (2013).
9. Kapoor, N., Badiye, A.: Sex differences in the thumbprint ridge density in a central Indian population. *Egyptian Journal of Forensic Sciences* 5(1), 23-29 (2015).
10. Sánchez-Andrés, A., et al.: Impact of aging on fingerprint ridge density: Anthropometry and forensic implications in sex inference. *Science & Justice* 58(5), 323-334 (2018).
11. Gnanasivam, P., Vijayarajan, R.: Gender classification from fingerprint ridge count and fingertip size using optimal score assignment. *Complex & Intelligent Systems* 5(3), 343-352 (2019).
12. Verma, M., Agarwal, S.: Fingerprint based male-female classification. *Cisis* 53, 251-257 (2008).
13. Badawi, A. et al.: Fingerprint based gender classification. *The international conference on image processing. Computer vision and pattern recognition* 1, 41-46 (2006).
14. Tannenbaum, C. et al.: Sex and gender analysis improves science and engineering. *Nature*, 575(7781), 137-146 (2019).
15. Patil, V., Ingle, D. R.: An association between fingerprint patterns with blood group and lifestyle based diseases: a review. *Artif Intell Rev* 54, 1803–1839 (2020).
16. Henry, E. R.: *Classification and uses of fingerprints*. George Rutledge and Sons Ltd, London 1900.
17. Galton, F.: *Fingerprints*. Macmillan, London (1892).
18. MATLAB. version 7.10.0 (R2010a). Natick, Massachusetts: The MathWorks Inc. (2010).
19. Fahmy, M. F., Thabet, M. A.: A fingerprint segmentation technique based on morphological processing. In *IEEE International Symposium on Signal Processing and Information Technology*, 000215-20 (2013).
20. Haojun, J.: *Fingerprint Image Enhancement*, <https://github.com/jianghaojun/Fingerprint-Image-Enhancement>, last accessed 2020/03/11.
21. Bahgat, G. A. et al.: Fast and accurate algorithm for core point detection in fingerprint images. *Egyptian Informatics Journal* 14, 15-25 (2013).
22. Becker, S. et al.: Comprehensive Framework for High Resolution Image-based 3D Modeling and Documentation of Crime Scenes and Disaster Sites. *International Journal On Advances in Systems and Measurements*, 11 (2), 1-12, 2018

A kernel driver modification to visualize and reconstruct data transfer between computer and USB mass storage devices

Joshua Zöllner,¹ Dmitry Petschke,² Prof. Dr. Alexander Schinner,³ Prof. Dr. Kristin Weber,⁴ Prof. Dr. Manuel Mayer⁵

Abstract: The aim of this work is to create a completely new method for analysing the physical access to USB mass storage devices and to reconstruct the file access from the logged data. This is achieved by replacing a real USB stick with a full software simulation based on a Raspberry PI Zero using USB gadget mode. To achieve full information, we extended the logging capabilities of the Linux kernel driver. This allows to log position and size of each reading operation at the lowest possible level. For write operation, the written data is logged, too. This enables logging completely independent of the operating system or file system and allows a forensic image to be calculated that has time as an additional dimension. Further advantages of this method are that it is completely undetectable from the host computer and random accesses bypassing a file system can also be logged. A reconstruction of the original file access is shown and the possibilities for new attack vectors are discussed.

Keywords: Digital forensics; Data Transfer; File System Analysis; Protocol Reverse Engineering; Forensic Image; USB Software-Defined Mass Storage Device

1 Introduction

Since Stuxnet it is generally known that USB mass storage devices can pose a threat to computers and networks. It is also well known that malware, especially ransomware, tries to manipulate or encrypt USB mass storage devices. [La17] To detect if a company is attacked by this kind of malware there are currently two possibilities:

On machine acquisition. This method analyses the USB mass storage device during the data transfer. It requires the installation of additional monitoring software on the untrusted computer, which is also the first problem. It is possible for malware to bypass the software stack and directly read and write data on the USB mass storage device. Additionally, malware

¹ FHWS - Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt, Informatik und Wirtschaftsinformatik, Sanderheinrichsleitenweg 20, 97074 Würzburg, Germany zoellner.joshua@gmail.com

² FOM – Hochschule für Oekonomie und Management, Wirtschaftsinformatik, Arnulfstraße 30, 80335 München, Germany dmitry.petschke@gmail.com

³ FHWS - Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt, Informatik und Wirtschaftsinformatik, Sanderheinrichsleitenweg 20, 97074 Würzburg, Germany alexander.schinner@lehrbeauftragte.fhws.de

⁴ FHWS - Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt, Informatik und Wirtschaftsinformatik, Sanderheinrichsleitenweg 20, 97074 Würzburg, Germany kristin.weber@fhws.de

⁵ FOM – Hochschule für Oekonomie und Management, Wirtschaftsinformatik, Arnulfstraße 30, 80335 München, Germany manuel.mayer@fom.de

can detect, attack, and manipulate the software which is supposed to monitor the data transfer. An advantage is that, if the monitoring software works, every change from the untrusted computer to the USB mass storage device can be logged and analysed. This means that the behaviour and the different states of the mass storage device during read and write operations can be analysed. A possible, but not yet realised, method of investigation would be to continuously monitor the entropy of the written blocks and compare it with the entropy of the previously read blocks. If there is a significant difference here, it can be concluded that data on the storage gets encrypted. From an IT forensics point of view this method of acquiring data is bad practice, because the installation of additional software implies that changes must be made to evidence which makes it not forensically sound. In addition, there is the question of how and where the logged data could be stored.[Ca05]

Classical image acquisition. The second method analyses a forensic image of the USB mass storage device before (T_0) and after N unknown operations (T_N) while it has been connected to the untrusted computer. This has the advantage of analysing the transferred payload without the possibility of interference from the malware. In addition, the untrusted computer is unchanged. Problems occur when the behaviour of the malware is to be resolved in a chronological manner. This method can only provide two states of the mass storage device before (S_{T_0}) and after (S_{T_N}). This means that changes on the mass storage device cannot be monitored live. Therefore, it cannot be detected when a block is changed and later gets overwritten by the original content. Due to the static nature of image acquisition the changes this block is neither visible in S_{T_0} and S_{T_N} , which results in missing a possible indicator of compromise (IOC). [Ca05]

New approach. We propose a new method that combines the advantages of the two methods mentioned above and at the same time avoids the disadvantages. We replace the physical layer with a software-defined mass storage device (SD-MSD). With this it is possible to log and analyse every read and write action executed. It is also possible to reconstruct all states $S_{T_0\dots N}$ of the USB mass storage device.

Since a Raspberry Pi Zero supports USB On The Go (OTG) and by using the appropriate kernel module we can emulate a USB mass storage device which we can use as base for our SD-MSD. This makes it possible to present an image file, located on the Raspberry Pi Zero, to the untrusted computer as an USB drive and to log any kind of access to the Raspberry Pi. The major advantage is, that this does not need any interaction from the untrusted computer and the log can be analysed on-the-fly in a secure environment. [Zo21] To display the possibilities of the new approach we summarize the strengths, weaknesses, opportunities, and threats in the following:

Strengths: Our approach is a new way to analyse mass storage devices, that combines the advantages of current approaches. No modification through third party tools is necessary. It is easy to use and allows an on-the-fly analysis of the device in a secured environment.

Weakness: Dedicated hardware is needed to use this method and some preparation is required.

Opportunities: There is a possibility to use this method as an attack tool, but further analysis

is required.

Threat: An attacker, aware of this monitoring method might try to flood the protocol with a huge number of irrelevant operations (so called “snow blinding“)

2 Fundamentals

To create the SD-MSD with all necessary features we either had to look for an already existing module to simulate USB or to code one ourselves. For this purpose, we used `g_mass_storage` which is built from the Linux kernel source file `f_mass_storage.c`. After the verification, that USB simulation with `g_mass_storage` works we analysed the kernel and the module for a method to protocol the content of the transferred data. Unfortunately, kernel debugging options and the modules itself did not yield the result we wished for and just displayed the meta data, but not the actual content. Now we investigated the source code in `f_mass_storage.c` in search of access to transferred data. We identified two functions `do_read()` and `do_write()` which executed reads and writes respectively. In these functions we found the code snippets shown in the listings 1 and 2.

```
/* Perform the read */
file_offset_tmp = file_offset;
nread = kernel_read(curlun->filp, bh->buf, amount,
                    &file_offset_tmp);
VLDBG(curlun, "file read %u @ %llu -> %d\n", amount,
        (unsigned long long)file_offset, (int)nread);
```

List. 1: Code snippet from `do_read`

```
/* Perform the write */
file_offset_tmp = file_offset;
nwritten = kernel_write(curlun->filp, bh->buf, amount,
                        &file_offset_tmp);
VLDBG(curlun, "file write %u @ %llu -> %d\n", amount,
        (unsigned long long)file_offset, (int)nwritten);
```

List. 2: Code snippet from `do_write`

Tab. 1: Overview on the available parameters

Parameter	Description	Relevant
<code>file_offset_tmp</code>	current location of the address on the simulated mass storage device	Yes
<code>nread</code>	how many bytes have been read	No
<code>nwritten</code>	how many bytes have been written	No
<code>curlun->filp</code>	file pointer to the current logical unit	No
<code>bh->buf</code>	content of the data	Yes
<code>amount</code>	length of the current block	Yes

The parameters in table 1 represent everything which is needed to achieve the logging of the data. In our implementation the captured data is written to the kernel log and can be read from the file `/var/log/syslog`. An alternative approach was to read dynamically with the tool `dmesg`. We found this approach unstable as limited buffer and timing issues resulted in fragmented lines.

The format of the log had a few challenges. First, we had to decide how the bytes are going to be represented in the log. We decided to convert the binary encoded data to hexadecimal data. An alternative would have been the representation as Base64 encoded data, but an advantage was not recognisable here. Another challenge was how many bytes should be printed per line. We decided to limit the bytes to 16 per line. Another, theoretically possible, problem could be that several different write or read operations occur in parallel. However, we could not observe this in our investigations. The solution would be simple: each process would have to be assigned an individual ID that is output with each log line.

The final format begins with the determination if it is a read or write command, which was executed. Within this determination there are three different fragmentation states as shown in table 2. In the design we orientated us on the IP-package fragmentation. [In81]

Tab. 2: Fragmentation states for read and write operations

Log code	Action	Explanation
RD	Read	Start of a new buffer
WR	Write	Start of a new buffer
RD+	Read	Continuation of a started buffer
WR+	Write	Continuation of a started buffer
RD+b	Read	End of a buffer
WR+b	Write	End of a buffer

Afterwards the range of the address is shown like this: `[Starting address : End address]`. It describes at which address on the image file the data can be located. We decided to keep it short and simple, so it is easy to control during the development. An extension we implemented is a simple run length encoding (RLE) which summarizes lines starting with zero. It is used to make the log more manageable to analyse. Possible trivial extensions include a real RLE, which, however, would increase CPU usage and latency. To reduce the space required for logging, it would be possible to use a compression algorithm and change the encoding to Base64. Both would have slowed down development because the protocol would not have been as easy to read.

The implementation is shown in listing 3 which is added in both functions `do_read()` and `do_write()`.

```
....
VLDBG(...);
print_buf(...);
```

List. 3: Added logging function in `do_read()` and `do_write()`

The structure of `print_buf()` can be seen in listing 4.

```
static void print_buf(char* buff, loff_t file_offset, u32 amount,
char proc[2]){
/* Initialising variables */
for ( i = 0; i < amount; i++ ){
    /* Handle of RLE */
    /* Handle of fragmentation state */
    for (x = i; i < x+16; i++){ /* Limit to 16 bytes per line */
        printk(KERN_CONT " %02x", buff[i] );
    }
}
i = i - 1 ; /* Reduce the counter i by one to get the right
addresses in the log */
}
```

List. 4: Structure of `print_buf()`

3 Proof of concept

Setup. In our development setup the untrusted computer is connected to the Raspberry Pi over the network and physically to the USB port of the Raspberry Pi. This physical setup can be seen in figure 1.

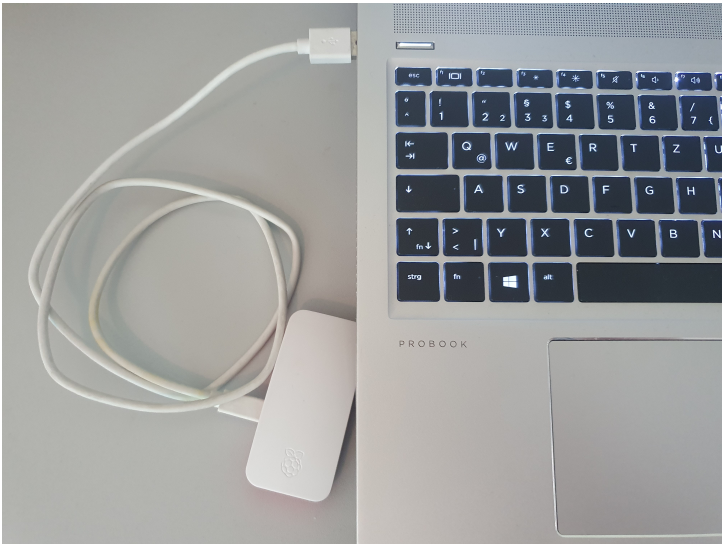


Fig. 1: Physical setup with Raspberry Pi Zero connected to a HP Probook over USB

On the Raspberry Pi Zero we created a partitioned and formatted image file for the modified `g_mass_storage` to use. After loading the module into the kernel of the Raspberry Pi Zero data is transferred from the computer to the Raspberry Pi via USB. During and after the data transfer the log can be read via the network connection.

With this modified kernel module, we can log the data transfer from the computer to the SD-MSD. To validate and evaluate our Approach, we performed the following tests. The first test was conducted to proof, that the transferred data can be logged, this can be seen in listing 5. The listing is a snippet of a log, created after the data transfer. If the HEX representation of the read processes is translated back to ASCII, it can be seen that the file system of the image was read (FAT). The write processes show that "MSDOS5.0" is written to the file system. This shows that the SD-MSD was used by a Windows machine.

```
RD: [1314816:1314831] 46 41 54 20 20 20 20 20 20 20 08 00 00 00 00
RD+: [1314832:1314837] RLE : 6*0
RD+: [1314838:1314853] 41 4a 4a 52 00 00 00 00 00 00 00 00 00 00 00
RD+b: [1314854:1318911] RLE : 4058*0
WR: [1048576:1048591] eb 3c 90 4d 53 44 4f 53 35 2e 30 00 02 40 08 00
WR+: [1048592:1048607] 02 00 02 00 00 f8 00 01 3f 00 ff 00 00 08 00 00
```

List. 5: Snippet of the log

The second test was conducted to proof, that the data is logged in the correct order. The content of this file were the numbers 15 to 31999 counted in steps of 16 and with no spaces or carriage returns. Additionally, there existed leading zeros until 16 numbers were filled up. Listing 6 shows the content of the file, which was written to the SD-MSD.

```
000000000000000015000000000000000000000031000000000000000470000...
...000000000000000031967000000000000000003198300000000000031999
```

List. 6: Content of the file

Afterwards we looked at the created log and could verify, that the addresses were at the correct locations. This means that the allocation of the addresses is like it is displayed in figure 2.

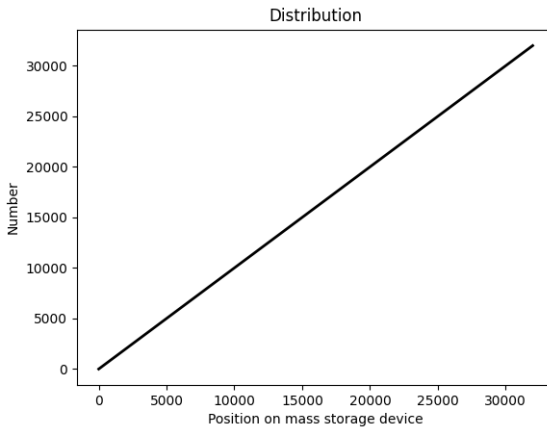


Fig. 2: Graph of the distribution of numbers on the image

The third test was conducted to prove, that deleted files can be recreated similarly to Autopsy. For this we started our testing environment and created a file from the computer on the `in_g_mass_storage` hosted image file. First, we modified it and then we deleted it. Afterwards, we analysed the image file in autopsy and compared it to our created log. The result was that the log had the same result as autopsy. With this it is proven that the written data is logged correctly.

The fourth test was conducted to show that file modification on an image can be recreated. To recreate the image, we used the unmodified image as a base. Then we applied the write operations, read from the log to replay the operations on the image. The recreated image has the same md5 sum as the modified image, this means that the recreation was successful. This can be seen in figure 3.

```
$ md5sum.exe image_2021-05-30_16_20_43.img test.img_16:19:57_30_05_2021
4ea14ae5b0af48fc233b575e8314fce8 *image_2021-05-30_16_20_43.img
4ea14ae5b0af48fc233b575e8314fce8 *test.img_16:19:57_30_05_2021
```

Fig. 3: md5 sum comparison

The last test was conducted to prove, that overridden data can be restored. After we verified that our method could recreate the image and locate the documents, similar to how it is done by Autopsy. The next step was to overwrite the whole partition of the SD-MSD with zeros, with the command found in listing 7.

```
dd if=/dev/zero of=/dev/sdb1 status=progress
```

List. 7: dd command used to overwrite the partition

After wiping the image, we can see in figure 4 that the content of the file was replaced with zeros. Our goal is to restore the image to a point before the wipe happened. To achieve this, we need to take a closer look at the logfile. In this it we can see that the entire content has been overridden by zeros, indicated by the output from listing 8.

```
WR+b: [1055232:1055743] RLE: 512 * 0
WR+b: [1055744:1056255] RLE: 512 * 0
WR+b: [1056256:1056767] RLE: 512 * 0
WR+b: [1056768:1057279] RLE: 512 * 0
WR+b: [1057280:1057791] RLE: 512 * 0
WR+b: [1057792:1058303] RLE: 512 * 0
WR+b: [1058304:1058815] RLE: 512 * 0
WR+b: [1058816:1059327] RLE: 512 * 0
```

List. 8: Snippet of the log during the wipe

When we want to restore the image to a state before it was overridden by zeros, we must execute all former logged write commands until the wipe happens on the base image. This can be seen in figure 5, where the overridden file from figure 4 has been recreated.

```
001457F0 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 .....
00145800 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 .....
00145810 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 .....
```

Fig. 4: Part of the image where our secret file was stored

```
001457F0 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 .....
00145800 54 68 69 73 20 69 73 20 61 20 73 65 63 72 65 74 This is a secret
00145810 21 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 !.....
```

Fig. 5: Part of the image where our secret file was successfully restored

4 Conclusion

In this paper we showed a new approach of analysing mass storage devices. We proved step by step that we can log the transfer of data, that the calculated addresses are correct, that we have the same capabilities as autopsy and that we can recreate every state of an image, even if it was overridden. Additionally, this keeps the evidence unmodified and allows us to analyse the data transfer on the fly. With this we can detect when data is encrypted or compressed while it is happening and we can see changes to the content of the mass storage device. This could also be the basis for a completely new attack vector, as the contents of files can change unnoticed by the operating system between two read operations.

Bibliography

[Ca05] Carrier, Brian: File System Forensic Analysis. Pearson Education, Inc, March 2005.

- [In81] Information Sciences Institute University of Southern California: , RFC 791, September 1981.
- [La17] Langer, Ralph: , Stuxnet und die Folgen, August 2017.
- [Ma] Mass Storage Gadget (MSG).
- [TB14] Tanenbaum, Andrew S.; Bos, Herbert: Modern operating systems, Forth Edition. Prentice Hall, March 2014.
- [Zo21] Zoellner, Joshua: , Ueberwachung des Datenaustausches zwischen Wechseldatentraegern und Computern, March 2021.

A Method for Evaluating and Selecting Software Tools for Remote Forensics

Maurice Meyer¹, Gunnar Auth ² and Alexander Schinner³


Abstract: In today's networked system environments, remote access to possibly involved IT system components is a fundamental requirement for digital forensics. For conducting professional remote forensics investigations in large system landscapes a growing number of software tools, both commercial and open source, is available today. On the other hand, reviews and comparisons of this special type of software tools are scarce. In support of finding the best-fitting remote forensics tool among the available solutions based on individual requirements and preconditions, this article presents a method for a criteria-based evaluation and selection process. While the method construction generally builds on established procedures for software evaluation and selection, the according criteria catalog including measurement procedures and weightings was derived from literature as well as considerations with experts from the IT security subsidiary of a large German telecom group. Furthermore, the method is demonstrated and validated by applying it to three selected software tools: Cynet, GRR Rapid Response and Velociraptor.

Keywords: Digital Forensics, Remote Forensics, Indicator of Compromise, Software Evaluation, Software Selection.

1 Introduction

During recent years, the general risk awareness of companies and other goal-oriented organizations concerning threats to information security has been constantly growing. Following best practice, many organizations have implemented information security management systems together with preventive measures on both organizational and technical level. Despite this positive development, the risk of cybersecurity breaches remains high. According to a recent study by Accenture [Ac20], indirect cyberattacks upon weaknesses in the supply chain or business ecosystem have been increasing strongly. To enable a targeted incident response, searching for traces and securing of evidence are crucial tasks at the outset of the process. Digital forensics aims at reconstructing and ultimately proving the course of events that lead to a detected security intrusion into digital

¹ Deutsche Telekom Security GmbH, Bonner Talweg 100, 53113 Bonn, Germany, maurice.meyer@t-systems.com

² Meißen University of Applied Sciences, Herbert-Böhme-Str. 11, 01662 Meißen, Germany, gunnar.auth@hsf.sachsen.de,  <https://orcid.org/0000-0002-3013-2739>

³ Deutsche Telekom Security GmbH, Bonner Talweg 100, 53113 Bonn, Germany, alexander.schinner@t-systems.com

systems by searching, analyzing, and securing data traces as evidence [KK12], [KK17], [Kh20]. Due to the fast dispersion and ongoing development of digital technology as well as the according growth of data volume (big data), the required effort for and complexity of digital forensics have also grown in a way that makes the use of professional software tools virtually indispensable. In today's networked system environments, remote access to involved IT system components is a fundamental requirement for digital forensics. The term remote forensics emphasizes this characteristic of forensics in large system landscapes and implies the use of dedicated software tools.

In response to the demand, a growing number of software tools, both commercial and open source, is available today. On the other hand, reviews and comparisons of this special type of software tools are scarce and lack a prescriptive method description [Kh20], [Pr16], [Fe21]. In support of finding the best-fitting remote forensics tool among the available solutions based on individual requirements and preconditions, this article presents a method for a multi-criteria evaluation and selection process. While evaluation of technology-based resources is a common problem in information systems (IS) research and IS management practice [Ir02], the soundness, validity, and usefulness of the evaluation result depends strongly on the method applied including appropriateness of the specific criteria in relation to the evaluation objective [HRS14]. Based on the problem described above, the evaluation objective for our method is to identify the best-fitting tool for the application domain remote forensics. Thus, the specific criteria catalog including measurement procedures and weightings was derived from relevant literature as well as considerations with professional experts from the IT security subsidiary of a large German telecom group. Furthermore, the method is demonstrated and validated by applying it to three selected software tools: Cynet, GRR Rapid Response and Velociraptor.

Since we did not aim for a comprehensive review of remote forensic tools but for the construction of a specific evaluation method for practical use, the selection was made based on the preferences of the involved IT security company. Through the collaboration with the IT security company, it was possible to use a test environment created from real-world IT infrastructure, network services, applications and evidences. This way, we were able to collect meaningful data on the appropriateness of the method for its application in practice.

In the discussion of the results, some limitations and remaining problems that occurred during the application of the method are shortly examined. The article closes with a conclusion and a short outlook on further research.

2 Related Work

In this section, we present an overview of related work on the digital forensics process and research to support tool evaluation and selection in the domain of digital forensics.

2.1 The Digital Forensics Process

Since software tools aim to support tasks in a certain application domain, we take the digital forensics process (DFP) as a starting point for a structured determination of evaluation criteria. In literature, several prescriptive models for describing the productive tasks and their sequence for conducting a digital forensics investigation have been proposed [Al09], [Bu11], [KK12], [Kh20]. Although there are differences in the details, the basic steps conform in most models: 1) identify incident, 2) preserve incident scene, 3) collect incident data, 4) examine incident data, 5) analyze incident data, and 6) document & report investigation. Moreover, several authors (e.g., [Fl17], [Va17]) emphasize the iterative character of the process. The best tool support for remote forensics should cover the whole process.

2.2 Support for Forensics Tool Selection

Virtually all steps of the DFP require appropriate tools to access, extract, process and store data. In comparison to other application fields for data analysis methods and techniques, in forensics the speed of examination often plays a crucial role. A detected intrusion may be only the beginning of an ongoing attack where the intruders are still inside the system. This leads to a high demand of automating the examination and analysis steps with specialized tools [Fl17].

Casey & Stanley [CS04] reviewed two tools (EnCase Enterprise and ProDiscover IR) for remote forensic with a focus on preservation and examination by comparing them based on a list of features. A technical procedural guide for using open source tools along the DFP was authored by Altheide & Carvey [AC11]. Rafique & Khan [RK13] provide a list of 38 tools which support static and/or live digital forensics. Van Baar et al. [vbb14] present a service-based approach for processing and investigating digital artifacts in forensics based on a closed-source, noncommercial product, developed at the Netherlands Forensics Institute. Vidas, Kaplan & Geiger [VKG14] have developed a forensics tool especially designed for unexperienced investigators without formal education on digital. A scarce example for a contribution dealing with the systematic selection of cybersecurity solutions is represented through the approach by Llansó, McNeil & Noteboom [LMN19] but focusses on defensive solutions mitigation cybersecurity risks.

The only work we found concentrating on methodological aspects of forensics tool selection is from 2005. O'Connor [Oc05] proposes 19 generic criteria grouped by the three categories technical, managerial, and legal. Most of his criteria cover general software product quality characteristics which are today specified through the software product quality model and requirements of the international standards ISO 25010 [ISO11] and 25051 [ISO14] (e.g., usability, reliability, speed/performance). For the five criteria specifically related to forensic tasks (“audit trails”, “forensic integrity” and the three legal criteria), a connection to tool functionality is not described.

In summary, our literature review has shown that most of the existing works concentrate

on reviewing or comparing ready to use software products. Although, this type of work provides valuable information for using tools in digital forensics, the information tends to become outdated in relatively short time with the ongoing development of the considered tools and the emergence of new tools.

3 Evaluation Process and Criteria for Remote Forensics

The evaluation process was designed following the recommendations of Heinrich, Riedl & Stelzer [HRS14], starting with the definition of evaluation objectives as stated in section 1. The according evaluation objects are remote forensics tools available as product. The central element of our method is a set of evaluation criteria for remote forensics. Aligned to the evaluation objective, the scope of the criteria is on supportive functions for specific tasks in the preserve, collect, examine, and analyze steps of the DFP. The criteria were first extracted from literature and then validated and complemented through five experts of the IT security company. An overview of the experts and their professional experience is given in Table 1:

ID	Job title	Focus	Experience
E1	Squad Lead Incident Response Service	Forensics	16 years
E2	Senior Incident Handler	Forensics	5 years
E3	Incident Handler and IT Forensic Analyst	Forensics	5 years
E4	Expert Incident Response and Forensics	Forensics	3 years
E5	IT Security Architect	Security	10 years

Tab. 1: Overview on involved forensics experts

Another important source for criteria was the series of “Cyber Letter” (Cyber-Brief) published annually by the German Federal Office for the Protection of the Constitution (Bundesamt für Verfassungsschutz, BfV). These letters report on current methods of attack on organizations and provide information about the detected indicators of compromise (IoC). IoC are data artifacts that indicate the use of malware [Ch18]. The IoC provide important information for the targeted application of search procedures that a remote forensics tool should support. The IoC in the cyber letters of the last five years were evaluated and formed the first basis for the catalog of criteria.

In order to be able to compare the different criteria, a scale of measure was defined. A nominal scale was used because it is easier to understand and takes less time. It is divided into three value categories: 1 = not fulfilled, 2 = partially fulfilled, 3 = fulfilled. In comparison to the often used “must have”/“nice to have” distinction [Li07], these three categories allow for a finer differentiation and could even be extended if necessary. Each criterion to be considered individually gets assigned a value as soon as the result is determined.

Furthermore, the weighting of the criteria serves to illustrate their importance. A fulfilled but optional criterion should play a smaller role in the decision than a fulfilled but

necessary criterion. Equating optional criteria with necessary criteria would lead to a shift in the results. This leads to a stronger significance of the results after weighting the different criteria. The weighting used here is based on a scale from 1 to 3 (1 = optional, 2 = important, 3 = necessary). During expert discussions, practical estimations of the individual experts were asked regarding the weighting of the criteria. In order to balance out the subjective estimates of the experts, the arithmetic mean of the various estimates was calculated for each criterion. Since the results show only a slight difference between the unrounded and rounded values, the rounded values are used. This results in the final weighting (w) for each criterion (j), as can be seen in table 3. The respective criteria categories are not weighted as they are not considered individually.

4 Demonstration of the Method with an Example of Use

In the selection of the forensic tools for demonstrating the method we followed the preferences of the involved forensics expert who emphasized the requirement of strong remote capabilities. Table 2 lists the three selected tools.

Name	Vendor	Website
Velociraptor	Velocidex	https://www.velocidex.com/
Cynet XDR	Cynet	https://www.cynet.com/platform/
GRR Rapid Response	GRR community	https://github.com/google/grr

Tab. 2: Selected remote forensics tools

In Table 3 the evaluation results (n_{ij}) of all three remote forensics tools (i) for the respective criteria (j) are shown. For better illustration, Harvey balls are used in the table instead of numerical values. To enable results to be calculated, a numerical value is defined for each Harvey ball used:

- 1) Empty Harvey ball: ○ = 1 (criterion not met)
- 2) Half-filled Harvey ball: ◐ = 2 (criterion partly met)
- 3) Filled Harvey ball: ● = 3 (criterion fully met)

No.	Category/criterion	Description	w _j	Velo.	Cynet	GRR
1	YARA rules	YARA rules are used to identify and classify malware [Rö18]				
1.1	YARA for arbitrary files	YARA rules applicable to arbitrary files [Co17]	2	●	○	○
1.2	YARA for arbitrary processes	YARA rules applicable to arbitrary processes [Co17]	2	●	○	●
2	File system & memory	Support of data preservation from file system and main memory				
2.1	Volume shadow copy service (VSS)	Preservation of system status and application data [Mi19]	3	●	○	◐
2.2	Alternate data streams (ADS)	Can be used to store and exchange hidden information [KCS18]	2	◐	○	◐
3	Registry key search	System database of Windows OS containing data like last used applications [Ca05], [Xi12]	3	●	◐	●
4	Memory scanning	Collecting live process data from RAM [Li14]	2	●	◐	●
5	Memory dump	Saving the current status from main memory into a copy to test it in a controlled environment later [Li14]	2	●	○	◐
6	Process data	Identify running processes and collect timestamps [Li14]	3	●	●	●
7	Hash values	Support of different hash value formats to search for known malicious hashes [Vi20]				
7.1	Authentihash	Authentication of digital certificates of software [Ri19]	1	○	○	○
7.2	Imphash	Comparing of hash values generated during import of executable program files [Na19]	2	○	○	○
7.3	MD5	Widely used for unique identification [ML18]	3	◐	●	◐
7.4	SHA-1	Widely used for unique identification [ML18]	3	◐	○	◐
7.5	SHA-256	Widely used for unique identification and well suited as evidence in court [ML18]	3	◐	●	◐
7.6	SSDEEP	Identifying similar hash values [MB19]	1	○	◐	○
7.7	Vhash	Message authentication [MB19]	1	○	○	○
8	Permanent monitoring	Defining search parameters for permanently monitoring the client system [Bu18b]	3	●	●	●

9	Log files	Data on executed actions, program crashes, etc. including time stamps [Li14]	3	●	○	●
10	Communication logs	Incoming and outgoing communication with associated IP addresses, ports, time stamps, etc. [Li14]	2	●	●	●
11	Metadata	Extracting file metadata, e.g., name, format, size, timestamps etc. [Ba20]	3	●	●	●
12	Downloads	Download traceability [Bu18a]	2	◐	○	◐
13	Program execution	Verifying program execution (e.g., with Prefetch, Shimcache, Amcache) [Bu18a]	3	●	●	●
14	Session data	Active and past logins [De17]	3	●	●	●
15	DNS entries	Active Directory information (host name, domain, domain users, privileges, etc.) [E4]	2	●	●	●
16	User data	Retracing whether users were active via Active Directory or locally [E1]	2	●	●	●
17	File search	Searching for files by name, size, etc. [Bu18a]	3	●	●	●
18	System hardware information	Basic information on hardware architecture and components [E1]	1	●	●	●
19	System configuration	Basic system configuration, e.g., installed OS [E1]	3	●	●	●
20	Application programming interface (API)	API support for connecting other forensics tools [E3]	2	●	●	●
21	Export formats	Exporting the results in different formats [E4]				
21.1	Comma separated values (CSV)	Simple and popular data format [E4]	3	●	●	●
21.2	Extensible Markup Language (XML)	Simple and popular data format [E4]	2	○	○	○
21.3	JavaScript Object Notation (JSON)	Simple and popular data format [E4]	2	◐	○	○
22	Display options	Displaying and enclosing values and results [E1]				
22.1	Filter options	Displaying filtered results only [E1]	2	◐	●	◐
22.2	Configurable dashboard	User defined overview of selected information [E1]	1	○	○	○
23	Enterprise level support	Criteria for using the tool in large organizations [E1]				

23.1	Web client	Tool can be used via web browser [E5]	2	●	●	●
23.2	Rich client	Tool can be used via locally installed client [E5]	2	○	○	○
23.3	Authentication and access control	Analyzing user authentication and access control data [SS97]	3	●	●	●
23.4	Multi-tenancy	Support of multi-tenancy architectures [E5]	2	○	●	○
24	Configurable search	Search can be adapted to user's needs [E2]	3	●	◐	◐
25	Search queries	Support of user defined search queries [E2]	2	●	●	◐
26	Search query export	Search queries can be exported [E2]	2	◐	○	○

Tab. 3: Evaluation criteria catalog and results for remote forensics tools

After all evaluation objects (i) have been tested and a degree of fulfillment (F_{ij}) has been assigned to each criterion, the calculation with the criterion weights per criterion (w_j) was performed according to Riedl [Ri06]. In deviation from [Ri06], we decided against normalizing the weightings to support our pragmatic method design.

$$S_i = \sum_{j=1}^m w_j \cdot F_{ij}$$

The application of the formula resulted in the overall score (S_i) with consideration of the criteria weights. The total weighted values achieved for each software product are stated in Table 4. The maximum to be achieved is 273.

Measure	Velociraptor	Cynet	GRR
Overall score S_i (of max. 273)	227	200	209
Relative score (percentage)	83,5 %	73,3 %	76,6 %

Tab. 4: Overall evaluation results for the three forensics tools

5 Discussion

As part of the evaluation process, there are some limitations that have arisen due to a lack of resources and the laboratory character of our research setting. These are discussed in more detail here.

The tests carried out for the evaluation relate to exactly one client. This allowed the findings presented to be obtained, but the three tools considered are designed for use in extensive IT system landscapes. Since the resources to simulate a representative IT system landscape with several hundred or thousands of systems were not available, no specific

statement based on test results can be made about certain capabilities. On the one hand, there was not enough test data available for analysis and, on the other hand, in the case of the mentioned scalability an assessment was only made based on the documentation and literature. Here, we see a need for further research. In forensic examinations of a large number of systems, scalability is a criterion that should not be underestimated [CMC15], but we did not include it in the catalog of criteria due to the lack of test options. Further criteria that were not included in our consideration are the implementation effort, as a reliable determination was not possible based on the use of only one system, as well as the maintenance effort. These two criteria have a significant influence on the total cost of ownership, in which the costs are determined from a life cycle perspective in a comprehensive manner.

In addition to the consideration of other supplementary criteria, the implementation of more in-depth analyzes is also noted. Heinrich, Riedl and Stelzer [HRS14] recommend, among other things, that a sensitivity analysis should be carried out. This provides information on how the ascertained ranking reacts to changes in various parameters (e.g., the criteria weighting). The decision-making risk is determined based on this analysis. The smaller the deviations, the lower the risk of a decision based on the determined values.

Moreover, the fineness of the scales should be adjusted for using the method in an individual application scenario. By clearly defining finer scale values and by carrying out pair comparisons, for example, it is possible to find a more suitable value for the individual requirements already when weighting the criteria. In the selection process, this also enables a more detailed comparison and thus a better basis for decision-making.

Also, in this context, the use of the “partially fulfilled” valuation should be decided. In many cases, “fully met” and “not met” will be enough. Nevertheless, in some cases it can be useful to have a finer distinction available. In such cases, the Harvey ball symbol should be complemented with a qualitative description of the valuation. An example are the hash value criteria (no. 7 in table 3). We used “partially fulfilled” here when the tool did not offer a search function for a certain hash but could show it when files are extracted from the analyzed system. While we did not include such additional information in this paper for lack of space, it should be specified for selecting forensics tools in a real-world setting.

6 Conclusion

The method presented in this paper allows for a systematic evaluation of remote forensics tools. The evaluation results provide qualitative and quantitative data as a decision basis for selecting the best-fitting tool for the application domain remote forensics. While the evaluation process is based on a generally accepted method, the specific criteria catalog was developed based on literature as well as expert knowledge. The participation of professional forensics and security experts also contributed to the appropriateness of the method for practical use. Therefore, the method allows for an easy tailoring to individual forensics environments through changes of criteria and weighting as well as including

additional criteria or omitting existing ones.

After evaluating the three remote forensics tools, it remains to be seen which of them is objectively the optimal one. Each has different strengths and weaknesses. As is often the case, the choice depends on individual circumstances and preferences. Therefore, it is not possible to arrive at a universally valid recommendation. Nevertheless, the evaluation results provide useful impressions of the respective capabilities of modern tools.

In the course of a critical reflection on the implementation of the evaluation process and the results, additional need for research has become apparent. The stated limitations of the conducted evaluation and the given restrictions compared to a real environment are primary starting points for further research.

Bibliography

- [AC11] Altheide, C.; Carvey, H. A.: Digital forensics with open source tools. Syngress, Burlington, MA, 2011.
- [Ac20] Accenture: Innovate for Cyber Resilience - Third Annual State of Cyber Resilience. https://www.accenture.com/_acnmedia/PDF-116/Accenture-Cybersecurity-Report-2020.pdf, accessed 15 Apr 2021.
- [Al09] Almulhem, A.: Network forensics: Notions and challenges: 2009 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT). IEEE, Ajman, United Arab Emirates, pp. 463–466, 2009.
- [Ba20] Baun, C.: Betriebssysteme kompakt – Grundlagen, Daten, Speicher, Dateien, Prozesse und Kommunikation. Springer-Verlag GmbH, 2020.
- [Bu11] Bundesamt für Sicherheit in der Informationstechnik: Guide to IT Forensics (Leitfaden „IT-Forensik“) Version 1.0.1 (in German). https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Cyber-Sicherheit/Themen/Leitfaden_IT-Forensik.pdf, accessed 3 Aug 2020.
- [Bu18a] Bundesamt für Verfassungsschutz: BfV Cyber Letter (Cyber-Brief) Nr. 02/2018 (in German). <https://www.verfassungsschutz.de/SharedDocs/publikationen/DE/2018/bfv-cyber-brief-2018-2.pdf>, accessed 1 Jul 2021.
- [Bu18b] Bundesamt für Verfassungsschutz: BfV Cyber Letter (Cyber-Brief) Nr. 01/2018 (in German). <https://www.verfassungsschutz.de/SharedDocs/publikationen/DE/2018/bfv-cyber-brief-2018-1.pdf>, accessed 1 Jul 2021.
- [Ca05] Carvey, H.: The Windows Registry as a forensic resource. *Digital Investigation* 3/2, pp. 201–205, 2005.
- [Ch18] Cho, H. et al.: Cyber Threat Analysis and Quantification Based on Indicator of Compromise: 2018 International Conference on Platform Technology and Service (PlatCon). IEEE, Jeju, pp. 26–31, 2018.
- [CMC15] Cruz, F.; Moser, A.; Cohen, M.: A scalable file based data store for forensic analysis.

- Digital Investigation 12, S90-S101, 2015.
- [Co17] Cohen, M.: Scanning memory with Yara. *Digital Investigation* 20, pp. 34–43, 2017.
- [CS04] Casey, E.; Stanley, A.: Tool review – remote forensic preservation and examination tools. *Digital Investigation* 4/1, pp. 284–297, 2004.
- [De17] Delija, D.: Remote digital forensics practices. *International Journal of Digital Technology & Economy* 1/2, pp. 27–36, 2017.
- [Fe21] Fernando, V.: Cyber Forensics Tools: A Review on Mechanism and Emerging Challenges: 2021 11th IFIP International Conference on New Technologies, Mobility and Security (NTMS). IEEE, Paris, France, pp. 1–7, 2021.
- [Fl17] Flaglien, A. O.: The Digital Forensics Process. In (Årnes, A. Ed.): *Digital Forensics*. John Wiley & Sons, Ltd, Chichester, UK, pp. 13–49, 2017.
- [HRS14] Heinrich, L. J.; Riedl, R.; Stelzer, D.: *Informationsmanagement: Grundlagen, Aufgaben, Methoden*. De Gruyter Oldenbourg, Berlin, 2014.
- [Ir02] Irani, Z.: Information systems evaluation: navigating through the problem domain. *Information & Management* 1/40, pp. 11–24, 2002.
- [ISO11] ISO: ISO/IEC 25010:2011 — Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — System and software quality models, 2011.
- [ISO14] ISO: ISO/IEC 25051:2014 — Software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Requirements for quality of Ready to Use Software Product (RUSP) and instructions for testing, 2014.
- [KCS18] Kao, D.-Y.; Chen, Y.-P.; Shih, N.-H.: Reconstructing ADS data hiding in windows NTFS: A temporal analysis. *Digital Investigation* 26, S137, 2018.
- [Kh20] Khera, V.: Introduction to Digital Forensics. <https://cyberprotection-magazine.com/introduction-to-digital-forensics/>, accessed 3 Jun 2021.
- [KK12] Kaur, R.; Kaur, A.: Digital Forensics. *International Journal of Computer Applications* 5/50, pp. 5–9, 2012.
- [KK17] Karie, N.; Karume, S.: Digital Forensic Readiness in Organizations: Issues and Challenges. *Journal of Digital Forensics, Security and Law* 12, Article 5, 2017.
- [Li07] Lin, H. et al.: COTS Software Selection Process: Sixth International IEEE Conference on Commercial-off-the-Shelf (COTS)-Based Software Systems (ICCBSS'07). IEEE, Banff, Alta., pp. 114–122, 2007.
- [Li14] Ligh, M. H. et al.: *The art of memory forensics: detecting malware and threats in Windows, Linux, and Mac memory*. Wiley, Indianapolis, IN, 2014.
- [LMN19] Llanso, T.; McNeil, M.; Noteboom, C.: Multi-Criteria Selection of Capability-Based Cybersecurity Solutions: Proc. of the 52nd Hawaii International Conference on System Sciences, pp. 7322–7330, 2019.
- [MB19] Miller, P.; Bryce, C.: *Learning Python for forensics: leverage the power of Python in forensic investigations*. Packt Publishing, Birmingham, UK, 2019.

- [Mi19] Microsoft: Volume Shadow Copy Service. <https://docs.microsoft.com/en-us/windows-server/storage/file-server/volume-shadow-copy-service>, accessed 3 Jun 2021.
- [ML18] Muniz, J.; Lakhani, A.: Investigating the Cyber Breach: the digital forensics guide for the network engineer. Cisco Press, 2018.
- [Na19] Naik, N. et al.: Cyberthreat Hunting - Part 1: Triaging Ransomware using Fuzzy Hashing, Import Hashing and YARA Rules: Proc. of the 2019 IEEE International Conference on Fuzzy Systems, pp. 1–6, 2019.
- [Oc05] O'Connor, R. V.: Software selection: towards an understanding of forensic software tool selection in industrial practice. *International Journal of Technology, Policy and Management* 4/5, p. 311, 2005.
- [Pr16] Prasanthi, B. V.: Cyber Forensic Tools: A Review. *International Journal of Engineering Trends and Technology* 5/41, pp. 266–271, 2016.
- [Ri06] Riedl, R.: Analytischer Hierarchieprozess vs. Nutzwertanalyse: Eine vergleichende Gegenüberstellung zweier multiattributiver Auswahlverfahren am Beispiel Application Service Providing. In (Fink, K.; Ploder, C. Eds.): *Wirtschaftsinformatik als Schlüssel zum Unternehmenserfolg*. DUV, Wiesbaden, pp. 99–127, 2006.
- [Ri19] Rivera, R. et al.: Costly freeware: a systematic analysis of abuse in download portals. *IET Information Security* 1/13, pp. 27–35, 2019.
- [RK13] Rafique, M.; Khan, M. N. A.: Exploring Static and Live Digital Forensics: Methods, Practices and Tools. *International Journal of Scientific & Engineering* 10/4, pp. 1048–1056, 2013.
- [Rö18] Röcher, D.-J.: Cyber Threat Intelligence 101. *Datenschutz und Datensicherheit - DuD* 10/42, pp. 623–628, 2018.
- [SS97] Sandhu, R. S.; Samarati, P.: *Authentication, Access Control, and Intrusion Detection*. CRC Press, pp. 40–48, 1997.
- [Va17] Varol, A.: Review of Evidence Collection and Protection Phases in Digital Forensics Process. *International Journal of Information Security Science* 4/6, pp. 39–46, 2017.
- [Vi20] VirusTotal: Result details for suspicious file analysis. <https://www.virustotal.com/gui/file/fc96e2e703152e74f74ed39252f7e7a67b1398693fd82f5a645a3dc835ffd518/details>, accessed 4 Jun 2021.
- [VKG14] Vidas, T.; Kaplan, B.; Geiger, M.: OpenLV: Empowering investigators and first-responders in the digital forensics process. *Digital Investigation* 11, S45-S53, 2014.
- [vzv14] van Baar, R. B.; van Beek, H.; van Eijk, E. J.: Digital Forensics as a Service: A game changer. *Digital Investigation* 11, S54-S62, 2014.
- [Xi12] Xie, H. et al.: Forensic Analysis of Windows Registry Against Intrusion. *International Journal of Network Security & Its Applications (IJNSA)* 2/4, pp. 121–134, 2012.

Herausforderungen beim Einsatz Künstlicher Intelligenz zur Erkennung von im Zusammenhang mit Kinderpornografie stehenden Mediendateien

Nicole Garbers¹, Michael Brodthage¹

Abstract: Der Einsatz Neuronaler Netze bietet deutlich bessere Erkennungsraten bei kinderpornografischem Bild- und Videomaterial als beispielsweise Hautanteilfilter und stellt daher im Rahmen der Strafverfolgung ein wichtiges Werkzeug zur Vorselektierung von Massendaten dar. Trotz aller Erfolge gibt es beim Einsatz in der Praxis Herausforderungen, die noch detaillierter betrachtet werden müssen. Diese Veröffentlichung beschäftigt sich mit verschiedenen Aspekten in der Datenzusammenstellung des Trainingsmaterials und der Datenvorverarbeitung und zeigt anhand von Beispielen deren jeweiligen Einfluss auf die spätere Klassifizierungsleistung des Neuronalen Netzes unter den besonderen Herausforderungen bei der Verfolgung von im Zusammenhang mit Kinderpornografie stehenden Straftaten.

Keywords: Neuronale Netze, Datenvorverarbeitung, Erkennung von Kinderpornografie

1 Einleitung

In den vergangenen Jahren ist durch die fortschreitende Digitalisierung die Datenmenge sichergestellter Asservate in allen Phänomenbereichen kontinuierlich stark angestiegen, ganz besonders davon betroffen ist der Bereich der im Zusammenhang mit Kinder- und Jugendpornografie stehenden Straftaten. Im Jahr 2020 haben allein die Ermittlungsbehörden in Niedersachsen 2373 Terabyte sichergestellt [An21].

Als Kinderpornografie wird dabei im Sinne des §184b Strafgesetzbuch eine fotorealistische Darstellung sexueller Handlungen von, an oder vor einer Person unter 14 Jahren (Kind) bzw. die Wiedergabe eines ganz oder teilweise unbedeckten Kindes in aufreizend geschlechtsbetonter Körperhaltung oder die sexuell aufreizende Wiedergabe der unbedeckten Genitalien/des Geschlechts bezeichnet. Es ist sowohl die Verbreitung, der Erwerb als auch der Besitz kinderpornografischer Inhalte strafbar. Bei jugendpornografischen Inhalten, ist die Person 14 Jahre oder älter, aber noch nicht 18 Jahre alt.

Da jede solcher Darstellung einen tatsächlichen Missbrauch zeigt, ist neben einer Verbreitungshandlung auch immer zu untersuchen, ob möglicherweise ein aktueller Missbrauch vorliegt. Solche Fälle dürfen auf keinen Fall im sichergestellten Datenvolumen „untergehen“ – auch nicht im Fall immer steigender Datenmengen.

¹ Landeskriminalamt Niedersachsen, Am Waterlooplatz 11, 30169 Hannover

So sind inzwischen in den meisten Smart Devices im Haus oder im Smartphone zahlreiche Serviceleistungen vorhanden, die auf den Einsatz Künstlicher Intelligenz bzw. Neuronaler Netze beruhen. Da liegt es nahe, zu überprüfen, inwieweit solche technischen Ansätze zur Bewertung großer unstrukturierter Datenmengen auch bei der Strafverfolgung im Zusammenhang mit Kinder- und Jugendpornographie stehenden Straftaten unterstützen können.

In der Vergangenheit eingesetzte Hautanteilfilter liefern in vielen Fällen keine geeignete Vorselektierung. Nicht nur spielt die Hautfarbe der abgebildeten Person fälschlicherweise eine entscheidende Rolle wie in [KHH21] gezeigt wurde, sondern es kommt auch zu unvermeidbaren Fehlinterpretationen bei vielen häufig vorkommenden Bildszenen innerhalb des gesicherten Datenmaterials. So werden Selfies mit einer Nahaufnahme des eigenen Gesichts fälschlicherweise als pornografisch erkannt, während kinderpornografische Szenen bei schlechter Beleuchtung (z. B. Eigenaufnahmen der Kinder in die Unterhose, rote Beleuchtung) oder Aufnahmen mit nur einem kleinen Hautanteil (z.B. verkleidetes Kind vor buntdekoriertem Hintergrund mit teilweise entblößtem Genitalbereich) nicht erkannt werden.

Deutlich bessere Ergebnisse verspricht der Einsatz Neuronaler Netze in diesem Bereich [Ga2019]. Doch trotz erster Erfolge im Rahmen wissenschaftlicher Betrachtungen sind bei der Übertragung in die polizeiliche Praxis noch Besonderheiten zu beachten, die in dieser Veröffentlichung näher beleuchtet und dargelegt werden.

2 Neuronale Netze zur Erkennung von Kinderpornografie

Automatische Klassifizierungsansätze zur Erkennung von Kinder-/Jugendpornografie haben zu einem die Aufgabe, pornografische Szenen zu erkennen, zum anderen, für die abgebildeten Personen zu bestimmen ob diese unter 14 Jahren bzw. 14 Jahre und älter aber jünger als 18 Jahre sind. In der Literatur findet man dazu unterschiedliche Ansätze, die anfangs auf Hautanteilfiltern basierten, inzwischen aber zumeist auf einem oder mehreren Neuronalen Netzen aufbauen. Dabei kann die Klassifikation in die beiden oben beschriebenen Aufgaben aufgeteilt und dann zusammengeführt werden, oder aber ein Neuronales Netz verwendet werden, welches sowohl die Erkennung von Pornografie als auch die Altersschätzung gelernt hat.

Einen Workflow zur Erkennung von Kinderpornografie, der aus einer Kombination von Gesichtserkennung zur Altersschätzung und einer Berechnung des Hauttonanteils besteht, stellen beispielsweise Moser et al. in [MTH15] vor. Zur Altersschätzung wird sich dabei auf bestimmte Regionen der Augen, der Nase und des Mundes konzentriert. Bestimmte geometrische Verhältnisse in diesem Bereich sollen es ermöglichen, eine Unterscheidung zwischen „Kind“ und „Erwachsener“ zu treffen [IT12]. Allerdings zeigen Moser et al., dass dieses Vorgehen auf den dortigen Testdaten keinen Erfolg aufweist: 2/3 der Gesichter wurden falsch klassifiziert.

anschließend durchgeführten Filterung nach sichtbarer Haut verwendet. Hier konnte eine Erkennung von Pornografie in 84 % der Fälle erreicht werden, die Falsch-Positiv-Rate wurde dabei jedoch nicht genannt. In Summe bewerten die Autoren ihr Vorgehen für fehleranfällig und nicht für die Praxis geeignet.

Der Einsatz Neuronaler Netze bietet hier deutlich bessere Ergebnisse. In [Ga2019] wurden auf einem Testdatensatz aus 42.000 Bildern die Klassifizierungsleistung eines HSV- und eines YcbCr-Filters mit derer von drei verschiedenen Neuronalen Netzarchitekturen (VGG-19, ResNet152, DenseNet) verglichen. Dabei konnte durch die Verwendung eines Hautanteilfilters Erwachsenenpornografie zu 63.25% (YcbCr) bzw. 63.06% (HSV) erkannt werden. Ein leicht schlechteres Ergebnis konnte ebenfalls für Jugendpornografie erreicht werden (59.12% (YcbCr), 62.44% (HSV)), während kinderpornografisches Material nur zu 47.71% (YcbCr) bzw. 50.38% (HSV) erkannt wurde. Die Falsch-Positiv-Rate betrug dabei 14.25 % (YcbCr) bzw. 18.25% (HSV).

Für das DenseNet als beispielhafter Vertreter der untersuchten Neuronalen Netze zeigten sich deutlich bessere Ergebnisse. So konnten 99.46 % der jugendpornografischen Bilder, 98.80% der kinderpornografischen Bilder und 98.38 % der pornografischen Bilder als pornografisch erkannt werden. Gleichzeitig betrug die Falsch-Positiv-Rate auf den Alltagsbildern 15.13%.

Ähnlich gute Ergebnisse wurden auch in [APR17] bestätigt.

In [MCD] wurde ebenfalls ein aus Altersschätzung und Pornografieerkennung bestehender Workflow vorgestellt. Dieser bestand jedoch im Gegensatz zu [MTH15] aus zwei Neuronalen Netzen. Damit erreichen die Autoren bei der Unterscheidung „Kind“ „Erwachsener“ eine Genauigkeit von über 94 % auf den mehr als 500.000 Bilder umfassenden Datensatz IMDB-WIKI-101 und auf dem RCPD Datensatz (region-based annotated child pornography dataset [8]) der brasilianischen Polizei mit kinderpornografischen Bildern eine Genauigkeit von ca. 82%. Ähnlich gute Ergebnisse konnten bei der Erkennung von Pornografie in dem RCPD Datensatz erreicht werden. Kombiniert man jedoch diese beiden Netze, um tatsächlich Kinderpornografie im RCPD Datensatz zu erkennen, beträgt die Falsch-Positiv-Rate 68% und die Falsch-Negativ-Rate 64.61%.

Dies zeigt deutlich, dass es bei der Übertragung von wissenschaftlichen Erkenntnissen in den praktischen Wirkbetrieb noch Problemstellungen gibt, die besonders betrachtet werden müssen.

3 Besondere Problemstellungen im Phänomenbereich Kinderpornografie

Dieses Kapitel beleuchtet die verschiedenen Aspekte bezüglich der Auswahl, des Trainings und des Einsatzes im Wirkbetrieb von Neuronalen Netzen im Phänomenbereich Kinderpornografie. Es werden die Besonderheiten und die sich daraus ergebenden Anforderungen sowie noch offenen Problemstellungen dargestellt.

Aus strafrechtlichen Aspekten ist die Erkennung von kinder- und jugendpornografischer Bilder und Videos relevant. Dabei spricht man in Deutschland von Kinderpornografie, wenn eine abgebildete Person jünger als 14 Jahre ist. Diese Definition ist international jedoch nicht einheitlich, so gilt beispielsweise in Australien je nach Territorium eine Altersgrenze von 16, 17 oder 18 Jahren und in den USA eine Altersgrenze von 18 Jahren. Eine Differenzierung in Kinder- und Jugendpornografie wie in Deutschland findet somit nicht statt.

Auch bei der Bewertung der Bildszene gibt es nationale Unterschiede, so fällt in Deutschland auch ein nur teilweise unbekleidetes Kind in aufreizend geschlechtsbetonter Körperhaltung unter den Begriff der Kinderpornografie, während in Österreich eine sexuelle Handlung, Genitalien oder die Schamgegend sichtbar sein müssen.

Dies führt dazu, dass Neuronale Netze zu dieser Fragestellung aufgrund der unterschiedlichen Anforderungen nicht ohne weiteres von internationalen Forschungsgruppen übernommen werden können.

Die korrekte Altersschätzung stellt dabei ein besonderes Problem da. Auch wenn es inzwischen auch für diese Fragestellung eine Vielzahl von Neuronalen Netzen gibt, so weisen sie alle nicht die benötigte Genauigkeit auf. Bereits eine Varianz von +/- 3 Jahren [PGK], [Pa16], die im Erwachsenenbereich häufig ausreichend ist, würde bei der hier vorliegenden Fragestellung zu einer falschen Eingruppierung der abgebildeten Person führen. Erschwerend kommt hinzu, dass die meisten Neuronalen Netze entweder Gesichtszüge interpretieren oder das Alter anhand von Körperproportionen schätzen. Während Kleinkinder beim Verhältnis Kopf zu Körper ein Verhältnis von 1:5 aufweisen, ändert sich dies für Schulkinder zu 1:6, bevor zu Beginn der Pubertät ein vermehrtes Längenwachstum der Gliedmaßen einsetzt. In kinder- bzw. jugendpornographischen Aufnahmen sind jedoch häufig weder Gesichtszüge noch ausreichend Körperproportionen sichtbar, um darauf basierend mit bereits existierenden Lösungen eine Altersschätzung durchführen zu können.

Aus technischen Gründen stellt sich außerdem die Frage, wie es sich am besten realisieren lässt, drei eher ähnliche Kategorien von einer sehr breit gefassten Kategorie zu trennen, siehe auch Kapitel 3.3. So ist der Übergang von Kinder- zu Jugendpornografie bzw. Jugend- zu Erwachsenenpornografie als eher fließend zu betrachten, während irrelevante Szenen deutlich andere Themenfelder zeigen. Schwierig hiervon zu trennen, sind jedoch wiederum kinderpornografische Szenen, in denen das Kind beispielsweise vollständig bekleidet ist und „verkleidet“ sexuell anregend posieren muss. Diese Betrachtungen zeigen, dass typische „NudeNetze“ zur Pornografieerkennung [Ba19], [Mo15], [JWT18] bei dieser Fragestellung nicht eingesetzt werden können.

Bei kinderpornografische Szenen, in denen das Kind beispielsweise vollständig bekleidet ist und „verkleidet“ sexuell anregend posieren muss, ist eine Abgrenzung zu dem Alltagsbereich als schwierig einzustufen, da es nicht ausreicht, wenn das Neuronale Netz Bilder dieser Art anhand eines vorhandenen bzw. nicht vorhandenen Haut-anteils

dass ein hier keine Differenzierung über die Summe der insgesamt vorliegenden Pixel mit Hautfarbe erfolgen kann. In [Ya2021] untersuchten die Autoren die Verbreitung von Gesichtern über die verschiedenen Kategorien des Imagenet-Datensatzes und zeigten, dass in 106 Kategorien, mehr als die Hälfte der Bilder dieser Kategorie Gesichter beinhalteten. Dies verdeutlicht, dass, in einem diversen Datensatz unterschiedlicher Themenbereiche, Gesichter immer eine dominierende Rolle spielen und ein Hautanteilfilter, wie bereits in Abschnitt 2 beschrieben, kein geeigneten Ansatz liefert.

3.2 Objekterkennung vs. Szenenerkennung

Prinzipiell bietet sich zwei unterschiedliche Ansätze zur Umsetzung an. Es ist sowohl möglich, ein Neuronales Netz zu trainieren, bestimmte Objekte wie beispielsweise Geschlechtsteile oder Brustwarzen zu erkennen (Objekterkennung) oder ein Bild als komplette Szene zu bewerten (Bildklassifikation). Im letzteren Fall lernt das Neuronale Netz eigenständig, relevante Objekte oder Szenenbereich im Rahmen der Auswertung zu identifizieren.

Der Vorteil einer objektbasierten Klassifizierung ist die Robustheit gegenüber der Vielzahl möglicher Szenen, d.h. Hintergründe, Anzahl abgebildeter Personen oder Aufnahmeerperspektiven haben deutlich weniger Einfluss, da nur nach dem Vorhandensein bestimmter Objekte gesucht wird. In [Fu05] wurden dafür beispielsweise als Objekte „männliche Geschlechtsorgane“, „weibliche Geschlechtsorgane“, „weibliche Brust“ und „Gesäß“ gewählt. Problematisch für eine Übertragung dieser Kategorien in den Phänomenbereich Kinderpornografie ist, dass die Geschlechtsmerkmale altersbedingt noch nicht soweit ausgeprägt sind. So müsste die Klasse „weibliche Brust“ mindestens auf die Definition „Brustwarzen“ erweitert werden, da keine Geschlechtsunterscheidung möglich ist. Dies wurde beispielsweise in [Ya2021] für die Erkennung von pornografischem Bildmaterial betrachtet.

Dieser Ansatz weist neben einer Sensitivität hinsichtlich Fehlklassifikationen ähnlicher Objekte und der Tatsache, dass die genannten Objekte sich hinsichtlich Größe und Form stark unterscheiden können, vor allem das Problem auf, dass die Objekte nicht unbedingt in der Bildszene präsent sein müssen. Wie bereits in Abschnitt 3.1 dargestellt, können kinderpornografische Aufnahmen auch bekleidete Kinder zeigen, die entsprechend posieren, oder aber die sexuelle Handlung ergibt sich nur durch die Position zweier Personen zueinander. Solche Szenen könnten, unabhängig von der Zusammenstellung der Trainingsdaten, niemals erkannt werden. Daher ist in jedem Fall eine szenenbasierte Klassifizierung einer objektbasierten vorzuziehen.

3.3 Unausgewogene Datenklassen

Unabhängig von den tatsächlich gewählten Kategorien, in die differenziert werden soll (z.B. Kinder-/Jugendpornografie vs. Erwachsenenpornografie vs. irrelevant oder Kinder-/Jugendpornografie vs. irrelevant), existiert in jedem Fall eine Klasse „irrelevant“, die eine

einer Klassifikation in Hundebild oder ein Katzenbild o.ä. vor, d.h. alle zu identifizierende Klassen lassen sich, auch für sich betrachtet, eindeutig abgrenzen. Bei der Fragestellung „Porno“ vs. „Nicht-Porno“ o.ä. definiert sich eine Klasse nur durch das Nicht-Vorhandensein der anderen Klasse und kann damit nahezu unendlich viele verschiedene Bildszenen enthalten. Diese Schwierigkeit versucht man im Allgemeinen durch die Auswahl eines entsprechend großen und diversen Trainingsdatensatzes zu lösen. In [Fu05] wurde als Alternative dazu vorgeschlagen, einen sogenannten Deep One-Class mit Pornographie-Erkennung zu verwenden. Ob dieser Ansatz in abgewandelter Form tatsächlich auch in dem hier vorliegenden Themenfeld Anwendung finden kann, muss noch untersucht werden.

3.4 Zusammenstellung Datenmaterial

Eine weitere Schwierigkeit liegt in der Zusammenstellung von geeignetem Trainingsmaterial. Da in Deutschland der Besitz kinder- und jugendpornografischer Schriften auch zu wissenschaftlichen Zwecken verboten ist, bestehen nur die Möglichkeiten, ausländische Kooperationspartner zu beauftragen, die Daten „verändert“, aber für das Neuronale Netz interpretierbar zur Verfügung zu stellen oder die Neuronalen Netze innerhalb der Polizeibehörden zu entwickeln. Je nach Wahl der Möglichkeit und deren Umsetzung kann dies zu der Situation führen, dass polizeiliche Ermittlungspersonen das Material vorauswählen und deren Bildinhalt aber während der Entwicklung des Neuronalen Netzes für die wissenschaftlichen Entwickler nicht einsehbar ist. Dadurch kann eine mögliche Unausgewogenheit z.B. bei der Szenenwahl in den Eingabedaten nicht mehr erkannt werden. Dies müsste bereits im Vorfeld bei der Zusammenstellung durch die Ermittlungsbehörde verhindert werden. Es gilt hierbei beispielsweise darauf zu achten, dass in jeder Klasse auch Aufnahmen von Personen jeglichen Alters dieser Klasse zu sehen sind, ebenso wie die Gleichverteilung bezüglich Geschlechter, Ethnie, Einzel vs. Gruppenaufnahmen, verschiedener Hintergründe wie Auto, Wald, Wiese oder abgedunkeltes Schlafzimmer, bearbeiteter Bilder mit Logos oder Schriftzug und vor allem der möglichen Szenen, die sich in Aufnahmewinkel (Nahaufnahmen, Eigenaufnahmen, professionelle Studioaufnahmen) und Aufnahmemotiv deutlich unterscheiden können, um nur einige Aspekte zu nennen. Ebenso sollte sich nicht nur auf hochwertige Aufnahmen oder „Standardszenen“ konzentriert werden, denn dies beinhaltet die große Gefahr, dass gerade nicht „professional“ erstellte Bilder/Videos nicht richtig klassifiziert werden. Diese Bilder/Videos zeigen jedoch überproportional oft den aktuellen Missbrauch und sollten daher besonders sensitiv erkannt werden.

Auf der anderen Seite dürfen die Bilder keine Gemeinsamkeiten enthalten, die nicht für die zugehörige Klasse Repräsentant sind. Würden beispielsweise nur Bilder eines bestimmten Fotostudios mit demselben blauen Hintergrund als Trainingsmaterial für die Klasse Kinderpornografie gewählt werden, so bestünde die Gefahr, dass das Neuronale Netz zukünftig bei jedem blauen Hintergrund Kinderpornografie erkennen würde, wenn nicht auch in den anderen Klassen entsprechend oft ein blauer Hintergrund im Trainingsmaterial enthalten ist.

die Gefahr, dass die Definition bzw. Abgrenzung der Klassen für das Training andere sind, als die unter denen die Daten zusammengestellt werden bzw. wurden. So beginnen Serienaufnahmen beispielsweise oft mit einem Alltagsbild wie z.B. einem Kind auf einem Fahrrad, welches sich erst auf späteren Aufnahmen entkleidet. Da diese Serien auf jedem Bild mit dem Logo versehen sein können und als Ganzes verbreitet werden, ist es aus polizeitaktischen Gründen sinnvoll, bereits das erste Alltags-bild mit diesem Logo als Indiz für das mögliche Vorhandensein inkriminierter Dateien zu werden, auch wenn es keine strafrechtliche Bedeutung hat. Für das Training eines Neuronalen Netzes (welches nicht das Erkennen solcher Schriftzüge zum Ziel hat), wäre die Einsortierung des ersten Alltagsbildes in beispielsweise Kinderpornografie hingegen eine Fehlklassifikation, die, falls sie in ausreichender Menge im Datensatz vorkommen, die Klassifizierungsqualität des Neuronalen Netzes erheblich negativ beeinflussen können.

Welche Auswirkungen eine Fehlklassifikation bei der manuellen Bestimmung der Trainingsdaten haben kann, wurde in [Ch20] ausführlich am Beispiel eines Hunde/Katzen- bzw. eines bös-/gutartigen Hautmerkmal-Klassifikator untersucht. So wurde im Fall einer falsch vergebenen Kategorie im Trainingsdatensatz gezeigt, dass bereits eine Fehlklassifikation von 5 % zu einer Reduktion der Erkennungsrate um bis zu 9% führte.

Im Fall weiterer Klassen, die sich teilweise nicht so eindeutig separieren lassen, wie beispielsweise die Kategorien Kinder-, Jugend- und Erwachsenenpornografie, ist zu erwarten das eine falsche Klassifizierung während der Trainingsphase zu einer noch höheren Reduktion der Erkennungsrate führen.

Würde man diese Bilder aufgrund ihrer Fehlklassifikation als nicht-inkriminiertes Datenmaterial aus Gründen der notwendigen Datenreduktion bei Massendaten nicht mehr manuell sichten, so birgt jedes dieser Bilder die Gefahr, dass ein sich noch fortsetzender Missbrauch nicht erkannt und damit nicht beendet werden kann. Hinzukommt, dass in [Ch20] nur binäre Klassifizierungen betrachtet wurden.

3.5 Verzerrungen

Die meisten Neuronalen Netze benötigen eine feste Größe der Bilder, hier wird standardmäßig häufig 224 x 224 gewählt. Dies führt jedoch in den allermeisten Fällen dazu, dass sich die Seitenverhältnisse im Vergleich zum Original verändern und es somit zu Verzerrungen der geometrischen Formen der abgebildeten Objekte kommt. In vielen Anwendungsbereichen der Bildklassifikation lässt sich die richtige Bildklasse an mehreren abgebildeten Objekten erkennen, z.B. Autos, Ampeln oder Straßenschilder für eine „Straßenszene“ vs. Palmen, Meer oder Sand für eine „Strandzene“. Bei der Unterscheidung eines inkriminierten Bildes von einem nicht strafbaren Bild hingegen kann ein einzelnes kleines Detail bzw. Objekt entscheidend sein, z.B. ein spielendes Kind in seinem Kinderzimmer, einmal jedoch wurde ihm der Hosenstall geöffnet und sein Geschlechtsteil wird „präsentiert“. Wird letztes nicht richtig erkannt, da es aufgrund von Verzerrungen beispielsweise statt einer länglichen Form eher elliptisch geworden ist, so

ebenfalls sehr schnell an seine Grenzen, da er nur funktioniert, wenn die eigentliche Information auf dem Bild noch gut erkennbar bleibt. Eine automatische Data Augmentation ohne visuelle Kontrolle kann dazu führen, dass Bilder, die bereits vor der Verzerrung nur grenzwertig eindeutig einer Kategorie zugeordnet werden konnten, die dazu vorliegenden Kriterien verlieren (z.B. Genitalbereich ist rasiert/hat noch keinen Haaransatz) und einer anderen Kategorie zugeordnet werden müssten.

Neben einer Skalierung gibt es noch die Möglichkeit, nur jeweils einen 224 x 224 Pixel Ausschnitt des Bildes zu betrachten. Dieser kann zufällig gewählt werden oder aber die kürzeste Bildseite wird unter Beibehaltung der Seitenverhältnisse auf 224 Pixel skaliert und anschließend mittig ein Ausschnitt gebildet. Letzteres hat den Vorteil, dass das Bild nur an zwei Rändern abgeschnitten wird. Trotzdem bleibt die Gefahr bestehen, dass das „entscheidende Detail“ für die Erkennung von Kinderporno-grafie genau in diesem Bereich zu finden gewesen wäre. Eine noch höhere Bildabdeckung kann man durch die Verwendung mehrerer Ausschnitte pro Bild erreichen.

In [Ch20] haben die Autoren daher als Lösungsansatz das sogenannte Scale Constraint Pooling vorgeschlagen, dass Eingabedaten verschiedener Größe in eine einheitliche Ausgabegröße konvertiert.

3.6 Vorschaubilder

Alle Betriebssysteme und viele Anwendungsprogramme erstellen im zur schnelleren Darstellung Vorschaubilder, die auch noch existent sein können, wenn das Original-bild bereits gelöscht wurde. Daher spielen sie bei der Strafverfolgung eine entscheidende Rolle. Die Bilder liegen dabei häufig mit einer Farbauflösung von 32-bit und einer Größe von 32 x 32 bis 256 x 256 Pixeln (teilweise auch größer) vor. Bei dieser Größe ist eine Erkennung von Details und damit eine treffsichere Klassifizierung erschwert. Dies zeigen auch Vergleiche aus anderen Anwendungsbereichen, wie beispielsweise die Tiny ImageNet Challenge im Vergleich zur ImageNet Challenge. Der Datensatz der Tiny ImageNet Challenge besteht aus 64x64 Bildern, während die Bilder im ImageNet Datensatz durchschnittlich 482x418 Pixel groß sind. Mit einem Ansatz aus der Sprachverarbeitung konnten Dosovitsky et al. [Do20] 2020 eine Steigerung der Erkennungsrate für den Tiny ImageNet Datensatz auf immerhin 84.65 % erreichen, während der gleiche Ansatz auf dem ImageNet Datensatz jedoch 88.55 % lieferte. Andere Neuronale Netze und Ansätze zeigen ähnliche Schwierigkeiten.

In diesem Zusammenhang soll auf eine weitere Besonderheit bei der Anwendung Neuronaler Netz im strafrechtlichen Kontext verwiesen werden: Es darf hier keine Bewertung eines Bildinhaltes erfolgen, der „künstlich“ aufgrund von höchstmöglicher Wahrscheinlichkeit im Rahmen einer Auswertung erzeugt wurde. Dadurch ist beispielsweise die Verwendung sogenannter Generative Adversarial Networks, wie sie bei anderen Anwendungen häufig zur Erstellung photorealistischer Bilder bzw. zur Bildbearbeitung im Einsatz sind, ausgeschlossen.

vermeiden und sich ggf. aufgrund der Bildqualität auf die Unterscheidung pornografisch bzw. nicht-pornografisch zu beschränken und keine Altersschätzung durchzuführen. Da am Ende der Beweisführung ein menschlicher Betrachter jederzeit in der Lage sein muss, aus dem vorhandenen Material dieselben Rückschlüsse ziehen zu können, sind hier gewissen Grenzen vorgegeben.

3.7 Teildefekte Bilder

Bei wissenschaftlichen Fragestellungen wird vielfach mit genau definierten Trainings- und auch Testdatensätzen gearbeitet, wie beispielsweise ImageNet oder MS-COCO. Diese Bilder zeigen häufig nicht nur „Hochglanz“-Motive ohne Aufnahmefehler oder schlechter Beleuchtung, sondern weisen auch keine Beschädigungen im Format oder ihrer Darstellung auf. Im Anwendungsfall der Strafverfolgung ist dies jedoch häufig nicht gegeben. Im Rahmen der Datenwiederherstellung können in vielen Fällen nur noch Teilbereiche des Bildes wiederhergestellt werden, fehlende Zwischenstücke werden durch die forensische Aufbereitungssoftware „beliebig“ ergänzt, entsprechende Beispiele sind in Abb. 1 zu finden. Für das Neuronale Netz besteht vor der korrekten Interpretation der gezeigten Szene hier zuerst einmal die Schwierigkeit, die Szene von den Füllpixeln korrekt zu separieren. Im Fall, dass die eine einheitliche Füllfarbe vorliegt, könnte dieser Bereich als dominant prägend für die Klasse interpretiert werden und das Bild ähnlicher einer einheitlichen Farbkachel bewertet wird. Andernfalls, wenn verschiedenfarbige Pixel vorliegen, besteht die Gefahr, dass diese möglicherweise als eine Art Rauschen interpretiert werden und damit die Klassifikation stören können.

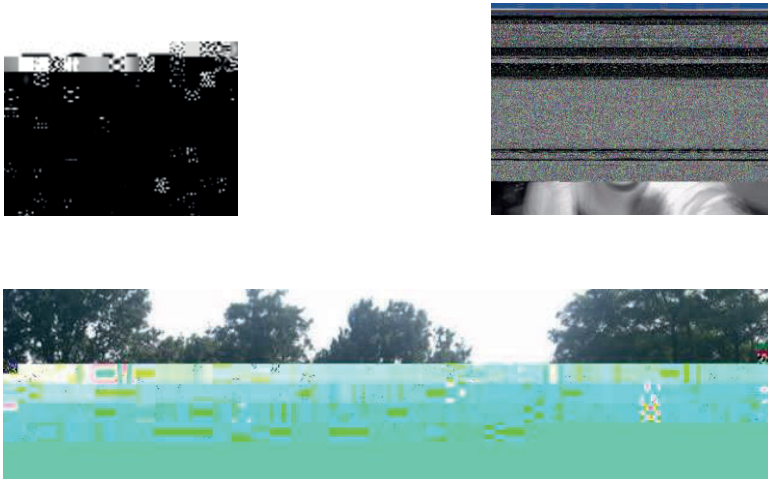


Abb. 1. Beispiele für wiederhergestellte Dateien mit fehlenden Bildbereichen

Es wurde gezeigt, dass die Zusammenstellung des Trainingsmaterials detaillierte Kenntnisse sowohl des Phänomenbereichs Kinderpornografie bedarf als auch Kenntnisse aus der Entwicklung Neuronaler Netze, um mögliche unbeabsichtigte Korrelationen zwischen den Bildaufnahmen zu erkennen und ein Fehltraining des Neuronalen Netzes zu verhindern.

Eine weitere Herausforderung stellt die sehr differenziale Bildauflösung des zu untersuchenden Datenmaterials in der Praxis dar, beginnend vom Vorschaubild bis zu hochauflösenden Aufnahmen. Im Rahmen Datenvorverarbeitung kann es außerdem zu ungünstigen Verzerrungen von für die richtige Klassifizierung relevanten Objekten kommen, die die Erkennung erschweren. Ebenso stellt sich die Frage, inwieweit Bilddateien sehr verschiedener Auflösungen mit einem einzigen Neuronalen Netz und einheitlichen Kategorien sinnvoll behandelt werden können.

Nach Wiederherstellung gelöschter Bilddateien kann eine Beschädigung des Bild-inhalts vorliegen, so dass nur noch ein Teil des Originalbildes sichtbar ist. Der fehlende Bereich wird dabei im Rahmen der Wiederherstellung ein- oder mehrfarbig motivlos gepixelt ergänzt. Da gerade diese Bilder im nicht-zerstörtem Bereich eine strafrechtliche Relevanz aufweisen können, sollten sie entweder ebenso korrekt wie ein vollständig vorliegendes Bild automatisch erkannt oder andernfalls als „nicht-bewertbar“ markiert werden.

Literaturverzeichnis

- [An21] Ansprechstelle Kinderpornografie LKA Niedersachsen: interne Erhebung, 2021
- [APR17] Avila, S., Perez, M., Rocha, A.: Leverage Deep Neural Networks to Fight Child Pornography in the Age of Social Media, *Journal of Visual Communication and Image Representation*, Bd. 50, S. 303-3013, 2017
- [Ba19] Banaeeyan, R., Karim, H. A., Lye, H., Fauzi, M. F. A., Mansor, S., See, J.: Automated Nudity Recognition using Very Deep Residual Learning Network, *International Journal of Recent Technology and Engineering* 8.3S, S.136-141, 2019
- [Br21] Brazilian Federal Police, <http://patreo.dcc.ufmg.br/datasets/rcpd/>, 10.07.2021
- [Ch20] Chen, J., Liang, G., He, W., Xu, Ch., Yang, J., Liu, R.: A Pornographic Images Recognition Model based on Deep One-Class Classification With Visual Attention Mechanism, *IEEE Access* 8, S. 122709-122721, 2020
- [Do20] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., Houlsby, N.: An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale, *arXiv:2010.11929*, 2020
- [Fu05] Fuangkhon, P.: Neural Network-Based Nipple Detection for Obscene Pictures, *WSEAS Transactions on Computers* ,4, S.1138-1145, 2005
- [Ga2019] Garbers, N.: Erkennung inkriminierter Bilder: Neuronale Netzarchitekturen und Hautanteilfilter im Vergleich, *Polizei-Informatik*, S. 82 – 992, 2019
- [IT12] Izadpanahi, S. and Toygar, O.: Geometric feature based age classification using facial images, *Conference on Image Processing (IPR)*, S. 1-5, 2012

- [KHH21] Kaplan, S.; Handelman D., Handelman, A.: Sensitivity of neural networks to corruption of image classification, *AI Ethics*, S. 1-10, 2021
- [MTH15] Moser, A., Temper, M., Haslinger, D.: Challenges and Limitations Concerning Automatic Child Pornography: Classification, 10th International Conference on Computer Vision Theory and Applications, 2015
- [MCD] Macedo, J., Costa, F., dos Santos, J. A: A Benchmark Methodology for Child Pornography Detection, 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), S. 455-462, 2018
- [Mo15] Moustafa, M.: Applying deep learning to classify pornographic images and videos, *arXiv:1511.08899*, 2015
- [Pa16] Panis, G., Lanitis, A, Tsapatsoulis, N., Cootes T.F.: Overview of research on facial ageing using the FG-NET ageing database. *IET Biom* 5, S. 37–46, 2016
- [PGK] Punyani, P., Gupta, R. & Kumar, A. Neural networks for facial age estimation: a survey on recent advances. *Artif Intell Rev* 53, S. 3299–3347, 2020.
- [Ya2021] Yang, K., Yau, J., Fei-Fei, L., Deng, J., Russakovsky, O.: A Study of Face Obfuscation in ImageNet. *arXiv preprint arXiv:2103.06191*, 2021

AI Reading, or Automatic Semantic Decomposition into Knowledge Graphs and Symbolic reasoning through Marker Passing

Prof. Dr. Johannes Fähndrich¹ Dr. Frank Trollmann²

Abstract: Marker passing algorithms have been applied to solve problems in artificial intelligence related to the semantic of written words. Such approaches could also prove to be useful in digital forensics, e.g., to reduce the effort of extracting evidence from confiscated data. We call creating semantic or even pragmatic understanding of text: AI reading. With this, we show that the aggregation of knowledge out of heterogeneous information sources can be a combination of symbolic and connectionist approaches. With that, the extraction of knowledge graphs can be automated. This approach has the benefit, when used correctly, that both the creation and the use of the knowledge graph through Marker Passing stay explainable. In this paper we describe a tool chain of a Marker Passing approach from the point of view of digital forensics and discuss challenges and opportunities arising from the application of such an approach.

Keywords: Semantic Decomposition; Digital Investigation; Marker Passing.

1 Introduction

Understanding language is a basic ability of humans, separating us from other animals. Language is developed in the first years of our life but is continually learned over the entire life [BI98]. Understanding language is also an important part of artificial intelligence (AI). Here the first question which arises is the meaning of understanding [Gr68]. The application of methods of AI to digital forensic is not a new idea e.g. [CDO19] approaches digital forensic with answer set programming, and [TR14] analyze supervised multi class discrimination. Language understanding can also be of utility to the law enforcement community. Sighting large amounts of textual data for relevant details can require considerable human resources, especially as the amount of seized data increases [QC14]. However, using AI for this purpose also comes with significant challenges [Vi16].

In the scope of AI, we have worked on an approach to combine Demantic Decomposition and Marker Passing [Fä18] to solve tasks that require language understanding. This approach has been to domain like semantic descriptions of software agents [FWA16] and solving Winograd schemes [FWK18]. We argue that this approach will also prove useful in the scope of digital forensics.

¹ Hochschule für Polizei Baden-Württemberg, Germany, JohannesFaehndrich@hfpol-bw.de

² CODE University of Applied Sciences, frank.trollmann@code.berlin

In this paper we consider this approach from the point of view of digital forensics to point out possible applications and challenges for future work that arise from this domain. Among these challenges is the requirement to provide a traceable reasoning chain that can be understood and verified by a human law enforcement expert.

The paper is divided into two parts. The first part (cf. Section 2), describes the components of the approach and illustrates. The second part (cf. Section 3) analyzes this algorithm from the point of view of digital forensics and discuss strengths and weaknesses to derive future research avenues. Section 4 concludes the paper.

2 Marker Passing

This section describes the Marker Passing algorithm and its four main components. While we use examples from the domain of law enforcement as illustrations, we keep the descriptions independent of this domain. These four components will be picked up in Section 3 where we focus more closely on the requirements in digital forensics.

Marker Passing is an algorithm that is able to answer a question about an information source. As an example, the information could be all data contained on a persons hard drive and the question could be "which communication did a person have over which channels regarding a certain topic". Usually, answering the question isn't easily possible, e.g., because the data spread over multiple files with different formats. To tackle this problem, the Marker Passing algorithm employs a graph-based representation of this text, called a Semantic Graph. This representation is generated from the original information source via a process we call Semantic Decomposition. The Semantic Graph represents important entities from the original data source as well as their relationships. The relations in this Semantic Graph may extend beyond the relations captured by the original text. They may, for example, relate names that occur in multiple data sources.

Marker Passing operates on this Semantic Graph. It propagates data (so called markers) along the graph to detect connections between pieces of information. The algorithm does this by placing markers with information onto the graph and propagating these markers along its edges. The resulting placement of markers can be interpreted to answer the original question.

This algorithm can be subdivided into four components:

1. *Semantic Graph*: A graph structure that is appropriate to capture the entities and relationships relevant in the application domain.
2. *Semantic Decomposition*: An algorithm that generates a Semantic Graph out of information sources.
3. *Marker Passing*: An algorithm that propagates a marking along the edges of the Semantic Graph.

4. *Result Interpretation*: An algorithm that interprets the result of the Marker Passing to answer the original question.

These four components are discussed in more detail in the following subsections.

2.1 Semantic Graph

The first step of adapting the Marker Passing algorithm to a new domain is defining the Semantic Graph. The goal is to capture entities that are relevant to the application domain and relationships that are relevant to the types of questions that should be answered. Traditionally, a graph consists of nodes and edges, where each edge connects two nodes. In the example of communication, we could choose to model topics of interest and persons as nodes and connect them via edges whenever we discover a message about a topic that is sent from one person to another. As discussed in more detail in [Fä18], a more complex graph structure is usually needed to represent AI Reading problems. For this reason, the following extensions have been made:

- **n-ary edges**: can connect more than two nodes. They can represent relations between multiple entities (e.g., a message from a sender to a recipient about a topic).
- **edges connecting edges**: can point towards other edges. This enables the representation of meta-information. E.g., we can use these edges to keep track of which file a message edge has been extracted from.
- **node attributes**: can represent nodes with additional attributes. E.g., a person can be annotated with first and last name.
- **typing mechanism**: can be used to treat nodes and edges differently based on which kind of information they represent. E.g., we may decide to treat nodes representing persons differently than nodes representing topics.
- **type inheritance**: can be used to represent similarities on a type level. E.g., we can express a hierarchy of topics such as “sexual assault is a specific type of assault”.

Figure 1 shows an example of a semantic graph modeling the offense of “sexual assault” as a 3-ary relation. It connects an “offender” as the source, a “victim” as the target and a “weapon” as the object of the sexual assault. This is enabled by the typing concept in our Semantic Graph. The example also illustrates edges connecting edges. The edge “is a” represents the fact that a “sexual assault” is a special type of “assault”.

The application of the Marker Passing algorithm to a new domain requires the definition of a type graph. This requires a list of node and edge types and their possible attributes. This is a one-time effort: once this type graph has been defined it can be used to run the algorithm on all problems with a similar structure. It is also often possible to use and adapt existing type graphs for similar problems, as was done, e.g., in [FWK18].

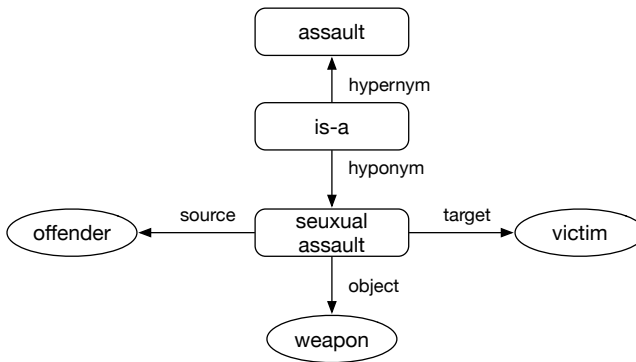


Fig. 1: Example of a typed graph.

2.2 Semantic Decomposition

Semantic Decomposition is the process of extracting the Semantic Graph from the original information source. The idea of breaking complex concepts down into less complex concepts is a well known solution strategy in explaining complex concepts.

While the algorithm for decomposition can be custom-made for specific data sources and file formats, a lot of the data available on a computer will be in natural language. For this type of data a variety of syntactic and semantic approaches already exist and can be integrated into such an algorithm. Riemer [Ri15] describes a syntactical decomposition of parts. For example, sentences can be decomposed into less complex syntactical structures. A first approach on a Semantic Decomposition algorithm has been introduced in [FAA14] and was detailed in [FWK18]. This algorithm has been improved and made available to the community³.

Fig. 2 depicts a simplified version of the graph created from different data sources. Beyond the original data files, Semantic Decomposition can include additional online data sources, e.g., a dictionary to better understand words and their synonyms [FAA15]. With these additional data sources (here called dictionaries), it is possible to establish relations that go beyond uses of the same word in multiple files. The decomposition can take custom ontologies or domain specific databases like mobile device forensics, special information retrieval or structured data sources into account, which makes it customizable for most contexts. The domain DB in Fig. 2 is an example of typical information extracted from confiscated phones.

³ <https://github.com/Datenverlust/SemanticDecomposition> Please contact the first author if the access is still not public on the date of publication. Access will be granted.

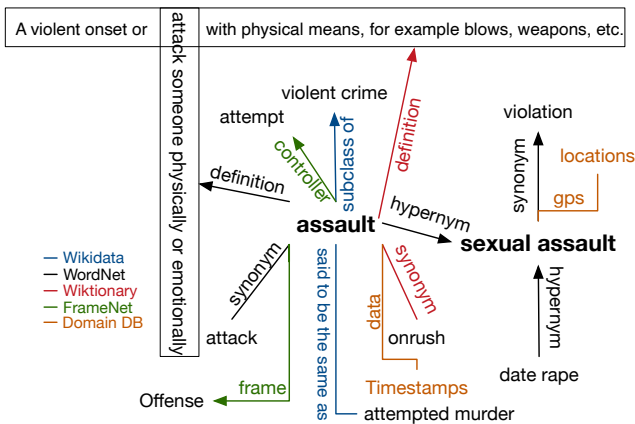


Fig. 2: Example of a Semantic Graph created by a decomposition.

2.3 Marker Passing Algorithm

The Marker Passing algorithm has been published in [FWA16] and has been detailed in [Fä18]. It generalizes the idea of activation spreading, like it is used in neuronal networks, to more structured information. Marker Passing has been used in NLP tasks as a model of thought e.g. [Me19] With that we want to achieve an amalgamation of connectionist and symbolic reasoning like formulated in the vision of Charniak [Ch86a; Ch86b; DF; He89]. An implementation has been made available to the community ⁴.

The information (e.g., a number, a location, a Person) is represented by so-called “markers” that are associated to the nodes of the graph. They denote the attention of the Marker Passing algorithm and annotate additional information. This marking is not static. It changes over time according to the rules of the Marker Passing algorithm.

The initial placement of markers and the Marker Passing behavior are problem dependent. The initial placement of markers is based on the question asked. E.g., if our question is to establish who knew about a sexual assault, we could place markers on nodes representing this topic.

The Marker Passing algorithm moves these markers through the Semantic Graph. The rules according to which markers move are problem specific. This has, e.g., been applied to implement a semantic similarity measure [FWA16]. In our conversation example, the algorithm may spread markers from the topic of sexual assault to all persons that knew about it according to chat conversations to identify the offender or possible victims.

⁴ <https://github.com/Datenverlust/MarkerPassingAlgorithm>

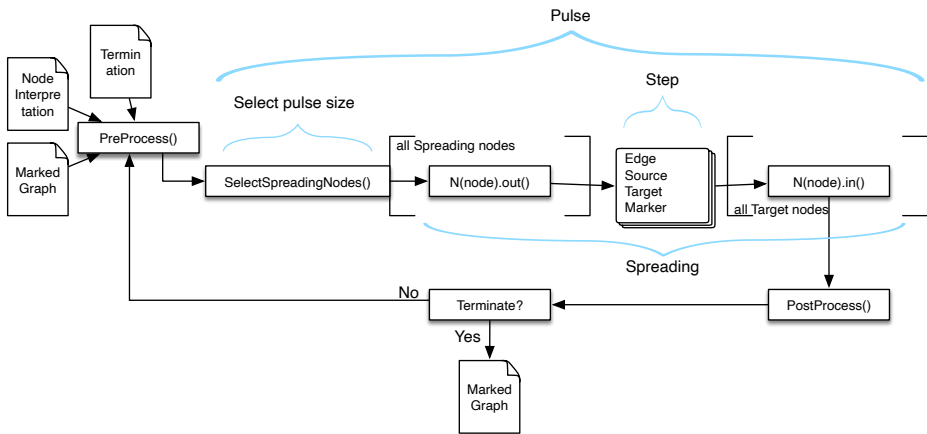


Fig. 3: Abstract description of our Marker Passing algorithm.

Figure 3 shows an abstract representation of the algorithm. The input of the algorithm consists of a marked graph, a node interpretation and a termination condition. These serve the following purposes:

Marked Graph This is a Semantic Graph as described in Section 2.1. It is annotated with markers that represent the initial marking.

Node Interpretation The Node Interpretation describes the Marker Passing behavior of specific node types. It consists of a set of concept interpretations (representing the node types). Each node type defines how it processes markers passed to it, when it is activated, and how it passes markers to other concepts. These functions allow each concept type to react to different markers differently by specifying.

Termination condition This condition defines the stopping point of the spreading. If no condition is given, the activation spreads until no concept is activated. This only accomplishes the goal of Marker Passing in some cases [Be09].

The output of the Marker Passing algorithm is a marked graph. The structure of the Semantic Graph does not change during the Marker Passing, thus the output graph differs from the input graph only in the location and value of markers.

The Marker Passing algorithm is divided into a sequence of **pulses**. Each pulse activates some concepts and spreads their activation to neighbors. A common visualization is to imagine the activated concepts to light up and create a pulsing light which “wanders” through the graph. This visualization gives the Marker Passing pulse its name.

The pulse has two intermediary steps: the **Pre-** and **Post-Processing**. They can be used to integrate general tasks such as cleaning up or normalizing the activation.

For each pulse, the pulsing concepts are selected during **Pulse Size Selection**. This step selects some activated concepts to spread their activation. While the intuition of the reader may be to pulse all activated concepts at once, this turns out to be impractical in reality. The ability to select a smaller pulse size enables us to avoid situations in which the termination condition becomes fulfilled by activation of one concept and then negated by activation of another concept. Furthermore, it enables more fine-grained control in cases where the order of activation influences the result.

After selecting these concepts, the **spreading** of activation takes place. This step consists of three stages:

Out-function All pulsing concepts determine which of the connected relations to spread information over and which markers they should receive by using their node interpretation. These markers are then forwarded to the respective relations.

Spreading Step The relations take these incoming markers and process them. They also select which of the connected concepts receive which markers. The markers are then forwarded to these concepts.

In-function The receiving concepts get their respective markers via their in-function. This function determines how to integrate the new markers with the already placed markers. Again, the node interpretation has an influence on this step.

Confiscated data rarely contains relevant facts directly. With the large amount of data collected, the data triage needs to be efficient but reduce false negatives. In our running example, Marker Passing could extract a Semantic Graph on which information has been known to whom (e.g., by spreading over email and messaging protocols) and when (e.g., via the associated timestamps). Since the graph will include the same concepts if they appear in different contexts, the marker can reveal connections between information sources.

The interpretation of the result Result is still a manual task. The output of the Marker Passing algorithm is the final position of the propagated markers in the Semantic Graph. The goal of the result interpretation is to lift this representation back to the level of the original question.

Since the Semantic Graph and the algorithm have been defined with this type of question in mind, this step can be trivial - e.g., it may involve reading the result directly from the markers. However, more complex methods - e.g., making statements about the distribution of markers or their movement during the algorithm, are also possible and are highly dependent on the question to be answered.

3 Applications in Digital Forensics

In this section we discuss possible uses of Marker Passing in the field of digital forensics and highlight challenges and extensions derived from these uses. We start with an example of using AI reading to process mass data. With that we want to highlight design decisions in the parameters of the Marker Passing. Some of these parameters can also be learned experimentally and are subject to fine-tuning.

The application to digital forensics starts with the creation of a format for the Semantic Graph and an algorithm for the Semantic Decomposition. Previous work on Semantic Decomposition has been based on public OSINT dictionaries. This work needs to be extended to include data found in private police databases. This customization can go further: The case file and the collected data in one case make a good starting point. Depending on the investigated crime, specific data sources can be integrated e.g. car registration, password or username lists, social networks, cellphone GIS data or traffic information. The reasoning capabilities of the Marker Passing are heavily influenced by the information included in this stage.

The Marker Passing needs to be configured to suit the reasoning use case. While this parameter selection is context dependent, we argue for a few rules to keep the marker passing deterministic and explainable.

Marked Graph We argue to keep the graph structure constant while the Marker Passing is running. This assures that the graph structure is only influenced by the initial data sources and can be traced back to them.⁵

Node Interpretation Node interpretation gives the Marker Passing the ability to change passing behavior depending on the node type. E.g. secured facts can have another passing behavior than non-secured facts. Another node dependent interpretation might be dependent on the data source the node came from⁶.

Termination condition The termination condition should be deterministic. Here an optimization needs to be found, when the markers should stop passing. This stop on the activation prevents markers to move away from relevant concepts.

Out-function The out-function should not be probabilistic. It should always be reproducible which markers it passes, given a specific marking of the concept.

SelectSpreadingNotes The selection of the spreading nodes can be modeled in many ways. Some of these options (e.g. inspired of natural observation, like in neuronal networks) are modelled with e.g. inhibition or decay over time. These are dependent on system

⁵ If relations are found that were not in the graph, one might be tempted to include these new relations into the graph: e.g. a ownership relation of a piece of evidence.

⁶ e.g. the information in a DMV might be more trustworthy than some testimony out of a case file.

time, thus depend on the underlying machine and its performance, which makes the results only reproducible in a controlled environment. To be explainable, such mechanisms should be avoided.

In-function The in-function should be deterministic, meaning the resulting marking of the concept should only be dependent on the previous marking and the newly passed markers.

We ignored the pre- and post-processing function, because they are thought for clean up and special (out of the algorithm) modification of the graph and the markers. However, these functions should be written in a way that they don't manipulate the relevant information in the graph.

One important concern in law enforcement is traceability - the ability of humans to understand the results of the algorithm and to trace it back to the original data sources. An algorithm that does not fulfill this requirement would not be usable as it hurdles police and court proceedings to provide conclusive evidence.

In our Marker Passing algorithm, this requirement for traceability means that we need to be able to trace the answer back to the original data sources and how they are processed by the algorithm. This means, all components of the algorithm need to support traceability. In specific (working backward from the answer) the following requirements need to be fulfilled by the respective components:

Result Interpretation: the process of creating the answer from the final marking of the Semantic Graph needs to be traceable. It needs to be clear, which of the markers present in the network were involved in creating the answer, and how so.

Marker Passing: The path of the markers through the Semantic Graph needs to be traceable. Even if the algorithm itself is deterministic, it is not generally possible to understand the path of tokens from the final marking. Traceability of Marker Passing requires a recording of the marking history, either by having tokens keep track of their path or by recording the states of the activation spreading and enabling a replay. Depending on the use case, more specialized methods could be implemented that only keep track of information relevant for the end result. In addition to keeping track of the history, visualization techniques may play an important role in enabling an investigator to understand the evolution of the marking over time.

Semantic Decomposition: Any concept in the Semantic Graph needs to be traceable to the original data sources. This can be guaranteed by a Semantic Decomposition that keeps track of meta-data regarding the origin of its concepts (e.g., pointer to file and line number). This needs to be possible both for concepts and relations.

Semantic Graph: The Semantic Graph is static and thus there is no need for traceability. However, it needs to provide the appropriate edges, concepts and data types for keeping track of the above-mentioned traceability information.

If these requirements are fulfilled, it is possible to trace back the answer to the final marking of the graph, understand which concepts / relations have been involved in producing it and find the respective information in their original files. This makes it possible to understand and validate the reasoning taken by the algorithm.

Other properties that should be discussed are correctness and completeness to assess the reliability of the approach if it is to be used in law enforcement. Correctness means the answer of the algorithm is correct (e.g., a social network connection pointed out by the algorithm is indeed a provable connection). Completeness means the algorithm detects all answers that could be given (e.g., that it finds all connections contained in an email dump). This can be measured in false positives and false negatives. Both correctness and completeness are dependent on the specific implementation of the algorithm, and their importance may also depend on the specific use case.

4 Conclusion

In conclusion, the definition of a Semantic Graph enables us to describe knowledge representations in a connectionist way. The Semantic Decomposition algorithm can be used to create such graphs from text sources like Wikipedia or WordNet. Additional information sources can be integrated by defining new dictionaries to extract concepts and relations. The application of the Semantic Decomposition can be manifold.

Based on such a graph, the Marker Passing algorithm allows a symbolic reasoning on such a graph. With the interpretation of the resulting markers, many graph based algorithms can be simulated (PetriNets, Artificial Neuronal Networks, or other state transition systems). The idea of this combination of connectionist graphs and symbolic reasoning could be applied to many challenging tasks. The mechanism realizes the fuzzy understanding of text while reading, thus the title of this paper. Utilizing text to extract concepts and relations, and then perform reasoning upon those extracted concepts and relations, could be seen as reading.

To become explainable, two things need to be parametrized in this approach:

- Decomposition with source specific concepts and relations
- Marker passing with history on the nodes or the markers.

By specifying where concepts and relations come from, the text, which has been used to create the Semantic Graph, can be traced back and analyzed. With the history, the concepts

and relations that have participated in the algorithm can be reconstructed and the reasoning of the algorithm can be traced. The result of this approach thereby is fully explainable, if certain conditions hold. E.g. the graph should not be modified by the Marker Passing algorithm. Or the decomposition is done before the Marker Passing is started.

The challenge of bringing methods of AI to digital forensics is that the approach needs to be validated on real world data, before it can be applied to any real data. Additionally, the inner working, the selected parameters for the marker passing and the used data sources need to be discussed with judges and prosecutors so that the resulting graph including the markers can be interpreted.

Our future work will include a modeling of "focus" to investigate in large text data like confiscated files, which are, because of their size, no longer processable by humans. The data is used to create graphs connecting concepts with relations, and the Marker Passing algorithm needs to be parametrized to "focus" on relevant data. Additionally we have to validate the approach on real world data, with help of investigators to select the parameters accordingly.

References

- [Be09] Berthold, M. R.; Brandes, U.; Kötter, T.; Mader, M.; Nagel, U.; Thiel, K.: Pure spreading activation is pointless. In: Proceedings of the 18th ACM conference on Information and knowledge management. Pp. 1915–1918, 2009.
- [Bl98] Bloom, L.: Language acquisition in its developmental context./, 1998.
- [CDO19] Costantini, S.; De Gasperis, G.; Olivieri, R.: Digital forensics and investigations meet artificial intelligence. *Annals of Mathematics and Artificial Intelligence* 86/1, pp. 193–229, 2019.
- [Ch86a] Charniak, E.: A Neat Theory of Marker Passing. *Association for the Advancement of Artificial Intelligence*/, pp. 584–588, 1986.
- [Ch86b] Charniak, E.: A Neat Theory of Marker Passing. In: *AAAI*. Pp. 584–588, 1986.
- [DF] Dyer, M. G.; Flowers, M.: Integrating Marker Passing and Connectionism for Handling Conceptual and Structural Ambiguities./.
- [Fä18] Fähndrich, J.: Semantic decomposition and marker passing in an artificial representation of meaning, Doctoral Thesis, Berlin: Technische Universität Berlin, 2018.
- [FAA14] Fähndrich, J.; Ahrndt, S.; Albayrak, S.: Formal Language Decomposition into Semantic Primes. *ADCAIJ: Advances in distributed Computing and AI Journal*/, 2014.
- [FAA15] Fähndrich, J.; Ahrndt, S.; Albayrak, S.: Self-Explanation through Semantic Annotation and (automated) Ontology Creation: A Survey. In: *10th International Symposium Advances in Artificial Intelligence and Applications*. 2015.

- [FWA16] Fährdrich, J.; Weber, S.; Ahrndt, S.: Design and Use of a Semantic Similarity Measure for Interoperability Among Agents. In (Klusch, M.; Unland, R.; Shehory, O.; Pokahr, A.; Ahrndt, S., eds.): *Multiagent System Technologies: 14th German Conference, MATES 2016*, Klagenfurt, Österreich, September 27-30, 2016. Proceedings. Springer International Publishing, Cham, pp. 41–57, 2016, ISBN: 978-3-319-45889-2.
- [FWK18] Fährdrich, J.; Weber, S.; Kanthak, H.: A Marker Passing Approach to Winograd Schemas. In (Ichise, R.; Lecue, F.; Kawamura, T.; Zhao, D.; Muggleton, S.; Kozaki, K., eds.): *Semantic Technology*. Springer International Publishing, Cham, pp. 165–181, 2018, ISBN: 978-3-030-04284-4.
- [Gr68] Grice, H. P.: Utterans meaning, sentence-meaning, and word-meaning. In: *Philosophy, Language, and Artificial Intelligence*. Springer, pp. 49–66, 1968.
- [He89] Hendler, J. A.: Marker-passing over microfeatures: Towards a hybrid symbolic/-connectionist model. *Cognitive Science* 13/1, pp. 79–106, 1989.
- [Me19] Meyer, R.: *Computerlinguistische Ansätze zur Repräsentation und Verarbeitung von Wortbedeutungen*. De Gruyter, 2019.
- [QC14] Quick, D.; Choo, K.-K. R.: Impacts of increasing volume of digital forensic data: A survey and future research challenges. *Digital Investigation* 11/4, pp. 273–294, 2014.
- [Ri15] Riemer, N.: *The Routledge Handbook of Semantics*. Routledge, 2015, ISBN: 1317412451.
- [TR14] Tallón-Ballesteros, A. J.; Riquelme, J. C.: Data mining methods applied to a digital forensics task for supervised machine learning. In: *Computational Intelligence in Digital Forensics: Forensic Investigation and Applications*. Springer, pp. 413–428, 2014.
- [Vi16] Vincze, E. A.: Challenges in digital forensics. *Police Practice and Research* 17/2, pp. 183–194, 2016.

**Workshop: Post-Quanten
Kryptographie in der Praxis (PQKP)**

Requirements Engineering of the Cryptography within the Application Landscape of the German Banking Industry

Wasilij Beskorovajnov,¹ Roland Gröll²

Extended Abstract

The standardization of PQC is going to be the beginning of an enormous effort by the industry incorporating new cryptography into the current application landscape. Moreover, the German federal office for information security (BSI³) recommends that applications with high security requirements should assume that in 2030 a cryptanalytic relevant quantum computer will exist. This suggests that early migration considerations are quintessential now. The German banking industry is heavily reliant upon cryptography. A natural question is therefore: What requirements does the application landscape of the German banking industry impose on cryptography and which PQC schemes meet these requirements. We attempt to answer this question by closely reviewing the application and cryptography landscape of the German banking industry in accordance with a classical requirements engineering (RE) process that is being taught by the IREB Foundation Level course⁴, which we assume to be the most pervasive in the industry. We describe how this methodology was applied in the project QuantumLeap⁵ and present an overview of the cryptography being used in the German banking industry together with promising PQC replacement candidates.

RE comprises four activities, i.e. requirements elicitation, documentation, verification and management. We went for a top-down approach and restricted ourselves to two application scenarios, i.e. online banking and money withdrawal/point of sale (PoS). Every scenario comprises several systems (human/technical). The application landscape of the banking industry is strictly regulated by various standards, which generally are an important source of (non-)functional requirements. Expert interviews on the other hand were a crucial tool in prioritizing elicited requirements as the specifications do not tell the current pervasiveness of the protocols. The application scenario of online banking comprises different goals, such as client (multi-factor) authentication, authorization of account-information retrieval by

¹ Forschungszentrum Informatik FZI, Kompetenzzentrum IT-Sicherheit, Haid-und-Neu-Str. 10-14, 76139 Karlsruhe, Germany Beskorovjanov@fzi.de

² Forschungszentrum Informatik FZI, Kompetenzzentrum IT-Sicherheit, Haid-und-Neu-Str. 10-14, 76139 Karlsruhe, Germany groell@fzi.de

³ Bundesamt für Sicherheit in der Informationstechnik

⁴ <https://www.ireb.org/en/cpre/foundation/>

⁵ <https://projekt-quantumleap.de/>

third-parties⁶, authorization and confidentiality of business transactions and transport layer security (TLS). Every goal has one or more protocols specified. Most protocols within the application of money withdrawal or PoS are described in the EMV specifications⁷. The protocols describe what, how and to what end cryptography is being used, which allowed us to extract requirements regarding the formal security such as EUF-CMA for digital signatures or IND-CCA2 for encryption. It is important to note that specifications usually do not mention what kind of formal security is required. Currently, throughout the landscape of all applications PKCS#1 RSA⁸ is used for digital signatures and hybrid encryption in conjunction with AES-CBC, which is specified under the name RAH⁹ or E002¹⁰. The derivation of transaction numbers (TAN) is a AES-CBC MAC and hashing is done via SHA-2.

The main source for the elicitation of non-functional requirements are the interacting systems within the protocols. We have identified embedded systems such as Integrated Circuit Cards (ICC) with one or more applications and card reader, end-user devices, back-end systems and human actors. As cryptography is implemented across all systems we took the ICC as the system with the most substantial constraints. We have identified that the RAM, which is according to current data sheets of NXP appx. 10kB¹¹, is going to be the biggest constraint. As current benchmarks suggest¹² we see Kyber and Saber as the most promising candidates for the hybrid encryption schemes. The digital signatures on the other hand are much more expensive in almost all aspects. Greconici et al. [GKS21] showed different implementation strategies, where one of the strategies allows for low RAM-usage during the signature process, which is perfect for systems such as ICCs, and another one favors signature cycles count, which is perfect for user-end devices. Therefore, we expect Dilithium to be a promising candidate for a digital signature scheme. Regarding the symmetric cryptography it is currently unclear which construction is a good candidate to replace the currently broken AES-CBC MAC due to Kaplan et al. [Ka16].

References

- [GKS21] Greconici, D. O.; Kannwischer, M. J.; Sprenkels, D.: Compact Dilithium Implementations on Cortex-M3 and Cortex-M4. *IACR Transactions on Cryptographic Hardware and Embedded Systems*, pp. 1–24, 2021.
- [Ka16] Kaplan, M.; Leurent, G.; Leverrier, A.; Naya-Plasencia, M.: Breaking symmetric cryptosystems using quantum period finding. In: *Annual international cryptography conference*. Springer, pp. 207–237, 2016.

⁶ <https://www.berlin-group.org/nextgenpsd2-downloads>

⁷ <https://www.emvco.com/>

⁸ <https://datatracker.ietf.org/doc/html/rfc8017>

⁹ https://www.hbci-zka.de/spec/4_1.htm

¹⁰ <https://www.ebics.de/de/ebics-standard>

¹¹ <https://www.cardlogix.com/downloads/support/SmartMX3-family-P71D320-datasheet.pdf>

¹² <https://github.com/mupq/pqm4/blob/master/benchmarks.md>

On the State of Post-Quantum Cryptography Migration

Nouri Alnahawi,¹ Alexander Wiesmaier,¹ Tobias Grasmeyer,¹ Julian Geißler,¹ Alexander Zeier,² Pia Bauspieß,¹ Andreas Heinemann¹

Abstract: Safeguarding current and future IT security from quantum computers implies more than the mere development of Post-Quantum Cryptography (PQC) algorithms. Much work in this respect is currently being conducted, making it hard to keep track of the many individual challenges and respective solutions so far identified. In consequence, it is difficult to judge, whether all (known) challenges have been sufficiently addressed, and whether suitable solutions have been provided. We present results of a literature survey and discuss discovered challenges and solutions categorized into different areas and utilize our findings to evaluate the state of readiness for a full scale PQC migration. We use our findings as starting point to initiate an open community project in the form of a website³ to keep track of the ongoing efforts and the state of the art in PQC research. Thereby, we offer a single entry-point for the community into the subject reflecting the current state in a timely manner.

Keywords: system security; network security; post-quantum cryptography; integration; migration

1 Introduction and Related Work

Most cryptographic algorithms currently used are subject to security decay over time. This risk has become significantly greater in the light of the expected rise of quantum computers, as established asymmetric schemes like Rivest–Shamir–Adleman (RSA), Digital Signature Algorithm (DSA), and Elliptic-Curve Diffie–Hellman (ECDH) will be broken [Gr96; Sh97]. To maintain security against future and unknown threats, a migration from widely used and established cryptographic standards in IT-systems and infrastructures towards quantum-secure cryptographic schemes must be prepared and executed as soon as possible. This task involves many aspects that need to be taken into consideration and is a transition that has to occur on numerous levels. It is thus more than a simple exchange or replacement of a cryptographic scheme or system [Mo15; OP19]. In this paper we highlight all aspects of the current state of PQC in practice and discuss the main challenges regarding its integration and migration. We focus on research efforts and open questions around PQC development, integration, and adaptation to provide an overview of challenges already solved, challenges currently investigated, challenges recognized yet still unsolved, as well as current blind spots. Through a systematic literature review, using references quoted in [OP19] as a starting point for a keyword-based search, followed by classification and forward/backward citation chasing, we were able to extend and refine the categorization scheme by [OP19] and include

¹ Darmstadt University of Applied Sciences; first.last@h-da.de

² MTG AG; azeier@mtg.de

³ <https://fbi.h-da.de/cma>, last accessed 2021-07-30

the respective state of the art. The adapted categorization scheme is presented in Fig. 1 and is also reflected in the structure of this paper. The corresponding overview in Table 1 highlights the current state and provides brief commentary on the major challenges and breakthroughs respectively. Thus, we present a comprehensive literature survey of relevant papers and the current state of research and readiness. Based on this, we discuss our findings, draw our conclusions and derive future work. As this can only be a snapshot of the current state, we initiated a website (<https://fbi.h-da.de/cma>) that we keep updating and invite the community to support us in keeping track of the current state over time.

Ott et al. [OP19] point out various contexts in which the migration towards PQC will take place and refer to many known requirements and challenges, as well as possible scenarios and contexts of the migration process on a higher level of abstraction. Campagna et al. [Ca15] provide recommendations on how to make common security standards quantum-secure, such as Transport Layer Security (TLS), Internet Key Exchange Protocol Version 2 (IKEv2), Secure/Multipurpose internet Mail Extensions (S/MIME), Secure Shell (SSH), and X.509. Furthermore, they describe important use cases for cryptography and potential migration strategies to transition to PQC. To implement PQC in practical applications, [NW17] name twelve challenges that need to be solved, including migration (hybrid approaches) and agility (agile protocols and update mechanisms). The challenges faced by the industry in the PQC transition process are discussed in [KNW18], and seven recommendations for next generation cryptography are offered. Armknecht et al. [Ar19] discuss a variety of problems in regard to achieving IT security and possible strategies to solve these. Aspects addressed include attacker models, TLS implementations and certificates. A survey of future challenges for PQC in the internet of things (IoT) is provided in [Fe20], as well as an extensive comparison of the performance of PQC algorithms and useful guidelines for post-quantum (PQ) blockchain security [FF20]. Barker et al. [BPS20] present PQC adoption challenges and thoughts on migrating to PQC after the standardization process of the NIST is concluded. Examples are a migration playbook and a way to get an inventory of used cryptography in an IT-system or infrastructure. Recommendations for action on migration to PQC are given in [BS20b], encouraging the use of crypto-agile hybrid solutions and the corresponding adaptation of cryptographic protocols. Following on the same footsteps with similar recommendations, [He21b; HKW21] offer a brief evaluation of the current state of both PQ and quantum cryptography and highlight the chances and limitations of quantum cryptography. A technical report by ETSI [ET20] provides strategies and recommendations for quantum-safe schemes defining a staged approach for a successful migration.

2 Towards New Standards

The focus of the standardization efforts towards PQC lies within developing new cryptographic primitives and algorithms, as these are considered the cornerstone of cryptographic systems. Nonetheless, efforts have been made in scheme integration, as well as in Requests for Comments (RFC).

NIST Candidate Algorithms PQC algorithms are based on mathematical problems that cannot be efficiently solved by quantum or classical computers [Be09]. These fall mainly into five categories [GAB19; NW17; OP19]: Multivariate (quadratic polynomial equation), lattice-based (a grid as a discrete subset of an n -dimensional real vector space), code-based (the problem of decoding general error correcting codes), isogeny-based (algebraic geometry), and hash-based.

The state of PQC standardization is represented by the respective ongoing NIST process [NI16]. It was initiated in 2016 [Ch16] with 82 algorithm submissions, from which 69 candidates were accepted into the project, and out of these, 26 advanced to the second round [Al19]. Round three was announced with 15 finalists [Mo20]. Including the alternatives, these are nine public-key encryption and key-establishment algorithms, and six digital signatures. Following the third NIST PQC conference, it is planned to release draft standards between 2022 and 2023, and hopefully finalize the process by 2024. An overview of all algorithm candidates is provided in Table 1.

As the process continues, improvements for the PQC schemes are still being studied and proposed. Vectorizing time-consuming primitives [KKP20] is used to improve Kyber using KEM SHA3 and AES variants [Av21]. A fast and first-order secure Kyber implementation optimized for ARM Cortex-M4 is presented in [He21a]. Number-theoretic transform (NTT) is another approach used to increase the performance significantly for Kyber [BKS19], NTRU [Al21a; Ch19] and Saber [Ch21b; DA19] on Cortex-M4 chips. Using advanced single-instruction multiple-data instructions (NEON) [DK21], optimized constant-time software implementations for Kyber, NTRU, and Saber are provided. Techniques to optimize memory requirements in computing discrete logarithms in SIKE [Co20] are presented in [HKP21]. Reducing the input of the hash function leads to noticeable improvements in lattice-based key-encapsulation mechanisms (KEMs), improving Kyber and Saber [Ju21] using generic Fujisaki-Okamoto (FO) transformation. A side-channel resistant implementation of Saber is provided in [Ab21], showing a novel primitive for masked logical shifting on arithmetic shares. Masking the underlying zero-knowledge proof system of Picnic [Ar21] should improve its security against SCAs. Sampling discrete Gaussians through arbitrary-centered sampling over Integers [ZSS20] and generating discrete Gaussians with arbitrary center and standard deviation [Ho20] offer improvements for lattices and Falcon [Fo20]. A speed up in the key generation of Rainbow [DS05] is proposed in [Pe20], eliminating some drawbacks of the scheme. Through efficient statistical zero-knowledge proofs for module/ring learning with errors (LWE) and short integer solution (SIS) relations [Bo20], it is shown that the security of lattice-based schemes can be improved. Costello [Co21] provides a thorough analysis of SIKE and highlights several aspects of its yet unbroken security.

Scheme Integration The cryptographic library liboqs, developed by the Open Quantum Safe project [SM16], provides quantum resistant cryptographic primitives, and compares implementations of the NIST candidates (except GEMSS [Ca19b]) integrated using OpenSSL within applications such as Apache web server and Chromium. They offer PQ and hybrid

implementations for TLS, SSH, S/MIME and x.509 certificates [Mi21]. Early usage of PQ KE cipher-suites for the TLS protocol based on R-LWE [Bo15] is demonstrated in an integration into the OpenSSL library using the Apache web server showing that secure PQ KE is practical. Classic McEliece is used in a quantum safe MACsec [CS21], offering hybrid mode, ephemeral key-exchange (KE), and an end-to-end encryption in Ethernet networks. Suitable NIST candidates are evaluated for use in Domain Name System Security Extensions (DNSSEC) within certain constraints [Mü20]. However, most candidates require larger signature and public key (PK) sizes compared to classical algorithms. Other efforts focus on replacing the DH-based handshake using PQ secure KEMs or hybrid KE [Az21; Hü20] and using DH type authenticated KE protocols from supersingular isogenies [Ba21b; Fu18; KK18]. Moreover, a PQ handshake is successfully integrated in the WireGuard VPN using Kyber and Rainbow [MAY21].

RFC Drafts RFC8391 [Hu18] defines the eXtended Merkle Signature Scheme XMSS, a quantum-safe hash-based signature scheme. RFC8554 [MCF19] describes the hash-based Leighton-Micali signature system. It provides asymmetric authentication and can achieve a high security level, which is secure against a quantum computer. RFC drafts for hybrid KE schemes for TLS 1.2 and 1.3 are proposed in [CC20; SFG20; Wh17]. Older drafts for TLS include the use of additional shared secrets and the Quantum-Safe Hybrid (QSH) cipher-suite [SS17; SWZ16]. Another draft proposed in [Ka20] defines hybrid KE and PQ PK authentication methods for SSH. RFC8784 [Fl20] describes an extension of IKEv2 that is quantum-resistant by using pre-shared keys. A draft to integrate PQ KE into IKEv2 is proposed in [Tj19], using multiple payloads in conjunction with the existing (EC)DH payload. The intermediate exchange for IKEv2 is defined in [Sm21], for transferring large data in the IKEv2 Security Association (SA) establishment. This should help avoiding IP fragmentation of large IKE messages, but cannot be used in the initial IKEv2 exchange.

Other Algorithms Some other novel approaches seem very promising and practical, even though they are not part of the NIST standardization process. A new class of post-quantum signature schemes based on symmetric primitives is presented in [Ch17]. Supersingular EC [CD20] show a significant speed-up in commutative supersingular isogeny DH (CSIDH). Instead of using hidden field equations (HFE), PK encryption schemes can be constructed using multivariate quadratic polynomials [Sa20], based on the generalization of isomorphism of polynomials with two secrets (IP2S) problem to obtain a DH like structure. Implementing Gaussian pre-image sampling on module lattices [Be21] seems practical for signature schemes and for advanced constructions using trapdoors such as identity-based encryption. The PERN multivariate encryption scheme is proposed in [YWT20], based on solving nonlinear equations over real numbers. Mikata [Th21], a signature scheme over NTRU lattices, is arguably a more versatile variant of Falcon. Another variant of Falcon, Zalcon [Fo21], claims more suitable for constraint environments such as smart cards. Lattice-based KINDI-KEM, a former NIST candidate, is formally specified in DIN SPEC 91444 [Dr21].

3 Performance Considerations

Multiple research efforts deal with the performance of PQC algorithms, mostly focusing on hardware implementations and network performance.

Algorithm Performance As indicated by NIST [NI16], the algorithms can be classified into three categories. The first is for a 128-bit security level, category three is for 192 bits, and category 5 is for 256 bits. A thorough overview of all NIST candidates and their security levels is provided in [Fe20]. Benchmarks and comparisons of round two NIST candidates [Ba19; Da20] state that a higher security level increases latency and timing overhead. They indicate that Kyber, FrodoKEM [Al21b], Saber, and SIKE have both high-speed and lightweight implementations. Kyber also shows advantages in terms of execution times, power consumption, and energy usage. Dilithium stands out for its superior signature generation, in terms of speed and memory consumption. In general, lattice- and code-based algorithms are more feasible than isogeny-based algorithms due to the huge performance overhead of the latter [BRP20; PS20; We20]. Stateless hash-based symmetric signature schemes, such as SPHINCS+ [Au20], are generally less efficient than lattice-based signature schemes like Dilithium [Ba21a] or Falcon. They are mainly interesting for applications without high latency requirements, such as offline code signing or certificate signing. Test results show great advantage for SPHINCS+ over similar schemes, such as Picnic [Za20], in terms of speed, signature size, and security. Algorithms that require too much memory or depend on external libraries such as Classic McEliece [Ch20], BIKE [Ar20] and RAINBOW pose a serious challenge on the implementation level.

Hardware Performance Many research efforts highlight the challenges and benefits of implementing PQC algorithms on special hardware and IoT devices, compare their performance, and address required measures and recommendations [Fe20; Ga18; Su20]. They also benchmark PQC schemes such as XMSS, qTesla [Al20c], SIKE, Classic McEliece, Dilithium, Rainbow, and Kyber providing real-time performance evaluation on specific devices and in embedded systems, such as the ARM Cortex-M4, ARM Cortex-R5, IBM Z, and IoT Trust-Anchor chips [CC21; CKY21; Jo21; Ka19; Ma20; MK19]. Some algorithms are still not suitable for such hardware platforms. However, improvements, optimizations, and outlook for further optimization potential are provided. The suitability of PQC on resource constrained devices, such as automotive hardware security modules (HSM) and in vehicle to vehicle communication (V2V), is also analyzed and evaluated [Go21; Ni21; WS20]. Although it is considered feasible, the deployment of algorithms with multiple parameter sets and large key/packet sizes is still work in progress and requires several adaptations. On common IoT platforms, optimizations are offered [AS20], and it is encouraged to focus on smaller RAM usage vs smaller code-size [De21]. Moreover, it is recommended to choose Falcon over Dilithium for signatures, and that Saber or Kyber be selected as a KEM. Kyber is recommended in mobile applications and is suitable for usage on smartphones [Si21]. Saber

is still being checked regarding its suitability for mobile devices [Ri21]. In general, lattice-based schemes such as NTRU-HRSS [Ch19; Hü17] and SPHINCS+ are recommended of IoT settings [Ne19]. However, a decreased security level needs to be considered. In many cases, Field Programmable Gate Arrays (FPGA) are used to investigate the possible performance benefits [Da20; Ko17; Ku20a; ZGF20]. Several high-performance hardware architectures, low level implementation mappings, as well as hybrid KE methods on FPGAs are proposed and evaluated for PQ algorithms, such as SIKE, SPHINCS, Kyber, NTRU, and Saber [Ba19; DMG21; HLX21; KAK18; Ku17]. In general, executing cryptographic algorithms on optimized hardware leads to a significant performance boost and is essential for the usage of these algorithms, especially for less performant PQ algorithms in performance critical use cases [TLW19].

Network Performance With larger signatures and keys, more data must be transferred within networks. This leads to packet fragmentation, which may result in high delays in lossy networks [PST19]. In general, protocols such as TLS, DTLS, and QUIC are able to handle bigger signature sizes, and while the overhead is significant, it is acceptable in most modern use cases [Ka18a; SKD20a; SM16]. The key establishment with Kyber-based TLS performs well compared to ECDH, whereas using SPHINCS+ signatures is challenging on embedded PQC-TLS servers, due to signature sizes and signing time [Bü20]. Additionally, the execution times for PQC algorithms are usually higher than for classical ones and a longer loading time for websites is a problem, because it may lead to bad user experience [Ad19; Kw19; SKD20b]. Use cases such as VPN are less impacted (establishing a connection only once in a while) [KS19]. Experiments in simulated [PST19] as well as in real networks [He19a; KV19] show which algorithms are suitable for these use cases. Some works evaluate the usage of PQC in Blockchain networks [AK18; Ca19a] and propose using lattice-based schemes as a replacement for ECDSA.

4 Security Considerations

The new PQ algorithms open up new undiscovered attack surfaces and introduce additional mathematical concepts into the cryptographic world. New risks that may come with these concepts have to be investigated.

Cryptanalysis Cryptanalysis examines the security of cryptographic schemes given a sophisticated attacker that is able to perform analytical and/or practical attacks. Many of the candidates that were submitted to the first round of the NIST challenge have fallen to various cryptanalysis attacks [Di20a]. In general, either the imposed security of the attacked algorithms dropped below the requirements of NIST, or the scheme was completely broken and the attackers were able to retrieve the private key [Sa19]. Some schemes seem to be secure under classical computing preliminaries only, but fail to resist quantum attacks

[DLW19]. In addition, PQC can not only be used for KE and signatures, other areas can be pseudo-random number generators and encryption [Ku20b]. On the other hand, researcher extracted indications for secure designs from those attacks. Methods to assure security against a certain type of attacks are presented in [Cr20; PS13] and [JS19] present models that enable direct comparisons between classical and quantum algorithms, improving security assumptions for PQC algorithms. Moreover, several new attacks targeting the foundations of the developed PQC schemes are introduced. A new key recovery technique based on MinRank can break multivariate primitives [SV20], whereas another on HFEs used in GEMSS shows that the proposed parameters of the scheme are not as secure as claimed [TPD21]. A new adaptive attack can recover the secret key for a realistic choice of parameters on homomorphic encryption schemes [FHR21]. Hybrid attacks on NTRU can be threatening and require revising its security estimates [Ng21]. Some security issues may occur due to state synchronization failures, which may be avoided through using hybrid stateless/stateful schemes with smaller signatures and faster signing [GFL19; Mc16]. The correctness of NIST candidates is also formally re-assessed [BS20a], and security notions, such as the anonymity and indistinguishability under (chosen/adaptive) ciphertext attacks (IND-CCA) [Al20a; GMP21], qIND-qCPA [GKS20], and random oracles in a PQ mode [KS20] are evaluated and improved. Moreover, new lower bounds on decoding attacks for lattice and code-based schemes are derived [KL21], whereas new upper bounds for rank attacks HFEs in multivariate schemes [ØSV21] are proven.

Side-Channel Attacks Side-channel attacks (SCA) target the specific circumstances of algorithm execution like timing and power consumption. PQC algorithms introduce new attack surfaces in this regard and have to be analyzed for weaknesses not only in a general regard of the algorithm but also in specific hardware circumstances. For all categories of PQC algorithms, SCAs were found, and at the same time effective countermeasures for these attacks have been introduced [Zh20a; Zh20b]. However, the research states that SCAs on PQC algorithms and their countermeasures are still in a very early state. Many SCAs on PQC are not yet evaluated and the current countermeasures usually use ad-hoc designs that protect against individual SCAs but remain vulnerable against others. Tendayi et al. [Te21] provide a multi-target and multi-tool platform to conduct test vector leakage assessment for NIST candidates. Chowdhury et al. [Ch21a] give an overview of several PQC-related SCAs. They discuss SCAs for the different categories of PQC algorithms, random number generators and physically non-cloneable functions.

There are also efforts to minimize the attack surface in general [DA20; Kh18; Zh20b], and other studies focusing on special types of attacks. Discrete Gaussian samplers are shown to be vulnerable to SCAs in many cases, such as timing information leakage with FPGA-based cumulative distribution tables (CDT) for discrete Gaussian samplers [Kh16], and the so-called flush, Gauss, and reload attack [Gr16] using CDT or Bernoulli sampling. Some of the most threatening discovered attacks include fault attacks on lattice-based schemes, such as (first-order) randomizing, zeroing, and skipping faults [BBK16]; the

singularity attack on multivariate PK schemes [Di20b] using a system of linear equations; horizontal SCA on lattice-based algorithms [Ay18] estimating the secret key from a single power measurement with very high success rates; differential power analysis [GSE20; Ka18b], which shows that the security of a hardware implementation of SPHINCS-256 is compromised, due to its stateless construction, and that SC information can successfully recover the entire secret key from fewer than 30 observed Picnic signatures; correlation power analysis (CPA) [Pa18], through which a secret leakage on Rainbow-like multi-layered schemes is identified; improved torsion point attacks on SIDH variants [KP21] constructing backdoors for curves and base field primes, which implies that some modifications for SIKE should be considered; single-trace SCAs [Am20; AR21] recovering the entire secret key on NewHope [Al20b], and a large portion of it on NTRU; theoretical fault injection attacks [Ta21], which can threaten SIKE. Moreover, electromagnetic (EM) SC assisted CCAs over multiple LWE/LWR-based PKEs and KEMs [Pr21] lead to complete key-recovery on lattice-based schemes. Additionally, the spatial dependencies, such as the dependency of the secret key and certain registers, of cryptographic algorithms can be exploited with the help of near-field micro-probes [An19]. However, multivariate signature schemes offer good protection against fault attacks, as potentially harmful attack vectors only partially succeed in recovering the key [KL19].

However, such attacks are not restricted to the new PQC algorithms, as they can still target some of the standard building blocks in many of the old and new schemes alike. Single trace soft-analytical SCA on KECCAK [KPP20] shows that is possible to exploit some vulnerabilities on 8-bit microcontrollers. Countermeasures such as masking and hiding are therefore recommended.

Downgrade Attacks Security protocols such as TLS may suffer from possible downgrade attacks leading the communicating parties to fall back to earlier versions of the protocol [LSH20]. This type of attacks could be especially exploited in the case of a downgrade from PQC to classical cryptography. Crockett et al. [CPS19] distinguish between three main cases, where the client, the server, or both are either aware or unaware of the new protocols. This implies a possible downgrade on the side that is already using a new protocol.

5 Migration Process

We distinguish between two levels regarding the process of migration towards PQC. Those being the migration of algorithms into existing cryptographic applications and security protocols, and the migration of IT-systems and infrastructures.

Algorithm Migration Several approaches show special interest in hybrid KE methods and hybrid combiners. The primary goal of these is to ensure that the desired security

property holds as long as one of the component schemes remains unbroken [CPS19]. These approaches use two or more independent algorithms chosen from both PQ, and classical schemes. This way, it is neither necessary to fully trust the potentially outdated classical scheme, nor the relatively new PQ scheme until it had proven secure enough.

An experiment with a hybrid approach on Google servers is presented in [Br16]. The solution combines EC with NewHope. Another series of experiments presented in [KV19; La16] compare hybrid PQ to non-PQ key agreement in TLS. First implementations and case studies [CPS19; ZWH21] using TLS and SSH in OpenSSL reveal some challenges regarding message size for the protocol itself. Crockett et al. [CPS19] suggest conveying the cryptographic data either by extending the format of the protocol message, or through concatenating the data into a single value. The method of combining also affects the security and performance of the protocol.

A similar approach is proposed for IKEv2 [He19b; Zi15], showing the feasibility of hybrid KE with NIST candidates. The use of lattice-based PQC algorithms with the Apache Kafka Software using a hybrid approach [We20] shows that the performance of the hybrid solution is as good as the slowest algorithm used within. The use of lattice-based PQC algorithms for the industrial protocol Open Platform Communications Unified Architecture (OPC UA) is examined in [PS20]. The approaches are feasible but result in additional performance and communication overhead. Parallel usage of two independent cryptographic schemes within public key infrastructures is also possible with X.509-compliant hybrid certificates [Bi19a]. This approach enables a step-wise transition to PQ secure and hybrid algorithms. There are also new models for KEMs and authenticated KE protocols [Bi19b], proposing several combiners and a provably sound design for hybrid KE using KEMs. Fischlin et al. [FLP14] state a notion for robust multi-property combiners and implement a combiner that satisfies the strongest notion by preserving every property of the input hash functions. An underrepresented approach is possible through the usage of composite key structures for keys and signatures in PKCS#10, CMP, X.509, and CMS structures [OP20].

System Migration The migration process is not simply limited to algorithm exchange or replacement. Updating a set of software applications, their implemented algorithms or changing their underlying infrastructure requires a tedious process, which involves thorough planning and execution over a specific period of time. This is especially challenging when the systems to be migrated need to stay online and/or must maintain interoperability with other systems that may or may not (yet) be upgraded [OP19]. Therefore, one cannot assume a successful migration without considering the challenges related to the planned approach. According to [OP19], one must first identify the domains and their respective priorities in the planned migration. Moreover, there are interrelated dependencies that may not be easy to resolve. Thus, the migration could prove difficult in IT-infrastructures with security vulnerabilities, making them prone to failures. Several works recommend the use of a cryptographic inventory and automated discovery tools as the first step of a planned migration [BPS20; KNW18], which is also suggested by ETSI in their three-stage migration

approach [ET20]. Fischer und El Bansarkhani [Qu21] propose a four-step process for the PQC migration. It starts by detecting the affected components and assessing their PQC readiness. Then, a risk analysis is conducted to prioritize the elements to be transitioned, followed by the actual migration of the individual components to PQC.

6 Discussion, Conclusion and Future Work

We discuss our major findings and address open issues. We follow the same categorization scheme presented in the structure of this paper and draw final thoughts and conclusions.

Towards New Standards The efforts undertaken towards establishing new cryptographic standards are promising, yet still focus more on low level cryptographic primitives and schemes. Following the third NIST PQC conference in June 2021, several algorithms have been, and are still being continuously analyzed and benchmarked. Given the rigorous selection process with worldwide participation and a current set of 15 finalists and alternative candidates, there should be little doubt that we will soon have a set of vetted PQC algorithms at our disposal. Many candidates have already proved acceptable replacements for different classical schemes and showed capable of integration within specific cryptographic protocols and applications utilizing them. However, this implies rather having a set of multiple established algorithms to be used in parallel. Consequently, algorithm selection depends greatly on other factors and future developments. Therefore, carefully selecting algorithms for specific use cases will be an important topic in the future. Still, there is enough room for further development to adapt to special requirements. However, there seem to be a lack of interest in the strategies required to execute the actual migration on a more sophisticated level of abstraction. The efforts made for testing, benchmarking or integration in new or existing protocols consist mainly of independent research and draft proposals, which is not sufficient for the overall migration process on a broader scale.

Performance Considerations As correctly assumed by NIST [NI16], the performance assessment of cryptographic algorithms on specialized hardware platforms has shown to be an important research question for the candidates of the standardization process. Depending on the choice of algorithm and field, PQC can very well be a replacement for classic algorithms. However, it is not very clear, which algorithms perform best in which situation, moreover on which hardware platforms or for which application scenarios. Generally, lattice and code-based algorithms proved more efficient in terms of speed than isogeny-based ones. However, code-based algorithms need more memory. Lattice-based signature schemes also proved more efficient in terms of speed than stateless hash-based schemes in scenarios with strong latency. Special FPGAs are expected to improve the overall performance significantly, but their usage is not always a valid option. Network devices have to handle the overhead introduced by PQC and still provide acceptable latency. This also impacts end user devices,

e.g. smartphones, which have limited power resources and are used in mobile networks. For example, the overhead introduced by PQC is manageable in protocols such as TLS and DTLS. On the other hand, larger key sizes lead to longer loading times, negatively impacting user experience in lossy networks and web related use cases. This however, is not the case for VPN, which is less impacted by PQC. Thus, the efficiency of PQC must allow for usage within established and new application scenarios. Notably, IoT, which is still an underrepresented field, is of special interest due to its restricted resources. So far, lattice-based algorithms seem to be the best choice for such systems.

Security Considerations Due to the described performance challenges that come with the new PQ algorithms, the choice of the algorithm parameters and key sizes play a vital role in establishing security of a cryptographic scheme. Depending on the use case different algorithms and/or set of parameters may be suitable. It is important to understand the trade-off between security and algorithm requirements to be able to choose the right algorithms and/or parameters for different use cases. Moreover, since some PQC algorithm aspects are still fairly new and didn't face enough scrutiny, the long-term security of these algorithms remains unclear. Especially attack vectors like side channel attacks based on timing or power consumption are still in an early state and pose a risk. A very special technique that has not gotten enough research so far is the exploitation of location-based leakages. Some of the biggest issues regarding the security are related to design aspects. Cipher-suite negotiation determines an algorithm that all parties are willing to use. Downgrade attacks target these negotiations to force a weaker security. Many vulnerabilities have been found and fixed, but a crucial question remains, if less costly mitigation measures can be implemented that impact the overall performance as little as possible.

Migration Process The most common difficulties arise from the integration of PQC algorithms into existing protocols. The existing migration approaches address issues on the implementation and communication levels and offer promising solutions for many challenges; such as defining hybrid formats, algorithm negotiation and parameters. Many of the NIST candidates prove suitable for usage in a hybrid scheme; nonetheless, some protocols have size constraints that prevent some schemes from being used. A trade-off between increasing the size limits and performance needs further investigation for more precise results. Depending on the scenario, different design decisions are required.

On a broader scale, a full migration of IT-systems and infrastructures involves more complex issues. Not all systems will be updated to PQ algorithms in time or cannot be updated at all. Therefore, strategies on how to deal with these systems have to be developed. These could be automation frameworks that are used transparently within protocols or systems to manage their migration into PQC or ensure their security. Automated software is needed in order to deploy PQC algorithms and protocols with minimum human interaction. This is especially the case in larger IT-infrastructure such as data centers and cloud-based

applications. Such software should be able to identify the currently used cryptographic components. Automated tools could also support real-time analysis of an ongoing migration, check for any weaknesses and verify the security of new PQC libraries. The aforementioned issues clearly point out the importance of developing frameworks capable of managing, executing and testing the various aspects within a migration process on different scales and implementation levels. Our survey shows a lack of such automated software. Too few projects and research dealing with these questions were found.

Conclusion Considering the expected rise of quantum computing, challenges posed upon classical cryptography and the IT-systems using them will become tangible threats jeopardizing the safety and security of IT-systems, applications and communication alike. In this paper, we present a survey of the ongoing research efforts towards realizing PQC in practice. Our findings indicate a major focus of the ongoing research on algorithms per se, and the feasibility of migrating different cryptographic schemes and security protocols with reasonable trade-offs regarding algorithm performance and communication efficiency. Considerable work on the standardization and integration into established standards, such as TLS, SSH, and X.509, has been done. It is nevertheless clear that the tasks involved in the entirety of the PQC migration stretch over a wide range of sub-fields and PQC-related categories. There are still many challenges and open issues that need to be addressed, especially on levels above algorithm development. These include protocol design, migration and deployment strategies, legacy systems, testing frameworks, and process automation.

Outlook The overview and findings at hand provide a starting point for providing answers to the yet untouched issues. Identifying the challenges and open issues is, however, only the first step in our research. Theoretical and practical case studies can be conducted to come up with new ideas and solutions. We initiated an online project (<https://fbi.h-da.de/cma>), inviting researchers to join us continuing the work started with the paper at hand. We are also working on the development of solutions aiding the community in the transition to PQC and achieving this migration (<https://fbi.h-da.de/pqc>). A first prototype of an easy-to-use cryptographic interface called eUCRITE (<https://fbi.h-da.de/eucrite>), providing minimal knowledge abstractions for both conventional and PQC functionalities is evolved. Other projects under development include an automated cryptographic infrastructure detection tool, a PQC testing framework, TLS 1.3 PQC integration evaluation, and a code-binding generator for PQC algorithms based on an abstract meta-description language. The next step of our endeavor is to utilize the gathered findings, knowledge and experience throughout our research on the way towards usable and agile PQC.

Acknowledgment: This research work has been funded by the German Federal Ministry of Education and Research and the Hessian Ministry of Higher Education, Research, Science and the Arts within their joint support of the National Research Center for Applied Cybersecurity ATHENE.

Appendix

Categorization Scheme

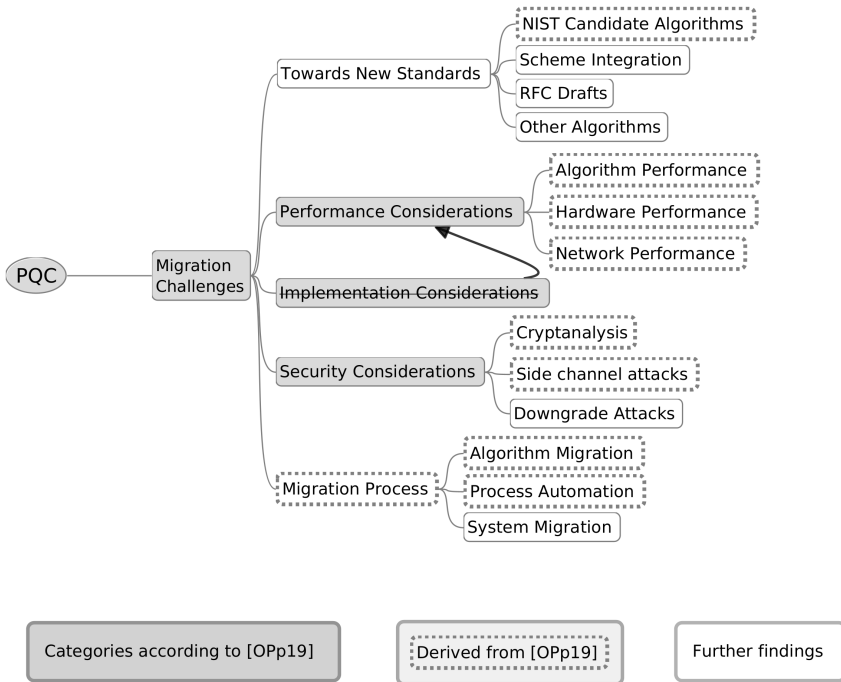


Fig. 1: Categories of PQC Migration as defined by [OP19] (grey). Extensions derived by us from [OP19] (dotted line). Further extensions by us (white and strikethrough).

Survey Overview

Tab. 1: Survey Overview

NIST Candidate Algorithms	
Basics of Post-Quantum Cryptography	[Be09]
Current state of NIST standardization process	[A119; Ch16; Mo20; NI16]
Categories of PQC algorithms	[GAB19; NW17; OP19]
Improvements for NIST finalists	[Ab21; BKS19; Bo20; Co21; DK21; He21a; HKP21; Ho20; Ju21; KKP20; Pe20; ZSS20]
ALL NIST PQC finalists	[Al20c; Al21a; Ar20; Au20; Av21; Ba21a; Ca19b; Ch19; Ch20; Ch21b; Co20; DA19; DS05; Fo20; Za20]
Scheme Integration	
Integration into existing tools	[Bo15; Mi21; SM16]
Quantum safe MACsec	[CS21]
Integration into DNSSEC	[Mü20]
Handshakes with PQC	[Az21; Ba21b; Fu18; Hü20; KK18; MAY21]
RFC Drafts	
Hash Signatures	[Hu18; MCF19]
Hybrid KE schemes for TLS	[CC20; SFG20; Wh17]
Shared secrets and QSH cipher-suite	[SS17; SWZ16]
Hybrid key-exchange	[CC20; Ka20]
IKEv2	[Fl20; Sm21; Tj19]
Other Algorithms	
New class of signature schemes	[Ch17]
Speed-up for CSIDH	[CD20]
PKE using multivariate quadratic polynomials	[Sa20]
Signatures and identity-based encryption	[Be21]
PERN multivariate encryption scheme	[YWT20]
Falcon Variants	[Fo21; Th21]
Lattice-based KINDI-KEM	[Dr21]
Algorithm Performance	
Security levels and impact	[Ba19; Da20; Fe20]
Lattice and code superior to isogeny	[BRP20; PS20; We20]

Hardware Performance	
Hardware Implementations	[CC21; CKY21; Da20; Fe20; Ga18; Go21; Jo21; Ka19; Ko17; Ku20a; Ma20; MK19; Ni21; Su20; WS20]
FPGAs	[Ba19; Da20; DMG21; HLX21; KAK18; Ko17; Ku17; Ku20a; TLW19; ZGF20]
Mobile Devices	[Ri21; Si21]
IoT	[AS20; De21; Hü17; Ne19]
Network Performance	
Latency increased but acceptable	[Ad19; Ka18a; Kw19; PST19; SKD20a; SKD20b; SM16]
Hash-based crypto on embedded devices	[Bü20]
VPN less impacted by PQC	[KS19]
Experiments with various algorithms	[He19a; KV19; PST19]
Lattice for blockchain networks	[AK18; Ca19a]
Cryptanalysis	
Broken NIST candidates	[Di20a; DLW19; Sa19]
Code-based PRNG	[Ku20b]
Attack mitigations	[Cr20; JS19; PS13]
Avoiding synchronization failures	[GFL19; Mc16]
Formal assessments	[Al20a; BS20a; GKS20; GMP21; KS20]
Lower bounds in lattice and code PQC	[KL21]
Upper bounds for multivariate PQC	[ØSV21]
Key recovery	[SV20; TPD21]
Adaptive attack	[FHR21]
Hybrid attacks on NTRU	[Ng21]
Side-Channel Attacks	
Attacks and their countermeasures	[Ar21; BBK16; Ch21a; DA20; Kh16; Kh18; Zh20a; Zh20b]
Successful attacks	[Am20; An19; AR21; Ay18; Di20b; Gr16; GSE20; Ka18b; KPP20; Pa18; Pr21]
Potential threads	[Al20b; KL19; KP21; Ta21; Te21]

Downgrade Attacks	
Possible attack vectors	[LSH20]
Missing protocol awareness	[CPS19]

Algorithm Migration	
Large scale experiments and benchmarks	[Br16; KV19; La16]
Protocol Challenges	[CPS19; ZWH21]
Hybrid and Combiners	[Bi19a; Bi19b; FLP14; OP20; We20]
IKEv2 feasible for PQC	[He19b; Zi15]
Lattice-based PQC for OPC UA	[PS20]

System Migration	
Crypto inventory and automated discovery	[BPS20; ET20; KNW18]
Stages in the process of the PQC migration	[ET20; Qu21]

References

- [Ab21] Abubakr, A.; Kamyar, M.; Viet Ba, D.; Jens-Peter, K.; Kris, G.: A Lightweight Implementation of Saber Resistant Against Side-Channel Attacks, Conference Name: Third PQC Standardization Conference, 2021, URL: <https://csrc.nist.gov/CSRC/media/Events/third-pqc-standardization-conference/documents/accepted-papers/abdulgadir-lightweight-implementation-gmu-pqc2021.pdf>, visited on: 07/23/2021.
- [Ad19] Adam Langley: Real-world measurements of structured-lattices and supersingular isogenies in TLS, Oct. 2019, URL: <https://www.imperialviolet.org/2019/10/30/pqsivssl.html>, visited on: 09/17/2020.
- [AK18] An, H.; Kim, K.: QChain: Quantum-resistant and Decentralized PKI using Blockchain, 2018, URL: https://caislab.kaist.ac.kr/publication/paper_files/2018/SCIS'18_HC_BC.pdf, visited on: 07/14/2021.
- [Al19] Alagic, G.; Alperin-Sheriff, J.; Apon, D.; Cooper, D.; Dang, Q.; Liu, Y.-K.; Miller, C.; Moody, D.; Peralta, R.; Perlner, R.; Robinson, A.; Smith-Tone, D.: Status report on the first round of the NIST post-quantum cryptography standardization process, tech. rep. NIST IR 8240, Gaithersburg, MD: National Institute of Standards and Technology, Jan. 2019, NIST IR 8240, URL: <https://nvlpubs.nist.gov/nistpubs/ir/2019/NIST.IR.8240.pdf>, visited on: 02/24/2021.
- [Al20a] Alagic, G.; Jeffery, S.; Ozols, M.; Poremba, A.: On Quantum Chosen-Ciphertext Attacks and Learning with Errors. en, *Cryptography* 4/1, p. 10, Mar. 2020, ISSN: 2410-387X, URL: <https://www.mdpi.com/2410-387X/4/1/10>, visited on: 07/13/2021.
- [Al20b] Alkim, E.: NewHope - Algorithm Specifications and Supporting Documentation, Apr. 2020, URL: https://newhopecrypto.org/data/NewHope_2020_04_10.pdf, visited on: 07/19/2021.
- [Al20c] Alkim, E.; Barreto, P. S. L. M.; Bindel, N.; Krämer, J.; Longa, P.; Ricardini, J. E.: The Lattice-Based Digital Signature Scheme qTESLA. In (Conti, M.; Zhou, J.; Casalicchio, E.; Spognardi, A., eds.): *Applied Cryptography and Network Security*. Vol. 12146, Series Title: Lecture Notes in Computer Science, Springer International Publishing, Cham, pp. 441–460, 2020, ISBN: 978-3-030-57807-7 978-3-030-57808-4, URL: http://link.springer.com/10.1007/978-3-030-57808-4_22, visited on: 03/18/2021.
- [Al21a] Alkim, E.; Cheng, D. Y.-L.; Chung, C.-M. M.; Evkan, H.; Huang, L. W.-L.; Hwang, V.; Li, C.-L. T.; Niederhagen, R.; Shih, C.-J.; Wälde, J.; Yang, B.-Y.: Polynomial Multiplication in NTRU Prime: en, *IACR Transactions on Cryptographic Hardware and Embedded Systems*, pp. 217–238, 2021, ISSN: 2569-2925, URL: <https://tches.iacr.org/index.php/TCHES/article/view/8733>, visited on: 03/18/2021.

- [Al21b] Alkim, E.; W. Bos, J.; Ducas, L.; Longa, P.; Mironov, I.; Naehrig, M.; Nikolaenko, V.; Peikert, C.; Raghunathan, A.; Stebila, D.: FrodoKEM - Learning With Errors Key Encapsulation, June 2021, URL: <https://frodokem.org/files/FrodoKEM-specification-20210604.pdf>, visited on: 07/14/2021.
- [Am20] Amiet, D.; Curiger, A.; Leuenberger, L.; Zbinden, P.: Defeating NewHope with a Single Trace. IACR Cryptol. ePrint Arch. 2020/, p. 368, 2020.
- [An19] Andrikos, C.; Batina, L.; Chmielewski, L.; Lerman, L.; Mavroudis, V.; Papagiannopoulos, K.; Perin, G.; Rassias, G.; Sonnino, A.: Location, Location, Location: Revisiting Modeling and Exploitation for Location-Based Side Channel Leakages. In (Galbraith, S. D.; Moriai, S., eds.): Advances in Cryptology – ASIACRYPT 2019. Vol. 11923, Series Title: Lecture Notes in Computer Science, Springer International Publishing, Cham, pp. 285–314, 2019, ISBN: 978-3-030-34617-1 978-3-030-34618-8, URL: http://link.springer.com/10.1007/978-3-030-34618-8_10, visited on: 03/18/2021.
- [Ar19] Armknecht, F.; Verbauwhede, I.; Volkamer, M.; Yung, M.: Biggest Failures in Security. en/, p. 23, 2019.
- [Ar20] Aragon, N.: BIKE: Bit Flipping Key Encapsulation Round 3 Submission./, Oct. 2020, URL: https://bikesuite.org/files/v4.1/BIKE_Spec.2020.10.22.1.pdf.
- [AR21] Askeland, A.; Rønjom, S.: A Side-Channel Assisted Attack on NTRU, Published: Cryptology ePrint Archive, Report 2021/790, 2021.
- [Ar21] Aranha, D. F.; Berndt, S.; Eisenbarth, T.; Seker, O.; Takahashi, A.; Wilke, L.; Zaverucha, G.: Side-Channel Protections for Picnic Signatures, Published: Cryptology ePrint Archive, Report 2021/735, 2021.
- [AS20] An, S.; Seo, S. C.: Efficient Parallel Implementations of LWE-Based Post-Quantum Cryptosystems on Graphics Processing Units. en, Mathematics 8/10, Number: 10 Publisher: Multidisciplinary Digital Publishing Institute, p. 1781, Oct. 2020, URL: <https://www.mdpi.com/2227-7390/8/10/1781>, visited on: 07/13/2021.
- [Au20] Aumasson, J.-P.: SPHINCS+ Submission to the NIST post-quantum project, v.3, Oct. 2020, URL: <https://sphincs.org/data/sphincs+-round3-specification.pdf>, visited on: 07/19/2021.
- [Av21] Avanzi, R.: CRYSTALS-Kyber Algorithm Specifications And Supporting Documentation(version 3.01)./, Jan. 2021, URL: <https://pq-crystals.org/kyber/data/kyber-specification-round3-20210131.pdf>.
- [Ay18] Aysu, A.; Tobah, Y.; Tiwari, M.; Gerstlauer, A.; Orshansky, M.: Horizontal side-channel vulnerabilities of post-quantum key exchange protocols. In: 2018 IEEE International Symposium on Hardware Oriented Security and Trust (HOST). Pp. 81–88, Apr. 2018.

- [Az21] Azarderakhsh, R.; Khatib, R. E.; Koziel, B.; Langenberg, B.: Hardware Deployment of Hybrid PQC, Published: Cryptology ePrint Archive, Report 2021/541, 2021.
- [Ba19] Basu, K.; Soni, D.; Nabeel, M.; Karri, R.: NIST Post-Quantum Cryptography - A Hardware Evaluation Study, tech. rep. 047, 2019, URL: <https://eprint.iacr.org/2019/047>, visited on: 04/02/2020.
- [Ba21a] Bai, S.; Ducas, L.; Kiltz, E.; Lepoint, T.; Lyubashevsky, V.; Schwabe, P.; Seiler, G.; Stehlé, D.: CRYSTALS-Dilithium - Algorithm Specifications and Supporting Documentation, Feb. 2021, URL: <https://pq-crystals.org/dilithium/data/dilithium-specification-round3-20210208.pdf>, visited on: 07/14/2021.
- [Ba21b] Banegas, G.; Bernstein, D. J.; Campos, F.; Chou, T.; Lange, T.; Meyer, M.; Smith, B.; Sotáková, J.: CTIDH: faster constant-time CSIDH, Published: Cryptology ePrint Archive, Report 2021/633, 2021.
- [BBK16] Bindel, N.; Buchmann, J.; Krämer, J.: Lattice-Based Signature Schemes and their Sensitivity to Fault Attacks, Published: Cryptology ePrint Archive, Report 2016/415, 2016.
- [Be09] Bernstein, D. J.: Introduction to post-quantum cryptography. In (Bernstein, D. J.; Buchmann, J.; Dahmen, E., eds.): Post-Quantum Cryptography. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1–14, 2009, ISBN: 978-3-540-88701-0 978-3-540-88702-7, URL: http://link.springer.com/10.1007/978-3-540-88702-7_1, visited on: 07/17/2021.
- [Be21] Bert, P.; Eberhart, G.; Prabel, L.; Roux-Langlois, A.; Sabt, M.: Implementation of Lattice Trapdoors on Modules and Applications./, July 9, 2021.
- [Bi19a] Bindel, N.; Braun, J.; Gladiator, L.; Stöckert, T.; Wirth, J.: X.509-Compliant Hybrid Certificates for the Post-Quantum Transition. Journal of Open Source Software 4/40, p. 1606, Aug. 2019, ISSN: 2475-9066, URL: <https://joss.theoj.org/papers/10.21105/joss.01606>, visited on: 05/14/2021.
- [Bi19b] Bindel, N.; Brendel, J.; Fischlin, M.; Goncalves, B.; Stebila, D.: Hybrid Key Encapsulation Mechanisms and Authenticated Key Exchange. In (Ding, J.; Steinwandt, R., eds.): Post-Quantum Cryptography. Vol. 11505, Series Title: Lecture Notes in Computer Science, Springer International Publishing, Cham, pp. 206–226, 2019, ISBN: 978-3-030-25509-1 978-3-030-25510-7, URL: http://link.springer.com/10.1007/978-3-030-25510-7_12, visited on: 03/12/2020.
- [BKS19] Botros, L.; Kannwischer, M. J.; Schwabe, P.: Memory-Efficient High-Speed Implementation of Kyber on Cortex-M4. In (Buchmann, J.; Nitaj, A.; Rachidi, T., eds.): Progress in Cryptology – AFRICACRYPT 2019. Vol. 11627, Series Title: Lecture Notes in Computer Science, Springer International Publishing, Cham, pp. 209–228, 2019, ISBN: 978-3-030-23695-3 978-3-030-23696-0,

URL: http://link.springer.com/10.1007/978-3-030-23696-0_11, visited on: 03/17/2021.

- [Bo15] Bos, J. W.; Costello, C.; Naehrig, M.; Stebila, D.: Post-Quantum Key Exchange for the TLS Protocol from the Ring Learning with Errors Problem. In: 2015 IEEE Symposium on Security and Privacy. ISSN: 2375-1207, pp. 553–570, May 2015.
- [Bo20] Boschini, C.; Camenisch, J.; Ovsiankin, M.; Spooner, N.: Efficient Post-quantum SNARKs for RSIS and RLWE and Their Applications to Privacy. PQCrypto 12100/, pp. 247–267, 2020.
- [BPS20] Barker, W.; Polk, W.; Souppaya, M.: Getting Ready for Post-Quantum Cryptography:: Explore Challenges Associated with Adoption and Use of Post-Quantum Cryptographic Algorithms, en, preprint, May 2020, URL: <https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.05262020-draft.pdf>, visited on: 09/23/2020.
- [Br16] Braithwaite, M.: Experimenting with Post-Quantum Cryptography, en, July 2016, URL: <https://security.googleblog.com/2016/07/experimenting-with-post-quantum.html>, visited on: 04/03/2020.
- [BRP20] Borges, F.; Reis, P. R.; Pereira, D.: A Comparison of Security and its Performance for Key Agreements in Post-Quantum Cryptography. IEEE Access 8/, pp. 142413–142422, 2020.
- [BS20a] Bindel, N.; Schanck, J. M.: Decryption Failure Is More Likely After Success. In: PQCrypto. Pp. 206–225, 2020.
- [BS20b] BSI: Migration zu Post-Quanten-Kryptografie - Handlungsempfehlungen des BSI. de/, p. 9, Aug. 2020.
- [Bü20] Bürstinghaus-Steinbach, K.; Krauß, C.; Niederhagen, R.; Schneider, M.: Post-Quantum TLS on Embedded Systems, tech. rep. 308, 2020, URL: <https://eprint.iacr.org/2020/308>, visited on: 04/03/2020.
- [Ca15] Campagna, M.; Chen, L.; Dagdelen, O.; Ding, J.; Fernick, J.; Gisin, N.; Hayford, D.; Jennewein, T.; Lütkenhaus, N.; Mosca, M.: Quantum Safe Cryptography and Security: An introduction, benefits, enablers and challenges. European Telecommunications Standards Institute ETSI White Paper/8, pp. 1–64, June 2015, URL: <https://www.etsi.org/images/files/ETSIWhitePapers/QuantumSafeWhitepaper.pdf>.
- [Ca19a] Campbell, R.: Evaluation of Post-Quantum Distributed Ledger Cryptography. The Journal of the British Blockchain Association 2/1, pp. 1–8, May 2019, ISSN: 25163949, 25163957, URL: <https://jbba.scholasticahq.com/article/7679-evaluation-of-post-quantum-distributed-ledger-cryptography>, visited on: 07/13/2021.

- [Ca19b] Casanova, A.: GeMSS: A Great Multivariate Short Signature. NIST Round 2/, 2019, URL: https://www-polysys.lip6.fr/Links/NIST/GeMSS_specification.pdf.
- [CC20] Crockett, E.; Campagna, M.: Internet-Draft: Hybrid Post-Quantum Key Encapsulation Methods (PQ KEM) for Transport Layer Security 1.2 (TLS), Mar. 2020, URL: <https://datatracker.ietf.org/doc/html/draft-campagna-tls-bike-sike-hybrid-03>.
- [CC21] Chen, M.-S.; Chou, T.: Classic McEliece on the ARM Cortex-M4, Published: Cryptology ePrint Archive, Report 2021/492, 2021.
- [CD20] Castryck, W.; Decru, T.: CSIDH on the surface. In: International Conference on Post-Quantum Cryptography (PQCrypto 2020). Vol. 12100, Springer, pp. 111–129, 2020.
- [Ch16] Chen, L.; Jordan, S.; Liu, Y.-K.; Moody, D.; Peralta, R.; Perlner, R.; Smith-Tone, D.: Report on Post-Quantum Cryptography, tech. rep. NIST IR 8105, National Institute of Standards and Technology, Apr. 2016, NIST IR 8105, URL: <https://nvlpubs.nist.gov/nistpubs/ir/2016/NIST.IR.8105.pdf>, visited on: 02/24/2021.
- [Ch17] Chase, M.; Derler, D.; Goldfeder, S.; Orlandi, C.; Ramacher, S.; Rechberger, C.; Slamanig, D.; Zaverucha, G.: Post-Quantum Zero-Knowledge and Signatures from Symmetric-Key Primitives, Published: Cryptology ePrint Archive, Report 2017/279, 2017.
- [Ch19] Chen, C.; Danba, O.; Hostein, J.; Hülsing, A.; Rijneveld, J.; M. Schanck, J.; Schwabe, P.; Whyte, W.; Zhang, Z.: NTRU - Algorithm Specifications And Supporting Documentation, Mar. 2019, URL: <https://ntru.org/f/ntru-20190330.pdf>, visited on: 07/14/2021.
- [Ch20] Chou, T.; Cid, C.; UiB, S.; Gilcher, J.; Lange, T.; Maram, V.; Misoczki, R.; Niederhagen, R.; Paterson, K. G.; Persichetti, E., et al.: Classic McEliece: conservative code-based cryptography./, Oct. 2020.
- [Ch21a] Chowdhury, S.; Covic, A.; Acharya, R. Y.; Dupee, S.; Ganji, F.; Forte, D.: Physical security in the post-quantum era: A survey on side-channel analysis, random number generators, and physically unclonable functions. en, *Journal of Cryptographic Engineering*/, Feb. 2021, ISSN: 2190-8508, 2190-8516, URL: <http://link.springer.com/10.1007/s13389-021-00255-w>, visited on: 03/18/2021.
- [Ch21b] Chung, C.-M. M.; Hwang, V.; Kannwischer, M. J.; Seiler, G.; Shih, C.-J.; Yang, B.-Y.: NTT Multiplication for NTT-unfriendly Rings: en, *IACR Transactions on Cryptographic Hardware and Embedded Systems*/, pp. 159–188, Feb. 2021, ISSN: 2569-2925, URL: <https://tches.iacr.org/index.php/TCHES/article/view/8791>, visited on: 03/18/2021.

- [CKY21] Chou, T.; Kannwischer, M. J.; Yang, B.-Y.: Rainbow on Cortex-M4, Published: Cryptology ePrint Archive, Report 2021/532, 2021.
- [Co20] Costello, C.; De Feo, L.; Jao, D.; Longa, P.; Naehrig, M.; Renes, J.: Supersingular Isogeny Key Encapsulation, Oct. 2020, URL: <https://sike.org/files/SIDH-spec.pdf>, visited on: 07/14/2021.
- [Co21] Costello, C.: The Case for SIKE: A Decade of the Supersingular Isogeny Problem, Published: Cryptology ePrint Archive, Report 2021/543, 2021.
- [CPS19] Crockett, E.; Paquin, C.; Stebila, D.: Prototyping post-quantum and hybrid key exchange and authentication in TLS and SSH. In: NIST, p. 24, 2019.
- [Cr20] Cremers, C.; Düzl , S.; Fiedler, R.; Fischlin, M.; Janson, C.: BUFFing signature schemes beyond unforgeability and the case of post-quantum signatures, Published: Cryptology ePrint Archive, Report 2020/1525, 2020.
- [CS21] Cho, J. Y.; Sergeev, A.: Post-quantum MACsec in Ethernet Networks. Journal of Cyber Security and Mobility/, Mar. 2021, ISSN: 2245-4578, 2245-1439, URL: <https://journals.riverpublishers.com/index.php/JCSANDM/article/view/5973>, visited on: 07/13/2021.
- [DA19] D’Anvers, J.-P.: SABER: Mod-LWR based KEM (Round 3 Submission). NIST/, 2019, URL: <https://www.esat.kuleuven.be/cosic/pqcrypto/saber/files/saberspecround3.pdf>.
- [DA20] D’Anvers, J.-P.: A Side-channel Resistant Implementation of SABER. IACR Cryptol. ePrint Arch 733/, 2020, URL: <https://eprint.iacr.org/2020/733.pdf>.
- [Da20] Dang, V. B.; Farahmand, F.; Andrzejczak, M.; Mohajerani, K.; Nguyen, D. T.; Gaj, K.: Implementation and Benchmarking of Round 2 Candidates in the NIST Post-Quantum Cryptography Standardization Process Using Hardware and Software/Hardware Co-design Approaches. Published: Cryptology ePrint Archive, Report 2020/795, 2020.
- [De21] Derek, A.: Requirements for Post-Quantum Cryptography on Embedded Devices in the IoT, Conference Name: Third PQC Standardization Conference, 2021, URL: <https://csrc.nist.gov/CSRC/media/Events/third-pqc-standardization-conference/documents/accepted-papers/atkins-requirements-pqc-iot-pqc2021.pdf>, visited on: 07/23/2021.
- [Di20a] Ding, J.; Deaton, J.; Schmidt, K.; Vishakha; Zhang, Z.: Cryptanalysis of the Lifted Unbalanced Oil Vinegar Signature Scheme. In (Micciancio, D.; Ristenpart, T., eds.): Advances in Cryptology – CRYPTO 2020. Springer International Publishing, Cham, pp. 279–298, 2020, ISBN: 978-3-030-56877-1.

- [Di20b] Ding, J.; Zhang, Z.; Deaton, J.; Wang, L.-C.: A Complete Cryptanalysis of the Post-Quantum Multivariate Signature Scheme Himq-3. In (Meng, W.; Gollmann, D.; Jensen, C. D.; Zhou, J., eds.): Information and Communications Security. Springer International Publishing, Cham, pp. 422–440, 2020, ISBN: 978-3-030-61078-4.
- [DK21] Duc Tri, N.; Kris, G.: Optimized Software Implementations of CRYSTALS-Kyber, NTRU, and Saber Using NEON-Based Special Instructions of ARMv8, Conference Name: Third PQC Standardization Conference, 2021, URL: <https://csrc.nist.gov/CSRC/media/Events/third-pqc-standardization-conference/documents/accepted-papers/hess-fast-quantum-safe-pqc2021.pdf>, visited on: 07/23/2021.
- [DLW19] Dong, X.; Li, Z.; Wang, X.: Quantum cryptanalysis on some generalized Feistel schemes. *en, Science China Information Sciences* 62/2, p. 22501, Feb. 2019, ISSN: 1674-733X, 1869-1919, URL: <http://link.springer.com/10.1007/s11432-017-9436-7>, visited on: 11/16/2020.
- [DMG21] Dang, V. B.; Mohajerani, K.; Gaj, K.: High-Speed Hardware Architectures and Fair FPGA Benchmarking of CRYSTALS-Kyber, NTRU, and Saber, 2021, URL: <https://csrc.nist.gov/CSRC/media/Events/third-pqc-standardization-conference/documents/accepted-papers/gaj-high-speed-hardware-gmu-pqc2021.pdf>, visited on: 07/23/2021.
- [Dr21] Dr. Rachid, El Bansarkhani; Hans-Peter, Fischer; Andreas, Schwab; Dr. Michael, Riecker; Dr. Juliane, Krämer: Definition of a quantum computer resistant encryption scheme, Apr. 2021, URL: <https://www.din.de/en/innovation-and-research/din-spec-en/current-din-specs/wdc-beuth:din21:336773230>, visited on: 07/24/2021.
- [DS05] Ding, J.; Schmidt, D.: Rainbow, a New Multivariable Polynomial Signature Scheme. In (Ioannidis, J.; Keromytis, A.; Yung, M., eds.): Applied Cryptography and Network Security. Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, pp. 164–175, 2005, ISBN: 978-3-540-31542-1.
- [ET20] ETSI: Migration strategies and recommendations to Quantum Safe schemes, July 2020, URL: https://www.etsi.org/deliver/etsi_tr/103600_103699/103619/01.01.01_60/tr_103619v010101p.pdf, visited on: 07/28/2021.
- [Fe20] Fernández-Caramés, T. M.: From Pre-Quantum to Post-Quantum IoT Security: A Survey on Quantum-Resistant Cryptosystems for the Internet of Things. *IEEE Internet of Things Journal* 7/7, pp. 6457–6480, 2020, URL: <https://ieeexplore.ieee.org/abstract/document/8932459>.
- [FF20] Fernández-Caramès, T. M.; Fraga-Lamas, P.: Towards Post-Quantum Blockchain: A Review on Blockchain Cryptography Resistant to Quantum Computing Attacks. *IEEE Access* 8/, pp. 21091–21116, 2020, ISSN: 2169-3536.

- [FHR21] Fauzi, P.; Hovd, M. N.; Raddum, H.: A Practical Adaptive Key Recovery Attack on the LGM (GSW-like) Cryptosystem. In (Cheon, J. H.; Tillich, J.-P., eds.): *Post-Quantum Cryptography*. Springer International Publishing, Cham, pp. 483–498, 2021, ISBN: 978-3-030-81293-5.
- [FI20] Fluhrer, S.; Kampanakis, P.; McGrew, D.; Smysov, V.: *Mixing Preshared Keys in the Internet Key Exchange Protocol Version 2 (IKEv2) for Post-quantum Security*, June 2020, URL: <https://rfc-editor.org/rfc/rfc8784.txt>.
- [FLP14] Fischlin, M.; Lehmann, A.; Pietrzak, K.: Robust Multi-Property Combiners for Hash Functions. *en, Journal of Cryptology* 27/3, pp. 397–428, July 2014, ISSN: 0933-2790, 1432-1378, URL: <http://link.springer.com/10.1007/s00145-013-9148-7>, visited on: 03/12/2020.
- [Fo20] Fouque, P.-A.: *Falcon: Fast-Fourier Lattice-based Compact Signatures over NTRU Specification v1.2.*, Jan. 2020, URL: <https://falcon-sign.info/falcon.pdf>.
- [Fo21] Fouque, P.-A.; Gérard, F.; Rossi, M.; Yu, Y.: *Zalcon: an alternative FPA-free NTRU sampler for Falcon*, 2021, URL: <https://csrc.nist.gov/CSRC/media/Events/third-pqc-standardization-conference/documents/accepted-papers/yang-zalcon-pqc2021.pdf>, visited on: 07/23/2021.
- [Fu18] Fujioka, A.; Takashima, K.; Terada, S.; Yoneyama, K.: *Supersingular Isogeny Diffie-Hellman Authenticated Key Exchange*, Published: *Cryptology ePrint Archive*, Report 2018/730, 2018.
- [Ga18] Gaj, K.: *Challenges and Rewards of Implementing and Benchmarking Post-Quantum Cryptography in Hardware*. In: *Proceedings of the 2018 on Great Lakes Symposium on VLSI*. ACM, Chicago IL USA, pp. 359–364, May 2018, ISBN: 978-1-4503-5724-1, URL: <https://dl.acm.org/doi/10.1145/3194554.3194615>, visited on: 01/17/2021.
- [GAB19] Grote, O.; Ahrens, A.; Benavente-Peces, C.: *A Review of Post-quantum Cryptography and Crypto-agility Strategies*. In: *2019 International Interdisciplinary PhD Workshop (IIPhDW)*. IEEE, Wismar, Germany, pp. 115–120, May 2019, ISBN: 978-1-72810-423-2, URL: <https://ieeexplore.ieee.org/document/8755433/>, visited on: 07/20/2021.
- [GFL19] Gazdag, S.-L.; Friedl, M.; Loebenberger, D.: *Post-Quantum Software Updates*. *en.*, ISBN: 9783885796886 Publisher: Gesellschaft für Informatik e.V., 2019, ISSN: 1617-5468, URL: <http://dl.gi.de/handle/20.500.12116/25014>, visited on: 07/18/2021.
- [GKS20] Gagliardoni, T.; Krämer, J.; Struck, P.: *Quantum indistinguishability for public key encryption*. *arXiv preprint arXiv:2003.00578*, 2020.
- [GMP21] Grubbs, P.; Maram, V.; Paterson, K. G.: *Anonymous, Robust Post-Quantum Public Key Encryption*, Published: *Cryptology ePrint Archive*, Report 2021/708, 2021.

- [Go21] Gonzalez, R.; Hülsing, A.; Kannwischer, M. J.; Krämer, J.; Lange, T.; Stöttinger, M.; Waitz, E.; Wiggers, T.; Yang, B.-Y.: Verifying Post-Quantum Signatures in 8 kB of RAM, Published: Cryptology ePrint Archive, Report 2021/662, 2021.
- [Gr16] Groot Bruinderink, L.; Hülsing, A.; Lange, T.; Yarom, Y.: Flush, Gauss, and Reload – A Cache Attack on the BLISS Lattice-Based Signature Scheme. In (Gierlichs, B.; Poschmann, A. Y., eds.): Cryptographic Hardware and Embedded Systems – CHES 2016. Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, pp. 323–345, 2016, ISBN: 978-3-662-53140-2.
- [Gr96] Grover, L. K.: A fast quantum mechanical algorithm for database search. In: Proceedings of the twenty-eighth annual ACM symposium on Theory of computing - STOC '96. ACM Press, Philadelphia, Pennsylvania, United States, pp. 212–219, 1996, ISBN: 978-0-89791-785-8, URL: <http://portal.acm.org/citation.cfm?doid=237814.237866>, visited on: 07/16/2021.
- [GSE20] Gellersen, T.; Seker, O.; Eisenbarth, T.: Differential Power Analysis of the Picnic Signature Scheme. IACR Cryptol. ePrint Arch. 2020/, p. 267, 2020.
- [He19a] Heesch, M. v.; Adrichem, N. v.; Attema, T.; Veugen, T.: Towards Quantum-Safe VPNs and Internet, Published: Cryptology ePrint Archive, Report 2019/1277, 2019.
- [He19b] Heider, T.: Towards a Verifiably Secure Quantum-Resistant Key Exchange in IKEv2, PhD thesis, Ludwig-Maximilians-Universität München, Oct. 2019, URL: <https://www.nm.ifi.lmu.de/pub/Diplomarbeiten/heid19/PDF-Version/heid19.pdf>, visited on: 07/14/2021.
- [He21a] Heinz, D.; Kannwischer, M. J.; Land, G.; Schwabe, P.; Sprenkels, D.: First-Order Masked Kyber on ARM Cortex-M4, 2021, URL: <https://csrc.nist.gov/CSRC/media/Events/third-pqc-standardization-conference/documents/accepted-papers/heinz-first-order-pqc2021.pdf>, visited on: 07/23/2021.
- [He21b] Hemmert, D. T.; Lochter, M.; Loebenberger, D.; Margraf, Marian; Reinhardt, S.; Sigl, G.: Quantencomputerresistente Kryptografie: Aktuelle Aktivitäten und Fragestellungen. In: Tagungsband zum 17. Deutschen IT-Sicherheitskongress. SecuMedia Verlag, Ingelheim, Germany, German Federal Office for Information Security (BSI), pp. 367–380, Feb. 2021.
- [HKP21] Hutchinson, A.; Karabina, K.; Pereira, G.: Memory Optimization Techniques for Computing Discrete Logarithms in Compressed SIKE. IACR Cryptol. ePrint Arch. 2021/, p. 368, 2021.
- [HKW21] Hagemeyer, D. H.; Kousidis, D. S.; Wunderer, D. T.: Standardisierung von Post-Quanten-Kryptografie und Empfehlungen des BSI. In: Tagungsband zum 17. Deutschen IT-Sicherheitskongress. SecuMedia Verlag, Ingelheim, Germany, German Federal Office for Information Security (BSI), pp. 382–294, Feb. 2021, ISBN: 978-3-922746-83-6.

- [HLX21] He, P.; Lee, C.-Y.; Xie, J.: Compact Coprocessor for KEM Saber: Novel Scalable Matrix Originated Processing, 2021, URL: <https://csrc.nist.gov/CSRC/media/Events/third-pqc-standardization-conference/documents/accepted-papers/xie-compact-coprocessor-pqc2021.pdf>, visited on: 07/23/2021.
- [Ho20] Howe, J.; Prest, T.; Ricosset, T.; Rossi, M.: Isochronous Gaussian Sampling: From Inception to Implementation. In: PQCrypto. Pp. 53–71, 2020.
- [Hü17] Hülsing, A.; Rijneveld, J.; Schanck, J.; Schwabe, P.: High-Speed Key Encapsulation from NTRU. In (Fischer, W.; Homma, N., eds.): Cryptographic Hardware and Embedded Systems – CHES 2017. Springer International Publishing, Cham, pp. 232–252, 2017, ISBN: 978-3-319-66787-4.
- [Hu18] Huelsing, A.; Butin, D.; Gazdag, S.-L.; Rijneveld, J.; Mohaisen, A.: XMSS: eXtended Merkle Signature Scheme, Issue: 8391 Num Pages: 74 Series: Request for Comments Published: RFC 8391, May 2018, URL: <https://rfc-editor.org/rfc/rfc8391.txt>.
- [Hü20] Hülsing, A.; Ning, K.-C.; Schwabe, P.; Weber, F.; Zimmermann, P.R.: Post-quantum WireGuard, Published: Cryptology ePrint Archive, Report 2020/379, 2020.
- [Jo21] Jonathan Bradbury, B. H.: Fast Quantum-Safe Cryptography on IBM Z, Conference Name: Third PQC Standardization Conference, 2021, URL: <https://csrc.nist.gov/CSRC/media/Events/third-pqc-standardization-conference/documents/accepted-papers/hess-fast-quantum-safe-pqc2021.pdf>, visited on: 07/23/2021.
- [JS19] Jaques, S.; Schanck, J. M.: Quantum Cryptanalysis in the RAM Model: Claw-Finding Attacks on SIKE. In (Boldyreva, A.; Micciancio, D., eds.): Advances in Cryptology – CRYPTO 2019. Vol. 11692, Series Title: Lecture Notes in Computer Science, Springer International Publishing, Cham, pp. 32–61, 2019, ISBN: 978-3-030-26947-0 978-3-030-26948-7, URL: http://link.springer.com/10.1007/978-3-030-26948-7_2, visited on: 11/16/2020.
- [Ju21] Julien, D.; Kathrin, H.; Eike, K.; Vadim, L.; Gregor, S.: Faster Kyber and Saber via a Generic Fujisaki-Okamoto Transform for Multi-User Security in the QROM, Conference Name: Third PQC Standardization Conference, 2021, URL: <https://csrc.nist.gov/CSRC/media/Events/third-pqc-standardization-conference/documents/accepted-papers/duman-faster-kyber-pqc2021.pdf>, visited on: 07/23/2021.
- [Ka18a] Kampanakis, P.; Panburana, P.; Daw, E.; Geest, D. V.: The Viability of Post-quantum X.509 Certificates, tech. rep. 063, 2018, URL: <http://eprint.iacr.org/2018/063>, visited on: 03/10/2020.

- [Ka18b] Kannwischer, M. J.; Genêt, A.; Butin, D.; Krämer, J.; Buchmann, J.: Differential Power Analysis of XMSS and SPHINCS. In (Fan, J.; Gierlichs, B., eds.): Constructive Side-Channel Analysis and Secure Design. Vol. 10815, Series Title: Lecture Notes in Computer Science, Springer International Publishing, Cham, pp. 168–188, 2018, ISBN: 978-3-319-89640-3 978-3-319-89641-0, URL: http://link.springer.com/10.1007/978-3-319-89641-0_10, visited on: 08/26/2020.
- [Ka19] Kannwischer, M. J.; Rijneveld, J.; Schwabe, P.; Stoffelen, K.: pqm4: Testing and Benchmarking NIST PQC on ARM Cortex-M4. Published: Cryptology ePrint Archive, Report 2019/844, 2019.
- [Ka20] Kampanakis, P.; Steblia, D.; Friedl, M.; Hansen, T.; Sikeridis, D.: Post-quantum public key algorithms for the Secure Shell (SSH) protocol, Internet-Draft draft-kampanakis-curdle-pq-ssh-00, Backup Publisher: Internet Engineering Task Force Num Pages: 13, Internet Engineering Task Force, Oct. 2020, URL: <https://datatracker.ietf.org/doc/html/draft-kampanakis-curdle-pq-ssh-00>.
- [KAK18] Koziel, B.; Azarderakhsh, R.; Kermani, M. M.: A High-Performance and Scalable Hardware Architecture for Isogeny-Based Cryptography. *IEEE Transactions on Computers* 67/11, pp. 1594–1609, Nov. 2018, ISSN: 1557-9956.
- [Kh16] Khalid, A.; Howe, J.; Rafferty, C.; O’Neill, M.: Time-independent discrete Gaussian sampling for post-quantum cryptography. In: 2016 International Conference on Field-Programmable Technology (FPT). IEEE, Xi’an, China, pp. 241–244, Dec. 2016, ISBN: 978-1-5090-5602-6, URL: <http://ieeexplore.ieee.org/document/7929543/>, visited on: 07/20/2021.
- [Kh18] Khalid, A.; Oder, T.; Valencia, F.; O’Neill, M.; Güneysu, T.; Regazzoni, F.: Physical Protection of Lattice-Based Cryptography: Challenges and Solutions. In: Proceedings of the 2018 on Great Lakes Symposium on VLSI. ACM, Chicago IL USA, pp. 365–370, May 2018, ISBN: 978-1-4503-5724-1, URL: <https://dl.acm.org/doi/10.1145/3194554.3194616>, visited on: 01/17/2021.
- [KK18] Kiefer, F.; Kwiatkowski, K.: Hybrid ECDHE-SIDH Key Exchange for TLS, Internet-Draft draft-kiefer-tls-ecdhe-sidh-00, Backup Publisher: Internet Engineering Task Force Num Pages: 13, Internet Engineering Task Force, Nov. 2018, URL: <https://datatracker.ietf.org/doc/html/draft-kiefer-tls-ecdhe-sidh-00>.
- [KKP20] Koteswara, S.; Kumar, M.; Pattnaik, P.: Performance Optimization of Lattice Post-Quantum Cryptographic Algorithms on Many-Core Processors. In: 2020 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS). Pp. 223–225, Aug. 2020.

- [KL19] Krämer, J.; Loiero, M.: Fault Attacks on UOV and Rainbow. In (Polian, I.; Stöttinger, M., eds.): *Constructive Side-Channel Analysis and Secure Design*. Vol. 11421, Series Title: *Lecture Notes in Computer Science*, Springer International Publishing, Cham, pp. 193–214, 2019, ISBN: 978-3-030-16349-5 978-3-030-16350-1, URL: http://link.springer.com/10.1007/978-3-030-16350-1_11, visited on: 11/20/2020.
- [KL21] Kirshanova, E.; Laarhoven, T.: Lower bounds on lattice sieving and information set decoding, Published: *Cryptology ePrint Archive*, Report 2021/785, 2021.
- [KNW18] Kreutzer, M.; Niederhagen, R.; Waidner, M.: Eberbacher Gespräch: Next Generation Crypto, de, Jan. 2018, URL: <https://www.sit.fraunhofer.de/de/eberbach-crypto/>, visited on: 07/17/2021.
- [Ko17] Koziel, B.; Azarderakhsh, R.; Mozaffari Kermani, M.; Jao, D.: Post-Quantum Cryptography on FPGA Based on Isogenies on Elliptic Curves. *IEEE Transactions on Circuits and Systems I: Regular Papers* 64/1, pp. 86–99, Jan. 2017, ISSN: 1558-0806.
- [KP21] Kutas, P.; Petit, C.: Torsion point attacks on “SIDH-like” cryptosystems, Conference Name: *Third PQC Standardization Conference Place: Birmingham, Bruxelles*, 2021, URL: <https://csrc.nist.gov/CSRC/media/Events/third-pqc-standardization-conference/documents/accepted-papers/kutas-torsion-point-pqc2021.pdf>, visited on: 07/23/2021.
- [KPP20] Kannwischer, M. J.; Pessl, P.; Primas, R.: Single-Trace Attacks on Keccak. *IACR Transactions on Cryptographic Hardware and Embedded Systems*/, pp. 243–268, June 2020, ISSN: 2569-2925, URL: <https://tches.iacr.org/index.php/TCHES/article/view/8590>, visited on: 02/10/2021.
- [KS19] Kampanakis, P.; Sikeridis, D.: Two PQ Signature Use-cases: Non-issues, challenges and potential solutions, tech. rep. 1276, 2019, URL: <https://eprint.iacr.org/2019/1276>, visited on: 04/02/2020.
- [KS20] Krämer, J.; Struck, P.: Encryption Schemes Using Random Oracles: From Classical to Post-Quantum Security. In: *PQCrypto*. Pp. 539–558, 2020.
- [Ku17] Kuo, P.-C.; Li, W.-D.; Chen, Y.-W.; Hsu, Y.-C.; Peng, B.-Y.; Cheng, C.-M.; Yang, B.-Y.: High Performance Post-Quantum Key Exchange on FPGAs, tech. rep. 690, 2017, URL: <https://eprint.iacr.org/2017/690>, visited on: 04/03/2020.
- [Ku20a] Kumar, V. B. Y.; Gupta, N.; Chattopadhyay, A.; Kasper, M.; Krauß, C.; Niederhagen, R.: Post-Quantum Secure Boot. In: *2020 Design, Automation Test in Europe Conference Exhibition (DATE)*. Pp. 1582–1585, 2020.

- [Ku20b] Kuznetsov, A.; Kiian, A.; Smirnov, O.; Cherep, A.; Kanabekova, M.; Chepurko, I.: Testing of Code-Based Pseudorandom Number Generators for Post-Quantum Application. In: 2020 IEEE 11th International Conference on Dependable Systems, Services and Technologies (DESSERT). Pp. 172–177, 2020.
- [KV19] Kwiatkowski, K.; Valenta, L.: The TLS Post-Quantum Experiment, en, Oct. 2019, URL: <https://blog.cloudflare.com/the-tls-post-quantum-experiment/>, visited on: 04/06/2020.
- [Kw19] Kwiatkowski, K.; Sullivan, N.; Langley, A.; Levin, D.; Mislove, A.: Measuring TLS key exchange with post-quantum KEM. In: Workshop Record of the Second PQC Standardization Conference. <https://csrc.nist.gov/CSRC/media/Events/Second-PQC-Standardization-Conference/documents/accepted-papers/kwiatkowski-measuring-tls.pdf>. 2019.
- [La16] Langley, A.: ImperialViolet - CECpq1 results, Nov. 2016, URL: <https://www.imperialviolet.org/2016/11/28/cecpq1.html>, visited on: 06/26/2020.
- [LSH20] Lee, S.; Shin, Y.; Hur, J.: Return of version downgrade attack in the era of TLS 1.3. In: Proceedings of the 16th International Conference on emerging Networking EXperiments and Technologies. ACM, Barcelona Spain, pp. 157–168, Nov. 2020, ISBN: 978-1-4503-7948-9, URL: <https://dl.acm.org/doi/10.1145/3386367.3431310>, visited on: 02/23/2021.
- [Ma20] Mayes, K.: Performance Evaluation and Optimisation for Kyber on the MOLTOS IoT Trust-Anchor. In: 2020 IEEE International Conference on Smart Internet of Things (SmartIoT). Pp. 1–8, Aug. 2020.
- [MAY21] Mathilde, R.; Aymeric, G.; Yolán, R.: PQ-WireGuard: we did it again, Conference Name: Third PQC Standardization Conference, 2021, URL: <https://csrc.nist.gov/CSRC/media/Events/third-pqc-standardization-conference/documents/accepted-papers/raynal-pq-wireguard-pqc2021.pdf>, visited on: 07/23/2021.
- [Mc16] McGrew, D.; Kampanakis, P.; Fluhrer, S.; Gazdag, S.-L.; Butin, D.; Buchmann, J.: State Management for Hash-Based Signatures. In (Chen, L.; McGrew, D.; Mitchell, C., eds.): Security Standardisation Research. Springer International Publishing, Cham, pp. 244–260, 2016, ISBN: 978-3-319-49100-4.
- [MCF19] McGrew, D.; Curcio, M.; Fluhrer, S.: Leighton-Micali Hash-Based Signatures, Issue: 8554 Num Pages: 61 Series: Request for Comments Published: RFC 8554, Apr. 2019, URL: <https://rfc-editor.org/rfc/rfc8554.txt>.

- [Mi21] Michael, B.; Vlad, G.; Basil, H.; Christian, P.; John, S.; Douglas, S.; Goutam, T.: Updates from the Open Quantum Safe Project, Conference Name: Third PQC Standardization Conference, Apr. 2021, URL: <https://csrc.nist.gov/CSRC/media/Events/third-pqc-standardization-conference/documents/accepted-papers/schanck-open-quantum-safe-project-pqc2021.pdf>, visited on: 07/23/2021.
- [MK19] Marzougui, S.; Krämer, J.: Post-Quantum Cryptography in Embedded Systems. In: Proceedings of the 14th International Conference on Availability, Reliability and Security. ACM, Canterbury CA United Kingdom, pp. 1–7, Aug. 2019, ISBN: 978-1-4503-7164-3, URL: <https://dl.acm.org/doi/10.1145/3339252.3341475>, visited on: 07/18/2021.
- [Mo15] Mosca, M.: Cybersecurity in an era with quantum computers: will we be ready?, Published: Cryptology ePrint Archive, Report 2015/1075, 2015.
- [Mo20] Moody, D.; Alagic, G.; Apon, D. C.; Cooper, D. A.; Dang, Q. H.; Kelsey, J. M.; Liu, Y.-K.; Müller, C. A.; Peralta, R. C.; Perlner, R. A.; Robinson, A. Y.; Smith-Tone, D. C.; Alperin-Sheriff, J.: Status report on the second round of the NIST post-quantum cryptography standardization process, tech. rep. NIST IR 8309, Gaithersburg, MD: National Institute of Standards and Technology, July 2020, NIST IR 8309, URL: <https://nvlpubs.nist.gov/nistpubs/ir/2020/NIST.IR.8309.pdf>, visited on: 01/17/2021.
- [Mü20] Müller, M.; de Jong, J.; van Heesch, M.; Overeinder, B.; van Rijswijk-Deij, R.: Retrofitting post-quantum cryptography in internet protocols: a case study of DNSSEC. en, ACM SIGCOMM Computer Communication Review 50/4, pp. 49–57, Oct. 2020, ISSN: 0146-4833, URL: <https://dl.acm.org/doi/10.1145/3431832.3431838>, visited on: 01/17/2021.
- [Ne19] Nejatollahi, H.; Dutt, N.; Ray, S.; Regazzoni, F.; Banerjee, I.; Cammarota, R.: Post-Quantum Lattice-Based Cryptography Implementations: A Survey. ACM Computing Surveys 51/6, Jan. 2019, URL: <https://doi.org/10.1145/3292548>, visited on: 04/02/2020.
- [Ng21] Nguyen, P. Q.: Boosting the Hybrid Attack on NTRU: Torus LSH, Permuted HNF and Boxed Sphere, Conference Name: Third PQC Standardization Conference, 2021, URL: <https://csrc.nist.gov/CSRC/media/Events/third-pqc-standardization-conference/documents/accepted-papers/nguyen-boosting-hybridboost-pqc2021.pdf>, visited on: 07/23/2021.
- [NI16] NIST: Submission Requirements and Evaluation Criteria for the Post-Quantum Cryptography Standardization Process, Dec. 2016, URL: <https://csrc.nist.gov/groups/ST/post-quantum-crypto/documents/call-for-proposals-final-dec-2016.pdf>, visited on: 07/19/2021.

- [Ni21] Nina, B.; Sarah, M.; Hanif, R.; Geoff, T.: Suitability of 3rd Round Signature Candidates for Vehicle-to-Vehicle Communication –Extended Abstract, Conference Name: Third PQC Standardization Conference, 2021, URL: <https://csrc.nist.gov/CSRC/media/Events/third-pqc-standardization-conference/documents/accepted-papers/bindel-suitability-abstract-pqc2021.pdf>, visited on: 07/23/2021.
- [NW17] Niederhagen, R.; Waidner, M.: Practical Post-Quantum Cryptography. Fraunhofer SIT, 2017.
- [OP19] Ott, D.; Peikert, C.; participants other workshop, o. w.: Identifying Research Challenges in Post Quantum Cryptography Migration and Cryptographic Agility. arXiv:1909.07353 [cs]/, arXiv: 1909.07353, Sept. 2019, URL: <http://arxiv.org/abs/1909.07353>, visited on: 02/06/2020.
- [OP20] Ounsworth, M.; Pala, M.: Composite Keys and Signatures For Use In Internet PKI, Internet-Draft draft-ounsworth-pq-composite-sigs-03, Backup Publisher: Internet Engineering Task Force Num Pages: 18, Internet Engineering Task Force, July 2020, URL: <https://datatracker.ietf.org/doc/html/draft-ounsworth-pq-composite-sigs-03>.
- [ØSV21] Øygarden, M.; Smith-Tone, D.; Verbel, J.: On the Effect of Projection on Rank Attacks in Multivariate Cryptography, Published: Cryptology ePrint Archive, Report 2021/655, 2021.
- [Pa18] Park, A.; Shim, K.-A.; Koo, N.; Han, D.-G.: Side-Channel Attacks on Post-Quantum Signature Schemes based on Multivariate Quadratic Equations. en, IACR Transactions on Cryptographic Hardware and Embedded Systems/, pp. 500–523, Aug. 2018, ISSN: 2569-2925, URL: <https://tches.iacr.org/index.php/TCHES/article/view/7284>, visited on: 09/25/2020.
- [Pe20] Petzoldt, A.: Efficient key generation for rainbow. In: International Conference on Post-Quantum Cryptography. Springer, pp. 92–107, 2020.
- [Pr21] Prasanna, R.; Martianus Frederic, E.; Shivam, B.; Anupam, C.; Sujoy Sinha, R.: On Generic Side-Channel Assisted Chosen Ciphertext Attacks on Lattice-based PKE/KEMs, Conference Name: Third PQC Standardization Conference, 2021, URL: <https://csrc.nist.gov/CSRC/media/Events/third-pqc-standardization-conference/documents/accepted-papers/ravi-generic-side-channel-pqc2021.pdf>, visited on: 07/23/2021.
- [PS13] Perlner, R.; Smith-Tone, D.: A Classification of Differential Invariants for Multivariate Post-quantum Cryptosystems. In (Hutchison, D.; Kanade, T.; Kittler, J.; Kleinberg, J. M.; Mattern, F.; Mitchell, J. C.; Naor, M.; Nierstrasz, O.; Pandu Rangan, C.; Steffen, B.; Sudan, M.; Terzopoulos, D.; Tygar, D.; Vardi, M. Y.; Weikum, G.; Gaborit, P., eds.): Post-Quantum Cryptography. Vol. 7932, Series Title: Lecture Notes in Computer Science, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 165–173, 2013, ISBN: 978-3-642-38615-2 978-3-642-38616-9,

URL: http://link.springer.com/10.1007/978-3-642-38616-9_11, visited on: 11/16/2020.

- [PS20] Paul, S.; Scheible, P.: Towards Post-Quantum Security for Cyber-Physical Systems: Integrating PQC into Industrial M2M Communication. In (Chen, L.; Li, N.; Liang, K.; Schneider, S., eds.): Computer Security – ESORICS 2020. Vol. 12309, Series Title: Lecture Notes in Computer Science, Springer International Publishing, Cham, pp. 295–316, 2020, ISBN: 978-3-030-59012-3 978-3-030-59013-0, URL: http://link.springer.com/10.1007/978-3-030-59013-0_15, visited on: 10/01/2020.
- [PST19] Paquin, C.; Stebila, D.; Tamvada, G.: Benchmarking Post-Quantum Cryptography in TLS, tech. rep. 1447, 2019, URL: <http://eprint.iacr.org/2019/1447>, visited on: 09/18/2020.
- [Qu21] QuantiCor: Sicherheitsrisiko Quantencomputer, 2021, URL: <https://quanticor-security.de/whitepaper/>, visited on: 07/21/2021.
- [Ri21] Ribeiro, L. A. D. S.; da Silva Lima, J. P.; de Queiroz, R. J. G. B.; Chagas, A. B.; Quintino, J. P.; da Silva, F. Q. B.; Santos, A. L. M.; Ribeiro, J. R. J.: Saber Post-Quantum Key Encapsulation Mechanism (KEM): Evaluating Performance in Mobile Devices and Suggesting Some Improvements, 2021, URL: <https://csrc.nist.gov/CSRC/media/Events/third-pqc-standardization-conference/documents/accepted-papers/ribeiro-saber-pq-key-pqc2021.pdf>, visited on: 07/23/2021.
- [Sa19] Samardjiska, S.; Santini, P.; Persichetti, E.; Banegas, G.: A Reaction Attack Against Cryptosystems Based on LRPC Codes. In (Schwabe, P.; Thériault, N., eds.): Progress in Cryptology – LATINCRYPT 2019. Springer International Publishing, Cham, pp. 197–216, 2019, ISBN: 978-3-030-30530-7, URL: https://link.springer.com/chapter/10.1007/978-3-030-30530-7_10.
- [Sa20] Santoso, B.: Generalization of Isomorphism of Polynomials with Two Secrets and Its Application to Public Key Encryption. In: International Conference on Post-Quantum Cryptography. Springer, pp. 340–359, 2020.
- [SFG20] Stebila, D.; Fluhrer, S.; Gueron, S.: Internet-Draft: Hybrid key exchange in TLS 1.3, Feb. 2020, URL: <https://tools.ietf.org/id/draft-stebila-tls-hybrid-design-03.html>.
- [Sh97] Shor, P. W.: Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer. en, SIAM Journal on Computing 26/5, pp. 1484–1509, Oct. 1997, ISSN: 0097-5397, 1095-7111, URL: <http://epubs.siam.org/doi/10.1137/S0097539795293172>, visited on: 07/16/2021.

- [Si21] da Silva Lima, J. P.; Ribeiro, L. A. D. S.; de Queiroz, R. J. G. B.; Quintino, J. P.; da Silva, F. Q. B.; Santos, A. L. M.; José, R.: Evaluating Kyber post-quantum KEM in a mobile application, 2021, URL: <https://csrc.nist.gov/CSRC/media/Events/third-pqc-standardization-conference/documents/accepted-papers/ribeiro-evaluating-kyber-pqc2021.pdf>, visited on: 07/23/2021.
- [SKD20a] Sikeridis, D.; Kampanakis, P.; Devetsikiotis, M.: Assessing the Overhead of Post-Quantum Cryptography in TLS 1.3 and SSH. In: Proceedings of the 16th International Conference on Emerging Networking Experiments and Technologies. CoNEXT '20, event-place: Barcelona, Spain, Association for Computing Machinery, New York, NY, USA, pp. 149–156, 2020, ISBN: 978-1-4503-7948-9, URL: <https://doi.org/10.1145/3386367.3431305>.
- [SKD20b] Sikeridis, D.; Kampanakis, P.; Devetsikiotis, M.: Post-Quantum Authentication in TLS 1.3: A Performance Study, tech. rep. 071, 2020, URL: <http://eprint.iacr.org/2020/071>, visited on: 03/10/2020.
- [SM16] Stebila, D.; Mosca, M.: Post-quantum Key Exchange for the Internet and the Open Quantum Safe Project. In (Avanzi, R.; Heys, H., eds.): Selected Areas in Cryptography – SAC 2016. Lecture Notes in Computer Science, <https://github.com/open-quantum-safe/liboqs>, Springer International Publishing, Cham, pp. 14–37, 2016, ISBN: 978-3-319-69453-5.
- [Sm21] Smyslov, V.: Intermediate Exchange in the IKEv2 Protocol, Internet-Draft draft-ietf-ipsecme-ikev2-intermediate-06, Backup Publisher: Internet Engineering Task Force Num Pages: 11, Internet Engineering Task Force, Mar. 2021, URL: <https://datatracker.ietf.org/doc/html/draft-ietf-ipsecme-ikev2-intermediate-06>.
- [SS17] Schanck, J. M.; Stebila, D.: A Transport Layer Security (TLS) Extension For Establishing An Additional Shared Secret, Internet-Draft draft-schanck-tls-additional-keyshare-00, Backup Publisher: Internet Engineering Task Force Num Pages: 10, Internet Engineering Task Force, Apr. 2017, URL: <https://datatracker.ietf.org/doc/html/draft-schanck-tls-additional-keyshare-00>.
- [Su20] Suhail, S.; Hussain, R.; Khan, A.; Hong, C. S.: On the Role of Hash-based Signatures in Quantum-Safe Internet of Things: Current Solutions and Future Directions. IEEE Internet of Things Journal 8/1, pp. 1–17, Apr. 2020, URL: <https://ieeexplore.ieee.org/document/9152977>, visited on: 08/12/2020.
- [SV20] Smith-Tone, D.; Verbel, J. A.: A Rank Attack Against Extension Field Cancellation. In: PQCrypto. Pp. 381–401, 2020.

- [SWZ16] Schanck, J. M.; Whyte, W.; Zhang, Z.: Quantum-Safe Hybrid (QSH) Ciphersuite for Transport Layer Security (TLS) version 1.2, Internet-Draft draft-whyte-qsh-tls12-02, Backup Publisher: Internet Engineering Task Force Num Pages: 19, Internet Engineering Task Force, July 2016, URL: <https://datatracker.ietf.org/doc/html/draft-whyte-qsh-tls12-02>.
- [Ta21] Tasso, É.; De Feo, L.; El Mrabet, N.; Pontié, S.: Resistance of Isogeny-Based Cryptographic Implementations to a Fault Attack. In: Constructive Side-Channel Analysis and Secure Design (COSADE) 2021. Lugano, Switzerland, Oct. 2021, URL: <https://hal-cea.archives-ouvertes.fr/cea-03266892>.
- [Te21] Tendayi, K.; Michael, F.; Tristen, T.; Alexander, N.; David, A.; Miaoqing, H.: Power-based Side Channel Attack Analysis on PQC Algorithms, Conference Name: Third PQC Standardization Conference, 2021, URL: <https://csrc.nist.gov/CSRC/media/Events/third-pqc-standardization-conference/documents/accepted-papers/kamucheka-power-based-pqc2021.pdf>, visited on: 07/23/2021.
- [Th21] Thomas, E.; Akira, T.; Mehdi, T.; Alexandre, W.: Mitaka: A Simpler, Parallelizable, Maskable Variant of Falcon, Conference Name: Third PQC Standardization Conference, 2021, URL: <https://csrc.nist.gov/CSRC/media/Events/third-pqc-standardization-conference/documents/accepted-papers/espitau-mitaka-pqc2021.pdf>, visited on: 07/23/2021.
- [Tj19] Tjhai, C.; Tomlinson, M.; grbartle@cisco.com; Fluhrer, S.; Geest, D. V.; Garcia-Morchon, O.; Smyslov, V.: Framework to Integrate Post-quantum Key Exchanges into Internet Key Exchange Protocol Version 2 (IKEv2), Internet-Draft draft-tjhai-ipsecme-hybrid-qske-ikev2-04, Backup Publisher: Internet Engineering Task Force Num Pages: 21, Internet Engineering Task Force, July 2019, URL: <https://datatracker.ietf.org/doc/html/draft-tjhai-ipsecme-hybrid-qske-ikev2-04>.
- [TLW19] Tian, J.; Lin, J.; Wang, Z.: Ultra-Fast Modular Multiplication Implementation for Isogeny-Based Post-Quantum Cryptography. In: 2019 IEEE International Workshop on Signal Processing Systems (SiPS). ISSN: 2374-7390, pp. 97–102, Oct. 2019.
- [TPD21] Tao, C.; Petzoldt, A.; Ding, J.: Efficient Key Recovery for all HFE Signature Variants. In: Springer-Verlag, 2021.
- [We20] Weller, D. L.: Incorporating post-quantum cryptography in a microservice environment. en/, p. 36, Feb. 2020.
- [Wh17] Whyte, W.; Zhang, Z.; Fluhrer, S.; Garcia-Morchon, O.: Internet-Draft: Quantum-Safe Hybrid (QSH) Key Exchange for Transport Layer Security (TLS) version 1.3, Oct. 2017, URL: <https://tools.ietf.org/html/draft-whyte-qsh-tls13-06>.

- [WS20] Wang, W.; Stöttinger, M.: Post-Quantum Secure Architectures for Automotive Hardware Secure Modules, tech. rep. 026, 2020, URL: <https://eprint.iacr.org/2020/026>, visited on: 05/14/2021.
- [YWT20] Yasuda, T.; Wang, Y.; Takagi, T.: Multivariate Encryption Schemes Based on Polynomial Equations over Real Numbers. In: PQCrypto. Pp. 402–421, 2020.
- [Za20] Zaverucha, G.: The Picnic Signature Algorithm Specification. NIST Round 3/, Apr. 2020, URL: <https://github.com/microsoft/Picnic/blob/master/spec/spec-v3.0.pdf>.
- [ZGF20] Zoni, D.; Galimberti, A.; Fornaciari, W.: Efficient and Scalable FPGA-Oriented Design of QC-LDPC Bit-Flipping Decoders for Post-Quantum Cryptography. IEEE Access 8/, pp. 163419–163433, 2020.
- [Zh20a] Zhang, C.; Liu, Z.; Chen, Y.; Lu, J.; Liu, D.: A Flexible and Generic Gaussian Sampler With Power Side-Channel Countermeasures for Quantum-Secure Internet of Things. IEEE Internet of Things Journal 7/9, Conference Name: IEEE Internet of Things Journal, pp. 8167–8177, Sept. 2020, ISSN: 2327-4662.
- [Zh20b] Zhang, F.; Yang, B.; Dong, X.; Guilley, S.; Liu, Z.; He, W.; Zhang, F.; Ren, K.: Side-Channel Analysis and Countermeasure Design on ARM-based Quantum-Resistant SIKE. IEEE Transactions on Computers/, Conference Name: IEEE Transactions on Computers, pp. 1–1, 2020, ISSN: 1557-9956.
- [Zi15] Zimmer, E.: Post-Quantum Kryptographie für IPsec, Feb. 2015, URL: <https://svs.informatik.uni-hamburg.de/publications/2015/2015-02-24-Zimmer-DFN-PQC-fuer-IPsec.pdf>, visited on: 07/14/2021.
- [ZSS20] Zhao, R. K.; Steinfeld, R.; Sakzad, A.: COSAC: COmpact and Scalable Arbitrary-Centered Discrete Gaussian Sampling over Integers. In: PQCrypto. Pp. 284–303, 2020.
- [ZWH21] Zeier, A.; Wiesmaier, A.; Heinemann, A.: Zur Integration von Post-Quantum Verfahren in bestehende Softwareprodukte. de, arXiv:2102.00157 [cs]/, arXiv: 2102.00157, Jan. 2021, URL: <http://arxiv.org/abs/2102.00157>, visited on: 03/15/2021.

Evaluating Error Mitigation Strategies for Entangled Quantum States on Near-Term Quantum Computers

Lennart Schulze,¹ Jan-Rainer Lahmann²

Abstract: Entanglement is one of the quantum mechanical properties to which recently emerging quantum computers attribute an exponential increase in computing power. However, these systems are subject to a set of noise-inducing physical processes and hardware-level imperfections that render the results from quantum circuits erroneous. Bridging the time until sufficient qubits are available to compensate for these effects, quantum error mitigation algorithms aim at improving the result accuracy on near-term quantum devices. This empirical investigation describes and compares customary fundamental approaches to error mitigation for quantum states in condition of entanglement on real quantum computers. It is demonstrated that two readily implementable techniques regarding circuit design and measurement error mitigation may lead to a considerable increase in the quality of results.

Keywords: quantum error mitigation; qubit mapping; transpiler optimisation; measurement error mitigation; quantum entanglement; Greenberger-Horne-Zeilinger state

1 Introduction

In the near future, quantum computers are expected to provide an exponential increase in computational capacity compared to classical systems. To realise this potential, quantum computing systems leverage fundamental quantum mechanical properties. Among these, entanglement is of utmost importance since it allows to introduce additional information to a quantum system of multiple qubits beyond the information carried by each constituent alone, achieving that they perform in a coordinated way [DL98]. Together with the effect of superposition, a large portion of the expected or already realised computational speedups can be attributed to this behaviour [Pr18]. Consequently, the fidelity with which fully entangled multi-qubit-system quantum states are produced, maintained, and measured is an important figure of merit for near-term quantum devices.

These systems, which are also referred to as noisy intermediate-scale quantum computers (NISQ) [Pr18], are subject to physical effects that cause the outputs from algorithms performed on them to deviate from the expectation, excluding the effect of quantum randomness. Specifically, quantum noise and errors introduced with the operations applied over the course of a quantum circuit amount to erroneous results [KLZ98]. Addressing this issue, the theory of quantum error correction prospectively allows to operate quantum

¹ IBM Germany, IBM-Allee 1, 71139 Ehningen, Germany, schulzeleennart@gmail.com

² IBM Germany, IBM-Allee 1, 71139 Ehningen, Germany, jan.lahmann@de.ibm.com

computers in a fault-tolerant manner [Sh96, Go96, Pr98]. However, the number of qubits required to encode logical qubits in multiple physical qubits envisioned by these approaches is beyond the capacity of current devices [Fo12]. Furthermore, the minimum error rates prerequisite to applying such framework have not been reached yet [KLZ98, Pr98, ABO08].

In consequence, research on quantum error mitigation, a set of quantum and classical algorithms feasible to reduce the effect of errors in computations on near-term devices, has received ample interest [TBG17, LB17, Pr18, Ka19, Su21, St20, En21]. The rapidly growing number of algorithms, nonetheless, render it challenging for the ordinary quantum algorithm developer to use its potential to the fullest.

Due to its importance for quantum computing, this study presents a comparison of four different quantum error mitigation techniques benchmarked on the seminal entangled Greenberger-Horne-Zeilinger (GHZ) quantum state [Gr90]. Consequently, it is aimed to characterise the differences in quantum algorithm result quality between ideal simulators and real quantum devices as well as to assess the capabilities of selected low-level techniques to improve the quality of algorithm results on these near-term devices. Comparing the effectiveness of these strategies, which are readily implementable using current software frameworks, a combined approach is concluded to improve the near-term-device execution of circuits, contributing to decoupling quantum algorithm development from the inherent limitations of quantum computing systems. As a result, directions for further research are pointed out.

To arrive at this objective, first, a review is provided of core concepts and recent progresses in the study of quantum noise and errors and quantum error mitigation as well as the characterisation of these effects in quantum computing systems. Second, an experimental research design is constructed introducing four customary error mitigation techniques on the circuit level and from the classical regime to be probed in eight experiments performed on different IBM Quantum computers. Third, the results are presented including their statistical significance and a discussion of the findings. Finally, this study concludes with a summary of the core implications and an outlook to future investigations.

2 Foundations and Related Work

Quantum noise describes the effect of uncontrollable physical interactions that arise from the quantum mechanical nature of quantum computing systems and which impact them in their operation. One of these effects that is referred to as decoherence, specifically, forces a coherent quantum state, which a quantum computer may assume, to collapse due to its interaction with the environment. Such decohered, or mixed, state has lost its quantum mechanical properties to an extent that it becomes futile for most computations. As the interactions stem from a variety of sources, such as unintended interactions between qubits, and accumulate, quantum computations are restricted in their duration. Conversely, longer executions, arising from circuits with more gates, are more prone to errors and, similarly,

more qubits in a system yield less accurate results from algorithms executed on them [KLZ98, NC10].

Furthermore, errors occur in quantum computations due to physical constraints in near-term devices that cause the individual operations performed as part of a quantum algorithm to be imprecise. Specifically, errors may occur in the preparation of the ground state of a qubit typically at the beginning of an algorithm, during application of an operation on a single qubit or between multiple qubits, or during the measurement of the state of a qubit typically at the end of an algorithm [ABO08, Di00].

Acknowledging these fundamental challenges, quantum error correction originated soon after the advent of quantum computing research [Sh96, Go96, Pr98, AG04]. To overcome the influence of noise and errors, conceptually, logical qubits may be encoded in a far higher number of physical qubits in patterns referred to as codes, which would allow to perpetually maintain a coherent quantum state. Fault-tolerant quantum computing describes how to use error-corrected quantum states for accurate quantum computations [ABO08].

In light of still unsatisfied requirements in the number of qubits and their inherent error rates for the application of fault-tolerant quantum computing, quantum error mitigation emerged. Summarising a varied set of quantum and classical routines, error mitigation algorithms commonly address one specific source of noise or error on near-term devices to reduce its influence in the computation, nonetheless, not allowing fully fault-tolerant computing [Pr18, En21, SET20]. Advanced standalone routines were proposed as probabilistic error cancellation [TBG17], dynamical decoupling [VL98, VKL99], zero-noise error extrapolation [LB17, TBG17], and more [EBL18, Su21]. An extensive review is provided in [En21].

In contrast, conceptual quantum error mitigation approaches may be targeted at the quantum circuit design, which corresponds to an algorithm [Qi21], considering that a higher number of operations yields worse results. Similarly, other algorithms address specific error types, such as measurement error mitigation [Ka19]. Thus, these low-level strategies are more targeted than advanced circuit-independent error mitigation routines, while still residing on the quantum circuit abstraction layer, as opposed to hardware-near optimisation [Mc18, AI20]. Due to their ability to provide a significant benefit concerning the individual error types, this study explores available error mitigation techniques to reduce gate and measurement errors.

To characterise quantum devices with respect to their sensitivity to noise and errors, different approaches exist [Wi17, NC10]. Randomised Benchmarking leverages partially randomised quantum circuits to characterise gate-level errors in quantum computers [Kn08, MGE11, MGE12]. Following a similar approach, Quantum Volume was introduced by [Mo18, Cr19] as a holistic metric that uses partially randomised circuits to provide a generalised single-number indication of the capacity of the quantum computer. It considers the number of available qubits of the quantum computing system together with properties impacting the accuracy of results such as the set of available hardware-level operations, the

parallelism of these operations, their associated error rates, the duration of coherence of quantum states, and the connectivity between qubits.

The last is referred to as qubit connectivity graph, or coupling map [Ko20], and determines which two-or-more-qubit gates may be performed directly without introduction of auxiliary gates. These are necessary to perform operations between unconnected qubits via physically connected ones. The coupling map property together with the error rates and decoherence times are combined in the error map representation of a quantum computing system.

Quantum fidelity measures describe quantum states and related entities of quantum information. The most established distance between two quantum states is the quantum state fidelity, introduced by [Hü92] for pure states and amended by [Jo94] for mixed quantum states. [Ni96, Ni02] defined an additional measurement of the maintenance of entangled states, the entanglement fidelity, to describe the distance between states in this condition.

Whereas the aforementioned regard states of quantum systems, a different way to characterise noise is by reviewing the results from a quantum computation. The Hellinger fidelity, introduced by [He09], is a metric that assesses the similarity of two probability distributions. It allows to measure the accuracy of the circuit execution and to deduce the influence of noise and errors in the system by comparing the mathematically provided ideal result distribution from a quantum circuit executed for a certain number of shots in terms of the classical computational basis states that shall be retrieved upon measurement against the distribution actually returned from its execution on a quantum computer. For discrete probability distributions $P = p_1, \dots, p_k$, $Q = q_1, \dots, q_k$, the Hellinger fidelity takes the form [He09]:

$$1 - H(P, Q) = 1 - \frac{1}{\sqrt{2}} \sqrt{\sum_j^k (\sqrt{p_j} - \sqrt{q_j})^2} \quad (1)$$

Hence, it may be used to examine the congruence of two result distributions from quantum algorithm executions, quantifying their similarity in a single number.

3 Methodology

It is aimed to compare the influence of noise and errors across several quantum computing systems and to demonstrate the effect of different error mitigation strategies for entangled quantum states on these devices. Hence, an experimental approach is employed.

3.1 Experimental Setup

The open-source quantum software framework Qiskit [A119] is chosen for implementation and execution of all experiments due to its integrated interface with IBM Quantum computers

as well as its integrated transpiler optimisation and measurement error mitigation (MEM) capabilities, which shall be examined. Alternative software frameworks include Cirq, PyQuil, or Q#. Extensive reference to diverse simulator software programs is provided in [G121, Mi21].

Subject of the experiments is a specific implementation of the entangled three-qubit GHZ state that is motivated by the so-called GHZ game introduced in [Va99]. Capturing the crucial condition of entanglement between multiple qubits, the GHZ state is highly relevant for characterising quantum computers, which depend on this effect, as proven by the plentitude of previous research leveraging it for this purpose [Bo99, We20]. In addition, due to a comparably small number of required qubits and gates, it is implementable on a variety of near-term quantum computers. Its default implementation is depicted in figure 1.

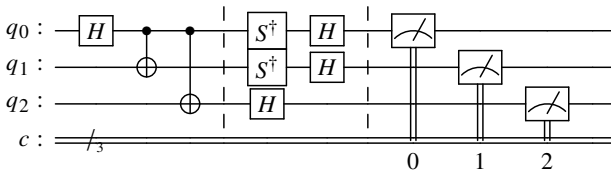


Fig. 1: Quantum circuit producing the three-qubit GHZ-YYX state

The Hadamard (H) gate creates an equal superposition between the computational basis states of qubit q_0 , which is extended to the qubits q_1 and q_2 through application of CNOT gates, resulting in the entangled GHZ state $\frac{1}{\sqrt{2}}(|000\rangle + |111\rangle)$. Subsequently, the measurement basis is changed from the computational basis Z to the X basis on qubit q_2 by adding an H gate and to the Y basis on qubits q_0 and q_1 by adding an S-dagger gate and an H gate. The resulting state

$$|GHZ_{YYX}\rangle = \frac{1}{2}(|001\rangle + |010\rangle + |100\rangle + |111\rangle) = [0 \ 0.5 \ 0.5 \ 0 \ 0.5 \ 0 \ 0 \ 0.5]^T, \quad (2)$$

where T is the transpose and $*$ is the complex conjugate, is expected to return the classical states 001, 010, 100, and 111 each with an equal portion of the total counts, subject to quantum randomness. Thus, the noise- and error-free ideal result probability distribution P_{ideal} over the computational basis state space of a three-qubit system is established as

$$P_{ideal}(X = y) = \begin{cases} 0.25 & \text{if } y = 001, 010, 100, \text{ or } 111 \\ 0 & \text{else.} \end{cases} \quad (3)$$

Based on this reference circuit, different experiments are performed. Each experiment consists of executing an adapted quantum circuit on the Qiskit Aer QASM simulator or an IBM Quantum computer, which are instances of superconducting transmon systems [Ko07]. The selection of the real devices is informed by their open accessibility at the time of execution. This criterion results in a group of six five-qubit systems, IBM Quantum Athens with a QV of 32, Ourense with QV8, Santiago with QV32, Valencia with QV16,

Vigo with QV16, and Yorktown with QV8, as well as the 15-qubit system Melbourne with QV8 in their configuration from November 2020 [IB20].

The algorithm of each experiment is executed with 8000 shots to account for the inherent randomness of quantum computations apart from errors and noise, a higher number returning a more balanced result set. It determines the number of executions of the quantum circuit on the system. Each execution results in one of the $2^n = 8$ classical computational basis states through projective measurement of the quantum state at the end of the circuit. From these measurements, the discrete probability distribution $P_{\text{experiment}}$ is constructed. The Hellinger fidelity is then computed between P_{ideal} and $P_{\text{experiment}}$ as $1 - H(P_{\text{experiment}}, P_{\text{ideal}})$, constituting the metric for the quality of the result from the experiment. A value close to 100% indicates no deviation from the expected result distribution beyond randomness, representing either an absence of noise and errors during execution or full error mitigation.

3.2 Hypotheses

It is conjectured that the effect of noise and errors will render the results obtained from real quantum computing systems deviating from those retrieved from the ideal Qiskit Aer simulator.

1. The quality of results from executing the quantum algorithm on an ideal quantum computing simulator and a real quantum computer differs.

In line with the motivation for this research and the findings discussed in section 2, it is further expected that each of four distinct quantum error mitigation techniques produces an improvement in the result quality compared to that from the reference implementation on the selected quantum computer, respectively. Furthermore, non-exclusive error mitigation techniques are supposed to behave mutually enhancing conditioned on addressing different error sources. From these assumptions, it is expected that the quality of the results from executing the quantum algorithm on a selected quantum computer will be increased

2. by mapping the qubits of the circuit to the physical qubits of the device such that it is optimised for the coupling map of the device,
3. by mapping the qubits of the circuit to the physical qubits of the device such that it is optimised for the coupling map and error rates, that is the error map, of the device,
4. by applying the Qiskit transpiler configured to the highest optimisation level on the circuit prior to execution,
5. by applying quantum measurement error mitigation on the results from the reference circuit,

6. by applying quantum measurement error mitigation on the results from an error map-based optimised circuit, and
7. by applying quantum measurement error mitigation on the results from the reference circuit that was transpiled by the Qiskit transpiler configured to the highest optimisation level.

Finally, all approaches are expected to yield results different from each other.

8. The quality of the results from executing the quantum algorithm on a real quantum computer differs between the different error mitigation techniques employed.

3.3 Error Mitigation Techniques

In the first step, all seven available quantum computing systems are compared in their result quality for the reference circuit and the two with the extreme Hellinger fidelity values are selected for application of the error mitigation techniques. Following from the hypotheses, four different quantum error mitigation techniques are compared on these systems independently and in combination, where applicable, which will be presented in the following. The quantum circuit that serves as reference to which the results from the subsequent experiments are compared implements the GHZ state in YYX readout configuration without any optimisation, as depicted in figure 1.

Coupling-map-adaptive qubit mapping. As demonstrated in section 4, the two systems selected for comparison are the 5-qubit systems Athens and Yorktown with a Quantum Volume of 32 and 8, respectively. They possess the coupling maps presented in figure 2.



Fig. 2: Qubit coupling map of IBM Quantum Athens (left) and Yorktown (right) [IB20]

When a multi-qubit operation is intended to be performed between qubits that do not possess a physical connection according to the coupling map, prior to the execution it is automatically replaced with multiple auxiliary operations that act on physically connected qubits such that the outcome of the original operation is preserved, for instance through

the usage of SWAPs. In Qiskit, this is performed by the Qiskit transpiler. To avoid the introduction of these gates, which amount to additional errors, the circuit can be remapped to qubits that match the coupling map natively. In the case of Athens, owing to the missing connection between qubit q_0 and q_2 , logical qubit q_0 from the reference circuit is mapped to physical qubit q_1 , q_1 to q_0 , and q_2 stays in the original position, as showcased in figure 3. As the coupling map of IBM Quantum Yorktown already matches the two-qubit gates of the logical quantum circuit, no remapping needs to be performed.

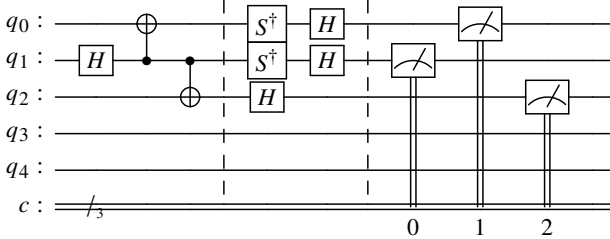


Fig. 3: GHZ-YYX circuit optimised for the coupling map of IBM Quantum Athens [IB20]

Error-map-adaptive qubit mapping. In addition to the coupling map, the error rates of the different operations in the quantum algorithm are considered to determine an optimal qubit mapping. Selecting the qubits with least sensitivity to state preparation errors, single-qubit operation (U) errors, multiple-qubit operation (CNOT) errors, and measurement errors is expected to yield a further increased result quality. Since the different error types compete in the minimisation objective, preference is given to minimising CNOT gate error rates as these account for a larger portion of the total error in the results than single-qubit gates. As second priority, measurement errors are selected.

Qubit	Readout error	Single-qubit U2 error rate	CNOT error rate
Q0	1.1500000000000066e-2	2.3843587783374748e-4	cx0_1: 7.544e-3
Q1	1.3800000000000034e-2	4.013488066819104e-4	cx1_0: 7.544e-3, cx1_2: 7.043e-3
Q2	1.2299999999999978e-2	2.1238602991590765e-4	cx2_1: 7.043e-3, cx2_3: 7.298e-3
Q3	1.5500000000000007e-2	2.0493668082695874e-4	cx3_2: 7.298e-3, cx3_4: 7.374e-3
Q4	1.6699999999999937e-2	2.95350050864255e-4	cx4_3: 7.374e-3

Tab. 1: Error rates of IBM Quantum Athens at time of the experiment [IB20]

Analysing the error rates in table 1 accordingly yields an error-map-adaptive mapping for Athens in which logical qubit q_0 is assigned to physical qubit q_2 , q_1 to q_1 , and q_2 to q_3 . Analogously, for Yorktown, the logical qubits q_0, q_1, q_2 are mapped to the physical qubits q_2, q_3, q_0 in response to the error rates presented in table 2 and the coupling map.

Qubit	Readout error	Single-qubit U2 error rate	CNOT error rate
Q0	6.0799999999999965e-2	1.1093439321828e-3	cx0_1: 1.802e-2, cx0_2: 1.649e-2
Q1	7.640000000000002e-2	1.1734373542497346e-3	cx1_0: 1.802e-2, cx1_2: 1.881e-2
Q2	3.649999999999998e-2	5.719412382501265e-4	cx2_0: 1.649e-2, cx2_1: 1.881e-2, cx2_3: 1.242e-2, cx2_4: 2.896e-2
Q3	4.810000000000003e-2	3.2626596012098067e-4	cx3_2: 1.242e-2, cx3_4: 1.891e-2
Q4	5.579999999999996e-2	1.3053013587863154e-3	cx4_2: 2.896e-2, cx4_3: 1.891e-2

Tab. 2: Error rates of IBM Quantum Yorktown at time of the experiment [IB20]

Qiskit transpiler optimisation. The previous two optimisation approaches are included in a similar form in the Qiskit transpiler engine, which prepares logical quantum circuits for execution on the target backend by selecting suitable physical qubits and adapting the circuit in a state-preserving manner, if necessary [Qi21]. The optimisation steps can be manually activated by setting the transpiler optimisation level. While an automatic coupling-map-adaptive mapping is contained in level 2, level 3 adds an automatic error-map-adaptive qubit selection, where the similarity to the manual approach is of concern.

Optimisation level	Description from Qiskit transpiler documentation
0	No explicit optimisation other than mapping to backend
1	Light optimisation by simple adjacent gate collapsing
2	Medium optimisation by initial layout selection and gate cancellation using commutativity rules
3	Heavy optimisation by noise adaptive qubit mapping and gate cancellation using commutativity rules and unitary synthesis

Tab. 3: Logic of Qiskit transpiler optimisation levels [Qi21]

As demonstrated in table 3, the transpiler optimisation on level 3 introduces an additional approach to error reduction in the form of compressing the number of gates to be executed on the device in a procedure denoted gate cancellation. To compare the manual against the automatic approach amended with this additional logic, this level is employed in the transpiling process of the reference circuit.

Measurement error mitigation. To moderate errors induced in the readout process of the quantum state during projective measurement, measurement error mitigation as implemented in Qiskit Ignis is performed as a classical routine on the results from the quantum circuit. Whereas the previous techniques are conducted to transform the quantum circuit before execution on the quantum computer, the measurement errors included in the results from the execution may be arithmetically mitigated after their retrieval. First, the system's readout error behaviour is characterised through execution of quantum circuits that prepare all 2^n possible computational basis states determined by the system's number of qubits n . The expected ideal values are subsequently compared to the results obtained. As general characteristic, the measurement error rates recorded from the deviation are fit into a matrix that can eventually be used to partially eliminate the effect of this class of errors in the results from any algorithm executed on that system, returning values closer to the expectation. The interested reader is referred to the rigorous introduction by [Ka19].

Details regarding the implementation of all four strategies may be accessed at [SL21].

3.4 Analysis

Each experiment is executed ten times on the applicable quantum computer within a time frame of 24 hours to account for recalibration effects. A hypothesis is supported when the results of the experiments confirm the prediction from the hypothesis. This is assumed to be the case when the corresponding mean Hellinger fidelity values differ at a significance level of 95% ($p \leq 0.05$). Otherwise, the hypothesis is rejected.

For comparison of two conditions, the two-sided independent samples T-test is employed, using Levene's statistic to assess similarity between the variances of the Hellinger fidelity distributions. For comparison of more than two conditions, for instance to compare multiple error mitigation techniques, one-way ANOVA tests are performed. To determine the pairwise relations, the Tukey post-hoc test is applied in case of an insignificant Levene test. In the opposite case, the ANOVA Welch test is conducted instead of the ANOVA test, followed by the Games-Howell post-hoc test.

4 Results and Discussion

In the following, the results from the experiments defined in section 3 are presented and discussed in order of the hypotheses.

4.1 Comparison between the Qiskit Aer simulator and IBM Quantum computers

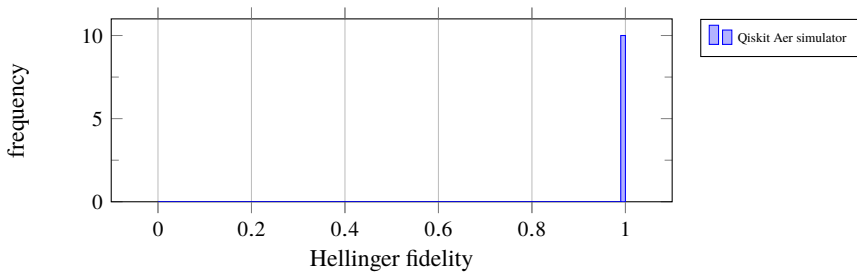


Fig. 4: Histogram of Hellinger fidelities from execution of the reference circuit on the noise- and error-free Qiskit Aer simulator

The mean Hellinger fidelity of the ideal execution on the noise and error-free Qiskit Aer simulator is 99.32% (00.23 % std). The deviation from 100.00% is expected due to the inherent randomness of quantum computations and is sufficiently small to indicate the ideal nature of the considered simulator. This value is used as benchmark for the results retrieved on the real systems.

The following mean Hellinger fidelities are recorded on the available quantum computers. The largest value of 81.03% (0.77% std) is obtained on Athens, followed by 76.09% (00.74% std) on Ourense, 75.96% (02.01% std) on Santiago, 71.89% (00.37% std) on Valencia, 71.52% (00.80% std) on Vigo, 71.15% (01.05% std) on Melbourne, and 61.68% (00.85% std) on Yorktown. These means of unequal variances (Levene, $F=7.560$, $p<0.001$) are significantly different (ANOVA Welch, $F(6, 27.2)=470.602$, $p<0.001$). Specifically, the mean of each device is significantly different from the mean of each other (Games-Howell, $p\leq 0.05$) except for Melbourne and Valencia, Melbourne and Vigo, Vigo and Valencia, and Ourense and Santiago.

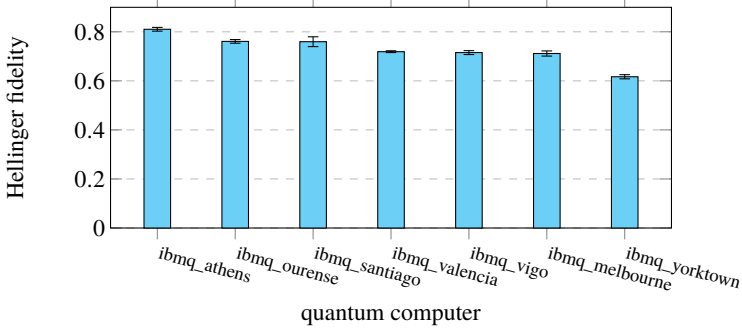


Fig. 5: Mean Hellinger fidelities from execution of the GHZ-YYX quantum circuit on IBM Quantum computers (std. dev.)

With an overall mean of 72.76% (05.70% std), the results from the real devices are significantly lower than the result from the simulator ($t(70.531) = -38.732$, $p < 0.001$) by 26.56 percent point (PP) in the 95% CI [25.19, 27.92], employing inhomogeneous variances (Levene, $F = 11.267$, $p = 0.001$). This deviation, given the presence of noise and errors in the quantum computers, is in line with the expectation and supports hypothesis 1. Based on the extreme values, Athens and Yorktown are selected for comparison of the quantum error mitigation techniques. For this purpose, the Hellinger fidelity values recorded in this reference experiment serve as benchmark for the subsequent experiments.

4.2 Coupling-map-adaptive qubit mapping

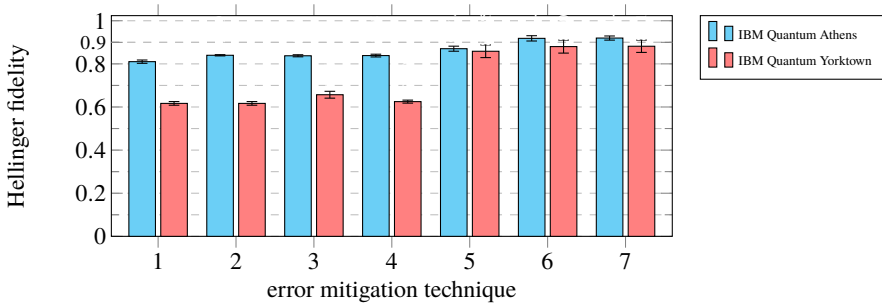


Fig. 6: Average Hellinger fidelities by error mitigation technique on IBM Quantum Athens and IBM Quantum Yorktown (std. dev.)

For Athens, the coupling-map-adaptive qubit mapping results in a mean Hellinger fidelity of 83.99% (00.29% std). Employing a variance similar to the reference (Levene, $F = 3.398$, $p = 0.082$), the mean of the distribution is significantly larger than that of the reference ($t(18) = -11.402$, $p < 0.001$) by 02.96 PP in the 95% CI [02.41, 03.50]. This supports hypothesis 2 and demonstrates the effectiveness of the technique induced by the elimination

of auxiliary gates. On Yorktown, no coupling-map-adaptive qubit mapping is applicable as the reference circuit is inherently aligned to the device's coupling map.

4.3 Error-map-adaptive qubit mapping

The Athens error-map-adaptive GHZ-YYX circuit yields a mean Hellinger fidelity of 83.76% (00.48% std). The improvement of 02.73 PP in the 95% CI [02.13, 03.33] compared to the reference is significant ($t(18)=-9.545$, $p<0.001$) between the two distributions with homogeneous variances (Levene, $F=0.659$, $p=0.427$). While similar, the mean is not significantly larger than the coupling-map-adaptive mapping, indicating that the additional consideration of error-rates did not yield an effect. Reason for this may lie in the implicit initial selection of qubits in experiment 4.2 already employing, on average, comparably low error rates so that the present selection can not decrease the overall error concentration beyond that level.

In the case of Yorktown, the error-map-adaptive mapping results in a mean Hellinger fidelity of 65.69% (01.59% std). This represents a significant improvement ($t(13.781)=-7.021$, $p<0.001$) of 04.02 PP in the 95% CI [02.79, 05.24] to the reference, comparing unequal variances (Levene, $F=6.709$, $p=0.018$). Since the reference circuit was already coupling-map-optimally mapped to the device, this increase demonstrates the influence of varying error rates within one system and the potential of decreasing the effect of errors through an optimised selection. This supports the associated hypothesis 3, even though in the case of Athens, it may be argued that the mere avoidance of additional gates decreased the amount of error in the results, similar to the coupling-map-based approach.

4.4 Transpiler optimisation

On Athens, the mean Hellinger fidelity returned from the circuit produced by the level-three-optimised transpiler is recorded as 83.85% (00.64% std). The appreciation of 02.82 PP in the 95% CI [02.15, 03.48] compared to the reference, both distributions possessing homogeneous variances (Levene, $F=0.001$, $p=0.971$), is significant ($t(18)=-8.913$, $p<0.001$). Confirming the two manual approaches of experiment 4.2 and 4.3 due the similar result level, this supports the corresponding hypothesis 4. The additional gate cancellation, however, is not found to produce a further decrease in the error level.

On Yorktown, the mean Hellinger fidelity is returned as 62.49% (00.70% std), improved by a magnitude of 00.81 PP in the 95% CI [00.77, 01.55] compared to the reference. This mean difference between two distributions of homogeneous variances (Levene, $F=0.562$, $p=0.463$) is significant ($t(18)=-2.321$, $p=0.032$). While the increase reinforces that the automated noise-adaptive mapping may reduce errors significantly beyond a sole coupling-map-adaptive approach, the inferiority of the value to that of experiment 4.3 indicates that the manually determined error-map-adaptive layout may have an even larger effect.

4.5 Measurement error mitigation

Measurement error mitigation applied on the results from the reference circuit on Athens produces a mean Hellinger fidelity of 87.04% (01.17% std). Compared to the reference result distribution of similar variance (Levene, $F=1.238$, $p=0.281$), the increase of 06.00 PP in the 95% CI [05.08, 06.93] is significant ($t(18)=-13.567$, $p<0.001$).

Applying measurement error mitigation on the results recorded for the reference circuit on Yorktown, the mean Hellinger fidelity is observed at 85.85% (02.94%). This constitutes a remarkable, significant improvement ($t(10.510)=-24.967$, $p<0.001$) of 24.17 PP in the 95% CI [22.03, 26.32] to the reference results, which yield a different variance (Levene, $F=22.096$, $p<0.001$).

For both backends, while overall supporting the corresponding hypothesis 5, comparing these values to the outcomes from the previous error mitigation techniques indicates the high effectiveness of measurement error mitigation and allows to deduce the strong presence of readout errors in the results from the experiments. For this reason, this technique is considered particularly powerful.

4.6 Error-map-adaptive qubit mapping and measurement error mitigation

Applying measurement error mitigation on the results from the error-map-adapted circuit executed on Athens yields the mean Hellinger fidelity value 91.84% (01.23% std). Employing a variance similar to the reference result distribution (Levene, $F=2.674$, $p=0.119$), the mean of this distribution is significantly higher ($t(18)=-23.595$, $p<0.001$) by 10.80 PP in the 95% CI [09.84, 11.77].

Similarly, the results returned on Yorktown produce a mean Hellinger fidelity of 88.02% (03.02% std), which surpasses the reference significantly ($t(10.433)=-26.550$, $p<0.001$) by 26.34 PP in the 95% CI [24.14, 28.54], comparing unequal variances (Levene, $F=42.364$, $p<0.001$). In addition to supporting the associated hypothesis 6, this further confirms the additive nature of the manual mapping approaches and measurement error mitigation.

4.7 Transpiler optimisation and measurement error mitigation

Mitigating the readout errors on the results obtained from the Qiskit transpiler-optimised GHZ-YYX circuit on Athens results in a mean Hellinger fidelity of 91.98% (00.96% std). This represents a significant improvement ($t(18)=-28.201$, $p<0.001$) of 10.94 PP in the 95% CI [10.13, 11.76] to the reference results with similar variance (Levene, $F=1.696$, $p=0.209$).

Measurement error mitigation applied on the results from the optimally transpiled reference circuit on Yorktown yields a mean quality of 88.18% (02.86% std). Compared to the

reference values of unequal variance (Levene, $F=9.616$, $p=0.006$), this marks a significant improvement ($t(10.590)=-28.043$, $p<0.001$) of 26.50 PP in the 95% CI [24.41, 28.59].

Similar to the previous technique, the increase on both backends supports hypothesis 7 and further indicates the additivity between transpiler optimisation and measurement error mitigation. The differences to the manual approach 4.6 are not significant for both devices.

4.8 Comparison of error mitigation techniques

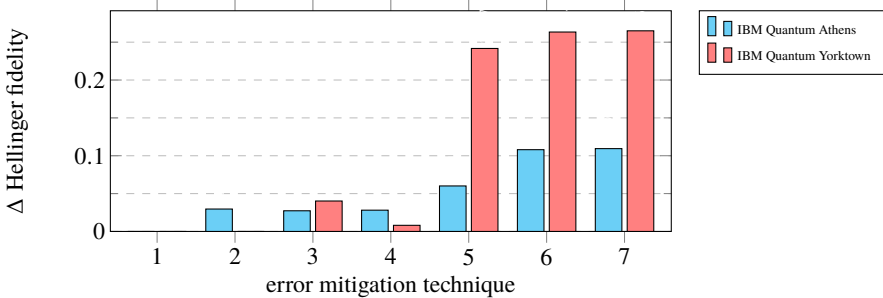


Fig. 7: Mean Hellinger fidelity improvement by error mitigation technique on IBM Quantum Athens and IBM Quantum Yorktown (std. dev.)

For Athens, the results from all techniques, which overall differ significantly (ANOVA Welch, $F(6, 27.047)=189.827$, $p<0.001$), employing inhomogeneous variances (Levene, $F(6, 63)=3.155$, $p=0.009$), are furthermore significantly different from one another (Games-Howell, $p\leq 0.05$) except for coupling-map-adaptive qubit mapping and error-map-adaptive qubit mapping, coupling-map-adaptive qubit mapping and transpiler optimisation, error-map-adaptive qubit mapping and transpiler optimisation as well as for measurement error mitigation combined with error-map-adaptive mapping and measurement error mitigation combined with transpiler optimisation. The ranking of the error mitigation techniques according to their Hellinger fidelity mean in decreasing order is provided as 7, 6, 5, 2, 4, 3, 1.

Similarly, on Yorktown, the mean quality values from distributions of inhomogeneous variances (Levene, $F(6, 63)=11.370$, $p<0.001$) are significantly different (ANOVA Welch, $F(6, 27.348)=312.392$, $p<0.001$). The mean from each technique, excluding the non-applicable coupling-map-adaptive mapping, is significantly different from the means of all others (Games-Howell, $p\leq 0.05$) except for the three strategies including measurement error mitigation compared to one another as well as for transpiler optimisation and the reference setup. The ranking of techniques is established as 7, 6, 5, 3, 4, 2, 1.

Overall partially rejecting the associated hypothesis 8, this strongly indicates that each independent approach included in the techniques introduces an improvement and that the manual implementation is similar to the automated approach from the Qiskit transpiler.

5 Conclusion

In this study, the effectiveness of applied quantum error mitigation strategies for the fully entangled quantum GHZ-YYX state produced on near-term quantum computers was investigated. On the available quantum computing systems by IBM Quantum, it was confirmed that the effect of quantum noise and errors decreases the result quality expressed as the Hellinger fidelity on average by 26.56 PP. The individual systems' quality values vary between 81.03% and 61.68% in a way only partially correlated to the number of qubits and the Quantum Volume. This finding renders the initial choice of backend on which to execute the algorithm a paramount one in the error reduction process.

Selecting two systems with inherently relatively high and low quality values from performance of the reference circuit, respectively, it was confirmed that the introduction of auxiliary gates in the transpiling process significantly reduces the quality. Thus, reducing the extent of this process through selection of physical qubits that optimally map the logical circuit regarding the device's coupling map was found to yield a considerable improvement in the result quality. A manual approach introduced for this objective was found as effective as the automated approach contained in the Qiskit transpiler.

Beyond the qubit connectivity, the consideration of individual error rates present in a system for the purpose of manually deriving an error-reducing qubit mapping scheme may introduce a significant improvement if the qubits are carefully selected. Automating this strategy through transpiler optimisation, however, was found to not, or only weakly, provide a significant benefit, allowing to deduce the non-applicability of gate cancellation in the given setup. These results imply that more advanced considerations may be required to be incorporated in the design of an advanced transpiling optimisation algorithm.

Among the considered individual techniques, measurement error mitigation resulted in the largest quality improvement on the two tested systems, providing evidence for a large presence of readout errors in the results. Specifically on the originally low-ranking system, this was found to be the case, where the method remarkably raised the quality to a level comparable to inherently higher-ranking systems. Combining this technique with error-map-adaptive qubit mapping, through a manual or automated implementation, yields the overall highest quality values - in the present case of 91.98% and 88.18% for the chosen systems. This is indicative of the additive relationship between non-exclusive quantum error mitigation approaches. Following this combined strategy, the large distance between the two devices in terms of the more holistic Quantum Volume measure is rendered void when regarding the result quality from the corresponding circuits.

While further investigations are required to be performed with larger circuits and advanced quantum error mitigation algorithms, this study concludes that circuit-level manual and automated error-map-adaptive qubit mapping approaches as well as measurement error mitigation, and their combination, are effective measures to increase the result quality from quantum algorithms producing entangled quantum states on near-term quantum computers.

Bibliography

- [ABO08] Aharonov, Dorit; Ben-Or, Michael: Fault-Tolerant Quantum Computation with Constant Error Rate. *SIAM Journal on Computing*, 38(4):1207–1282, 2008.
- [AG04] Aaronson, Scott; Gottesman, Daniel: Improved simulation of stabilizer circuits. *Physical Review A - Atomic, Molecular, and Optical Physics*, 70(5 A), 2004.
- [Al19] Aleksandrowicz, Gadi; Alexander, Thomas; Barkoutsos, Panagiotis; Bello, Luciano; Ben-Haim, Yael; Bucher, David; Cabrera-Hernández, Francisco Jose; Carballo-Franquis, Jorge; Chen, Adrian; Chen, Chun-Fu et al.: Qiskit: An open-source framework for quantum computing. 2019.
- [Al20] Alexander, Thomas; Kanazawa, Naoki; Egger, Daniel J; Capelluto, Lauren; Wood, Christopher J; Javadi-Abhari, Ali; C McKay, David: Qiskit pulse: programming quantum computers through the cloud with pulses. *Quantum Science and Technology*, 5(4):44006, 2020.
- [Bo99] Bouwmeester, Dik; Pan, Jian-Wei; Daniell, Matthew; Weinfurter, Harald; Zeilinger, Anton: Observation of Three-Photon Greenberger-Horne-Zeilinger Entanglement. *Physical Review Letters*, 82(7):1345–1349, 1999.
- [Cr19] Cross, Andrew; Bishop, Lev; Sheldon, Sarah; Nation, Paul; Gambetta, Jay: Validating quantum computers using randomized model circuits. *Physical Review A*, 100(3):32328, 2019.
- [Di00] DiVincenzo, David P.: The Physical Implementation of Quantum Computation. *Fortschritte der Physik*, 48(9-11):771–783, 2000.
- [DL98] DiVincenzo, David P; Loss, Daniel: Quantum information is physical. *Superlattices and Microstructures*, 23(3):419–432, 1998.
- [EBL18] Endo, Suguru; Benjamin, Simon C.; Li, Ying: Practical Quantum Error Mitigation for Near-Future Applications. *Physical Review X*, 8(3):31027, 2018.
- [En21] Endo, Suguru; Cai, Zhenyu; Benjamin, Simon C; Yuan, Xiao: Hybrid Quantum-Classical Algorithms and Quantum Error Mitigation. *Journal of the Physical Society of Japan*, 90(3):32001, 2021.
- [Fo12] Fowler, Austin G; Mariantoni, Matteo; Martinis, John M; Cleland, Andrew N: Surface codes: Towards practical large-scale quantum computation. *Physical Review A*, 86(3):32324, 2012.
- [Gl21] Glendinning, Ian: Quantum Programming Languages and Tools. <http://www.vcpc.univie.ac.at/~ian/hotlist/qc/programming.shtml>, accessed: 21/05/2021, 2021.
- [Go96] Gottesman, Daniel: Class of quantum error-correcting codes saturating the quantum Hamming bound. *Physical Review A*, 54(3):1862–1868, 1996.
- [Gr90] Greenberger, Daniel M; Horne, Michael A; Shimony, Abner; Zeilinger, Anton: Bell's theorem without inequalities. *American Journal of Physics*, 58(12):1131–1143, 1990.
- [He09] Hellinger, Ernst: Neue Begründung der Theorie quadratischer Formen von unendlichvielen Veränderlichen. *Journal für die reine und angewandte Mathematik*, 136:210–271, 1909.

- [Hü92] Hübner, Matthias: Explicit computation of the Bures distance for density matrices. *Physics letters. A*, 163(4):239–242, 1992.
- [IB20] IBM Corporation: IBM Quantum. <https://quantum-computing.ibm.com/>, accessed: 17/11/2020, 2020.
- [Jo94] Jozsa, Richard: Fidelity for Mixed Quantum States. *Journal of Modern Optics*, 41(12):2315–2323, 1994.
- [Ka19] Kandala, Abhinav; Temme, Kristan; Córcoles, Antonio D; Mezzacapo, Antonio; Chow, Jerry M; Gambetta, Jay M: Error mitigation extends the computational reach of a noisy quantum processor. *Nature*, 567(7749):491–495, 2019.
- [KLZ98] Knill, Emanuel; Laflamme, Raymond; Zurek, Wojciech H: Resilient Quantum Computation. *Science*, 279(5349):342 LP – 345, 1998.
- [Kn08] Knill, Emanuel; Leibfried, Dietrich; Reichle, Rolf; Britton, Joe; Blakestad, R Brad; Jost, John D; Langer, Chris; Ozeri, Roece; Seidelin, Signe; Wineland, David J: Randomized Benchmarking of Quantum Gates. *Physical Review A*, 77(1):12307, 2008.
- [Ko07] Koch, Jens; Terri, M Yu; Gambetta, Jay; Houck, Andrew A; Schuster, David I; Majer, Johannes; Blais, Alexandre; Devoret, Michel H; Girvin, Steven M; Schoelkopf, Robert J: Charge-insensitive qubit design derived from the Cooper pair box. *Physical Review A*, 76:042319, 2007.
- [Ko20] Kole, Abhoy; Hillmich, Stefan; Datta, Kamalika; Wille, Robert; Sengupta, Indranil: Improved Mapping of Quantum Circuits to IBM QX Architectures. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 39(10):2375–2383, 2020.
- [LB17] Li, Ying; Benjamin, Simon C: Efficient Variational Quantum Simulator Incorporating Active Error Minimization. *Physical Review X*, 7(2):21050, 2017.
- [Mc18] McKay, David C; Alexander, Thomas; Bello, Luciano; Biercuk, Michael J; Bishop, Lev; Chen, Jiayin; Chow, Jerry M; Córcoles, Antonio D; Egger, Daniel; Filipp, Stefan: Qiskit backend specifications for OpenQASM and OpenPulse experiments. *arXiv preprint arXiv:1809.03452*, 2018.
- [MGE11] Magesan, Easwar; Gambetta, Jay M; Emerson, Joseph: Scalable and Robust Randomized Benchmarking of Quantum Processes. *Physical Review Letters*, 106(18):180504, 2011.
- [MGE12] Magesan, Easwar; Gambetta, Jay M; Emerson, Joseph: Characterizing quantum gates via randomized benchmarking. *Physical Review A*, 85(4):42311, 2012.
- [Mi21] Miszczak, Jarek: List of QC simulators. <https://www.quantiki.org/wiki/list-qc-simulators>, accessed: 21/05/2021, 2021.
- [Mo18] Moll, Nikolaj; Barkoutsos, Panagiotis; Bishop, Lev S; Chow, Jerry M; Cross, Andrew; Egger, Daniel J; Filipp, Stefan; Fuhrer, Andreas; Gambetta, Jay M; Ganzhorn, Marc: Quantum optimization using variational algorithms on near-term quantum devices. *Quantum Science and Technology*, 3(3):30503, 2018.
- [NC10] Nielsen, Michael A.; Chuang, Isaac: *Quantum Computation and Quantum Information*, 10th Anniversary Edition. Cambridge University Press, New York City, New York, USA, 2010.

- [Ni96] Nielsen, Michael A: The entanglement fidelity and quantum error correction. arXiv preprint quant-ph/9606012, 1996.
- [Ni02] Nielsen, Michael A: A simple formula for the average gate fidelity of a quantum dynamical operation. *Physics Letters A*, 303(4):249–252, 2002.
- [Pr98] Preskill, John: Reliable quantum computers. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1969):385–410, 1998.
- [Pr18] Preskill, John: Quantum Computing in the NISQ era and beyond. *Quantum*, 2:79, 2018.
- [Qi21] Qiskit: Qiskit documentation. <https://qiskit.org/documentation/>, accessed: 27/05/2021, 2021.
- [SET20] Suzuki, Yasunari; Endo, Suguru; Tokunaga, Yuuki: Quantum error mitigation for fault-tolerant quantum computing. arXiv preprint arXiv:2010.03887, 2020.
- [Sh96] Shor, Peter W: Fault-tolerant quantum computation. In: *Proceedings of 37th Conference on Foundations of Computer Science*. pp. 56–65, 1996.
- [SL21] Schulze, Lennart; Lahmann, Jan-Rainer: GHZ Game on real devices. <https://github.com/JanLahmann/Fun-with-Quantum>, accessed: 27/05/2021, 2021.
- [St20] Strikis, Armands; Qin, Dayue; Chen, Yanzhu; Benjamin, Simon C; Li, Ying: Learning-based quantum error mitigation. arXiv preprint arXiv:2005.07601, pp. 1–19, 2020.
- [Su21] Suchsland, Philippe; Tacchino, Francesco; Fischer, Mark H.; Neupert, Titus; Barkoutsos, Panagiotis Kl.; Tavernelli, Ivano: Algorithmic Error Mitigation Scheme for Current Quantum Processors. *Quantum*, 5:492, 2021.
- [TBG17] Temme, Kristan; Bravyi, Sergey; Gambetta, Jay M: Error Mitigation for Short-Depth Quantum Circuits. *Physical Review Letters*, 119(18):180509, 2017.
- [Va99] Vaidman, Lev: Variations on the Theme of the Greenberger-Horne-Zeilinger Proof. *Foundations of Physics*, 29(4):615–630, 1999.
- [VKL99] Viola, Lorenza; Knill, Emanuel; Lloyd, Seth: Dynamical Decoupling of Open Quantum Systems. *Physical Review Letters*, 82(12):2417–2421, 1999.
- [VL98] Viola, Lorenza; Lloyd, Seth: Dynamical suppression of decoherence in two-state quantum systems. *Physical Review A*, 58(4):2733–2744, 1998.
- [We20] Wei, Ken X; Lauer, Isaac; Srinivasan, Srikanth; Sundaresan, Neereja; McClure, Douglas T; Toyli, David; McKay, David C; Gambetta, Jay M; Sheldon, Sarah: Verifying multipartite entangled Greenberger-Horne-Zeilinger states via multiple quantum coherences. *Physical Review A*, 101(3):32343, 2020.
- [Wi17] Wilhelm, Frank K.; Steinwandt, Rainer; Langenberg, Brandon; Liebermann, Per J.; Messinger, Anette; Peter K. Schuhmacher: Entwicklungsstand Quantencomputer. Technical report, Bundesamt für Sicherheit in der Informationstechnik, 2017.

Workshop: Datenschutz im Diskurs
(RuT2021)

Recht und Technik – Datenschutz im Diskurs

Rüdiger Grimm¹, Gerrit Hornung², Christoph Sorge³, Indra Spiecker genannt Döhmann⁴

Vorwort zu den Workshopbeiträgen

Der Workshop „Recht und Technik – Datenschutz im Diskurs“ ist nach acht Jahren ein kontinuierlicher und fester Bestandteil der Jahrestagung der Gesellschaft für Informatik. Letztes Jahr hatte er pandemiebedingt erstmals online stattgefunden. Zu unserer Freude wurde es trotz der Online-Distanz ein sehr lebendiger Workshop. Im Übrigen war es einer der bestbesuchten Workshops der ganzen Jahrestagung 2020.

Auch dieses Mal findet der Workshop wieder wie die ganze Jahrestagung 2021 online statt. Inhaltlich bleibt die Ausrichtung wie gewohnt: Wir bieten ein Forum für Beiträge von Informatiker/innen, Juristen/innen und Vertreter/innen benachbarter Fächer, die an Fragestellungen des technikbasierten Datenschutzes arbeiten. Wie schon in den Vorjahren werden Themen adressiert, die anwendungsorientiertes Potential für interdisziplinären Diskurs und Zusammenarbeit bieten und die Möglichkeiten aufzeigen, wie Datenschutz durch Technik präzisiert und umgesetzt werden kann.

Der Workshop setzt an den theoretischen und praktischen Aspekten des Schutzes von Daten und Privatheit und der europäischen Datenschutz-Grundverordnung an. Konkret werden datenschutzrechtliche und zugehörige technische Probleme folgender Anwendungsbereiche und Verarbeitungsmethoden thematisiert

- Digitale Wirtschaft
- Smart Cities
- Videoüberwachung
- Videokonferenzen
- Data Mining
- Datenextraktion von Webseiten
- Upload Filter
- Cookies

¹ Fraunhofer SIT Darmstadt und Universität Koblenz-Landau <grimm@uni-koblenz.de>

² Universität Kassel <Gerrit Hornung <gerrit.hornung@uni-kassel.de>

³ Universität des Saarlandes <christoph.sorge@uni-saarland.de>

⁴ Goethe-Universität Frankfurt am Main <spiecker@jur.uni-frankfurt.de>

Die Beiträge, die wir aus der Vielzahl an qualitätsvollen Einreichungen in einem *peer-reviewed*-Verfahren mit Unterstützung unserer Gutachter/innen auswählen konnten, zeigen die thematische und disziplinären Bandbreite der derzeit im Spannungsfeld von Recht und Technik diskutierten Themen; die Einreichungen gingen darüber noch hinaus. Die zunehmende Regulierung der Digitalisierung, die Konkretisierung der DSGVO-Anforderungen und ein insgesamt gestiegenes Bewusstsein dafür, dass nicht alles technisch Machbare auch gesellschaftlich wünschenswert ist, spiegeln sich in den Beiträgen wieder. Neben sehr konkreten Vorschlägen zur Bewältigung von Einzelproblemen sind auch Beiträge mit übergreifenden Einsichten repräsentiert.

Gerade die Entwicklung der Corona-App hat zudem gezeigt, dass gesellschaftliches Bewusstsein für den Datenschutz sehr wohl vorhanden ist und sich in erheblicher Einflussnahme auf den politischen Prozess niederschlägt. Das Vorgehen illustriert zudem, dass aus den Erkenntnissen von Recht und Technik sehr wohl konstruktive technische Lösungsmöglichkeiten für rechtliche Probleme erwachsen können, die gleichzeitig die Privatheit der Nutzer/innen stärken und einen Markt für neue Produkte kreieren.

In diesem Sinne freuen wir uns mit der Veröffentlichung der Beiträge des diesjährigen Workshops darauf, auch im nächsten Jahr wiederum „Recht und Technik – Datenschutz im Diskurs“ anzubieten.

Unser besonderer Dank gilt den Mitgliedern des Programmkomitees, unter denen in diesem Jahr folgende Personen in die Begutachtung eingebunden waren:

- Matthias Bäcker, Universität Mainz
- Franziska Boehm, KIT
- Jens-Matthias Bohli, Hochschule Mannheim
- Katharina Bräunlich, Diez
- Matthias Enzmann, Fraunhofer SIT, Darmstadt
- Christian Geminn, Universität Kassel
- Nils Gruschka, Universität Oslo
- Christoph Gusy, Universität Bielefeld
- Niko Härting, Rechtsanwalt, Berlin
- Walter Hötendorfer, Research Institute Digital Human Rights Center, Wien
- Thomas Kahler, DPOblog.eu
- Ronald Petric, TH Nürnberg
- Burkhard Schafer, Universität Edinburgh
- Tobias Singelstein, Ruhr-Universität Bochum
- Jürgen Taeger, Universität Oldenburg

Alles akzeptieren oder Einstellung(en) ändern? – zum Stand der Praxis bei der Nutzung von Cookies

Stefan Grombacher^{1,2}, Tobias Straub²

Abstract: Die Verwendung von Cookies auf Webseiten unterliegt der E-Privacy-Richtlinie bzw. ihrer Umsetzung in nationales Recht sowie der Datenschutz-Grundverordnung. Vorliegend wurden für die 500 am häufigsten aus Deutschland genutzten Webseiten die Dialoge untersucht, mit deren Hilfe um die nötige Einwilligung ersucht wird. Dabei zeigt sich deutlich, dass die Ausgestaltungen vielfach weniger die Bedürfnisse der Nutzenden berücksichtigen und eher versuchen, diese im Sinne der Webseitenbetreiber zu beeinflussen. Es wird diskutiert, inwiefern die in der Praxis verwendeten Mechanismen im Widerspruch zu den rechtlichen Vorgaben stehen könnten.

Keywords: Consent Management Platform, Cookie-Banner, Dark Patterns, DS-GVO, Einwilligung, E-Privacy-Richtlinie, erforderliche/nicht erforderliche Cookies, Telemediengesetz

1 Einleitung

Kaum eine Webseite kommt heute ohne einen Mechanismus aus, der über die jeweiligen Cookies informiert und um Zustimmung zu ihrer Verwendung ersucht. Von Nutzerinnen und Nutzern werden derartige Banner oftmals als störend wahrgenommen.³ Um schnell zum eigentlichen Inhalt der Webseite vorzustoßen, wird den Dialogen daher wenig Aufmerksamkeit [No20] geschenkt, auch gibt es einige Implementierungen von Browser Add-Ons, die versuchen, die Dialoge von vornherein zu unterdrücken [Ut19]. Aufgrund der Dialoggestaltung besteht der schnellste Weg häufig darin, vorgeschlagene Konfigurationen oder gleich alle Cookies zu akzeptieren – was dann aber eher den Präferenzen des Seitenbetreibers und der Werbewirtschaft entsprechen dürfte, als jenen der Nutzenden [MB20].

Auch die Verbraucherschutz- und Netzpolitik hat die Thematik jüngst im Entwurf der Bundesregierung für ein Gesetz zur Regelung des Datenschutzes und des Schutzes der Privatsphäre in der Telekommunikation und bei Telemedien (TTDSG-E) aufgegriffen.⁴ Dass angesichts der noch immer ausstehenden E-Privacy-Verordnung Unsicherheit und

¹ NTT DATA Deutschland GmbH, Hans-Döllgast-Straße 26, D-80807 München, stefan.grombacher@nttdata.com

² Duale Hochschule Baden-Württemberg, Center for Advanced Studies, Bildungscampus 13, 74076 Heilbronn, tobias.straub@dhbw-stuttgart.de

³ <https://www.bitkom.org/Presse/Presseinformation/Cookie-Banner-stoeren-Internetnutzer.html> (25.07.2018)

⁴ Gesetzentwurf der Bundesregierung, Stellungnahme im Bundesrat (26.03.2021) <https://www.hamburg.de/bjv/pressemeldungen/14984108/2021-03-25-bjv-cookie-banner-im-internet/> (25.03.2021), Gesetzentwurf: https://www.bundesrat.de/SharedDocs/drucksachen/2021/0101-0200/163-21.pdf?__blob=publicationFile&v=1, BT-Drs. 163/21, § 24

ein Klärungsbedarf für Betreiber besteht, zeigen im europäischen Kontext die bereits von der französischen Aufsichtsbehörde verhängten hohen Bußgelder aufgrund unzureichender Cookie-Banner [KI20].

Die Gestaltung der Dialoge und der Umgang mit ihnen war bereits Gegenstand von Feld-[Ut19] und Laborstudien mit Nutzenden [MB20, No20], Lösungsvorschläge in Form eines Best Practice-Modells wurden von [KTS20] entwickelt. Die Frage, inwieweit auf Webseiten inner- und außerhalb der EU mittels Cookies ein Tracking ohne Einwilligung erfolgt, wurde von [MBS20, Sa19, Tr19] mit Hilfe automatisierter Werkzeuge untersucht. Der vorliegende Beitrag kombiniert einen vergleichbaren Ansatz mit einer manuellen Analyse (ähnlich derer in [So20] für Online-Nachrichtenseiten), um die 500 am häufigsten aus Deutschland genutzten Webseiten zu untersuchen.⁵

2 Zusammenfassung der Rechtslage

Im Folgenden wird die Rechtslage, soweit für das Verständnis des Beitrags erforderlich, skizziert. Für eine ausführlichere Darstellung wird auf die rechtswissenschaftliche Literatur verwiesen; einen guten Überblick gibt etwa [RE21].

Ursprünglich war geplant, dass 2018 zeitgleich mit dem neuen Datenschutzrecht in Form der EU-Datenschutz-Grundverordnung (DS-GVO) auch eine E-Privacy-Verordnung in Kraft treten sollte. Das Gesetzgebungsverfahren zur E-Privacy-Verordnung, welche die *E-Privacy-Richtlinie* (2002/58/EG, zuletzt geändert durch die „Cookie-Richtlinie“ 2009/136/EG) ablösen sollte, ist aber ins Stocken geraten. Insofern sind aktuell die Regelungen der E-Privacy-Richtlinie (nachfolgend: RL) und ihre – lange in Frage stehende – Umsetzung in nationales Recht maßgeblich. Für eine Klarstellung hinsichtlich der Regelungen im Telemediengesetz (TMG) sorgte hierbei im Mai 2020 ein Grundsatzurteil des BGH.⁶

2.1 Erfordernis der Einwilligung

Cookies fallen unter Art. 5 Abs. 3 RL, da es sich dabei um Informationen handelt, die auf einem Endgerät eines Teilnehmers oder Nutzers gespeichert oder von dort gelesen werden. Diese Vorschrift ist wohlgemerkt unabhängig davon anzuwenden, ob die Informationen einen Personenbezug haben.

Nutzt eine Webseite Cookies, so muss der Betreiber gemäß der RL eine informierte Einwilligung einholen, es sei denn, es handelt sich um ein *technisch notwendiges* Cookie

⁵ Nach Einreichung dieses Beitrags wurde bekannt, dass NOYB eine Software entwickelt hat, um automatisiert Rechtsverstöße bei der Nutzung von Cookies zu erkennen und Beschwerden zu formulieren (<https://noyb.eu/en/project/cookie-banners>, Stand: 31.05.2021).

⁶ Az. I ZR 7/16; siehe etwa [Sp20] für eine ausführliche Darstellung.

(s.u.). Laut BGH ist die Formulierung⁷ in § 15 Abs. 3 TMG richtlinienkonform derart auszulegen, dass damit das Einwilligungserfordernis der RL als in nationales Recht umgesetzt gelten kann. Die RL verweist (indirekt⁸) auf die DS-GVO hinsichtlich der Informationspflichten (Art. 12 ff.) sowie der Ausgestaltung der Einwilligung (Art. 4 Nr. 11, Art. 6 Abs. 1 S.1 lit. a., Art. 7) und betont, dass insbesondere klar und umfassend über die Zwecke der Verarbeitung informiert werden muss.

2.2 Anforderungen an die Gestaltung der Interaktion

Zum Teil sind Cookie-Hinweise als Einblendung am Rand ausgestaltet, die sich ignorieren lassen und eine Nutzung der Webseite ohne Unterbrechung ermöglichen. Ein derartiges konkludentes Verhalten genügt aber nicht den Anforderungen an eine aktiv zu erteilende Einwilligung [Sp20]. Dem Urteil des BGH zufolge ist beim Ersuchen um eine Einwilligung insbesondere auch eine Dialogform unzulässig, bei der Checkboxen vorausgewählt sind und erst deaktiviert werden müssen, wenn eine Einwilligung verweigert werden soll (*Opt-out*).

Hinsichtlich der optischen und textuellen Gestaltung der Dialoge lässt sich eine hohe Variabilität beobachten, da sich z.B. auch die von dafür spezialisierten Dienstleistern bezogenen Consent Management Tools (s.u.) von den Webseitenbetreibern anpassen lassen. Teilweise wird durch als *Dark Patterns*⁹ bezeichneten Designprinzipien (in unterschiedlicher Abstufung) versucht, Nutzende zu Handlungen zu verleiten, die ihren eigenen Interessen zuwiderlaufen. Als unzulässig werden dabei Designs angesehen, die auf der ersten Ebene keine echte Wahl lassen und für die Ablehnung der Einwilligung einen zusätzlichen Klick erfordern, wohingegen strittig ist, inwieweit dies auch für die rein optische Hervorhebung der vom Betreiber bevorzugten Option gilt [RE21 mwN].

2.3 Zusammenspiel von RL und DS-GVO

Wie oben aufgeführt, sind im Kontext von Cookies hinsichtlich des „Wie“ der Einwilligung i.S.v. § 15 TMG die Vorschriften der DS-GVO anzuwenden.

Uneinheitlich wird in der Literatur dagegen das genaue Verhältnis von RL und DS-GVO beschrieben. Art. 95 DS-GVO wird der hM nach so verstanden, dass die Vorgaben der RL (bzw. ihrer Umsetzungen in nationales Recht) jene der DS-GVO verdrängen, wenn beide konkurrierende Pflichten enthalten, die dasselbe Ziel verfolgen [BH20, PK20]. Dagegen sehen [RE21] bei Cookies, die personenbezogene Daten verarbeiten, eine aufgrund von § 15 TMG nötige Einwilligung wegen der Verweisung auch als Rechtsgrundlage für die

⁷ entgegen dem Wortlaut „sofern der Nutzer dem nicht widerspricht“

⁸ Über Art. 94 Abs. 2 Satz 1 DS-GVO.

⁹ Eine Zusammenfassung der Taxonomien aus der Literatur enthält [LS21]. Beispiele und weiterführende Literatur finden sich z.B. beim Dark Pattern Detection Project (<http://www.dapde.de/>).

Verarbeitung gem. Art. 6 Abs. 1 S. 1 lit. a DS-GVO an, während bei technisch notwendigen Cookies auch eine andere Rechtsgrundlage (insb. Art. 6 Abs. 1 S. 1 lit. b oder lit. f) in Frage käme.

Verarbeitungen personenbezogener Daten durch Webseitenbetreiber unterfallen, sofern es nicht nur die in Art. 5 Abs. 3 RL genannten Verarbeitungsformen Speichern und Abrufen betrifft, den allgemeinen Vorgaben der DS-GVO [BH20], d.h. insbesondere den Rechtmäßigkeitsvoraussetzungen gem. Art. 6 Abs. 1 (bzw. Art. 9 Abs. 2, sofern es sich um sensitive Daten handelt).

2.4 „Erforderliche“ Cookies

Hinsichtlich der Frage, welche Cookies als erforderlich gelten können und somit unter den Ausnahmetatbestand von Art. 5 Abs. 3 Satz 2 RL fallen, wird häufig auf eine ältere Stellungnahme der Aufsichtsbehörden verwiesen, die folgende Arten umfasst [Gi20,RE21].

- Cookies zur Fehleranalyse
- Session-Cookies für Onlineformulare oder Warenkörbe
- Authentifizierungscookies für eine Browsersitzung
- Sicherheitscookies
- Multimedia-Player-Sitzungscookies
- Lastverteilungs-Sitzungscookies
- Cookies zur vom Nutzer angeforderten Anpassung der Benutzeroberfläche
- Content-Sharing-Cookies sozialer Plug-ins in Bezug auf beim Netzwerk angemeldete Nutzer

Nach Auffassung der französischen Aufsichtsbehörde (CNIL) können zu dieser Kategorie auch Cookies für eine „einfache“ Webanalyse zählen, bei der nur *anonyme* Auswertungen erstellt werden, wobei die Nutzenden informiert werden und die Möglichkeit zum Widerspruch haben müssen.¹⁰

3 Cookie-Banner in der Praxis

Auch wenn dies unter dem Aspekt der Transparenz und Benutzbarkeit erstrebenswert wäre, hat sich bislang kein einheitliches Format für Cookie-Dialoge etabliert. Nachfolgend werden daher zunächst Kriterien genannt, anhand derer sich die bestehenden Ansätze unterscheiden lassen. Das Gros der Webseiten bindet eine Lösung eines Consent Management Providers (CMP, auch Consent Management Platform) zum Verwalten der Einstellungen ein. Auf diese Systeme sowie die technische Spezifikation, auf der sie typischerweise basieren, geht der zweite Abschnitt ein.

¹⁰ Art. 5, Délibération n° 2019-093, 4.7.2019

3.1 Unterscheidungskriterien

Hinsichtlich des generellen *Interaktionsmusters* lassen sich die Cookie-Banner in zwei Hauptgruppen unterteilen:

1. Zunächst einmal gibt es reine *Info-Banner*, welche vor allem vor Inkrafttreten der DSGVO gebräuchlich waren und typischerweise darauf hinweisen, dass die jeweilige Webseite Cookies enthält, deren Verwendung man durch die weitere Nutzung der Seite zustimme.
2. Eine Möglichkeit zur echten Entscheidung zwischen Ablehnung und Zustimmung bieten dagegen erst *Consent-Banner*, wobei sich diese weiter unterscheiden lassen:
 - a) Vorschaltseiten und modale Dialoge verlangen eine unmittelbare Reaktion und machen es – zumindest ohne Manipulation der Seite – unmöglich, diese zu nutzen, bevor man mit dem Dialog interagiert hat und seine Zustimmung oder Ablehnung durch eine aktive Handlung zum Ausdruck gebracht hat.
 - b) Daneben gibt es auch nicht-modale Dialoge, welche aber bei manchen Webseiten wesentliche Teile des Inhalts verdecken und somit indirekt doch zu einer Handlung zwingen. In der Theorie sollte es jedoch so sein, dass ohne eine explizite Handlung auch keine einwilligungspflichtigen Cookies gesetzt werden dürfen, d.h. ein Ignorieren des Banners einer Ablehnung der Einwilligung gleichkommen muss.

Einen Einfluss auf den *Grad der Beeinflussung* von Nutzenden haben die vom Dialog vorgesehenen Interaktionspfade sowie die textuelle (Wortwahl, Umfang, Verständlichkeit) und visuelle Gestaltung der Oberfläche. Eine Beeinflussung der Nutzenden wird im Englischen auch als *Nudging* bezeichnet [Ut19]. Erfolgt eine solche entgegen der Interessen der Nutzenden, so spricht man von den erwähnten Dark Patterns. Bei einem neutralen Design sind die Annahme und Verweigerung der Einwilligung als gleichwertige Alternativen visualisiert.

3.2 Consent Management Provider

Um die Zustimmungssysteme hat sich eine eigene Dienstleistungsbranche von Anbietern gebildet, die Consent-Banner bereitstellen. Zu den bekanntesten gehören Cookiebot, OneTrust, Quantcast und SourcePoint. Das Funktionsprinzip ist dabei jeweils, dass ein Betreiber seine Webseite automatisch vom Anbieter auf die verwendeten Cookies scannen und diese vorkategorisieren lassen kann. Die Systeme erlauben ein Customizing und Branding der Dialoge gemäß den Vorstellungen des Webseitenbetreibers. Sie werden per JavaScript eingebunden, wobei der Betreiber selbst darauf zu achten hat, dass am Ende das Zusammenspiel mit eigenem Skript-Code und den tatsächlich gesetzten Cookies rechtskonform ist und insbesondere keine Cookies gesetzt werden, bevor der Dialog beantwortet wurde.

Das vom Interactive Advertising Bureau (IAB Europe), einem internationalen Wirtschaftsverband der Onlinewerbebranche, entwickelte *Transparency & Consent Framework* (TCF)¹¹ ist eine technische Spezifikation, gemäß der die Zustimmung von Nutzenden rechtskonform eingeholt, gespeichert und an Werbetreibende mitgeteilt werden kann. Dazu wurden CMPs als neue Beteiligte eingeführt, die sich, wie auch die Werbetreibenden, bei IAB Europe registrieren müssen. Das Framework definiert u.a. ein einheitliches Format eines *Consent String* zur Speicherung der von den Nutzenden geäußerten Präferenzen, welcher vom CMP verwaltet und Werbetreibenden über standardisierte Schnittstellen zugänglich gemacht wird. Die 2019 durchgeführte Untersuchung von [MBS20] ergab allerdings, dass etwa bei der Hälfte der Webseiten trotz des Einsatzes eines TCF-kompatiblen CMPs ein nicht richtlinienkonformes Verhalten vorlag.

4 Untersuchung

Ziel der empirischen Untersuchung war es festzustellen, wie die aus Deutschland am häufigsten genutzten Webseiten Cookie-Banner einsetzen. Konkret sollten dabei jeweils folgende Fragen aus Sicht von datenschutzsensiblen Nutzerinnen und Nutzern beantwortet werden:

1. Wie *aufwändig* ist das Ablehnen aller Cookies im Gegensatz zum Akzeptieren?
2. Erfolgt eine *Beeinflussung* der Nutzenden durch die Gestaltung des Banners?
3. (Wie einfach) ist es möglich, den *Einstellungs-Dialog* beim Besuch der Webseite später erneut aufzurufen?
4. Werden bereits *Cookies gesetzt*, ohne dass mit dem Banner interagiert wurde?

4.1 Methodik

Um eine einheitliche und zumindest teilautomatisierte Auswertung zu ermöglichen, wurden die Begriffe dabei wie folgt operationalisiert:

1. Der Aufwand wird in der mindestens erforderlichen Zahl von Mausaktionen, wozu Klicks oder Scrollvorgänge (unabhängig von deren Länge) zählen, für die Ablehnung aller (nicht notwendigen) Cookies gemessen.
2. Der Grad der Beeinflussung wird anhand der Farbgebung, Größe und Anordnung der Steuerelemente im Dialog untersucht.
3. Die Zahl der Mausaktionen zum Aufruf des Einstellungs-Dialogs wird ermittelt.

¹¹ <https://iabeurope.eu/tcf-2-0/> (Stand vom: 21.08.2019)

4. Es wird untersucht, ob 3rd Party Cookies gesetzt werden, ohne dass eine User Interaktion stattfindet.

Als Untersuchungsgegenstand wurden alle im Ranking von Alexa¹² (Stand: 06.01.21) unter den Top 500 gelisteten Webseiten betrachtet, unabhängig davon, ob es sich dabei um deutsch- oder rein fremdsprachige Angebote handelte. Am selben Tag wurden zunächst automatisiert Screenshots der Startseiten¹³ aller Webseiten in der Liste aufgenommen (Browser: Chrome, Bildschirmauflösung: 1920 ´ 1080 Pixel, standardmäßige Zoomstufe 100%). Dies erfolgte durch ein JavaScript-Programm, welches Webseitenaufrufe mittels Protractor,¹⁴ einem Framework für automatisierte Browser-Tests, simulierte.

Diese Screenshots wurden anschließend manuell untersucht, um festzustellen, welche Webseiten überhaupt Banner einsetzen. Die Webseiten, welche ein Consent-Banner einbinden, wurden dann einzeln manuell daraufhin untersucht, wie viele Mausklicks benötigt werden, um alle Cookies abzulehnen.

Gleichzeitig wurde für die Consent-Banner auch manuell vermerkt, ob und welche Techniken der Beeinflussung bei der Gestaltung der Steuerelemente (Links oder grafische Buttons) zum Einsatz kommen. Sofern ein Button zur Ablehnung (oder zum Akzeptieren einer Vorauswahl ausschließlich notwendiger Cookies) vorhanden war, wurde dieser mit dem Annahme-Button verglichen. War kein Ablehne-Button vorhanden, wurde der Annahme-Button mit demjenigen Button verglichen, welcher auf die zweite UI-Schicht führt. Falls das Ablehnen nur über die zweite UI-Schicht möglich war, wurde diese entsprechend untersucht. Es wurden in dem Fall wieder der Button zum Annehmen mit dem Button zum Ablehnen bzw. zum Akzeptieren der Vorauswahl ausschließlich notwendiger Cookies verglichen.

- Als tendenziös wurde eine *Farbgebung* gewertet, bei der sich der Button für die Annahme im Vergleich zu jenen für die Alternativen deutlich wahrnehmbarer vom Hintergrund des Banners abhob. Darunter fallen insbesondere Varianten, bei denen die Alternativen nur mit Grautönen oder gar nicht hinterlegt sind.
- Analog wurde auf das Vorliegen eines unausgewogenen *Größenverhältnisses* der Steuerelemente geprüft. Ein solches liegt v.a. dann vor, wenn eine der Alternativen nur als verlinkter Text realisiert wurde oder so erscheint, weil es sich zwar um einen Button handelt, dieser aber keinen erkennbaren Rahmen und die gleiche Farbe wie der Hintergrund des Banners hat.
- Eine Beeinflussung durch die *Positionierung* liegt vor, wenn sich das Steuerelement für das Ablehnen bzw. für das Betreten der zweiten UI-Schicht nicht an einer ver-

¹² <https://www.alexa.com/topsites/countries/DE>

¹³ Die URL wurde aus dem Domainnamen ohne Pfadangabe sowie dem Schema „http“ zusammengesetzt.

¹⁴ <https://www.protractortest.org>, verwendete Version: 5.4.2

gleichbaren Stelle befindet wie der Annahme-Button, beispielsweise wenn das Ablehnen nur durch einen in den Fließtext eingebetteten Link zu erreichen ist oder wenn auf der zweiten UI-Ebene das Steuerelement zum Akzeptieren aller Cookies sichtbar ist, während für die Wahl ausschließlich notwendiger Cookies ein Scrollen nötig ist.

Zur Untersuchung des dritten Kriteriums wurden jene Webseiten manuell untersucht, welche ein Consent-Banner eingebunden hatten. Es wurde dabei ermittelt, wie viele Mausaktionen erforderlich sind, lediglich um den Dialog wieder aufzurufen (ohne dann die getroffene Auswahl zu ändern). Falls kein Link mit einer gängigen Bezeichnung (etwa „Cookies“, „Cookie-Einstellungen“, „Einwilligung“, „Datenschutz“, „Privacy“) im Header oder Footer der Webseite oder deren Impressum vorhanden war und auch keine vergleichbare Funktion innerhalb von drei Minuten gefunden werden konnte, um den Dialog wieder zu öffnen, wurde der Versuch aufgegeben. Der Aufwand, den Nutzende hierfür zu investieren bereit sind, dürften in aller Regel recht begrenzt sein. Das Kriterium wurde nicht als erfüllt angesehen, wenn eine Webseite etwa Hinweise auf eine manuelle Löschung von Cookies über Browserfunktionen enthält oder darauf verweist, es aber selbst nicht möglich ist, den eigentlichen Dialog wieder aufzurufen.

Die Untersuchung, welche Cookies schon vor Interaktion mit einem Banner gesetzt werden, konnte wiederum vollautomatisiert ablaufen. Da aufgrund der Same-Origin Policy per JavaScript nicht auf die von einer dritten Seite gesetzten Cookies zugegriffen werden kann, wurde der Ansatz gewählt, stattdessen direkt die von Chrome zur Speicherung der Cookies verwendete SQLite-Datenbank auszuwerten, welche vor jedem Aufruf einer Webseite bereinigt wurde. Konkret wurden die einzelnen Webseiten wieder mittels Protractor aufgerufen, jedoch keine weitere Aktion veranlasst. Nach 50 Sekunden Wartezeit wurde vom Protractor-Skript die Chrome-Datenbank mit Hilfe des Packages `sqlite3`¹⁵ ausgelesen und die gefundenen 3rd Party Cookies zur späteren Auswertung in eine JSON-Datei geschrieben.

Eine Bewertung, ob es sich bei einem Cookie um ein notwendiges oder nicht-notwendiges im Sinne der RL handelt, ist auf automatisierte Weise schwerlich möglich. Als Annäherung wurde die vereinfachende Annahme getroffen, dass es sich bei 3rd Party Cookies, die keine Session Cookies sind, typischerweise um nicht-notwendige handeln dürfte. Bei den vorhandenen persistenten Cookies wurde auch die vom Server initial¹⁶ gesetzte Gültigkeitsdauer ermittelt.

4.2 Ergebnisse

Von den 500 berücksichtigten Seiten wiesen 219 ein Consent-Banner, 61 ein reines Info-Banner und 211 keines von beidem auf. Neun Seiten waren zum Testzeitpunkt nicht

¹⁵ <https://www.npmjs.com/package/sqlite3>, verwendete Version: 5.0.1

¹⁶ Der Standard (RFC6265, <https://tools.ietf.org/html/rfc6265>) lässt es prinzipiell zu, dass ein Server, an den ein Cookies zurückgeschickt wird, dessen Gültigkeitsdauer modifiziert.

erreichbar. Im Anhang ist in Tab. 1 für die Seiten mit Consent-Banner näher aufgeschlüsselt, wie groß der Aufwand für die Einwilligung bzw. ihre Verweigerung ist, welche irreführenden Designs zum Einsatz kommen und wie leicht sich die Einstellungen später verändern lassen.

Art des Banners

Die Tatsache, dass einige Webseiten nach wie vor kein Banner oder nur ein Info-Banner anzeigen, ist zunächst überraschend. Allerdings handelt es sich dabei in vielen Fällen um fremdsprachige Seiten aus dem Ausland wie z.B. aparat.com oder um solche mit pornografischen oder rechtlich fragwürdigen Inhalten (z.B. s.to). Nicht untersucht wurde, ob Seiten aus der Gruppe ohne jegliches Banner vielleicht nicht doch ausschließlich notwendige (oder gar keine) Cookies verwenden, aller Erfahrung nach dürfte dies aber eher die Ausnahme sein.

Aufwand

Die vorherrschende Ausgestaltung (70%; 154/219) des Consent-Banners als modaler Dialog, der die Webseite solange sperrt, bis eine Aktion der Nutzerin erfolgt ist, und die dabei häufige (82,8%; 127/154) Erschwernis der Ablehnung deuten darauf hin, dass Betreiber sich höhere Zustimmungsraten erwarten, indem sie Nutzenden die Möglichkeit nehmen, den Besuch der Seite mit geringem Aufwand fortzusetzen. Zu den Webseiten, die keine Möglichkeit bieten, Cookies abzulehnen, gehören Zeitungen und Nachrichtenmagazine wie z.B. spiegel.de, die aber einen alternativen Zugang gegen Bezahlung anbieten (so genannte *Tracking Wall*).

Dark Patterns im UI

Die große Mehrzahl (85,8%; 188/219) der Seiten mit Consent-Banner nutzt eine oder mehrere Techniken der UI-Gestaltung, um Nutzenden in ihrem Sinne zu beeinflussen. Dem gegenüber steht ein verschwindend geringer Anteil (3,2%) dieser Seiten, die dabei wenigstens eine 1-Klick-Ablehnung ermöglichen.

Wiederauffindbarkeit des Einstellungs-Dialogs

Nur bei 39,6% der Seiten genügen eine oder zwei Mausaktionen zum erneuten Aufruf des Dialogs, so dass sich Nutzende ihre beim ersten Seitenbesuch getroffenen Festlegungen ansehen und in der Folge ggf. anpassen können. Zwei Mausaktionen sind in dieser Hinsicht als akzeptabel anzusehen, da nicht erwartet werden kann, den Dialog von überall auf der Seite mit einem Klick zu erreichen und die erste Aktion meist im Scrollen ans Seitenende besteht, wo typischerweise z.B. auch ein Link zum Impressum zu finden ist.

Ohne Einwilligung gesetzte Cookies

Tab. 2 zeigt, wie viele persistente Cookies ohne Zutun des Nutzenden gesetzt werden. Da eine automatisierte Feststellung deren Zwecke nicht möglich war, werden hier nur 3rd Party Cookies gezählt, da diese viel eher für Werbung und Tracking eingesetzt werden als 1st Party Cookies. Zugunsten der Betreiber wird davon ausgegangen, dass eines der 3rd

Party Cookies vom Einsatz der Reichweitenmessung¹⁷ oder eines externen CMP-Tools herrührt und somit als „erforderlich“ gelten kann. Dennoch ist bei ca. einem Drittel der Seiten mit Consent-Banner das Verhalten hinsichtlich automatisch gesetzter 3rd Party Cookies zumindest auffällig und sollte näher überprüft werden.¹⁸ 3rd Party Cookies mit einer Gültigkeit von über einem Jahr wurden ohne Zutun nur in wenigen Fällen gesetzt, wenn Consent-Banner (7,8%; 17/219) bzw. Info-Banner (18%; 11/61) verwendet werden. Dagegen ist der Anteil von Seiten mit solch langlebigen Cookies mit 37,0% (78/211) deutlich höher in der Gruppe jener Seiten, die gar keine Banner enthalten.

5 Diskussion

Mit Hilfe eines teil-automatisierten Ansatzes wurden die Praktiken im Umgang mit Cookies der in Deutschland am häufigsten genutzten Webseiten untersucht. Wie erwähnt, sind die Ergebnisse mit einer gewissen Unschärfe behaftet. Einerseits wurde nicht überprüft, ob die Webseitenbetreiber tatsächlich dem Anwendungsbereich von RL bzw. DS-GVO unterliegen. Andererseits konnten die tatsächlichen Zwecke einzelner Cookies nicht festgestellt werden, so dass vereinfachende Annahmen getroffen werden mussten. Auch wurden ausschließlich Cookies als Technologie berücksichtigt und andere Methoden wie *Tracking Pixels* oder *Browser Fingerprinting* außer Acht gelassen. Dennoch lassen sich einige Erkenntnisse ableiten, die bei der Beurteilung der Frage, welche Gestaltungs- und Realisierungsvarianten rechtlich zulässig sind, Berücksichtigung finden sollten:

- Consent-Banner werden vorwiegend als **modale Dialoge** ausgestaltet, die die unmittelbare Aufmerksamkeit der Nutzenden verlangen. Ein Konflikt könnte sich dadurch mit *Erwägungsgrund 32 Satz 6* der DS-GVO ergeben, welcher fordert, dass bei Einwilligungen auf elektronischem Wege die Aufforderung hierzu ohne unnötige Unterbrechung des Dienstes zu erfolgen hat. Während die Nutzenden beim Besuch der Seite ohne Einschränkungen der Funktionalität auf die nicht erforderlichen Cookies verzichten könnten, wird ihnen der bequemste Weg, das Banner zu ignorieren, wenn sie keine Einwilligung erteilen möchten, verwehrt.
- In den meisten Fällen ist eine **Verweigerung der Einwilligung aufwändiger als ihre Erteilung**, da dafür zwingend die zweite UI-Schicht aufgerufen werden muss. Es ist zu vermuten, dass sich die Betreiber dadurch höhere Zustimmungsraten erhoffen, da die Annahme nur mit einem Klick die für die Nutzenden aufwandsärmste Methode darstellt.¹⁹ Angesichts der Tatsache, dass es technisch problemlos möglich

¹⁷ z.B. durch INFOnline

¹⁸ In [Tr19] wurde für im Zusammenhang mit der Identifikation von Tracking Cookies eine Heuristik verwendet, die darunter 3rd Party Cookies mit einer Gültigkeit von mindestens einem Monat versteht. Unter dieser Einschränkung reduziert sich im vorliegenden Datenmaterial der Anteil der Webseiten mit Consent-Banner und mindestens zwei 3rd Party Cookies auf lediglich 6%.

¹⁹ Ein solcher Effekt auf die Entscheidung der Nutzenden wurde etwa in [No20] nachgewiesen.

ist, eine gleichrangige Wahl zwischen der kompletten Ablehnung und Annahme sowie Detailsinstellungen²⁰ anzubieten, muss hinterfragt werden, inwieweit der bestehende Mechanismus den Grundsätzen von *Privacy by Design* und *Privacy by Default* (Art. 25 DS-GVO) genügt – wird doch die datenschutzfreundlichste Alternative bewusst verkompliziert und die für die Nutzenden ungünstigste als die am einfachsten zugängliche gestaltet. Ab wann der unbestreitbar größere (zeitliche und kognitive) Aufwand für den Nutzenden schon als Nachteil, der einer „echten und freien“ Wahl nach *Art. 4 Nr. 11 und Erwägungsgrund 42 Satz 5 DS-GVO* entgegensteht, zu werten ist, muss von der Rechtsprechung noch geklärt werden. In der Literatur und bei einigen Aufsichtsbehörden im europäischen Ausland wird diese Praxis [RE21] zufolge als unzulässig betrachtet. Die Analogie zum unerlaubterweise vorausgewählten Kästchen liegt jedenfalls nicht allzu fern (vgl. Abschnitt 2.2).

- Die sehr häufig festzustellende **irreführende optische Gestaltung** der Dialoge verleitet Besucherinnen und Besucher der Webseiten, eine Auswahl zu treffen, die ihren Interessen zuwiderläuft, wie auch Nutzerstudien gezeigt haben [Ut19, MB20]. Neben dem Grundsatz der Verarbeitung nach Treu und Glauben (*Art. 5 Abs. 1 lit. a DS-GVO*) berührt dies die grundlegenden Anforderungen aus *Art. 4 Nr. 11 DS-GVO*, dass eine Einwilligung informiert und unmissverständlich zu erfolgen hat.²¹
- Während in fast allen Fällen eine Einwilligung mit einem Klick erteilt werden kann, ist der **spätere Widerruf deutlich aufwändiger**, da mehrere Mausaktionen benötigt werden, um den Dialog erneut aufzurufen und dann dort die Einstellungen zu modifizieren. *Art. 7 Abs. 3 Satz 4 DS-GVO* fordert jedoch ausdrücklich, dass der Widerruf genauso einfach wie die Erteilung sein muss.
- Angesichts der **sehr großen Anzahl von 3rd Party Cookies**, die von manchen Webseiten automatisch gesetzt werden, ist es fraglich, ob all diese tatsächlich „erforderlich“ im Sinne von *Art. 5 Abs. 3 RL* sind und etwa einem der in Abschnitt 2.4 aufgeführten Zwecke dienen. Mit dem Grundsatz der Datenminimierung (*Art. 5 Abs. 1 lit. c DS-GVO*) dürfte jedenfalls eine exzessive Übermittlung (in einem Fall wurden auf einer Seite mit Consent-Banner über 130 persistente 3rd Party Cookies gesetzt, ohne die Nutzerinteraktion abzuwarten) personenbezogener Daten ebenfalls schwerlich in Einklang zu bringen sein. (Möglicherweise lässt sich das Phänomen im Einzelfall aber auch schlicht durch eine nachlässige Implementierung erklären, bei der versäumt wurde, das Setzen von Cookies bis zur erteilten Einwilligung zu unterdrücken.)

²⁰ Diese Möglichkeit könnte aufgrund von Erwägungsgrund 32 Satz 5 geboten sein.

²¹ Angesichts der sehr umfangreichen Informationen, die Dialoge bei manchen CMPs enthalten, drängt sich im Übrigen auch der Verdacht auf, dass damit der psychologische Effekt ausgenutzt werden soll, dass Nutzende im Lichte einer überwältigenden Zahl von Einstellungsmöglichkeiten oft darauf verzichten, überhaupt eine Entscheidung zu treffen [RS20].

6 Fazit

Unter den untersuchten Webseiten, die Consent-Banner einsetzen, verzichtet nur eine Minderheit auf den Einsatz von mehr oder minder subtilen Mechanismen der Beeinflussung von Nutzenden, indem sie die typischerweise zum Einsatz kommenden CMP-Tools entsprechend konfiguriert. Erschwerend kommt hinzu, dass es kein gängiges Format für die Einwilligungsdialoge gibt (wie es etwa der Fall ist bei Berechtigungsanfragen des Browsers für Standortabfragen oder Kamera- und Mikrofonfreigaben). Selbst wenn die Praktiken der Webseitenbetreiber – nach gegenwärtiger Rechtslage – noch als zulässig gelten können, sind die Notwendigkeit eines datenschutzfreundlichen Designs und der Bedarf für unterstützende Werkzeuge klar erkennbar, um den Bedürfnissen der Nutzenden gerecht zu werden. Als Ansätze in diese Richtung sind etwa der ab 2012 eingeführte, mittlerweile aber in der Umsetzung als gescheitert geltende, optionale „do not track“ (DNT)-Header im Browser²² zu nennen oder die bislang unberücksichtigten Vorschläge des EU-Parlaments in der Diskussion um die ePrivacy-Verordnung.²³

Die weitere technische Entwicklung sollte aber beobachtet werden, denn die Bedeutung von Cookies im Allgemeinen und 3rd Party Cookies im Besonderen könnte weiter nachlassen. Beispielsweise werden sie im Safari-Browser standardmäßig blockiert, während Google mit *FLoC* (Federated Learning of Cohorts) eine Technologie plant, bei der Nutzende aufgrund ihres Surfverhaltens durch den Browser selbst einer werberelevanten Zielgruppe zugeordnet werden sollen.

Literaturverzeichnis

- [BH20] Böhm, W.; Halim, V.: Cookies zwischen ePrivacy und DS-GVO – was gilt?, MMR, S. 651-656, 2020.
- [Gi20] Gierschmann, S.: Verwendung personenbezogener Daten – Cookie-Einwilligung II, MMR, S. 609-617. 2020.
- [KTS20] Kettner, S.; Thorun, C.; Spindler, G.: Innovatives Datenschutz-Einwilligungsmanagement, https://www.bmjv.de/SharedDocs/Downloads/DE/News/PM/090720_Datenschutz.html?nn=6705022, Stand: 7.9.2020.
- [Kl20] Kleinz, T.: Frankreich: Datenschützer verhängen Millionen-Bußgelder gegen Google und Amazon, heise online, <https://heise.de/-4985956>, News vom 10.12.2020.
- [LS21] Luguri, J.; Strahilevitz, L. J.: Shining a light on dark patterns. *Journal of Legal Analysis* 13.1, S. 43-109, 2021.
- [MB20] Machuletz, D.; Böhme, R.: Multiple Purposes, Multiple Problems: A User Study of Consent Dialogs after GDPR. *Proc. on Privacy Enhancing Technologies*, S. 481-498, 2020.

²² <https://www.w3.org/TR/tracking-dnt/>

²³ Amendment 106 ff. zu Artikel 10, https://www.europarl.europa.eu/doceo/document/A-8-2017-0324_EN.html (Stand: 20.10.2017)

- [MBS20] Matte, C.; Bielova, N.; Santos, C.: Do Cookie Banners Respect my Choice?: Measuring Legal Compliance of Banners from IAB Europe's Transparency and Consent Framework. Proc. IEEE Symposium on Security and Privacy, S. 791-809, 2020.
- [No20] Nouwens, M. et al.: Dark Patterns after the GDPR: Scraping Consent Pop-ups and Demonstrating their Influence. Proc. CHI Conference on Human Factors in Computing Systems, S. 194-206, 2020.
- [RE21] Rauer, N.; Ettig, D.: Update Cookies 2020 / Aktuelle Rechtslage und Entwicklungen, ZD, S. 18-24, 2021.
- [RS20] Rieger, S.; Sindere, C.: Dark Patterns: Design mit gesellschaftlichen Nebenwirkungen. Stiftung Neue Verantwortung, <https://www.stiftung-nv.de/sites/default/files/dark.patterns.pdf>, 13.05.2020.
- [Sa19] Sanchez-Rola, I. et al.: Can I Opt Out Yet? GDPR and the Global Illusion of Cookie Control. Proc. ACM Asia Conference on Computer and Communications Security, S. 340-351, 2019.
- [So20] Soe, T.H. et al.: Circumvention by design – dark patterns in cookie consent for online news outlets. Proc. Nordic Conference on Human-Computer Interaction, S. 1-12, 2020.
- [Sp20] Spindler, G.: Klarheit für Cookies, NJW, S. 2513-2517, 2020.
- [Tr19] Trevisan, M. et al.: 4 Years of EU Cookie Law: Results and Lessons Learned. Proc. on Privacy Enhancing Technologies, S. 126-145, 2019.
- [Ut19] Utz, C. et al.: (Un) informed consent: Studying GDPR consent notices in the field. Proc. ACM SIGSAC Conf. on Computer and Communications Security, S. 973-990, 2019.

Anhang

Aufwand		(absolut)
<i>Annehmen mit 1 Klick</i>	99,5%	218
<i>Ablehnen möglich</i>	95,4%	209
mit 1 Klick	20,1%	44
2 Klicks	53,0%	116
3 oder mehr Klicks	22,4%	49
<i>Ablehnen nicht möglich</i>	4,6%	10
Dark Patterns im UI		
<i>Farbe</i>	82,6%	181
<i>Größe</i>	36,1%	79
<i>Anordnung</i>	24,7%	54
mindestens eine Form der <i>Beeinflussung</i>	85,8%	188
Wiederauffindbarkeit des Einstellungs-Dialogs		
<i>möglich</i>	87,7%	192
mit 1 Mausaktion	4,6%	10
2 Mausaktionen	39,7%	87
3 Mausaktionen	16,9%	37
4 Mausaktionen	20,5%	45
5 oder mehr	5,9%	13
lediglich <i>Hinweis</i> auf Browser- Einstellungen oder Drittanbieter	11,0%	24
<i>keine Möglichkeit gefunden</i>	1,4%	3

Tab. 1: Ergebnisse für Webseiten mit Consent-Banner ($n = 219$).

Art des Banners	absolut	keine	Anzahl			
			1	2-4	5-20	20+
alle	491	40,5%	15,7%	18,5%	18,7%	6,5%
<i>Consent-Banner</i>	219	44,3%	22,4%	19,2%	10,5%	3,7%
modal	154	46,1%	22,1%	17,5%	10,4%	3,9%
nicht-modal	65	40,0%	23,1%	23,1%	10,8%	3,1%
<i>Info-Banner</i>	61	36,1%	9,8%	19,7%	21,3%	13,1%
modal	3	66,7%	0%	33,3%	0%	0%
nicht-modal	58	34,5%	10,3%	19,0%	22,4%	13,8%
ohne Banner	211	37,9%	10,4%	17,5%	26,5%	7,6%

Tab. 2: Ohne aktive Zustimmung gesetzte persistente 3rd Party Cookies.

Mithilfe von Privacy Nudging zu rechtsverträglichen Videokonferenztools

Sabrina Schomberg¹, Ernestine Dickhaut², Torben Jan Barev³ und Andreas Janson⁴

Abstract: Die Covid-19 Pandemie hat zu enormen Veränderungen im Arbeitsalltag geführt. Videokonferenztools gewannen durch die Zusammenarbeit über Distanz an Bedeutung. In kürzester Zeit kamen in vielen Unternehmen Videokonferenztools verschiedener Hersteller zum Einsatz; der Datenschutz stand dabei in der Regel nicht an erster Stelle und es kam zu zahlreichen Datenschutzverstößen. Im Laufe der Pandemie rückte der Datenschutz mehr in den Fokus und es wurde nachgebessert. Wir wollen mit unserem Beitrag anregen Videokonferenztools nicht nur datenschutzkonform, sondern rechtsverträglich zu gestalten. Daher haben wir die Benutzeroberfläche eines Tools (Zoom) an den Stellen angepasst, an denen datenschutzsensibles Verhalten besonders kritisch ist. Wir zeigen, wie kleine Änderungen in der Benutzeroberfläche und die Integration von Privacy Nudges zu datenschutzsensiblen Verhalten führen können. Um diese sogenannten Privacy Nudges für Entwickelnde von Videokonferenztools so praktisch und zugänglich wie möglich zu machen, greifen wir auf das Konzept der Design Pattern aus der Systementwicklung zurück.

Keywords: Videokonferenztools, Privacy Nudging, Design Pattern, Rechtsverträglichkeit.

1 Einleitung

Die Covid-19 Pandemie hat in vielen Bereichen zu einem enormen Fortschritt der Digitalisierung geführt. Viele Betriebe haben sich bereits während der ersten Welle dazu entschieden einen Großteil der Mitarbeitenden von einem auf den anderen Tag ins Homeoffice zu schicken oder zumindest diese Möglichkeit zu eröffnen. Damit einhergehend wurden in kürzester Zeit neue Wege zur Zusammenarbeit und Interaktion über Distanz notwendig [BQS21] und stellten die Unternehmen vor Herausforderungen. Videokonferenztools diverser Anbieter wie beispielsweise Microsoft Teams, Jitsi und Zoom gewannen durch neue Arbeitsverhältnisse über Nacht an Bedeutung [Wa20],

¹ Universität Kassel (DE), FG Öffentliches Recht, IT-Recht und Umweltrecht von Prof. Dr. Hornung, Henschelstraße 4 (K33), 34127 Kassel, sabrina.schomberg@uni-kassel.de.

² Universität Kassel (DE), FG Wirtschaftsinformatik von Prof. Dr. Leimeister, Pfannkuchstraße 1 (ITeG), 34121 Kassel, ernestine.dickhaut@uni-kassel.de.

³ Universität Kassel (DE), FG Wirtschaftsinformatik von Prof. Dr. Leimeister, Pfannkuchstraße 1 (ITeG), 34121 Kassel, torben.barev@uni-kassel.de.

⁴ Universität St.Gallen (CH), Institut für Wirtschaftsinformatik, Müller-Friedberg-Strasse 8 9000 St. Gallen, andreas.janson@unisg.ch.

[AA21]. Der Videokonferenzanbieter Zoom erwirtschaftet beispielsweise allein im ersten Quartal des Jahres 2021 etwa 328 Millionen US-Dollar, was eine Steigerung von 169 Prozent gegenüber dem Vorjahr bedeutet [Bo21]. Zahlen wie diese bestätigen die Relevanz und Abhängigkeit vieler Unternehmen von Anbietenden diverser Videokonferenztools.

Datenschutz war in der ad hoc Etablierung der Konferenztools oft nicht die erste Priorität. Es folgten negative Pressemeldungen und aufgedeckte Rechtsverstöße diverser Anbieter. Dadurch wurde auch der Datenschutz mehr diskutiert und es wurden Anpassungen vorgenommen. Lücken und datenschutzrechtliche Stolperfallen sind jedoch nach wie vor vorhanden [Be21], insbesondere wenn diese Tools im Rahmen der Arbeitstätigkeit genutzt werden. Die Arbeitgebenden sind jedoch auch im Homeoffice Verantwortlicher im Sinne der DSGVO [Wü20] und setzten sich der Gefahr von Bußgeldern aus, wenn sie diese Lücken nicht schließen. Prognosen zufolge werden Videokonferenzen auch nach der Pandemie langen Dienstreisen vorgezogen werden. Daher lohnt sich ein zweiter Blick auf vorhandene Videokonferenztools, die im Eiltempo auf den Markt gebracht wurden und bereits ersten rechtlichen Nachjustierungen unterlagen.

Um Videokonferenztools nachhaltig zu gestalten, nutzen wir das Konzept der Privacy Nudges, um diese Tools nicht bloß datenschutzkonform auszugestalten, sondern sogar rechtsverträglich [Ro93] zu designen. Ist das Ziel während der Systementwicklung, das Minimum rechtlicher Konformität gerade noch zu erreichen, um beispielsweise Sanktionen zu entgehen, spricht man von Rechtmäßigkeit. Liegt das Ziel jedoch darin die rechtlichen Anforderungen möglichst gut und umfassend zu erfüllen, wird dies mit Rechtsverträglichkeit bezeichnet. Das Konzept der Rechtsverträglichkeit hat das Ziel, nicht das kurzfristige Minimum, sondern das langfristige Optimum an Grundrechtsschutz zu gewährleisten [Th20]. Um Technik rechtsverträglich zu gestalten, werden dazu die relevanten Rahmenbedingungen analysiert und vorausschauend in die Systementwicklung integriert. Auf diese Weise soll die Zusammenarbeit über solche Tools für die Arbeitgebenden, und somit den datenschutzrechtlich Verantwortlichen, möglichst langfristig ohne rechtliche Bedenken unterstützt werden. Durch Nudging („Anstupsen“) sollen die Nutzenden der Videokonferenztools dazu gelenkt werden möglichst wenig Daten preiszugeben. Um diese sog. Privacy Nudges möglichst praktisch anwendbar und für Entwickelnde von Videokonferenztools verständlich zu machen, greifen wir auf das Konzept der Design Pattern aus der Systementwicklung zurück [A117] und erstellen Design Pattern für die rechtsverträgliche Gestaltung von Videokonferenztools. Mithilfe der Design Pattern wollen wir eine Weitergabe des rechtlichen Wissens in die praktische Systementwicklung ermöglichen. Damit ist das übergeordnete Ziel dieses Beitrags, im Rahmen eines interdisziplinären Ansatzes, an der Schnittstelle von Recht und Informatik, die Umsetzung rechtsverträglicher Videokonferenztools mit Privacy Nudging zu beschreiben und dieses Wissen in Design Pattern zu kodifizieren und Entwickelnden zugänglich zu machen.

2 Theoretischer Hintergrund

2.1 Beschäftigtendatenschutz und Videokonferenztools

Das Thema „Homeoffice“ hat durch die Pandemie einen enormen Aufwind erfahren. Dabei lassen sich arbeitsrechtlich weitere Unterscheidungen treffen zwischen häuslicher und alternierender Telearbeit (im Privatbereich der Beschäftigten fest eingerichteter Bildschirmarbeitsplatz i. S. d. § 2 Abs. 7 S. 1 ArbStättV, welcher für die gesamte oder Teile der vereinbarten wöchentlichen Arbeitszeit als Ort der Erbringung der Arbeitsleistung festgelegt wurde) sowie mobiler Arbeit (ortsungebundenes Arbeiten an wechselnden Orten wie z. B. im Zug, im Café oder am Küchentisch) [Mü20], [KRP20]. In der Pandemie wurden vielfach Beschäftigte, die nur für die mobile Arbeit ausgestattet sind, mit ihren Laptops nach Hause geschickt, weswegen der Begriff Homeoffice in der Pandemie oft alle dieser drei Formen umfasst [DH20].

Für das Recht oder die Pflicht der Beschäftigten, dienstliche Tätigkeiten von zu Hause zu erledigen, bedarf es einer wirksamen Rechtsgrundlage. Da jedoch auch viele Beschäftigte in der Pandemie ein Interesse daran haben von zu Hause zu arbeiten dürfte es überwiegend zu einvernehmlichen Regelungen, die auch ad hoc oder konkludent möglich sind, gekommen sein [Su20]. Gerade bei einer plötzlichen Umstellung auf Homeoffice durch die Pandemie, waren die technischen und organisatorischen Voraussetzungen oft noch nicht gegeben, obwohl die Beschäftigten längst zu Hause vor den mühsam zusammengesuchten und im Eilverfahren bestellten Laptops oder sogar privaten Endgeräten saßen [Su20]. Die DSGVO kennt jedoch keinen Ausnahmezustand, wengleich die Aufsichtsbehörden eine „gewisse Nachsicht“ für die Zeit der Umstellung versprochen [Ku20], [Wü20]. Mittlerweile prüfen die Datenschutzbehörden jedoch wieder ohne Nachsicht [Su20]. Und zwar auch bei der „Technikgestaltung“ von Softwareprodukten [KK21].

Auch im Homeoffice sind die Arbeitgebenden nach wie vor Verantwortlicher i. S. d. Art. 4 Nr. 7 DSGVO für jede Datenverarbeitung, die im Rahmen der Tätigkeit der Beschäftigten im Homeoffice ausgeführt wird [Wü20], [Su20]. Im Homeoffice gibt es sogar eine Reihe von zusätzlichen Stolperfallen des Beschäftigtendatenschutzes [VE20]. Dadurch, dass die Arbeitnehmenden von zu Hause, und somit aus ihrem Lebensmittelpunkt heraus, arbeiten, besteht ein erweitertes Risiko für die Privatheit der Beschäftigten. Betriebliche und private Sphären könnten vermischt und Aspekte der Privatheit unwillentlich mitverarbeitet werden. Die Gefahr der unfreiwilligen Preisgabe von Informationen über die Familie, Wohnungseinrichtung, Interessen durch Poster oder Bilder an der Wand besteht besonders durch Videokonferenztools ohne persönlichkeitschützende Maßnahmen [Su20], [Wü20].

Viele Unternehmen standen vor der Herausforderung sich schnell für ein Videokonferenztool entscheiden zu müssen, damit die Beschäftigten auch über die Distanz zusammenarbeiten können. Dabei wurde vermehrt zu schon bekannten und oft

amerikanischen Anbietern wie Microsoft Teams, Skype for Business oder Zoom gegriffen, welche Anforderungen der DSGVO regelmäßig nicht erfüllen [Be21], [Su20]. Außerdem entschied der EuGH im Juli 2020, dass der Beschluss der Kommission über die Angemessenheit des EU-US-Privacy-Shields ungültig ist und somit die Übermittlung personenbezogener Daten aus der EU an einen Datenempfänger in den USA mit sofortiger Wirkung unzulässig ist [Eu20]. Dies führte sogar zu verschiedenen Varianten von Videokonferenztools, wie z. B. einer DSGVO-konformen Variante speziell für den europäischen Raum, die sich hinsichtlich der Datenspeicherung und -verarbeitung unterscheidet. So hat unter anderem Zoom im Jahr 2020 auf Kritik reagiert und beispielsweise sprachliche Korrekturen hinsichtlich der fälschlichen Verwendung des Begriffs „Ende-zu-Ende-Verschlüsselung“ durchgeführt [Ro20], [To20].

Um ein Videokonferenztool im Homeoffice DSGVO-konform einzusetzen müssen verschiedene Anforderungen der DSGVO beachtet werden. Mit dem Diensteanbieter (hier Zoom), muss (je nach konkreter Ausgestaltung [PP21]) zunächst ein Auftragsverarbeitungsvertrag i. S. d. Art. 28 DSGVO geschlossen werden [St20], [Su20]. Dabei ist insbesondere ein angemessenes Datenschutzniveau bei Dienst anbietenden aus Drittländern zu gewährleisten. Seit dem Schrems-II Urteil des EuGH gibt es jedoch keinen Spielraum mehr für eine Übermittlung von Kundendaten in die USA auf Grundlage des Privacy Shield, da dieses wegen der dortigen Vorgaben zur Massenüberwachung für unzulässig erklärt wurde [Go20]. Bei der Nutzung von Videokonferenzlösungen sind darüber hinaus insb. die Anforderungen der Art. 24, 25 und 32 DSGVO einzuhalten [St20], [Wü20]. Zur Umsetzung der Vorgaben der DSGVO im Unternehmen gehört es dabei auch die Softwareprodukte so zu konfigurieren, dass diese für die konkrete Nutzung im Unternehmen alle Vorgaben einhalten [KK21].

Art. 25 DSGVO wird als Konkretisierung der Pflicht zur Umsetzung technischer und organisatorischer Maßnahmen durch den Verantwortlichen (Art. 24 DSGVO) verstanden [Ma18], [Ha18]. Adressat des Art. 25 DSGVO ist ausdrücklich nur der Verantwortliche, nicht jedoch der Hersteller von Verarbeitungstechnik. Für den Hersteller besteht daher grundsätzlich keine Pflicht zur datenschutzfreundlichen Ausgestaltung seiner Produkte. Er wird lediglich durch Erwägungsgrund 78 S. 4 dazu „ermutigt“. Aber die Verantwortlichen werden in der Regel nur solche Produkte kaufen, die den Anforderungen der DSGVO, insb. des Art. 25 DSGVO, gerecht werden, um sich nicht der Gefahr eines hohen Bußgeldes gem. Art. 83 Abs. 4 lit. a DSGVO auszusetzen. Denkbar ist insofern also eine mittelbare Wirkung, indem die Nachfrage nach datenschutzfreundlicher Technologie die Hersteller zur Umsetzung eben dieser Anforderungen zwingt [Ma18]. Ebenfalls denkbar ist, dass der Arbeitgeber den Softwareentwickler im Rahmen der Mängelgewehrleistung in Haftung nehmen kann, wenn die Software nicht den datenschutzrechtlichen Vorgaben entspricht [Dü19], [Ha20]. Privacy Nudges können helfen, den weiten Anwendungsbereich von Art. 25 DSGVO mit Leben zu füllen [Sc19].

Gem. Art. 32 DSGVO müssen außerdem technische und organisatorische Maßnahmen getroffen werden, die die Sicherheit der Verarbeitung gewährleisten. Die Norm richtet sich sowohl an den Verantwortlichen als auch an den Auftragsverarbeiter. Die

Nichteinhaltung ist gem. Art. 83 Abs. 4 lit. a DSGVO ebenfalls bußgeldbewehrt und gem. Art. 82 DSGVO tragen sowohl der Verantwortliche als auch der Auftragsverarbeiter, jeder für seinen Verantwortungsbereich, ein Haftungsrisiko gegenüber den betroffenen Personen [La19]. Konkretere Angaben zu den erforderlichen Maßnahmen enthalten lediglich Art. 32 Abs. 1 lit. a-d, Abs. 4 DSGVO und Erwägungsgrund 83. Art. 32 Abs. 1 lit. a DSGVO und Erwägungsgrund 83 nennen zunächst Verschlüsselung - und Art. 32 Abs. 1 lit. a DSGVO darüber hinaus auch Pseudonymisierung - als technische Maßnahme der Risikominimierung. Auch wenn Zoom selbst als Hersteller nur dazu „ermutigt wird“ (ErwG 78 S. 4 DSGVO) seine Produkte datenschutzfreundlich auszugestalten und ihn als (ggf.) Auftragsverarbeiter bzgl. der Anforderungen aus Art. 25 DSGVO eine mittelbare Verpflichtung trifft, so kann der Videokonferenzanbieter jedoch teilweise über Art. 32 DSGVO, welche eher die Bereiche IT-Sicherheit betrifft, direkt zur Rechenschaft gezogen werden. Dies erscheint jedoch auch sinnvoll, da der Verantwortliche in diesem Bereich keinerlei Möglichkeit hat durch einfaches Konfigurieren datenschutzrechtliche Lücken zu schließen. Ziel soll sein, dass ein datenschutzkonformes Homeoffice die Persönlichkeit der Beschäftigten nur so weit offenbart, wie schon deren betrieblicher Arbeitsplatz [Wü20]. Dies können Arbeitgebende maßgeblich erreichen, indem den zur Verfügung gestellten Geräten für das Homeoffice durch eine datenschutzfreundliche Technikgestaltung gem. Art. 25 und Art. 32 DSGVO „datenschutzrechtliche Scheuklappen“ aufgesetzt werden [Wü20].

Die allgemeinen Anforderungen der DSGVO und die hier besonders hervorgehobenen Anforderungen der Technikgestaltung müssen eingehalten werden. Dabei können die sehr abstrakten Normen zum „Datenschutz durch Technik“ für Unternehmen ein Risiko darstellen, da schwer sicher zu sagen ist, wann dem risikobasierten Ansatz der DSGVO genüge getan und ein Bußgeld sicher ausgeschlossen ist. Andererseits bieten die abstrakten Regelungen, gerade im Bereich der Technikgestaltung, aber auch Chancen. Es bleibt viel Raum für individuelle Umsetzungen und die Möglichkeit das Recht von morgen mit der Technikgestaltung von heute zu beeinflussen. Das Konzept der Rechtsverträglichkeit berücksichtigt dabei die Veränderung der Rechtsordnung in der Zukunft, am Maßstab heutiger Rechtsziele [Ro93].

2.2 Nudging

Individuen entscheiden sich oft irrational und zum eigenen Nachteil [Ac17]. Eine Möglichkeit Nutzende bei der Entscheidungsfindung in digitalen Arbeitssystemen zu unterstützen, sind sogenannte Nudges (vgl. auch [DR20]). *Thaler* und *Sunstein* definieren Nudging als unaufdringliche und die Entscheidungsfreiheit bewahrende Form der Entscheidungsbeeinflussung [TS08]. Was gewählt wird, hängt oft davon ab, wie die Entscheidungen präsentiert werden [WSV16]. Das Konzept Nudging kann demnach eine Vielzahl von Ansätzen beinhalten, um Entscheidungen zu beeinflussen. Laut *Kahnemanns* Dualprozess-Theorie nutzen Individuen zwei Denksysteme, um die Informationsfülle in der heutigen (digitalen) Welt besser zu bewerten und zielgerichtete Entscheidungen

treffen zu können. Das automatische Denksystem (System 1) repräsentiert unsere Intuitionen oder unseren unbewussten Autopiloten. Das reflektierte Denksystem (System 2) hingegen drückt sich durch unsere bewusste Planung und Kontrolle aus. Nudges können hingegen beide Denksysteme ansprechen [Ka03]. So können Nudges beispielsweise durch Aufmerksamkeitslenkungen oder Voreinstellungen, Individuen zu einem bestimmten Verhalten leiten und vorhersehbar steuern [Sc19].

Beim digitalen Nudging werden entsprechende Designelemente in der Benutzeroberfläche verwendet, um das Verhalten in digitalen Entscheidungsumgebungen zu leiten [WSV16]. Nudges verbieten hierbei keine Alternative oder verändern ökonomische Anreize einzelner Optionen [TS08]. Um als Nudge zu gelten, muss eine Intervention zudem leicht zu umgehen sein [TS08]. In der Software Zoom würden demnach alle Optionen zur Auswahl stehen. Das Individuum kann zu jeder Zeit frei entscheiden. Die Software würde jedoch die empfehlenswerte Option für das Individuum markieren oder vorauswählen. Durch diese Grundsätze sollen Nudges subtil zu Verhaltensänderungen führen und Individuen unterstützen. Nudges werden als eine freiheitsbewahrende (libertäre) Form des Paternalismus beschrieben, da versucht wird, gezielt Einfluss auf Entscheidungen von Menschen zu nehmen; den Entscheidenden allerdings nicht die Freiheit genommen wird, zu tun, was ihnen beliebt. Eine Unterform der digitalen Nudges sind hierbei die sogenannten Privacy Nudges. Privacy Nudging beschreibt eine gezielte Beeinflussung des Entscheidungsprozesses, um Menschen dazu zu bringen, dass diese „bessere“ Entscheidungen in Bezug auf deren Privatheit treffen und gleichzeitig ihre informationelle Selbstbestimmung berücksichtigen [Ac17], [SK18].

2.3 Design Pattern für die Weitergabe von rechtlichem Gestaltungswissen

Rechtliches Gestaltungswissen gewinnt in der Systementwicklung, insbesondere durch die DSGVO, immer mehr an Bedeutung. Heutzutage gehören rechtliche Anforderungen zu einem festen Bestandteil neuer Technologien. Die Allgegenwertigkeit heutiger Technologien ermöglicht die Sammlung und Verarbeitung einer großen Menge an personenbezogenen Daten, die besonders schützenswert sind. Jedoch fehlt Entwickelnden häufig rechtliches Fachwissen, um diese Anforderungen entsprechend umzusetzen. Hierbei können Design Pattern helfen, indem sie Lösungen für wiederkehrende rechtliche Gestaltungsherausforderungen bereitstellen und Entwickelnden bei der Umsetzung leiten [YSJ15], [Ro19], [Ha06].

Design Pattern haben ihren Ursprung in der Architektur, indem sie Lösungen für wiederkehrende Gestaltungsprobleme sammeln [Al77] und haben sich in den 90er Jahren ebenfalls in der Systementwicklung etabliert [Ga95]. Design Patterns bestehen aus Schablonen, in denen Informationen möglichst übersichtlich und einheitlich dargestellt werden. Neben der Bereitstellung von bewährten Lösungen, sind Design Pattern eine Möglichkeit, um komplexes Gestaltungswissen zugänglich und anwendbar zu machen [DJL21].

3 Praktische Anwendungsfälle

Um die Tauglichkeit des Privacy Nudging Konzepts aufzuzeigen, widmen wir uns im Folgenden der Gestaltung der Benutzeroberfläche von Zoom. Hierbei fokussieren wir die Stellen der Entscheidungsarchitektur, an denen ein datenschutzsensibles Verhalten besonders wichtig ist. Im Folgenden stellen wir zwei neue und angepasste Benutzeroberflächen vor, in denen digitale Privacy Nudges implementiert sind und die Nutzenden dazu angehalten werden weniger personenbezogene Daten preiszugeben.

3.1 Videovorschau mit Empfehlung der Nutzung eines virtuellen Hintergrundes

Art. 25 Abs. 2 DSGVO schreibt datenschutzfreundliche Voreinstellungen vor. Ton und Bild sollten daher bei Betreten eines Meetings standardmäßig ausgeschaltet sein, um die Beschäftigten davor zu schützen, unfreiwillige Eindrücke oder private Gespräche mit Familienmitgliedern zu teilen. Durch das bewusste Anschalten wird die Gefahr minimiert, dass die Beschäftigten zu spät bemerken, dass das Meeting bereits begonnen hat. Ein Meeting ganz ohne Bild und Ton eignet sich vor allem in kleinen Gruppen weniger, da Mimik und Gestik auch über die Distanz Nähe aufbauen und zu mehr Verständnis des Gegenübers führen können.

Doch auch mit diesem Schutz des bewussten Anschaltens kann der Konferenzhintergrund persönliche Informationen, wie z.B. Urlaubsfotos, Familienmitglieder oder individuelle Einrichtung der Beschäftigten preisgeben und sollte daher mit Bedacht gewählt werden. Bei der Einwahl in ein Zoom Meeting wird den Nutzenden in unserem angepassten Prozess daher als zusätzliche Schutzmaßnahme standardmäßig ein virtueller Hintergrund vorausgewählt (siehe Abbildung 1).

Dies kann ein neutraler Hintergrund sein, wie z.B. ein virtuelles Büro oder ein Unternehmenshintergrund, der verschiedene Marken Aspekte des Unternehmens berücksichtigt. In diesem Szenario werden daher sogenannte Default Privacy Nudges genutzt, um die Privatsphäre der Nutzer zu schützen. Default Privacy Nudges beschreiben Standardeinstellungen im System. Da Nutzer in digitalen Umgebungen die Privatsphäre-Einstellungen häufig nicht ihren Bedürfnissen anpassen, bleibt die voreingestellte Option (der Status quo) übermäßig bevorzugt und meist unverändert (Status quo Bias) [Ac17]. In diesem Fall wird der neutrale virtuelle Hintergrund in Zoom eingestellt. *Hummel* und *Maedche* bewerten Defaults tendenziell als die stärksten Nudges [HM19]. Diese Defaults werden als sehr angenehm wahrgenommen [Sc20].

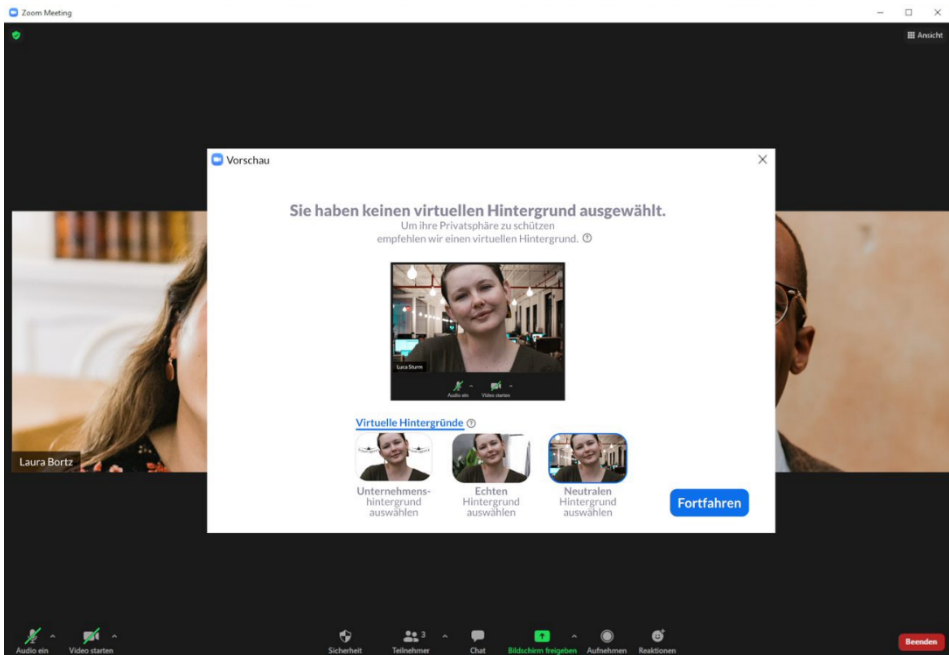


Abbildung 1. Empfehlung eines virtuellen Hintergrundes

In Bezug auf Privacy Nudging gelten Defaults als sehr effektiv, da sie in digitalen Arbeitssystemen standardmäßig das Maß der Datensparsamkeit vorgeben [Ac17]. Diese Standardeinstellungen und ähnliche leitende Maßnahmen könnten zu Reaktanz führen und dadurch eine gegenteilige Wirkung haben [BSJ21]. Reaktanz bezeichnet einen motivationalen Erregungszustand, der auftritt, wenn sich jemand in seiner Freiheit bedroht fühlt und führt zu dem Versuch die eigene Freiheit wiederherzustellen [WB75]. Einige Studienergebnisse suggerieren dies [Wa14], [Le16]. Daraus folgt die Forderung nach Transparenz. Auch für *Sunstein* gilt eine Handlung als manipulativ, wenn sie nicht transparent ist [Su15]. *Hausman* und *Welch* sind der Ansicht, dass Menschen zur Wahrung ihrer Autonomie und zum Schutz vor Missbrauch durch Nudges über solche Interventionen informiert werden sollten [HW10]. Nudges sollten demnach so gestaltet sein, dass es grundsätzlich jedem, der aufmerksam ist, möglich ist, den Nudge und die Intention der Entscheidungsarchitektur zu bemerken [Bo09]. In dem vorliegenden Szenario wurde daher ein Hinweis eingefügt, der besagt, dass die vorliegenden Einstellungen so gewählt wurden, um die Privatsphäre der Nutzer, die schützen. Außerdem fordert auch die DSGVO Transparenz (Art. 5 Abs. 1 lit. a DSGVO). Daneben gibt es auch generelle, verfassungsrechtliche Einwände gegenüber Privacy Nudges, da die informationelle Selbstbestimmung auch das Recht umfasst, möglichst viele Daten von sich selbst preiszugeben [SK18]. Eine mögliche Legitimation kann jedoch darin gesehen werden, dass die informationelle Selbstbestimmung eine Funktionsbedingung einer

demokratischen Gesellschaft und somit ein Schutzgut der Allgemeinheit ist [Sa15], [SK18]. Im arbeitsrechtlichen Kontext sind die Arbeitgebenden darüber hinaus gem. § 26 Abs. 5 BDSG zum Schutz der Daten der Beschäftigten verpflichtet [Su20].

3.2 Erweiterung der Videoaufzeichnung um pseudonyme Option

Bisher können Nutzende des Videokonferenztools Zoom der Aufzeichnung eines Meetings in der Regel nur zustimmen oder das Meeting verlassen. Die Optionen der Nutzenden könnten durch einen dritten Button erweitert werden, welcher es ermöglicht mit nur einem Klick Ton und Bild auszuschalten und den Namen zu pseudonymisieren. Im Rahmen der digitalen Arbeit wären sensible Daten nun ausschließlich für eine bestimmte Zielgruppe oder in diesem Fall, nur für das Individuum selbst zugänglich. Art. 25 Abs. 1 DSGVO nennt die Pseudonymisierung (in Art. 4 Nr. 5 DSGVO legaldefiniert), sogar als Beispiel für eine geeignete technische und organisatorische Maßnahme. Erwägungsgrund 78 führt weiterhin aus, dass der Verantwortliche interne Strategien festlegen und Maßnahmen zur Datenminimierung, schnellen Pseudonymisierung und Transparenz ergreifen soll. Der betroffenen Person soll es durch diese Maßnahmen außerdem ermöglicht werden, die Verarbeitung personenbezogener Daten zu überwachen, und der Verantwortliche soll in die Lage versetzt werden, Sicherheitsfunktionen zu schaffen und zu verbessern. Auch wenn diese pseudonyme Option keine notwendige Anforderung nach der DSGVO ist, sie doch eine geeignete Maßnahme den Datenschutz durch Technikgestaltung umzusetzen und entspricht dem Rechtsziel. Datenschutz durch Technikgestaltung ist dabei jedoch keine einmalige Maßnahme, sondern ein stetig fortlaufender Prozess, welcher an sich wandelnde tatsächliche oder technische Gegebenheiten angepasst werden muss [La19]. Alle Maßnahmen sind unter Berücksichtigung des Stands der Technik, der Implementierungskosten und der Art, des Umfangs, der Umstände und der Zwecke der Verarbeitung sowie der unterschiedlichen Eintrittswahrscheinlichkeit und Schwere der mit der Verarbeitung verbundenen Risiken für die Rechte und Freiheiten natürlicher Personen auszuwählen und zu treffen. Dies greift unter anderem den risikobasierten Ansatz der DSGVO auf und begrenzt die Auswahl geeigneter technischer Maßnahmen [BG17], [BH17]. Dabei lässt sich mit durchdachten Benutzeroberflächen oft leicht ein bisschen mehr Datenschutz umsetzen als zwingend notwendig.

Farbelemente können die Aufmerksamkeit der Nutzenden bewusst auf die pseudonyme Option lenken und so als Privacy Nudge in die neue Benutzeroberfläche eingearbeitet werden. Die Entscheidungsalternative „Still teilnehmen“ wird durch die farbliche Hinterlegung verstärkt hervorgehoben (siehe Abbildung 2). Im aufgezeigten Szenario, wenn der Host das Zoom Meeting aufnehmen möchte, steht den Nutzern zur Auswahl, wie sie daran teilnehmen möchten. Die Vorteile der Farbelemente zeigen sich vor allem in der einfachen und kostengünstigen Umsetzung solcher Nudges, die das Individuum schnell und effektiv dazu bewegen, die eigenen Entscheidungen bezüglich des

Datenschutzes und der Privatsphäre zu überdenken. Individuen können durch diese Nudges zu einem datensparsamen Verhalten geleitet werden.

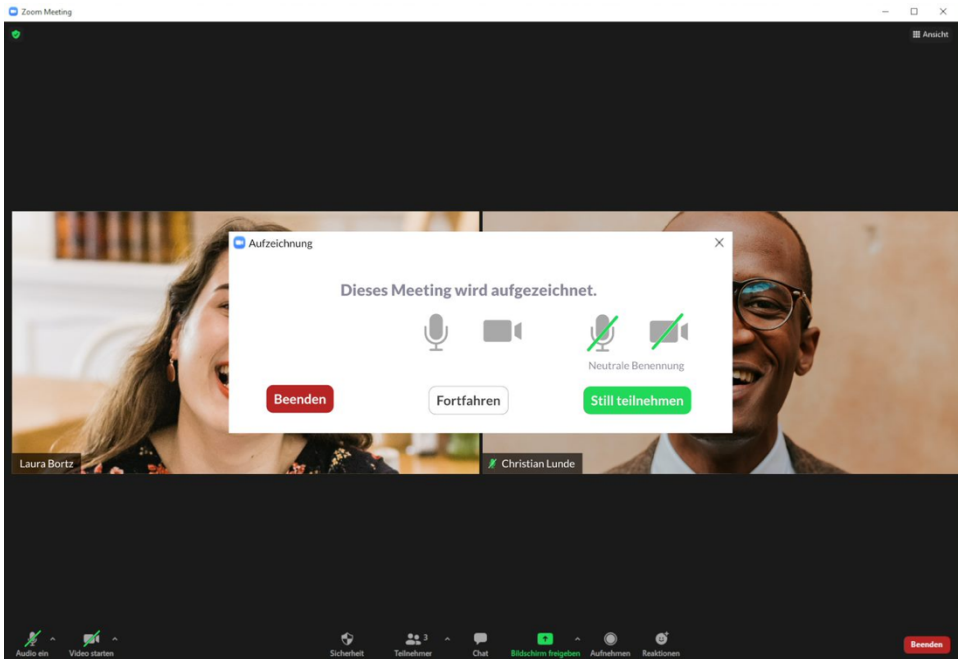


Abbildung 2. Anonyme Videoaufzeichnung

4 Erstellung von Design Pattern

Dieser Beitrag verfolgt, neben der Analyse und Empfehlung von Privacy Nudging, das Ziel dieses Wissen für die Praxis verständlich und anwendbar zu machen. In der Systementwicklung sind Design Pattern für die Kodifizierung und Weitergabe von Wissen ein bewährtes Mittel, auf welches hier zurückgegriffen wird. Die Erstellung der Design Pattern wird im Folgenden beispielhaft an dem Design Pattern „Virtueller Hintergrund“ vorgestellt, welches sich auf die Gestaltungslösungen in Abschnitt 3.1 bezieht.

Das Ziel der Design Pattern ist die Bereitstellung von Gestaltungslösungen für eine möglichst große Menge an Videokonferenztools. In Abschnitt 3 haben wir die Integration der Nudges an dem Anwendungsbeispiel Zoom veranschaulicht. Die erstellten Design Pattern können daher auch auf Zoom oder jedes weitere Konferenztool angewendet werden. Das Gestaltungswissen in den Design Pattern ist so abstrahiert, dass es keine plattformspezifischen Eigenschaften beinhaltet. Die Erstellung der Design Pattern erfolgte im Vergleich zu bekannten Methodiken der Softwareentwicklung nicht auf Basis

wiedererkennbarer Muster, sondern verfolgt das Ziel bewährtes Gestaltungswissen zu kodifizieren und somit anderen Personen zugänglich zu machen.

Das Design Pattern „Virtueller Hintergrund“ (siehe Abbildung 3) stellt die Gestaltung möglicher Nudging Elemente zur Auswahl eines virtuellen Hintergrundes vor, um ungewollte private Einblicke zu vermeiden. Wie für Design Pattern üblich hat auch dieses Pattern einen eindeutigen Namen, der bereits Assoziationen mit der Gestaltungslösung auslöst. Dies ist insbesondere dann wichtig, wenn Personen sich über verschiedene Design Pattern unterhalten, und erleichtert den Austausch, indem die Intention des Patterns nicht mehr definiert werden muss. Neben dem eindeutigen Namen wird der Einsatzzeitpunkt des Nudges während der Videokonferenz dargestellt. Im Falle des virtuellen Hintergrundes setzt der Nudge an, bevor das Bild der Kamera der teilnehmenden Person für alle sichtbar ist, indem diese dazu ermuntert wird bei Einschalten der Kamera in der Videokonferenz einen virtuellen Hintergrund auszuwählen. Der Zielzustand demonstriert die bestmögliche Situation der Gestaltungslösung nach Umsetzung des Design Patterns. So sollen die Nutzenden des Videokonferenztools die Möglichkeit haben an der Konferenz teilzunehmen, ohne dabei unnötige Daten ihres privaten Umfelds preiszugeben. Die Darstellung des Gestaltungsproblems verdeutlicht die Herausforderung, der sich das Design Pattern stellt. So kann durch Teilen der Mimik und Gestik die Videokonferenz an Interaktionsmehrwert gewinnen, der jedoch ungewollte Einblicke in das private Leben geben kann. Dem Problem gegenüber steht die Lösung. Die Lösungsbeschreibung gehört zu dem wichtigsten Element im Design Pattern und gibt auf abstrakte Weise vor wie mögliche Gestaltungslösungen aussehen können.








Virtueller Hintergrund 		Einsatzzeitpunkt 	
<input checked="" type="checkbox"/> Vor der Videokonferenz		<input type="checkbox"/> Während der Videokonferenz	
		<input type="checkbox"/> Nach der Videokonferenz	
Zielzustand 			
Als Nutzer möchte ich die Vorteile der Videokonferenzen nutzen, ohne dabei unnötige sensible Daten preiszugeben.			
Problem 			
Das Einschalten der Videokamera in Videokonferenzen fördert die digitale Zusammenarbeit, da sowohl die Mimik als auch die Gestik des Gegenübers wahrgenommen werden kann. Jedoch können durch Teilen der Umgebung private (oftmals sensible) Einblicke gegeben werden. Darunter können Fotos, andere Personen oder private Umgebungen, die Rückschlüsse schließen lassen, sein.			
Einflüsse 		Konsequenzen 	
Die automatische Erkennung der Person kann je nach Hintergrund dennoch vereinzelt Einblicke in den Hintergrund geben.		Die Auswahl des virtuellen Hintergrundes nimmt zu Beginn der Videokonferenz Zeit in Anspruch.	
Lösung 			
<ul style="list-style-type: none"> • Die Default-Einstellung schaltet sowohl die Kamera als auch den Ton aus • Bei Anschalten der Kamera wird die Nutzung eines virtuellen Hintergrundes vorgeschlagen, um keine kritischen Einblicke in die Umgebung zu gewähren • Bevor die Kamera eingeschaltet wird, wird die zu teilende Ansicht dargestellt, um im Falle der Entscheidung gegen einen virtuellen Hintergrund Bewusstsein über die Ansicht der Kamera zu zeigen • Farbliche Hinterlegungen lenken hierbei die Aufmerksamkeit auf ausgewählte Elemente, um bestimmte Entscheidungsalternativen verstärkt hervorzuheben. In verschiedenen Anwendungen kann beispielsweise die Schaltfläche zur Datenfreigabe farblich stärker betont werden. 			
Bestätigung der Umsetzung des Design Patterns		Datum	Unterschrift

Abbildung 3: Design Pattern „Virtueller Hintergrund“

Die dargelegte Gestaltungslösung verdeutlicht, wie sie auf verschiedene Videokonferenztools generalisiert wurde und so in möglichst vielen Szenarien Unterstützung bietet. Die Integration von Nudging Elementen hängt jedoch stark von dem Kontext des Systems ab. Daher werden im Feld „Einflüsse“ Hinweise dargestellt, die vorab bei der Gestaltung und Integration der Lösung zu berücksichtigen sind. Ebenso hat die Gestaltung der vorgeschlagenen Lösung Auswirkungen auf die Funktionsweise oder Darstellung des Systems. Im Falle des Nudges zur Nutzung eines virtuellen Hintergrundes sollte daher berücksichtigt werden, dass die Auswahl des virtuellen Hintergrundes zu Beginn der Videokonferenz zu Verspätungen der Teilnehmenden führen kann. In der Systementwicklung und insbesondere dann, wenn mehrere Entwickelnde gemeinsam arbeiten ist die Dokumentation ein wichtiger Bestandteil. Daher gibt es die Möglichkeit in einem Unterschriftenfeld zu vermerken, wenn das Design Pattern implementiert wurde. So kann das Design Pattern ergänzend zu weiterer Dokumentation verwendet werden.

5 Fazit

Der vorliegende Beitrag verfolgte das Ziel, Gestaltungswissen für rechtsverträgliche Videokonferenztools unter Berücksichtigung des Privacy Nudgings zu präsentieren. Hierfür hat der Beitrag zentrale Anforderungen aus rechtlicher Sicht aufgezeigt, um den rechtlichen Rahmen abzustecken und entsprechend zu beschreiben, wie hier Rechtsverträglichkeit über die reine Konformität hinausgeht. Zudem zeigt der Beitrag auf, wie das Konzept des Nudging die rechtlichen Anforderungen des Datenschutzes durch Technikgestaltung und durch datenschutzfreundliche Voreinstellungen gem. Art. 25 DSGVO aus einer Gestaltungsperspektive heraus erfüllen kann und wie Design Pattern ein geeignetes Instrument zur Kodifizierung von interdisziplinärem Gestaltungswissen sein können. Kernbeitrag ist demnach die praktische Darstellung zweier Lösungen die auf den Konzepten der „Defaults“ sowie „Framing“ aufbauen, um privatheitsfreundliche und damit rechtsverträgliche Videokonferenztools zu schaffen. Die erstellten Design Pattern lassen sich auf alle gängigen Videokonferenzplattformen einsetzen und stellen somit Lösungen für wiederkehrende Probleme dar. Schließlich zeigt der Beitrag exemplarisch auf, wie sich das Gestaltungswissen für den Fall des „virtuellen Hintergrundes“ für Praxis und Forschung kodifizieren lässt. Dieses Pattern kann dabei als Startpunkt für weitere Forschung rund um die interdisziplinäre Gestaltung von Videokonferenztools dienen. Die datenschutzgerechte Gestaltung von Videokonferenztools bietet in der zunehmend digitalen Welt großes Forschungspotential. Auf Basis der dargestellten Überlegungen sollten Nutzerstudien durchgeführt werden, die zum einen die Wirkungen der integrierten Privacy Nudges auf den Nutzenden sowie deren Verhalten untersuchen.

Danksagung

Dieser Artikel wurde im Rahmen des Projekts „Nudger“ (www.nudger.de; Förderkennzeichen: 16KIS0890K; 16KIS0891) unter der Projekträgerenschaft des VDI/VDE-IT erarbeitet und mit den Mitteln des Bundesministeriums für Bildung und Forschung gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autoren.

6 Literaturverzeichnis

- [AA21] Anthony Jnr, B.; Abbas Petersen, S.: Examining the digitalisation of virtual enterprises amidst the COVID-19 pandemic: a systematic and meta-analysis. *Enterprise Information Systems* 5/15, S. 617–650, 2021.
- [Ac17] Acquisti, A. et al.: Nudges for Privacy and Security. *ACM Computing Surveys* 3/50, S. 1–41, 2017.
- [Al17] Alexander, C. et al.: *A pattern language. Towns, buildings, construction.* Oxford Univ. Press, New York, NY, 2017.
- [Al77] Alexander, C. et al.: *A pattern language. Towns, buildings, construction.* Oxford University Press, New York, 1977.
- [Be21] Berliner Beauftragte für Datenschutz und Informationsfreiheit: Hinweise für Berliner Verantwortliche zu Anbietern von Videokonferenzdiensten, 2021.
- [BG17] Baumgartner, U.; Gausling, T.: Datenschutz durch Technikgestaltung und datenschutzfreundliche Voreinstellungen. Was Unternehmen jetzt nach der DS-GVO beachten müssen. *ZD*, S. 308–313, 2017.
- [BH17] Bieker, F.; Hansen, M.: Datenschutz "by Design" und "by Default" nach der neuen europäischen Datenschutz-Grundverordnung. *RDV* 4, S. 165–170, 2017.
- [Bo09] Bovens, L.: The Ethics of Nudge. In (Grüne-Yanoff, T.; Hansson, S. O. Hrsg.): *Preference change: Approaches from philosophy, economics and psychology.* Springer Science & Business Media, S. 207–220, 2009.
- [Bo21] Bocksch, R.: Zoom Video Communications. Zooms rapider Umsatzanstieg. <https://de.statista.com/infografik/21927/umsatzentwicklung-von-zoom/>, Stand: 07.07.2021.
- [BQS21] Bai, C.; Quayson, M.; Sarkis, J.: COVID-19 pandemic digitization lessons for sustainable development of micro-and small- enterprises. *Sustainable Production and Consumption* 27, S. 1989–2001, 2021.
- [BSJ21] Barev, T.; Schwede, M.; Janson, A.: The Dark Side of Privacy Nudging – An Experimental Study in the Context of a Digital Work Environment. In (Bui, T. Hrsg.): *Proceedings of the 54th Hawaii International Conference*

on System Sciences. Hawaii International Conference on System Sciences, 2021.

- [DH20] Dehmel, E.; Hartmann, N.: Das Coronavirus (COVID-19) auf dem Vormarsch. Die wichtigsten arbeitsrechtlichen Themen. BB (Betriebs-Berater), S. 885–891, 2020.
- [DJL21] Dickhaut, E.; Janson, A.; Leimeister, J. M.: Codifying Interdisciplinary Design Knowledge Through Patterns – The Case of Smart Personal Assistants. In (Hofmann, S.; Müller, O.; Rossi, M. Hrsg.): Extending the Boundaries of Design Science Theory and Practice. 15th. SPRINGER NATURE, [S.l.], S. 114–125, 2021.
- [DR20] Dominique Machuletz; Rainer Böhme: Multiple Purposes, Multiple Problems: A User Study of Consent Dialogs after GDPR. undefined, 2020.
- [Dü19] Dümeland, M.: Sachmangelhaftigkeit von Software bei nicht DSGVO-konformer Entwicklung. Kommunikation & Recht (K&R) 1, S. 22–25, 2019.
- [Eu20] EuGH: Data Protection Commissioner gegen Facebook Ireland Limited und Maximilian Schrems. Vorabentscheidungsersuchen des High Court (Irland), 2020.
- [Ga95] Gamma, E.: Design patterns: elements of reusable object-oriented software. Pearson Education India, 1995.
- [Go20] Golland, A.: Datenschutzrechtliche Anforderungen an internationale Datentransfers. NJW, S. 2593–2596, 2020.
- [Ha06] Hafiz, M.: A collection of privacy design patterns. In (Yoder, J. Hrsg.): Proceedings of the 2006 conference on Pattern languages of programs. ACM, New York, NY, S. 1, 2006.
- [Ha18] Hartung, J.: Art. 25. In (Kühling, J.; Buchner, B. Hrsg.): Datenschutz-Grundverordnung/BDSG. Kommentar. C.H. Beck, München, 2018.
- [Ha20] Hartung, J.: Art. 25: Datenschutz-Grundverordnung BDSG. Kommentar. C.H.Beck, München, 2020.
- [HM19] Hummel, D.; Maedche, A.: How effective is nudging? A quantitative review on the effect sizes and limits of empirical nudging studies. Journal of Behavioral and Experimental Economics 80, S. 47–58, 2019.
- [HW10] Hausman, D. M.; Welch, B.: Debate: To Nudge or Not to Nudge*. Journal of Political Philosophy 1/18, S. 123–136, 2010.
- [Ka03] Kahneman, D.: Maps of Bounded Rationality: Psychology for Behavioral Economics. American Economic Review 5/93, S. 1449–1475, 2003.

- [KK21] Klingbeil, T.; Kohm, S.: Datenschutzfreundliche Technikgestaltung und ihre vertraglichen Implikationen. Praxisnahe Anforderungen für Softwareprodukte. MMR, S. 3–8, 2021.
- [KRP20] Krieger, S.; Rudnik, T.; Povedano, A.: Homeoffice und Mobile Office in der Corona-Krise. NZA, S. 473–479, 2020.
- [Ku20] Kugelmann, D.: Gesundheitsnot kennt Datenschutzgebot. Fachinformationsdienst für internationale und interdisziplinäre Rechtsforschung, VerfBlog, 2020.
- [La19] Laue, P.: § 7. Technischer und organisatorischer Datenschutz. In (Laue, P.; Kremer, S. Hrsg.): Das neue Datenschutzrecht in der betrieblichen Praxis. Nomos, Baden-Baden, 2019.
- [Le16] Lehmann, B. A. et al.: Changing the default to promote influenza vaccination among health care workers. *Vaccine* 11/34, S. 1389–1392, 2016.
- [Ma18] Martini, M.: Art. 25. In (Paal, B. P.; Pauly, D. A. Hrsg.): Datenschutz-Grundverordnung, Bundesdatenschutzgesetz. C.H. Beck, München, 2018.
- [Mü20] Müller, S.: Homeoffice in der arbeitsrechtlichen Praxis. Rechtshandbuch für die Arbeit 4.0. Nomos, Baden-Baden, 2020.
- [PP21] Paal, B. P.; Pauly, D. A.: Einleitung: Datenschutz-Grundverordnung, Bundesdatenschutzgesetz. C.H. Beck, München, 2021.
- [Ro19] Rossi, A. et al.: Legal Design Patterns: Towards A New Language for Legal Information Design. 22nd International Legal Informatics Symposium IRIS 2019, 2019.
- [Ro20] Roßnagel, A.: Zoom und Datenschutz.
file:///C:/Users/Admin/AppData/Local/Temp/Zoom_und__Datenschutz_CIO_AR_23032020-1.pdf, Stand: 07.07.2021.
- [Ro93] Roßnagel, A.: Rechtswissenschaftliche Technikfolgenforschung. Umriss einer Forschungsdisziplin. Zugl.: Heidelberg, Univ., Habil.-Schr., 1992. Nomos-Verl.-Ges, Baden-Baden, 1993.
- [Sa15] Sandfuchs, B.: Privatheit wider Willen? Dissertation, 2015.
- [Sc19] Schomberg, S. et al.: Ansatz zur Umsetzung von Datenschutz nach der DSGVO im Arbeitsumfeld: Datenschutz durch Nudging. Datenschutz und Datensicherheit - DuD 12/43, S. 774–780, 2019.
- [Sc20] Schöbel, S. et al.: Understanding User Preferences of Digital Privacy Nudges – A Best-Worst Scaling Approach. In (Bui, T. Hrsg.): Proceedings of the 53rd Hawaii International Conference on System Sciences. Hawaii International Conference on System Sciences, 2020.

- [SK18] Sandfuchs, B.; Kapsner, A.: Privacy Nudges: Conceptual and Constitutional Problems. In (Bürk, S. et al. Hrsg.): *Privatheit in der digitalen Gesellschaft*. Duncker & Humblot, Berlin, S. 319–338, 2018.
- [St20] Stocklas, J.: *Datenschutz in Zeiten von Corona*. ZD-Aktuell, 2020.
- [Su15] Sunstein, C. R.: *Do People Like Nudges?* SSRN Electronic Journal, 2015.
- [Su20] Suwelack, F.: *Datenschutzrechtliche Vorgaben für Homeoffice und Remote Work. Nachhaltige und rechtssichere Umstellung - New Work. New Normal. New Problems?* ZD, S. 561–566, 2020.
- [Th20] Thies, L. F. et al.: *Konfliktäre Anforderungen an smarte persönliche Assistenten*. *Datenschutz und Datensicherheit - DuD 9/44*, S. 573–578, 2020.
- [To20] Townsend, K.: *Zoom's Security and Privacy Woes Violated GDPR*, Expert Says. *Zoom Security Risks, Privacy and GDPR Compliance*. <https://www.securityweek.com/zooms-security-and-privacy-woes-violated-gdpr-expert-says>, Stand: 07.07.2021.
- [TS08] Thaler, R. H.; Sunstein, C. R.: *Nudge. Improving decisions about health, wealth, and happiness*. Yale University Press, New Haven, 2008.
- [VE20] Verheyen, J.; Elgert, D.: *Datenschutz im Homeoffice - Ein Überblick*. *K&R*, S. 476–479, 2020.
- [Wa14] Wang, Y. et al.: *A field trial of privacy nudges for facebook*. In (Jones, M. et al. Hrsg.): *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*. ACM Press, New York, New York, USA, S. 2367–2376, 2014.
- [Wa20] Waizenegger, L. et al.: *An affordance perspective of team collaboration and enforced working from home during COVID-19*. *European Journal of Information Systems* 4/29, S. 429–442, 2020.
- [WB75] Wortman, C. B.; Brehm, J. W.: *Responses to Uncontrollable Outcomes: An Integration of Reactance Theory and the Learned Helplessness Model*. *Advances in Experimental Social Psychology* 8, S. 277–336, 1975.
- [WSV16] Weinmann, M.; Schneider, C.; Vom Brocke, J.: *Digital Nudging*. *Business & Information Systems Engineering* 6/58, S. 433–436, 2016.
- [Wü20] Wünschelbaum, M.: *COVID-19: Pandemiebewältigung und Datenschutz. Kollektivvereinbarungen als krisentaugliches DS-GVO-Instrument?* *NZA*, S. 612–616, 2020.
- [YSJ15] Yskout, K.; Scandariato, R.; Joosen, W.: *Do Security Patterns Really Help Designers?: Second ACM International Conference on Mobile Software*

Engineering and Systems - MOBILESoft 2015. May 16-17, 2015,
Florence, Italy. IEEE, Piscataway, NJ, S. 292–302, 2015.

Datenschutz durch Technikgestaltung und unternehmerische Strategie in der Digitalwirtschaft

Verhaltenssteuernde Elemente der DSGVO im Licht einer spieltheoretischen Analyse am Beispiel von Art. 25 Abs. 1

Mona Winau,¹ Florian Klaus Kaiser,² Marcus Wiens,³ Frank Schultmann,⁴ Indra Spiecker gen. Döhmann⁵

Abstract: Der Beitrag beleuchtet das Konzept des Datenschutzes durch Technikgestaltung nach Art. 25 Abs.1 DSGVO im Spannungsfeld zwischen einer notwendigerweise abstrakten Risikoregulierung und den Gefahren der daraus folgenden Konkretisierungsverantwortung datenverarbeitender Wirtschaftsunternehmen. Anhand einer spieltheoretischen Analyse wird nachvollzogen, wie sich Negativanreize der DSGVO aus Bußgeld- und Haftungsrisiken proaktiv nutzen lassen, um die Motivation datenschutzkonformer Firmen zu stärken, die tendenziell ein höheres Eigeninteresse an der Implementierung geeigneter technischer und organisatorischer Maßnahmen haben. Ökonomische Anreize einer Zertifizierung könnten eine datenschutzfreundliche Technikgestaltung für Unternehmen profitabel machen, wenn sich ein entsprechender Markt etabliert. Für die Unterscheidbarkeit datenschutzfreundlicher gestalteter Produkte von Angeboten mit geringem Datenschutzniveau können Zertifikate als Signal gegenüber dem Verbraucher dienen und so eine Grundlage für einen innovationsfördernden Wettbewerb für den Datenschutz durch Technikgestaltung schaffen.

Keywords: Rechtsanalyse; Spieltheorie; Regulierung durch Anreize

¹ Karlsruher Institut für Technologie, Kompetenzzentrum für angewandte Sicherheitstechnologie (KASTEL), Am Fasanengarten 5, 76131 Karlsruhe, Deutschland & Institut für Informations- und Wirtschaftsrecht (IIWR), Vincenz-Preißnitz-Straße 3, 76131 Karlsruhe, Deutschland, mona.winau@kit.edu

² Karlsruher Institut für Technologie, Kompetenzzentrum für angewandte Sicherheitstechnologie (KASTEL), Am Fasanengarten 5, 76131 Karlsruhe, Deutschland & Institut für Industrielehre und Industrielle Produktion (IIP), Hertzstraße 16, 76187 Karlsruhe, Deutschland, florian-klaus.kaiser@kit.edu

³ Karlsruher Institut für Technologie, Institut für Industrielehre und Industrielle Produktion, Hertzstraße 16, 76187 Karlsruhe, Deutschland, marcus.wiens@kit.edu

⁴ Karlsruher Institut für Technologie, Institut für Industrielehre und Industrielle Produktion, Hertzstraße 16, 76187 Karlsruhe, Deutschland & Adelaide Business School, University of Adelaide, frank.schultmann@kit.edu

⁵ Goethe-Universität Frankfurt am Main, Lehrstuhl für Öffentliches Recht, Informationsrecht, Umweltrecht und Verwaltungswissenschaft, Institut für Öffentliches Recht, Theodor-W.-Adorno-Platz 4, 60629 Frankfurt am Main, Deutschland, spiecker@jur.uni-frankfurt.de

1 Risikobasierte Datenschutzregulierung und Risiken des unternehmerischen Umsetzungsspielraum

Unternehmen, die personenbezogene Daten verarbeiten, müssen gem. Art. 25 Abs. 1 DSGVO⁶ dem konkreten Risiko angemessene Datenschutzmaßnahmen bereits in die technische und organisatorische Gestaltung des Bearbeitungsprozesses integrieren. Die technikneutral formulierte und an den abstrakten Begriff des Risikos gekoppelte Pflicht soll einen effektiven Schutz für die Rechte und Freiheiten natürlicher Personen gewährleisten und zugleich eine verhältnismäßige Anpassung der Pflichten an den jeweiligen Verarbeitungskontext beim Verantwortlichen sicherstellen.⁷ Allerdings begründen die Regelungsunschärfe und der erhebliche Spielraum des Verantwortlichen bei der Auswahl geeigneter Maßnahmen Rechtsunsicherheit und bergen die Gefahr, dass das Pflichtenprogramm zugunsten der wirtschaftlichen Vorteile einer unangemessenen Datennutzung in der Praxis ausgehöhlt wird.⁸ Die Datenschutz-Grundverordnung enthält wiederum diverse rechtliche und ökonomische Anreize, die eine datenschutzkonforme, oder sogar eine datenschutzfreundliche,⁹ Verarbeitung attraktiv machen sollen. Aus ihnen ergeben sich Parallelen zu einer Regulierung strategischen Firmenverhaltens, wie sie etwa in Modellen der Wettbewerbsregulierung und im öffentlichen Wirtschaftsrecht vorkommen.¹⁰

1.1 Kontextabhängigkeit der Datenverarbeitung und ihre abstrakte Regulierung

Informationen sind kontextabhängig. Eine einzelne Information kann zu vorhandenem Wissen hinzutreten und den Schlüssel zu einer Erkenntnis bilden oder eine einzelne Information kann ohne weiteren Nutzen bleiben. Ihr kann in diversen Kontexten verschiedene Bedeutung zukommen und sie kann zu unterschiedlichen Schlussfolgerungen führen. Wegen der Abhängigkeit des Nutzens einer Information von ihrer Art, ihrem Zustand, Kontext und Verarbeitungsmöglichkeiten ist die Bestimmung eines objektiven wirtschaftlichen Wertes für Daten mit erheblichen Schwierigkeiten verbunden.¹¹ Ebenso ist das Risiko, das mit der Preisgabe eines personenbezogenen Datums für die betreffende Person verbunden ist - etwa die Gefahr weitreichender oder falscher Schlüsse über dieselbe - von den genannten Faktoren abhängig und nicht abstrakt feststellbar.¹² Allgemein lässt sich aber feststellen, dass sowohl der ökonomische Nutzen von personenbezogenen Daten als auch das Risiko

⁶ Folgende Artikel ohne Kennzeichnung beziehen sich auf die DSGVO.

⁷ Art. 29-Gruppe, [Ar14], 2; CIPL, [Ce], 4.

⁸ Vgl. [Qu17], 14 ff; [Sc19], 505; [RG20], 40 ff; [By17], 117; [BTH14] formulieren in diesem Kontext "PbD expects [...] that the cat guards the milk", 71, und argumentieren, dass der Abstraktionsgrad von Privacy by Design-Regelungen wegen divergierender rechtlicher und technischer Privacy-Konzepte problematisch ist, 88 ff.

⁹ Im Sinne einer über vorgegebene Mindeststandards hinausgehenden datenschutzfreundlichen Verarbeitung.

¹⁰ M.w.N. zu Konzepten des funktionsfähigen Wettbewerbs, [Br10], 11 ff und zur Einbeziehung von Strategie für eine wirksame Anreizregulierung im öffentlichen Wirtschaftsrecht, 233 ff; zur Rolle des Wettbewerbs in der datenschutzrechtlichen Regulierung, [Po17], 220 ff.

¹¹ Vgl. [K6b], 305 f; [Hä16], 735 f.

¹² Vgl. EG 75, 76.

ihrer Verarbeitung mit der immens gesteigerten und zukünftig weiter steigenden Menge an Daten¹³ und den gigantischen, schnellen und automatisierten Möglichkeiten ihrer Speicherung, Kombination und Auswertung sehr hoch ist.¹⁴ Wie das Bundesverfassungsgericht bereits 1983 erkannte, ist das Recht auf informationelle Selbstbestimmung, in den Grenzen des Gemeinschaftsbezugs von Informationen, eine Voraussetzung für die freie Grundrechtsausübung des Einzelnen und gilt unter den Bedingungen automatisierter Verarbeitung für jedes beliebig inhaltsarme Datum, solange es einen Bezug zur Person zulässt.¹⁵ Je weiter die technischen Möglichkeiten fortschreiten und sich aus vorhandenen Informationen immer mehr und aussagekräftigere neue Informationen erschließen lassen, desto größer ist das Risiko, dass die Verarbeitung auch nur weniger und scheinbar trivialer Informationen über eine Person birgt. Wenn es bereits 1983 unter den Bedingungen automatisierter Verarbeitungstechniken keine „belanglosen“¹⁶, oder ökonomisch ausgedrückt, wertlosen Daten zu einer Person gab,¹⁷ gilt dies heute erst recht.¹⁸

1.2 Abstrakte Regulierung, Strategie und Anreizsteuerung

Die Technikneutralität und der risikobasierte Ansatz der Datenschutz-Grundverordnung begegnen dem sich fortlaufend erhöhenden, aber nicht allgemeingültig bestimmbareren Risiko der Verarbeitung personenbezogener Daten.¹⁹ Technikneutralität bedeutet, dass die Verordnungsregelungen für jede Verarbeitungstechnik (manuell oder automatisiert) gleichermaßen gelten.²⁰ Risikobasiert sind Pflichten der Verantwortlichen insoweit, wie ihre Geltung, ihr Umfang oder ihre Intensität vom Risiko der jeweiligen Verarbeitung abhängt.²¹ Eine Datenschutz-Folgeabschätzung etwa muss gem. Art. 35 Abs. 1 nur für solche Verarbeitungen durchgeführt werden, die voraussichtlich ein hohes Risiko für die Rechte und Freiheiten natürlicher Personen bergen. Und auch der Umfang und die Intensität von technischen oder organisatorischen Maßnahmen, die Verantwortliche gem. Art. 24 Abs. 1 zur Gewährleistung einer verordnungskonformen Verarbeitung und deren Nachweis treffen müssen, sind an das Risiko der Verarbeitung gekoppelt. Eine konkretere Ausgestaltung dieser Pflicht zu technischen und organisatorischen Schutzmaßnahmen

¹³ Vgl. [RGR18].

¹⁴ Dies wird anhand der rege geführten Debatten über Eigentums-, Besitz oder Nutzungsrechte an Daten und die wettbewerbsrechtliche Relevanz von Datenschutzrechtsverstößen deutlich, m.w.N. zu Eigentums-, Besitz- oder Nutzungsrechte an Daten, [K0]; [Ho19a];[Ad20]; zur kartellrechtlichen Debatte, [K6b]; [K6a]; [Ke16]; zur BGH-Entscheidung im Eilverfahren (Beschl. V. 23.6.2020 – KVR 69/19), [Po20], 1275. Nach der EU-Richtlinie über Verträge über digitale Inhalte und Dienstleistungen (RL (EU) 2019/770) wird die Bereitstellung personenbezogener Daten von Nutzern als Entgelt angesehen, [SS19], 418; kritisch [Hä16].

¹⁵ BVerfG, Urt. v. 15.12.1983 – 1 BvR 209/83 = NJW 1984, 421 f.

¹⁶ BVerfG, Urt. V. 15.12.1983 – 1 BvR 209/83 = NJW 1984, 419 (422).

¹⁷ Dies drückt sich insbesondere auch in der "formal indifference"[Zu15] der Extraktion von Daten aus.

¹⁸ Siehe auch [Sc17], 124.

¹⁹ Vgl.[BTH14], 65, 68;[Qu17], 4; Art. 29-Gruppe, [Ar14], 2; CIPL, [Ce], 3 f; zu Art. 25, [By17], 106.

²⁰ Vgl. EG 15; näher zum Prinzip der Technikneutralität, [Ve18], 73 ff; [BTH14] argumentieren, dass auch der europäische Datenschutzregulierung dennoch eine gewisse Technikausrichtung innewohnt, 77 ff.

²¹ Vgl. [Sc19], 503.

ist die hier betrachtete Pflicht zum Datenschutz durch Technikgestaltung gem. Art. 25 Abs. 1,²² oft auch - begrifflich unpräzise - als „Privacy by Design“ oder treffender „Data Protection by Design“ bezeichnet.²³ Die Bestimmung des Risikos und die Spezifizierung des Pflichtenprogramms für die konkrete Verarbeitung obliegt dem Verantwortlichen selbst.²⁴ Sie lässt in besonderem Maße Raum für eine strategische unternehmerische Entscheidung.

Aus ökonomischer Perspektive stehen die Potentiale neuer technischer Möglichkeiten und ihr gewinnbringender Einsatz im Vordergrund. Die meisten privaten Wirtschaftsunternehmen haben längst die automatisierte Informationsverarbeitung in ihre Prozesse integriert und stützen hierauf teilweise sogar die Finanzierung ihrer Produkte.²⁵ In diesem Zusammenhang wird vielfach von Daten als „Rohstoff des 21. Jahrhunderts“ gesprochen.²⁶ Diesen wertvollen Rohstoff in möglichst großem Umfang zu erwerben, um ihn gewinnbringend weiterzuverarbeiten oder verkaufen zu können, entspricht auf den ersten Blick einer auf Ertrag fokussierten, d.h. ökonomischen, Handlungsweise, während eine datensparsame, zweckgebundene, speicherbegrenzte und transparente Verarbeitung entsprechend den Grundsätzen aus Art. 5 Abs. 1 in Gegensatz zu ihr tritt.²⁷ Technische und organisatorische Datenschutzmaßnahmen sind außerdem mit zusätzlichen Kosten verbunden.²⁸ Damit sich eine solche Investition aus ökonomischer Sicht lohnt, muss sich aus ihnen ein Nutzen ergeben, der die Investitionskosten einschließlich des Verzichts auf einen möglichen Gewinn aus einer weitreichenden Verarbeitung personenbezogener Daten übersteigt.²⁹ Um den Nutzen zu erhöhen, hat der europäische Gesetzgeber die Implementierung von technischen und organisatorischen Gestaltungsmaßnahmen mit rechtlichen und ökonomischen Anreizen verknüpft.³⁰

2 Datenschutz durch Technikgestaltung gem. Art. 25 Abs. 1

Die Grundidee der Pflicht zum Datenschutz durch Technikgestaltung ist, dass Maßnahmen zum Schutz der Betroffenen unmittelbar in die verwendete Hard- und Software eingebaut und in organisatorische Abläufe der Verarbeitung personenbezogener Daten, über ihren gesamten Lebenszyklus hinweg, integriert werden.³¹ Es handelt sich also um eine Pflicht zur datenschutzrechtlichen Gestaltung von Verarbeitungsprozessen allgemein und nicht

²² Hartung, [KB20], Art. 24 Rn 9; Martini, [Pa21b], Art. 24 Rn 1; [By17], 114.

²³ Hartung, [KB20], Art. 25 Rn 1; Hansen, [SHSgD19], Art. 25 Rn 23; [By17], 115 f.

²⁴ [Sc19], 503 f.; [Qu17], 7, 9 ff.; [Qu18].

²⁵ Solche Geschäftsmodelle werden unter dem Begriff der „Datenwirtschaft“ diskutiert, [Ze15]; [Dr17].

²⁶ „Personal data is the new oil of the internet and the new currency of the digital world.“ Meglena Kuneva, ehemalige EU-Kommissarin für Verbraucherschutz nach [HG16], m.w.N. [Po19], 321; [KO], 24; eine Formulierung, die auch von der deutschen Bundeskanzlerin verwendet wird, FAZ, [Fr]; Bulletin der Bundesregierung, [Bu], 5.

²⁷ Zum Spannungsverhältnis zwischen der Ökonomisierung der Datenverarbeitung und Grundlagen der Datenschutzregulierung, [Hä16], 738 f.; zwischen Big Data und Grundprinzipien der DSGVO, [SgD17a]; [Za17].

²⁸ Siehe auch [BTH14], 75.

²⁹ Vgl. auch [RG20], 39.

³⁰ Mantz, [Sy18], Art. 25 Rn 78; [Th16], 715 f.; als unzureichend bewertend [RG20], 52 ff, 56; [By17], 116 f, 119.

³¹ Petri, [SHSgD19], Art. 24 Rn 2 ff.; [By17], 106.

nur der Technikgestaltung.³² Verantwortliche haben gem. Art. 25 Abs. 1 den Stand der Technik und die Implementierungskosten, aber auch Art, Umfang, Umstände und Zwecke der Verarbeitung, sowie die Eintrittswahrscheinlichkeit und die Schwere des Risikos für die Rechte und Freiheiten des Betroffenen bei der Entscheidung über geeignete Maßnahmen zur Gewährleistung einer datenschutzkonformen Verarbeitung einzubeziehen. Insbesondere die Grundsätze aus Art. 5 Abs. 1 sollen durch eine datenschutzkonforme Technikgestaltung unmittelbar bei der Verarbeitung personenbezogener Daten umgesetzt werden.³³ Die Risiken der jeweiligen Verarbeitung für die Rechte und Freiheiten der betroffenen Person sollen dadurch von den Verantwortlichen selbst begrenzt werden.³⁴

3 Konkretisierungsverantwortung und unternehmerische Strategie

Die Beurteilung des Risikos und die Auswahl geeigneter Maßnahmen obliegen in erster Linie den Verantwortlichen selbst. Eine Konkretisierung oder gar Definition des Risikobegriffs, über die genannten Faktoren der Eintrittswahrscheinlichkeit und der Schwere und die Beispiele in EG 75 hinaus, enthält die Verordnung nicht.³⁵ Die Datenschutz-Folgeabschätzung kann zumindest Anhaltspunkte zur Methodik der Risikobestimmung geben, ist aber nur für Verarbeitungen mit voraussichtlich hohen Risiken verpflichtend (Art. 35 Abs. 1).³⁶ Bei der Risikobewertung können sich die Verantwortlichen an konkretisierenden Leitlinien, Kurzpapieren und Empfehlungen von Datenschutzaufsichtsbehörden oder aus entsprechender Fachliteratur orientieren,³⁷ müssen den konkreten Fall letztlich aber eigenverantwortlich beurteilen. Dabei bleibt das Risiko der Verarbeitung allerdings nur einer, wenn auch ein zentraler, von mehreren Faktoren, die bei der Entscheidung über geeignete technische und organisatorische Maßnahmen gem. Art. 25 Abs. 1 zu berücksichtigen sind. Bei der konkreten Maßnahmenauswahl und der Beurteilung ihrer Geeignetheit anhand der genannten abstrakten Kriterien können Verantwortliche die Beispiele in EG 78 S. 3 sowie ebenfalls Maßnahmenvorschläge und Empfehlungen aus Öffentlichkeitsarbeit und Literatur heranziehen.³⁸ Für die Entscheidung verbleibt ihnen letztlich ein erheblicher Spielraum.³⁹ Trotz dieses Entscheidungsspielraums bei der Maßnahmenauswahl ist das Kriterium der Geeignetheit binär (geeignet - ungeeignet) und vollständig gerichtlich überprüfbar.⁴⁰ Aus den in Art.

³² Martini/Hansen, [SHSgD19], Art. 25 Rn 16.

³³ EDSA, [Eub], 7; Hartung, [KB20], Art. 25 Rn 10; [By17], 115.

³⁴ [Qu18], 503. Auch wenn die geforderten Maßnahmen sich klar auf die Umsetzung der Verordnungsregelungen beschränken und auch bei besonders hohen Risiken für andere Grundrechte, etwa durch Diskriminierungsgefahren oder Beeinträchtigungen der Meinungsfreiheit, keine darüberhinausgehenden Schutzvorkehrungen einschließen, kann die Kopplung von Umfang und Intensität der Datenschutzmaßnahmen an das jeweilige Verarbeitungsrisiko mittelbar auch einen Schutz für andere Rechte und Freiheiten entfalten. Das maßgebliche Risiko der Verarbeitung bezieht sich ausdrücklich nicht nur auf Datenschutz- und Privatheitsrechte (Art. 7 und 8), sondern allgemein auf Rechte und Freiheiten natürlicher Personen, EG 75; [Qu17]; DSK, [Un].

³⁵ [Sc19], 504.

³⁶ Vgl. auch [Sc19], 504 f; Martini, [Pa21b], Art. 25 Rn 37c.

³⁷ Z.B. Art. 29-Gruppe, [Ar]; DSK, [Un]; EDSA, [Eub], 11.

³⁸ Z.B. das Standard-Datenschutzmodell, DSK, [DS]; kritisch zur unspezifischen Regelung [RG20], 42 f.

³⁹ Hansen, [SHSgD19], Art. 25 Rn 36.

⁴⁰ Vgl. Martini, [Pa21b], Art. 25 Rn 36; zu grundrechtlichen Aspekten für eine Grenzziehung, [By17], 105 (109 ff).

25 Abs. 1 genannten Kriterien ergibt sich ein verhältnismäßiges Mindestpflichtenprogramm des Verantwortlichen als justiziable Untergrenze für den konkreten Fall.⁴¹ Darüber hinaus lässt der unternehmerische Entscheidungsspielraum bei der Auswahl der Maßnahmen eine graduelle Abstufung der Datenschutz-Technikgestaltung zu. Der Verantwortliche kann selbst strategisch entscheiden, ob er Schutzmaßnahmen integriert, die er in Anbetracht des ermittelten Risikos für gerade noch geeignet hält, oder solche, die darüber hinausgehen.⁴² So kann etwa der Transparenzgrundsatz (Art. 5 Abs. 1 lit. a) für Online-Dienste umgesetzt werden, indem lediglich eine Datenschutzerklärung zur Verfügung gestellt wird oder der Nutzer mithilfe eines Datenschutz-Dashboards die konkret zu seiner Person verarbeiteten Daten selbst überblicken und bestenfalls, zur Umsetzung weiterer Datenschutzgrundsätze, auch selbst berichtigen und löschen oder über ihre Verwendung entscheiden.⁴³

Entscheiden sich mehrere Unternehmen für eine datenschutzfreundliche Technikgestaltung,⁴⁴ könnten sich neue Märkte für datenschutzfreundliche Angebote etablieren, deren Wettbewerb die Innovation bei der Technikgestaltung fördern kann und die einen gewissen Anpassungsdruck auf Anbieter in bestehenden Märkten ausüben könnten.⁴⁵ Geht man davon aus, dass datenverarbeitende Unternehmen sich überwiegend fremdhergestellter technischer Systeme bedienen,⁴⁶ betrifft dies auch die Sphäre der Hersteller von Hard- und Softwareprodukten. Obwohl die Hersteller nicht Verantwortliche und damit nicht von der Regelung des Art. 25 Abs. 1 erfasst sind, könnten sie mit Produkten, die hinter den nachgefragten datenschutzrechtlichen Standards zurückbleiben, nicht mehr am Markt bestehen.⁴⁷ Da das datenschutzrechtliche Pflichtenprogramm aus Art. 25 Abs. 1 nicht statisch ist, sondern wiederum vom Stand der Technik⁴⁸ und den Implementierungskosten abhängt, kann sich auch die rechtliche Grenze für das Mindestmaß an geeigneten Schutzmaßnahmen verschieben. Eine Anpassung der eigenen Technikgestaltung an verfügbare datenschutzfreundliche Technologie-Standards kann demnach aus unternehmerischer Perspektive nicht nur für die eigene Wettbewerbsfähigkeit interessant sein, sondern auch datenschutzrechtlich für Art. 25 Abs. 1 relevant. Allerdings können neue technische Entwicklungen das Mindestmaß geeigneter Maßnahmen nicht linear steigern. Es ergibt sich unter Verhältnismäßigkeitsgesichtspunkten anhand der genannten Kriterien, von denen der Stand der Technik nur eines

⁴¹ Hansen, [SHSgD19], Art. 25 Rn 37.

⁴² Zur Differenzierung zwischen „hard PETs“ und „soft PETs“ [RG20], 49 ff. "Privacy Enhancing Technology" wird hierbei nicht von Privacy by Design-Maßnahmen unterschieden. Zur Differenzierung der Begriffe [BTH14], 69 ff. Allgemeiner zur Skalierbarkeit risikobasierter Pflichten, [Qu17], 9 ff.

⁴³ Martini, [Pa21b], Art. 25 Rn 29a.

⁴⁴ Die EU-Kommission geht zumindest davon aus, dass bislang einige Unternehmen solche Markt Vorteile erkannt und genutzt haben, KOM, [EUa], 3; Bewertung des Kommissionsbericht, [Ro21], 661.

⁴⁵ Vgl. [By17], 118 f, die Schwierigkeiten eines Markt datenschutzfreundlich gestalteter Produkte betonend.

⁴⁶ Vgl. auch [RG20], 40; [Bu20], 1.

⁴⁷ [KI21] zu Auswirkungen auf die schuldrechtlichen Verhältnisse zwischen Herstellern und Verwendern. Mit Verweis auf EG 78 S.4, wonach die Hersteller „ermutigt“ werden sollten, ihre Produkte anzupassen, Hartung, [KB20], Art. 25 Rn 13; Hansen, [SHSgD19], Art. 25 Rn 21. Tatsächliche Auswirkungen auf die Sphäre der Hersteller über die Marktentwicklung bezweifeln, [RG20], 43; [By17], 118 f. Auch die Datenethikkommission rät zu einer Inpflichtnahme der Hersteller, [St20]; DEK, [Da19], 74, 119.

⁴⁸ Darunter sind am Markt hinreichend verfügbare technische Standards; nicht der Stand der Wissenschaft, aber auch nicht der Stand der Praxis zu verstehen, Hartung, [KB20], Art. 25 Rn 21; Mantz, [Sy18], Art. 25 Rn 37 ff.

ist.⁴⁹ Mindestens muss der Verantwortliche aber die technischen Entwicklungen beobachten und fortlaufend in seine Entscheidung einbeziehen.⁵⁰

Entscheidet sich die überwiegende Zahl der datenverarbeitenden Wirtschaftsunternehmen hingegen für eine möglichst weitreichende Datenverarbeitung und treffen sie unzureichende oder nur solche Schutzmaßnahmen, die sie gerade noch für geeignet halten,⁵¹ ist die Wirksamkeit des risikobasierten Pflichtenprogramms stark eingeschränkt. Lediglich das obligatorische Mindestmaß, das aus Art. 25 Abs. 1 folgt, kann dann zwangsweise durch Aufsichtsbehörden im Rahmen ihrer begrenzten Kapazitäten durchgesetzt werden.⁵²

4 Anreize für eine datenschutzfreundliche Technikgestaltung

4.1 Verhaltenlenkung durch Anreize in der DSGVO

Das Prinzip der Verhaltenssteuerung durch Anreize ist einfach. Durch externe Einflüsse können positive oder negative Konsequenzen (Anreize) gesetzt werden und die Bewertung des damit verknüpften Verhaltens beeinflussen. Wird infolge des gewünschten Verhaltens eine Belohnung entrichtet oder werden Nachteile eingeschränkt, wird das gewünschte Verhalten im Bewusstsein der Person aufgewertet. Genauso kann durch den Umkehrfall negativer Konsequenzen (Strafe, Minderung von Vorteilen) ein nicht erwünschtes Verhalten für den Adressaten abgewertet werden. Der Einfluss solcher Anreize auf eine Entscheidung nach dem Kosten-Nutzen-Prinzip hängt unter anderem davon ab, inwieweit sich die gesetzten Anreize gegenüber anderen positiven oder negativen Effekten durchsetzen können.⁵³

Die folgende Darstellung ist auf die, für die anschließende ökonomische Analyse relevanten, Anreize aus möglichen monetären Einbußen (Bußgelder, Schadensersatz) und aus einem etwaigen ökonomischen Nutzen von Zertifikaten beschränkt.

4.2 Vollzugsmaßnahmen und Haftungsrisiko als Negativanreize

Die Überprüfung und gegebenenfalls die Durchsetzung der Pflicht aus Art. 25 Abs. 1 obliegt den Datenschutzaufsichtsaufsichtsbehörden (Art. 57 Abs. 1 lit. a). Bereits die Möglichkeit behördlicher Vollzugsmaßnahmen entfaltet eine abschreckende Wirkung, solange und soweit sie eine ernstzunehmende Gefahr für Unternehmen darstellt.⁵⁴ Sie ist Bestandteil der Anreizstruktur, die eine selbstständige Umsetzung der Verordnungsregelungen sicherstellen

⁴⁹ Hansen, [SHSgD19], Art. 25 Rn 27.

⁵⁰ [RG20], 37 (42); Mantz, [Sy18], Art. 25 Rn. 39 f; Ansätze für die Bewertung neuer Technikentwicklungen ergeben sich etwa aus einer Analyse zu Technikreife der ENISA, [Eu15].

⁵¹ Nur prozedurale und keine Maßnahmen der Technikarchitektur sind nach [RG20], 50 zu erwarten.

⁵² Zu strukturellen Defiziten des behördlichen Vollzugs im Datenschutzrecht, [Po17], 216 f; [SgD17b].

⁵³ [La20], 5 ff.

⁵⁴ [Gr19], 112, zitiert [WDH16], 4; vgl. auch [Gr16], 299.

soll.⁵⁵ Die Aufsichtsbehörden sind nach Art. 58 Abs. 2 lit. i, Art. 83 Abs. 4 lit. a verpflichtet, Verstöße gegen Art. 25 Abs. 1 mit einer Geldbuße bis zu einer Höhe von 10 000 000€ oder zwei Prozent des weltweit erzielten Jahresumsatzes zu ahnden.⁵⁶ Die Bußgelder sind gem. Art. 83 Abs. 1 im Einzelfall wirksam, verhältnismäßig, abschreckend und anhand der Kriterien des Abs. 2 zu bemessen. Letzterer umfasst wiederum einen Positivanreiz für die datenschutzfreundliche Implementierung technisch-organisatorischer Maßnahmen, indem der „Grad der Verantwortung“ für den Verstoß unter Berücksichtigung der getroffenen Maßnahmen einzubeziehen ist. Wurden geeignete Schutzmaßnahmen getroffen, wirkt sich dies positiv auf die Bußgeldbemessung aus.⁵⁷ Die Formulierung legt darüber hinaus nahe, dass eine besonders sorgfältige Vorsorge durch technisch-organisatorische Maßnahmen auch entsprechend mildernd zu berücksichtigen ist.⁵⁸ Aus den Kriterien der Bußgeldbemessung ergibt sich außerdem noch ein zusätzlicher Negativanreiz für eine Priorisierung kommerzieller Interessen zulasten geeigneter Datenschutzmaßnahmen. Denn sonstige erschwerende Umstände, vor allem finanzielle Vorteile aus der rechtswidrigen Verarbeitung sind gem. Art. 83 Abs. 2 lit. k ebenfalls bei der Bußgeldbemessung zu berücksichtigen.⁵⁹

Der Negativanreiz möglicher Schadensersatzpflichten ergibt sich aus der Haftung für materielle und immaterielle Schäden, die sich als voraussehbare Folge einer verordnungswidrigen Verarbeitung,⁶⁰ d.h. im hier betrachteten Fall einer Verarbeitung personenbezogener Daten trotz unzureichender Maßnahmen nach Art. 25 Abs. 1,⁶¹ ergeben (Art. 82 Abs. 1). Die Bemessung immaterieller Schäden ist bislang noch mit erheblicher Unsicherheit verbunden. Zwar formuliert EG 146 S. 3, dass der Schadensbegriff weit und den Zielen der Verordnung entsprechend auszulegen ist, woraus einhellig eine Abschreckungsfunktion des Schadensersatzes hergeleitet wird.⁶² Uneinigkeit besteht aber über die Folgen dieser Abschreckungswirkung für die konkrete Bemessung immaterieller Schäden.⁶³ In der Literatur wird vorgeschlagen die kommerzielle Nutzbarkeit der rechtswidrig verarbeiteten Daten zu berücksichtigen, um die Wirksamkeit der präventiven Schadensfunktion abzusichern, und die Kriterien des Art. 83 Abs. 2 als Orientierung heranzuziehen.⁶⁴ Die Priorisierung kommerzieller Interessen zulasten der Umsetzung der Pflichten aus Art. 25 Abs. 1 würde sich dann auch auf die Schadensersatzhaftung zusätzlich negativ auswirken, wäre also mit einer weiteren Abschreckungswirkung verbunden. In der deutschen Rechtsprechungspraxis

⁵⁵ Allgemeiner zum verhaltenslenkenden Potential von Geldbußen, [Gr16], 290 ff.

⁵⁶ Aus dem Wortlaut des Art. 83 Abs. 4 „werden [...] verhängt“ folgt eine gebundene Entscheidung, Boehm, [SHSgD19], Art. 83 Rn 15, Bergt, [KB20], Art. 83 Rn 30 ff.

⁵⁷ Hollaender, [WB20], Art. 83 Rn 36; zur Anreizwirkung siehe auch [RG20], 37 (56)

⁵⁸ Vgl. auch Boehm, [SHSgD19], Art. 83 Rn 28.

⁵⁹ Näher hierzu Bergt, [KB20], Art. 83 Rn 57.

⁶⁰ Näher hierzu Boehm, [SHSgD19], Art. 82 Rn 10, 14.

⁶¹ Nach a.A. handelt es sich bei Art. 25 nur um eine Verfahrensregelung, deren Nichtbeachtung für die Rechtmäßigkeit der Verarbeitung unerheblich ist und keine Schadensersatzpflicht begründet, Nolte/Werkmeister, [Go18], Art. 25 DSGVO Rn 3, 34.

⁶² M.w.N. Bergt, [KB20], Art. 82 Rn 17; [Ko21], 978 (979).

⁶³ Insbesondere über die Grenze zum „Strafschadensersatz“, vgl. Boehm, [SHSgD19], Art. 82 Rn 27. Eine mögliche strafende Wirkung zu hoher Schadensersatzposten betonen [We19], 295 und [Wy19], 3266 f.

⁶⁴ Boehm, [SHSgD19], Art. 82 Rn 27; Quaas, [WB20], Art. 82 Rn 31.

wurde der immaterielle Schadensersatz bislang eher zurückhaltend bemessen.⁶⁵ Europaweit soll inzwischen eine steigende Tendenz zu beobachten sein.⁶⁶ Über die Auslegung von Art. 82, insbesondere über die Frage nach einer Erheblichkeitsschwelle, wird demnächst der EuGH entscheiden müssen.⁶⁷ Wie die erforderliche Konkretisierung der Grundsätze zur immateriellen Schadensbemessung auch ausfallen wird, die Schadensersatzhaftung ist eine mehr oder weniger gewichtige negative Konsequenz, die unmittelbar aus einer Datenverarbeitung ohne geeignete Schutzmaßnahmen gem. Art. 25 Abs. 1 folgen kann. Ob sie im konkreten Fall als Folge eines DSGVO-Verstoßes eintritt, hängt indessen auch von Faktoren aus der Betroffenenosphäre, etwa der Kenntnis von und der Initiative zur Wahrnehmung von Schadensersatzansprüchen, der Beweislastverteilung und Beweismöglichkeiten sowie des zu erwartenden Kosten-Nutzen-Verhältnisses eines möglichen Prozesses ab. Während über Beweiserleichterungen oder sogar eine Beweislastumkehr zugunsten der Betroffenen Uneinigkeit besteht,⁶⁸ bezwecken die Erweiterungen der Vertretungs- und Verbandsklagerechte nach Art. 80 DSGVO zweifelsfrei eine Erleichterung der Klagemöglichkeiten für Betroffene⁶⁹ und werden zukünftig um die Möglichkeit kollektiver datenschutzrechtlicher Klagen aufgrund der kürzlich verabschiedeten EU-Verbandsklage-Richtlinie ergänzt werden.⁷⁰ Die praktische Relevanz der Schadensersatzhaftung für die Verantwortlichen könnte demnach durch erweiterte prozessuale Möglichkeiten datenschutzrechtlicher Massenverfahren und Sammelklagen weiter zunehmen.⁷¹ Gerade eine datenschutzrechtlich unzureichende Technikgestaltung, die in der Regel größere Personengruppen, beispielsweise den gesamten Kundenstamm eines datenverarbeitenden Unternehmens, betreffen, eignen sich für die geschäftsmäßige Durchsetzung von Ersatzansprüchen sowie Massen- und Sammelverfahren.⁷² Abseits der im Recht angelegten finanziellen Negativanreize treten zu erwartende Reputationsverluste, insbesondere bei öffentlichkeitswirksamen Prozessen, die eine Vielzahl an Personen betreffen, als zusätzlicher ökonomisch relevanter Kostenfaktor des DSGVO-Verstoßes hinzu.⁷³

⁶⁵ [KI20], 433; [Wy19], 3267 f; [Pa21a], 242 f, wobei die Arbeitsgerichtsbarkeit zur höheren Bußgeldbemessung tendiert, [Fu20], 565; [Ko21], 979 f; [Le21]; [Pa21a], 244.

⁶⁶ [Th20], 503; weniger deutlich bei der Rechtsprechungsauswertung anderer Mitgliedstaaten dagegen, [Pa21a], 244 f.

⁶⁷ Das BVerfG gab einer Verfassungsbeschwerde gegen das Urteil des AG Goslar, das einen Schadensersatzanspruch wegen Unerheblichkeit ablehnte, wegen einer Verletzung von Art. 101 Abs. 1 S.2 GG statt und verpflichtete das Gericht zur Vorlage beim EuGH, BVerfG Beschl. v. 14.01.2021 – 1 BvR 2853/19 = BeckRS 1962; dazu [Ko21], 978. Die Literatur spricht sich überwiegend gegen eine Bagatellgrenze aus, Quaas, [WB20], Art. 82 Rn 31; Bergt, [KB20], Art. 82 Rn 18a; Gola/Piltz, [Go18], Art. 82; Rn 13; a.A. Frenzel, [Pa21b], Art. 82 Rn 10; wohl auch [Wy19], 3265 (3266 ff); [Sp19], 476; differenzierend [Pa21a], 245 f.

⁶⁸ Jeweils m.w.N. für eine Beweislastumkehr, [Fu20], 565; dagegen [Wy19], 3268; [Wy21], 1191 f.

⁶⁹ Boehm, [SHSgD19], Art. 80 Rn.2.

⁷⁰ RL (EU) 2020/1828, <https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32020L1828&from=DE>; [Ue19], 702 f.

⁷¹ Zu den Grenzen der Schadensersatzdurchsetzung über UwG und UKlaG, [Ue19], insbesondere bzgl. Art. 25 Abs.1, 694 (700); zur Konstellation einer datenschutzrechtlichen Musterfeststellungsklage, [GS19], a.A. [Sp19], 477; zur Verbandsklage-RL [Pa21a], 246 f.

⁷² Vgl. zur Kommerzialisierung der Schadensersatzklagen in solchen Konstellationen, [Sp19]; [Wy21], 1190 f.

⁷³ vgl. [We19], 294; [GS19], 3417 f; bzgl. Bußgeldverfahren, [Gr16], 299 ff.

4.3 Zertifizierung als Positivanreiz

Eine wesentliche Erweiterung brachte die DSGVO für Möglichkeiten der sog. regulierten Selbstregulierung.⁷⁴ Neben der beschriebenen Konkretisierungsverantwortung der Verantwortlichen bezüglich ihrer eigenen Rechtspflichten, können sie sich freiwillig zur Einhaltung genehmigter Verhaltensregeln verpflichten (Art. 40 f) oder eine Zertifizierung (Art. 42 f) beantragen. Private Stellen werden dabei unter behördlicher Kontrolle in die Skalierung und Spezifizierung der Verordnungspflichten und in die Überprüfung ihrer Einhaltung eingebunden.⁷⁵ Für Unternehmen kann die Teilnahme an solchen Selbstregulierungsverfahren insbesondere vorteilhaft sein, um ein gewisses Maß an Rechtssicherheit zu erlangen,⁷⁶ und zumindest faktisch die Wahrscheinlichkeit einer behördlichen Überprüfung reduzieren.⁷⁷

Aus Art. 25 Abs. 3 folgt ein unmittelbarer rechtlicher Vorteil der Zertifizierung, indem sie als Faktor zum Nachweis geeigneter Maßnahmen anerkannt wird.⁷⁸ Verarbeitungsvorgänge, nicht die Technikgestaltung als solche oder ein Produkt selbst,⁷⁹ müssen mindestens als „datenschutzkonform“ zertifiziert werden, während eine Zertifizierung als „datenschutzfreundlich“ im Sinne eines abgestuften Konzepts nach einheitlich festgelegten höheren Standards nicht vorgesehen ist.⁸⁰ Enthalten die Zertifizierungskriterien konkrete Anforderungen an Maßnahmen nach Art. 25 Abs. 1, kann die erfolgreiche Zertifizierung deren Geeignetheit implizieren. Sie darf allerdings nicht als "Freischein" missverstanden werden, die den Verantwortlichen von einer kontinuierlichen Prüfung des Risikos und der getroffenen Maßnahmen entbinden könnte oder vor einer Überprüfung durch die zuständige Aufsichtsbehörde schützen würde.⁸¹ Der, wenn auch beschränkte, Anreiz der Nachweiserleichterung wird durch die Einbeziehung der Einhaltung von Zertifizierungskriterien bei der Bußgeldmessung gem. Art. 83 Abs. 2 lit. j ergänzt.

Von größerer Bedeutung für die unternehmerische Strategie dürften jedoch die potentiellen ökonomischen Vorteile einer Zertifizierung sein. Sie ermöglicht einen werbewirksamen Einsatz und damit eine Differenzierung datenschutzrechtlich geprüfter Angebote von solchen konkurrierender Marktteilnehmer. Zertifikate und Gütesiegel können in einem entstehenden Wettbewerb datenschutzfreundlicher Technikgestaltung als vertrauenswürdige Kommunikationsmittel gegenüber dem Verbraucher dienen und die Wettbewerbsfähigkeit

⁷⁴ [K6c], 448 (452); zum Konzept der regulierten Selbstregulierung, [Po17], 209 ff.

⁷⁵ Roßnagel, [SHSgD19], Art. 40 Rn 3; Art. 42 Rn 4; [Ka20], 1599 ff; bzgl. der Verhaltensregeln, [Re19], 305 f.

⁷⁶ EDSA, [ED], 8 ff; Scholz, [SHSgD19], Art. 42 Rn 5.

⁷⁷ Vgl. Scholz, [SHSgD19], Art. 42 Rn 33.

⁷⁸ Teilweise wird vertreten, dass dieser Vorteil auch für genehmigte Verhaltensregelungen gilt und es sich bei der ausschließlichen Nennung der Zertifizierung um ein Redaktionsversehen handelt, so Baumgartner, [Eh18], Art. 25 Rn 22; zurückhaltender, aber mit Verweis auf die zumindest praktische Nachweiserleichterung, Martini, [Pa21b], Art. 25 Rn 53a.

⁷⁹ Hansen, [SHSgD19], Art. 42 Rn 21 ff, 25 f.

⁸⁰ Solche höheren Standards können aber grundsätzlich im Rahmen der DSGVO-Zielvorgaben vorgesehen werden, [Ho19b], 219 ff.

⁸¹ Hansen, [SHSgD19], Art. 25 Rn 58.

eines Unternehmens verbessern.⁸² Ein ökonomischer Vorteil kann sich zudem bei der Vergabe öffentlicher Aufträge ergeben, für deren Ausschreibungen datenschutzfreundliche Technikgestaltung nach EG 78 S. 5 berücksichtigt werden sollten.⁸³

5 Spieltheoretisches Modell

Die Wirksamkeit der rechtlichen Steuerungselemente Sanktionierung, Haftung und Zertifizierung bei deren Interaktion mit dem unternehmerischen Gewinnkalkül zur Umsetzung von Art. 25 Abs. 1 sollen durch nachfolgendes spieltheoretisches Modell beleuchtet werden. Grundsätzlich besteht jedoch eine Übertragbarkeit auf andere DSGVO-Verstöße. Die Spieltheorie ist eine formal mathematische Methodik zur Analyse von Entscheidungssituationen. Insbesondere eignen sich spieltheoretische Untersuchungen, um die Auswirkungen und Interdependenzen von Entscheidungen unterschiedlicher Akteure – also Entscheidungen mit strategischem Charakter - abzubilden.⁸⁴ Hierbei ist eine Entscheidung als eine bewusste Wahl zwischen unterschiedlichen Handlungsalternativen gemeint. Spieltheoretische Analysen befassen sich demnach mit interaktiven und strategischen Entscheidungssituationen, bei denen mindestens zwei Entscheider zwischen mindestens zwei möglichen Handlungsalternativen auswählen können.⁸⁵ Ein Entscheider muss aufgrund der Interdependenzen antizipieren, wie sich sein gegenüber verhalten wird, um zu einer optimalen Entscheidung zu finden.

Definition 1 (Spiel) *Gegeben sei eine Entscheidungssituation, in der n Spieler aufeinander treffen. Jeder dieser Spieler sei durch die Menge S möglicher Strategien (als dem vollständigen Plan über alle Aktionen innerhalb einer Entscheidungssituation) sowie durch den Nutzen u des Entscheidungsträgers, welche abhängig von der getroffenen Strategie eine Auszahlung wiedergibt, beschreibbar. Auf diese Weise ist ein Spiel $\Gamma(N, S, u)$ durch die Spieler, ihre Strategien und ihre Nutzen eindeutig beschrieben.*

5.1 Basisspiel

Ein Unternehmen (F) ist im gegebenen Anwendungsfall in ein Simultanspiel⁸⁶ mit der Datenschutzaufsichtsbehörde bzw. ihrem staatlichen Rechtsträger (Land, Bund) (S) involviert. Das Unternehmen kann dabei entscheiden wie es die Spielräume bei der Umsetzung technisch implementierten Datenschutzes nutzt. Die Datenschutzaufsichtsbehörde determiniert die Kontrollintensität (z.B. Häufigkeit der Kontrollen) im Rahmen der ihr im Bundes- oder

⁸² Vgl. EG 100; Scholz, [SHSgD19], Art. 42 Rn 4; Bergt/Pesch, [KB20], Art. 42 Rn 1 ff; allgemeiner [Ho19b], 40 f; [Ri13], 250; [RBW16], 248 f.

⁸³ Siehe auch [RG20], 54.

⁸⁴ [BW20].

⁸⁵ [BW20].

⁸⁶ Vereinfacht gesagt, treffen die Spieler in einem Simultanspiel ihre Entscheidungen zeitgleich. Somit sind die Entscheidungen im Moment ihrer Entscheidung unbeobachtbar voneinander, [BW20].

Landeshaushalt zur Verfügung gestellten Ressourcen. Der Unternehmensgewinn setzt sich aus einem vom Datenschutzniveau unabhängigen Betriebsergebnis (g_0), einem Zusatzgewinn durch nicht legitimierte Datennutzung im Umfang von x ($a \cdot x$), Kosten für technisch wirksamen Datenschutz ($c_F \cdot (1 - x)$) sowie der zu erwartenden Strafe in Höhe von $q \cdot (S + H) \cdot x$ (Erwartungswert der Sanktionierung, Schadensersatzpflichten) zusammen. Die variable q beschreibt die Kontrollwahrscheinlichkeit, S die Sanktion sowie H den Erwartungswert der Schadensersatzpflichten. Letztere fallen dabei nur dann an, wenn eine Verletzung der DSGVO nachgewiesen wurde, d.h. diese können nur durch eine (erfolgreiche) Kontrolle entstehen. Aus der Nutzenfunktion lässt sich die optimale Datennutzungsstrategie x^* der Firma ableiten.

$$\Pi_F = g_0 + (a - q \cdot (S + H)) \cdot x - \frac{c_F \cdot (1 - x)^2}{2} \Rightarrow x^* = \frac{a + c_F - q \cdot (S + H)}{c_F} \quad (1)$$

Ein Unternehmen wird also weitreichender (personenbezogene) Daten verarbeiten, je höher der Nutzen daraus, je höher die Kosten der Implementierung datenschutzsteigernder Technologien und je kleiner die erwartete Strafe (Schadensersatzpflichten und Bußgelder). Die Behörde hat das Ziel sicherzustellen, dass sich das Unternehmen datenschutzkonform ($x = 0$) verhält. Es soll somit garantiert sein, dass vom Unternehmen geeignete Maßnahmen der Technikgestaltung getroffen wurden. Dabei spiegelt der Präferenzfaktor b wider, welchen Wert die Behörde einer wirksamen Durchsetzung des Datenschutzes durch Technikgestaltung beimisst. c_S seien die Kosten der Kontrolle. Es wird angenommen, dass die Kontrollkosten des Staates überproportional in der Kontrollintensität ansteigen, was in Opportunitätskosten der Behörde sowie ihrem begrenzten Budget begründet liegt. Deckt die Behörde einen Verstoß gegen Art. 25 Abs. 1 DSGVO auf, so entsteht ihr ein Nutzen in Höhe von V .

$$\Pi_S = b \cdot (1 - x) + x \cdot q \cdot V - \frac{c_A \cdot q^2}{2} \Rightarrow q^* = \frac{V \cdot x}{c_S} \quad (2)$$

Analog zum Unternehmen lässt sich für die Behörde die optimale Kontrollintensität q^* ableiten. Die Behörde wählt eine umso höhere Kontrollintensität, je höher der Nutzen eines entdeckten Regelverstoßes, je höher die (erwartete) Intensität des Datenmissbrauchs und je niedriger ihre Kontrollkosten sind. In einem Simultanspiel ergibt sich das sogenannte Nash-Gleichgewicht (NGG),⁸⁷ das die wechselseitig optimalen Strategien bestimmt, damit wie in Formel 3 dargestellt.

$$x_{NGG}^* = \frac{(a + c_F) \cdot c_S}{c_F \cdot c_S + V \cdot (H + S)}; q_{NGG}^* = \frac{(a + c_F) \cdot V}{c_F \cdot c_S + V \cdot (H + S)} \quad (3)$$

Die Parameter a und c_F treiben sowohl x als auch q in die Höhe. Darüber hinaus gilt, je höher die Kontrollkosten, umso größer x . Der Zusammenhang mit der Höhe des Nutzens

⁸⁷ Zum Lösungskonzept des NGG siehe [BW20].

eines entdeckten Regelverstoßes kann je nach Parameterkonstellation sowohl positiv als auch negativ sein. Abbildung 1(a) stellt die sogenannten Reaktionskurven der beiden Parteien dar. Diese stellen die jeweils optimale Strategie eines Spielers in Abhängigkeit der Strategie des anderen Spielers grafisch dar. Der Schnittpunkt der Kurven entspricht somit dem NGG des Spiels. Wie gezeigt, steigt die Kontrollintensität q , je höher die rechtswidrige Datenverarbeitung und umgekehrt, d.h. je höher die Kontrollintensität desto niedriger die DSGVO-Verstöße.

5.2 Unvollkommene Information

Wir unterscheiden nun zwischen zwei Typen von Firmen: Eine (+)Firma, die einen geringen Nutzen aus intensiver (bzw. missbräuchlicher) Datenverarbeitung zieht (bspw. hohe Sensitivität der Kunden und/oder niedriger Informationswert, niedriger Wert für a), und eine (-)Firma, die einen hohen Gewinn durch Datenverarbeitung erzielt (bspw. niedrige Sensitivität der Kunden, hoher Informationswert, hoher Wert für a). Die (-)Firma wird somit einen höheren x -Wert wählen als die (+)Firma. Die Behörde kann vor der Kontrolle aber nicht zwischen den Firmen unterscheiden und wählt daher eine mittlere Kontrollintensität q . Die optimale Kontrollintensität ist nun auch abhängig von den Erwartungen der Behörde, bei einer zufälligen Kontrolle eine (+)Firma oder eine (-)Firma anzutreffen (abhängig von den Wahrscheinlichkeiten p respektive $1 - p$).

Die generelle Struktur der Nutzenfunktion der Unternehmen bleibt unverändert, diese unterscheiden sich nur in der Höhe von a . Die Reaktionsfunktionen ergeben sich wie in Abbildung 1(b) dargestellt. Die Behörde orientiert sich an der mittleren Reaktionsfunktion der beiden Firmen-Typen und bestimmt so ihre optimale Kontrollintensität q . Die Firmen wiederum passen sich optimal an dieses mittlere Kontrollniveau an, woraus sich das NGG bei unvollkommener Information, das sog. Bayes-Nash-Gleichgewicht (BNGG),⁸⁸ ergibt.

$$q^* = \frac{(a_- + c_F - (a_- + a_+)p)S}{c_F c_S + S(H + S)}; x_+^* = \frac{c_F(a_+ + c_F)c_S + (a_- - a_+)(p - 1)V(H + S)}{c_F(c_F c_S + V(H + S))} \quad (4)$$

$$x_-^* = \frac{c_F(a_- + c_F)c_S + (a_- - a_+)pV(H + S)}{c_F(c_F c_S + V(H + S))}$$

Im BNGG kontrolliert die Behörde also beide Firmen mit der Kontrollwahrscheinlichkeit q^* . Die (+)-Firma wählt das (niedrigere) Datenmissbrauchslevel x_+^* und die (-) Firma wählt das (höhere) Datenmissbrauchslevel x_-^* . Die beiden schwarzen Punkte in der Grafik geben das BNGG an (optimale Reaktion von jedem Firmentypen auf die durchschnittliche Erwartung der Behörde). Die beiden grünen Punkte zeigen, welche NGG zustande kämen, wenn die Behörde beide Firmen (perfekt) unterscheiden könnte. An der Grafik wird erkennbar, dass die fehlende Unterscheidbarkeit der beiden Firmentypen durch die Behörde zu einem

⁸⁸ Zum Lösungskonzept des BNGG siehe [BW20].

hohen Effizienzverlust führt: Da die Behörde eine mittlere Kontrollintensität wählt, wird die (+)Firma zu stark und die (-)Firma vergleichsweise zu wenig kontrolliert. Genau an diesem Problem setzt der Gedanke der Zertifizierung an.

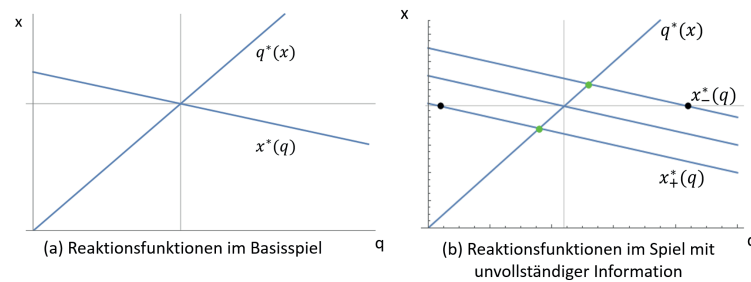


Abb. 1: Reaktionskurven der Behörde und des Unternehmens in beiden Spielen

5.3 Möglichkeiten der Zertifizierung und das Potential für die Herausbildung eines Marktes für datenschutzfreundliche Produkte

In der vorangegangenen spieltheoretischen Modellierung wurde aufgezeigt, dass allein die Wirkung der sanktionierenden Maßnahmen sowie die Haftungsansprüche dafür sorgen können, dass es für einige Unternehmen optimal ist geeignete technische Maßnahmen zum Schutz der Daten zu implementieren. Aufgrund der Intransparenz kann die behördliche Kontrolle die Firmen jedoch nicht zielgerichtet adressieren, was zu Effizienzverlusten und einem niedrigeren Datenschutzniveau führt. Die Zertifizierung bietet nun eine potentielle Lösung dieses Problems: Da es der (+)Firma leichter fällt als der (-)Firma, die Anforderungen an die Zertifizierung zu erbringen, können die beiden Firmen-Typen für die Behörde unterscheidbar werden. Der Vorteil für die (+)Firma liegt auf der Hand: Als zertifizierte Firma hat sie fortan mit einer niedrigeren Kontrollintensität zu rechnen, da sich die Behörde vor allem auf (-)Firmen konzentrieren kann. Darüber hinaus bietet die Aussicht auf diese günstige Position der (+)Firma weitere Anreize, in datenschutzfreundliche Verarbeitungsvorgänge und Technologien zu investieren. Dies wiederum würde Herstellern datenschutzfreundlicher Produkte den Anreiz geben, ihr Angebot auszubauen. Gleichzeitig hätte die zertifizierte (+)Firma die Chance, einen Standard für datenschutzkonforme Technikgestaltung zu setzen. Um jedoch sicherzustellen, dass die Zertifizierung ein glaubwürdiges Signal für alle Marktteilnehmer darstellt, müssen mehrere Voraussetzungen erfüllt sein. Zunächst muss sich für die betreffenden Unternehmen ein klarer Vorteil durch die Zertifizierung ergeben. Neben der bereits angesprochenen niedrigeren Kontrolle durch die Behörde kann auch die Wirkung auf die Kunden von Relevanz sein. Eine zertifizierte (+)Firma kann die Zahlungsbereitschaft oder die Loyalität der Kunden erhöhen, wodurch sich ihre Position am Markt verbessert. Inwiefern diese Marktvorteile für datenschutzfreundliche Produkte

jedoch so ausgeprägt sind, ist dabei Thema teils kontrovers geführter Diskussionen.⁸⁹ Die aus spieltheoretischer Sicht größere Hürde liegt jedoch bei der (-)Firma. Je größer der Vorteil einer Zertifizierung, desto eher wird auch die (-)Firma bemüht sein, zumindest einmalig aus taktischen Erwägungen heraus die Anforderungen an die Zertifizierung zu erfüllen, um anschließend wieder der opportunistischen Datennutzungsstrategie zu folgen. Die (-)Firma muss also entweder durch zu große Hürden (etwa der Pflicht zur regelmäßigen Erneuerung des Zertifikats oder adäquat hohe Kontrollintensität des ausstellenden Instituts) davon abgehalten werden, die Zertifizierung aus rein taktischen Gründen anzustreben oder es gelingt, mithilfe der (+)Firmen eine Marktdynamik in Richtung eines neuen Datenschutz-Standards anzustoßen, den auch die (-)Firma mitgehen muss. Hierzu könnte eine vertiefte Kooperation der Behörde mit zertifizierten (+)Firmen der erste Schritt sein.

6 Fazit und Ausblick

Für die datenschutzrechtliche risikobasierte Regulierung im Bereich der Technikgestaltung sind klare abschreckende Wirkungen durch Bußgeld- und Haftungsrisiken als Negativanreize grundsätzlich geeignet das Verhalten datenverarbeitender Unternehmen innerhalb ihres Umsetzungsspielraums zu lenken. Die Lenkungswirkung hängt maßgeblich von der Aufdeckungswahrscheinlichkeit eines Verstoßes, d.h. mittelbar von der behördlichen Kontrollintensität, sowie der Höhe der Bußgelder und Schadensersatzposten ab. Die hinreichende Ausstattung der Aufsichtsbehörden ist demnach zentral für die Wirksamkeit der Negativanreize. Auch wenn Aufgaben und Kompetenzen der Aufsichtsbehörden unberührt bleiben, können sichere Zertifizierungsprogramme außerdem zu einer effizienten und ressourcenschonenden behördlichen Praxis beitragen. Ökonomische Anreize einer Zertifizierung können geeignet sein datenschutzfreundliche Technikgestaltung zu fördern, wenn die Zertifizierung ihrerseits eine sichere Erkennbarkeit datenschutzkonformer und datenschutzfreundlicher Technikgestaltung für den Verbraucher ermöglicht. Die Aussagekraft von Zertifikaten nach dem gegenwärtigen Konzept ist angesichts abstrakter Vorgaben, mangelnder dauerhafter Überprüfung und Anpassung⁹⁰ und der Heterogenität nebeneinander bestehender Zertifikate- und Zertifizierungsverfahren eingeschränkt.⁹¹ Eine rechtssichere graduelle Abstufung von Zertifikaten nach der Datenschutzfreundlichkeit von Verarbeitungen böte eine optimale Entscheidungsgrundlage für Verbraucher und sollte für eine innovationsfördernde rechtliche Regulierung, die mit Marktanreizen arbeitet, bedacht werden. Hohe Standards, eine sichere Überprüfung der Zertifizierungsverfahren und Konzepte der "dynamischen Zertifizierung"⁹² sollten zudem als zentrale Faktoren für die Vermeidung taktischer Zertifizierungen, die die Glaubwürdigkeit von Zertifikaten beeinträchtigen und damit ihre Anreizwirkung für datenschutzfreundliche Technikgestaltung konterkarieren würden, sichergestellt werden.

⁸⁹ [BKP12], [GA07], [AJL13].

⁹⁰ [Ho19b], 290 ff.

⁹¹ [RBW16], 250 ff; [RG20], 53 f.

⁹² [Ho19b]

Literaturverzeichnis

- [Ad20] Adam, Simon: Daten als Rechtsobjekte. NJW, 73:2063–2068, 2020.
- [AJL13] Acquisti, Alessandro; John, Leslie K; Loewenstein, George: What is privacy worth? The Journal of Legal Studies, 42(2):249–274, 2013.
- [Ar] Guidelines in Data Protection Impact Assessment (DPIA) and determining whether processing is „likely to result in a high risk“ for the purposes of Regulation 2016/679 (WP 248 rev. 01). https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611236, Stand: 30.04.21.
- [Ar14] Statement on the role of a risk-based approach in data protection legal frameworks (WP 218). <http://www.privacy-regulation.eu/privazyplan/materialien/eu-artikel-29-gruppe-workingpaper/wp218%20EN%20Role%20of%20risk-based%20approach%20in%20data%20protection%20legal%20frameworks.pdf>, Stand: 30.04.21.
- [BKP12] Beresford, Alastair R; Kübler, Dorothea; Preibusch, Sören: Unwillingness to pay for privacy: A field experiment. Economics letters, 117(1):25–27, 2012.
- [Br10] Broemel, Roland: Strategisches Verhalten in der Regulierung. Mohr Siebeck, 2010.
- [BTH14] Birnhack, Michael; Toch, Erin; Hadar, Irit: Privacy Mindset, Technology Mindset. Jurimetrics, 55:55–114, 2014.
- [Bu] Rede der Bundeskanzlerin Dr. Angela Merkel beim Jahrestreffen des World Economics Forum am 24. Januar 2018 in Davos. <https://www.bundesregierung.de/resource/blob/975954/775758/16212f149c37f3bf3641872d146533ae/10-1-bkin-davos-data.pdf?download=1>, Stand: 30.04.21.
- [Bu20] Buss, Sebastian: Privacy by Design und Software. Berücksichtigung datenschutzfreundlicher Anforderungen bei der Softwarebeschaffung. CR, 35:1–6, 2020.
- [BW20] Bartholomae, Florian; Wiens, Marcus: Spieltheorie- Ein anwendungsorientiertes Lehrbuch. Springer Gabler, 2020.
- [By17] Bygrave, Lee A.: Data Protection by Design and Default: Deciphering the EU’s Legislative Requirements. Oslo Law Review, 4:105–120, 2017.
- [Ce] A Risk-based Approach to Privacy: Improving Effectiveness in Practice. https://www.huntonak.com/files/upload/Post-Paris_Risk_Paper_June_2014.pdf, Stand:30.04.21.
- [Da19] Gutachten der Datenethikkommission der Bundesregierung. https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.pdf?__blob=publicationFile&v=6, Stand:30.04.21.
- [Dr17] Drexl, Josef: Neue Regeln für die Europäische Datenwirtschaft. NZKart, 5:339–344, 2017.
- [DS] Standard-Datenschutzmodell, Version 2.0b. https://www.bfdi.bund.de/DE/Datenschutz/Themen/Technische_Anwendungen/TechnischeAnwendungenArtikel/Standard-Datenschutzmodell.html, Stand: 30.04.21.

- [ED] Leitlinien 1/2019 über Verhaltensregeln und Überwachungsstellen gemäß der Verordnung (EU)2016/679, Fassung 2.0. https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines_201901_v2.0_codesofconduct_de.pdf, Stand: 30.04.21.
- [Eh18] Ehmann, Eugen und Selmayr, Martin: DS-GVO. Datenschutz-Grundverordnung.Kommentar. C.H.Beck, 2. Auflage, 2018.
- [EUa] Mitteilung der Kommission an das Europäische Parlament und den Rat, COM(2020) 264 final. <https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:52020DC0264&from=EN>, Stand: 30.04.21.
- [Eub] Leitlinien 4/2019 zu Art. 25, Version 2. https://edpb.europa.eu/system/files/2021-04/edpb_guidelines_201904_dataprotection_by_design_and_by_default_v2.0_de.pdf, Stand: 30.04.21.
- [Eu15] Readiness Analysis for the Adoption and Evolution of Privacy Enhancing Technologies. https://www.enisa.europa.eu/publications/pets/at_download/fullReport, Stand: 30.04.21.
- [Fr] Merkel: Daten sind die Rohstoffe des 21. Jahrhunderts. <https://www.faz.net/aktuell/wirtschaft/cebit/angela-merkel-fordert-mehr-modernisierte-digitale-technologien-14120493.html>, Stand:30.04.21.
- [Fu20] Fuhlrott, Michael und Oltmanns, Sönke: Immaterieller Schadensersatz wegen Datenschutzverstoß: Höhe und Bemessungsfaktoren. *ArbAktuell*, 12:565–567, 2020.
- [GA07] Grossklags, Jens; Acquisti, Alessandro: When 25 Cents is Too Much: An Experiment on Willingness-To-Sell and Willingness-To-Protect Personal Information. In: *WEIS*. 2007.
- [Go18] Gola, Peter: Datenschutz-Grundverordnung. VO (EU) 2016/679. Kommentar. C.H.Beck, 2. Auflage, 2018.
- [Gr16] Grant, Hazel und Crowther, Hannah: How Effective are Fines in Enforcing Privacy? In (Wright, David und De Hert, Paul, Hrsg.): *Enforcing Privacy. Regulatory, Legal and Technology Approaches*. Springer International, S. 287–305, 2016.
- [Gr19] Greze, Benjamin: The extra-territorial enforcement of the GDPR: a genuine issue and the quest for alternatives. *International Data Privacy Law*, 9:109–128, 2019.
- [GS19] Geissler, Dennis; Stroebel, Lukas: Datenschutzrechtliche Schadensersatzansprüche im Musterfeststellungsverfahren. *NJW*, 72(47):3414–3418, 2019.
- [Hä16] Härting, Niko: Digital Goods und Datenschutz-Daten sparen oder monetarisieren? *CR*, 32(11):735, 2016.
- [HG16] Heuberger-Götsch, Olivier: Der Wert von Daten aus juristischer Sicht am Beispiel des Profiling. In: *Big Data*, S. 83–105. Springer, 2016.
- [Ho19a] Hoeren, Thomas: Datenbesitz statt Dateneigentum. *MMR*, 22:5–8, 2019.
- [Ho19b] Hofmann, Johanna M.: Dynamische Zertifizierung. Datenschutzrechtliche Zertifizierung nach der Datenschutz-Grundverordnung am Beispiel des Cloud Computing. *Nomos*, 2019.

- [K6a] Körber, Torsten: Ist Wissen Marktmacht? Überlegungen zum Verhältnis von Datenschutz, "Datenmacht und Kartellrecht - Teil 2. NZKart, 7:348–356, 2016.
- [K6b] Körber, Torsten: „Ist Wissen Marktmacht?“ Ueberlegungen zum Verhaeltnis von Datenschutz, „Datenmacht“ und Kartellrecht – Teil 1. NZKart, 4:303–310, 2016.
- [K6c] Kühling, Jürgen und Martini, Mario: Die Datenschutz-Grundverordnung: Revolution oder Evolution im europäischen und deutschen Datenschutzrecht? EuZW, 27:448–454, 2016.
- [K0] Kühling, Jürgen und Sackmann, Florian: Irrweg "Dateneigentum". ZD, 10:24–30, 2020.
- [Ka20] Kaminski, Margot E.: Binary Governance: Lessons from the GDPR's Approach to Algorithmic Accountability. California Law Review, 92:1529–1616, 2020.
- [KB20] Kühling, Jürgen; Buchner, Benedikt: Datenschutz-Grundverordnung. BDSG. C.H.Beck, 2020.
- [Ke16] Kerber, Wolfgang: Digital Markets, Data and Privacy: Competition Law, Consumer Law and Data Protection. GRUR Int., 65:639–647, 2016.
- [Kl20] Klein, Susanne: Immaterieller Schadensersatz nach der DS-GVO. GRUR-Prax, 12:433–435, 2020.
- [Kl21] Klingbeil, Thilo und Kohm, Simon: Datenschutzfreundliche Technikgestaltung und ihre vertraglichen Implikationen. MMR, 24:3–8, 2021.
- [Ko21] Korch, Stefan: Schadensersatz für Datenschutzverstöße. NJW, 12:978–981, 2021.
- [La20] Latzel, Clemens: Verhaltenssteuerung, Recht und Privatautonomie. Springer, 2020.
- [Le21] Leibold, Kevin: Gerichtliche Entscheidungen zum Schadensersatz nach Art. 82 Abs.1 DSGVO - ein Fall für den EuGH. ZD-Aktuell, 11:05043, 2021.
- [Pa21a] Paal, Boris und Aliprandi, Claudio: Immaterieller Schadensersatz bei Datenschutzverstößen. ZD, 11(5):241–247, 2021.
- [Pa21b] Paal, Boris und Pauly, Daniel: Datenschutz-Grundveordnung. Bundesdatenschutzgesetz. C.H.Beck, 3. Auflage, 2021.
- [Po17] Poll, Jens: Datenschutz in und durch Unternehmensgruppen im deutschen Datenschutzrecht. Nomos, 2017.
- [Po19] Podszun, Rupprecht und Kersting, Christian: Modernisierung des Wettbewerbsrechts und Digitalisierung. NJOZ, 19:321–325, 2019.
- [Po20] Podszun, Rupprecht: Der Verbraucher als Marktakteur: Kartellrecht und Datenschutz in der "Facebook Entscheidung des BGH. GRUR, 122:1268–1276, 2020.
- [Qu17] Quelle, Claudia: The 'risk revolution' in EU data protection law: We can't have our cake and eat it, too. In (Leenes, R.; van Brakel, R.; Gutwirth, S.; De Hert, P., Hrsg.): Data Protection and Privacy: The Age of Intelligent Machines. Tilburg Law School Legal Studies Research Paper Series, 2017.

- [Qu18] Quelle, Claudia: Enhancing Compliance under the General Data Protection Regulation: The Risky Upshot of the Accountability- and Risk-Based Approach. *European Journal of Risk Regulation*, 9:502–526, 2018.
- [RBW16] Rodrigues, Rowena; Barnand-Wills, Davis: The future of privacy certification in Europe: an exploration of options under article 42 of the GDPR. *International Review of Law, Computers and Technology*, 30:248–270, 2016.
- [Re19] Reifert, Natascha: Codes of Conducts nach der DSGVO. *ZD*, 9:305–310, 2019.
- [RG20] Rubinstein, Ira; Good, Nathaniel: The trouble with Article 25 (and how to fix it): The future of data protection by design and default. *International Data Privacy Law*, 10:37–54, 2020.
- [RGR18] Reinsel, David; Gantz, John; Rydning, John: The Digitilization of the World. From Edge to Core. IDC White Paper, 2018.
- [Ri13] Richter, Frederik: Die Stiftung Datenschutz als chancenreiche Ergänzung. *ZD*, 3:249–250, 2013.
- [Ro21] Roßnagel, Alexander: Die Evaluation der Datenschutz-Grundverordnung. *MMR*, 24:657–661, 2021.
- [Sc17] Schneider, Jens-Peter: Innovationsoffene Regulierung datenbasierter Dienste in der Informationsgesellschaft. *Datenschutz, Regulierung, Wettbewerb*. In (Körper, Torsten und Kühling, Juergen, Hrsg.): *Regulierung - Wettbewerb - Innovation*, S. 113–141. Nomos, 2017.
- [Sc19] Schröder, Markus: Der risikobasierte Ansatz in der DS-GVO. *ZD*, 9:503–506, 2019.
- [SgD17a] Spiecker gen. Döhmman, Indra: Big und Smart Data: Zweckbindung zwecklos? *Spektrum der Wissenschaft*, 1.17:56–62, 2017.
- [SgD17b] Spiecker gen. Döhmman, Indra: Stichwort Datenschutz. In: *Staatslexikon:Recht - Wirtschaft - Gesellschaft*, Jgg. 1. Goerres-Gesellschaft zur Pflege der Wissenschaft, 2017.
- [SHSgD19] Simitis, Spiros; Hornung, Gerrit; Spiecker gen. Doehmann, Indra: *Datenschutzrecht. DSGVO mit BDSG*. Nomos, 2019.
- [Sp19] Spittka, Jan: Die Kommerzialisierung von Schadensersatz unter der DSGVO. *GRUR-Prax*, 11(21):475–477, 2019.
- [SS19] Spindler, Gerald; Sein, Karin: Die endgültige Richtlinie über Verträge über digitale Inhalte und Dienstleistungen. *MMR*, 22:415–420, 2019.
- [St20] Stromberg, Sabrina: Empfehlungen der Datenethikkommission: Datenschutz "by design" "by default". *ZD - Aktuell*, 10:04397, 2020.
- [Sy18] Sydow, Gernot: *Europäische Datenschutz-Grundverordnung*. Nomos, 2. Auflage. Auflage, 2018.
- [Th16] Thode, Jan-Christoph: Die neuen Compliance- Pflichten nach der Datenschutz-Grundveordnung. *CR*, 31:714–721, 2016.
- [Th20] Thiel, Barabara und Wybitul, Tim: Bußgelder wegen Datenschutzverstößen - aus Sicht von Aufsichtsbehörden und Unternehmen. *ZD*, 10:3–7, 2020.

- [Ue19] Uebele, Fabian: Datenschutz vor Zivilgerichten. GRUR, 121(7):694–703, 2019.
- [Un] Kurzpapier Nr. 18, Risiko für die Rechte und Freiheiten natürlicher Personen. https://www.datenschutzzentrum.de/uploads/dsgvo/kurzpapiere/DSK_KP_Nr_18_Risiko.pdf, Stand: 30.04.21.
- [Ve18] Veerpalu, Anne: Decentralised Technology and Technology Neutrality in Legal Rules: An Analysis of De VoogD and Hedqvist. *Baltic Journal of Law and Politics*, 11:63–94, 2018.
- [WB20] Wolff, Amadeus; Brink, Stefan: BeckOK Datenschutzrecht. C.H.Beck, 35. Auflage, 2020.
- [WDH16] Wright, David; De Hert, Paul: Introduction to Enforcing Privacy. In (Wright, David; De Hert, Paul, Hrsg.): *Enforcing Privacy. Regulatory, Legal and Technological Approaches*. Springer International, S. 1–5, 2016.
- [We19] Wenzel, Michael und Wybitul, Tim: Vermeidung hoher DS-GVO-Bußgelder und Kooperation mit Datenschutzbehörden. *ZD*, 9:290–295, 2019.
- [Wy19] Wybitul, Tim: Immaterieller Schadensersatz wegen Datenschutzverstößen – Erste Rechtsprechung der Instanzgerichte. *NJW*, 9:3265–3269, 2019.
- [Wy21] Wybitul, Tim: Verteidigung gegen Schadensersatzklagen wegen Datenschutzverstößen. *NJW*, 74(17):1190–1194, 2021.
- [Za17] Zarsky, Tal Z.: Incompatible: The GDPR in the Age of Big Data. *Seton Hall Law Review*, 47:995–1018, 2017.
- [Ze15] Zech, Herbert: Industrie 4.0 Rechtsrahmen für eine Datenwirtschaft im digitalen Binnenmarkt. *GRUR*, 117:1151–1160, 2015.
- [Zu15] Zuboff, Shoshana: Big other: surveillance capitalism and the prospects of an information civilization. *JIT*, 30(1):75–89, 2015.

Konzept eines Modells zur ganzheitlichen Datenschutzbe- trachtung unter Anwendung von Data Mining

Optimierung des Vorgehens zur Dokumentation personenbezogener Daten und Verarbeitungstätigkeiten durch den Einsatz innovativer Technologien

Can Gümüş¹, Wolfgang Köhler², Prof. Dr. Christian Schultz³ und Prof. Dr. Christoph Rasche⁴

Abstract: Während die Digitalisierung weiter voranschreitet und immer größere Datenmengen verarbeitet werden, müssen zeitgleich steigende gesetzliche Anforderungen im Umgang mit Daten, insbesondere zum Schutz der Rechte und Freiheiten natürlicher Personen beachtet werden. Um die gesetzliche Konformität von Datenverarbeitungsprozessen sicherzustellen, sind Organisationen in der Pflicht, Transparenz über Verfahren zur Erfassung und Verarbeitung personenbezogener Daten herzustellen. Unternehmen greifen zunehmend auf innovative Datenanalytik-Technologien zurück, um Analysen großer Datenmengen durchführen zu können und Muster von oder Verbindungen zwischen Daten zu erkennen. Der Beitrag nimmt sich der Optimierung des Dokumentations- und Aktualisierungsprozesses von Verarbeitungstätigkeiten an und befasst sich mit der Entwicklung des Cross-Industry Standard Process for Data Mining Modells (CRISP-DM) zur Wahrung der Konformität durch den Einsatz von Data Mining.

Keywords: Datenschutz; Personenbezogene Daten; Verarbeitungstätigkeiten; Digitalisierung; Data Mining; Künstliche Intelligenz; Cross-Industry Standard Process for Data Mining

1 Einführung

Der fortschreitende Wandel hin zu digitalen Geschäftsmodellen und Arbeitsprozessen macht die Erhebung stetig wachsender Datenmengen notwendig. Gleichzeitig treten immer mehr Gesetze zum Schutz der Privatsphäre natürlicher Personen und deren personenbezogenen Daten in Kraft. Organisationen sehen sich zunehmend damit konfrontiert, detaillierte Informationen über Praktiken der Datenerfassung und -verarbeitung zu dokumentieren und Transparenz hinsichtlich interner Verfahren im Umgang mit Verarbeitungsprozessen sicherzustellen. So sind etwa Umfang und Zweck einer Verarbeitung vor Beginn der Datenverarbeitung gegenüber betroffenen Personen offenzulegen (Art. 13

¹ Burgstraße 29, 72213 Altensteig, can.guemues@gmx.de

² Universität Potsdam, Management, Professional Services & Sportökonomie, Karl-Liebknecht-Str. 24-25
14476 Potsdam, wkoehler@uni-potsdam.de

³ VICTORIA Internationale Hochschule, Business Studies, Bernburger Straße 24/25, 10963 Berlin, christian.schultz@victoria-hochschule.de

⁴ Universität Potsdam, Management, Professional Services & Sportökonomie, Karl-Liebknecht-Str. 24-25
14476 Potsdam, chrasche@uni-potsdam.de

DSGVO). Verarbeitungsprozesse sind in einem zentralen Verzeichnis der Verarbeitungstätigkeiten (VVT) festzuhalten (Art. 30 Abs. 1 DSGVO). Das Verzeichnis kann zu jeder Zeit von Datenschutzbehörden angefordert werden, wobei die Nachweispflicht dem Unternehmen obliegt (Art. 30 Abs. 4 DSGVO). Ausgenommen von den genannten Pflichten sind Organisationen mit weniger als 250 Beschäftigten – unter der Voraussetzung, dass durch die Verarbeitung kein Risiko für die Rechte und Freiheiten der Betroffenen besteht, die Verarbeitung nur gelegentlich erfolgt oder keine besonderen Datenkategorien gemäß Artikel 9 Absatz 1 oder Artikel 10 DSGVO verarbeitet werden (Art. 30 Abs. 5 DSGVO).

Die Praxis zeigt, dass insbesondere große, international operierende Organisationen die Umsetzung von Dokumentations- und Aktualisierungsanforderungen zu Verarbeitungstätigkeiten häufig auf Bereichs- oder Abteilungsebene herunterbrechen [KRS20]. Eine grundlegende Herausforderung betrifft die Sicherstellung von Vollständigkeit und Aktualität des zentralen VVT. Mangels fehlender Transparenz ist für viele Unternehmen die Gesamtheit der existierenden Verarbeitungsprozesse unbekannt. Somit sind weder die Ganzheitlichkeit noch der Erfüllungsgrad gesetzlicher Anforderungen in vollem Umfang überprüfbar. [KRS20]. Es fehlt eine Übersicht aller verarbeiteten personenbezogenen Daten, damit Verantwortliche der Datenverarbeitung die rechtskonforme Umsetzung datenschutzrechtlicher Vorgaben prüfen können. Überdies existieren meist keine einheitlichen Standards zur Dokumentation von Verarbeitungstätigkeiten. Daher sind Verarbeitungsprozesse oftmals heterogen organisiert und unvollständig oder fehlerhaft dokumentiert. Eine weitere Problemstellung resultiert aus der Dynamik und Vielfalt von Verarbeitungsprozessen. Während einige Verfahren beständig sind, befinden sich andere in einem ständigen Wandel. Ein regelmäßiger Turnus zur Aktualitäts- und Konformitätsprüfung ist daher nur bedingt geeignet [KRS20]. Sofern Datenkategorien verarbeitet werden, die ohne konkrete Kenntnis nicht als personenbezogene Daten identifizierbar sind, kann dies maßgeblich die Sicherstellung der Datenschutzkonformität beeinflussen [KRS20]. Werden personenbezogene Daten nicht als solche identifiziert, erfolgt auch keine Überführung und Zentralisierung der betroffenen Verarbeitungstätigkeiten im VVT. Im Falle einer behördlichen Prüfung drohen Unternehmen hohe Geldstrafen sowie Reputationsschäden.

Um den Problemstellungen entgegenzuwirken und Konformität zu gewährleisten, greifen Organisationen vermehrt auf Technologien wie etwa Data Mining zurück. Sie ermöglichen die nahtlose Analyse großer Datenmengen, transparente Visualisierungen und Überwachung von Datenströmen und lassen zusammenhängende Muster und Abhängigkeiten zwischen Daten erkennen [Ha16]. Im Fokus des vorliegenden Beitrags steht die Frage, wie Data Mining zur Einhaltung des Datenschutzes in Unternehmen entwickelt werden kann.

Ein besonderes Augenmerk liegt auf der Pflege des VVT, da dieses das zentrale Element der europäischen Datenschutzgrundverordnung (DSGVO) darstellt. Es wird geprüft, welche technologischen Anpassungen zur Lösung der Problemstellungen essentiell sind und wie der Einsatz von Data Mining effizient umgesetzt, die Komplexität der Arbeitsvorgänge verringert und die Flexibilität von Geschäftsprozessen erhöht werden kann.

2 Verwandte Arbeiten

Der folgende Abschnitt befasst sich mit einer Vorstellung bereits existierender Vorgehensmodelle zur Sicherstellung datenschutzrechtlicher Vorgaben, die auf dem Einsatz innovativer Technologien der Datenanalytik beruhen.

Im Beitrag von Becker und Buchkremer wird die Entwicklung eines agilen Vorgehensmodells erläutert, mit dessen Hilfe aufsichtsrechtliche Anforderungen durch Einsatz einer sogenannten Regulatory Technology Lösung implementierbar seien [BB18]. Die Autoren heben die Relevanz eines harmonischen Zusammenspiels zwischen Technologie und menschlichen Experten für agile Implementierungsprozesse hervor und kommen zu dem Schluss, dass iterative Vorgehen für die Analyse regulatorischer Anforderungen im Kontext des Datenschutzes erfolgsentscheidend sind.

Kittel beschreibt in einem Artikel, wie Agilität bei Geschäftsprozessen mit Datenschutzbezug sichergestellt werden kann [Ki13]. Es zeigt sich, dass Ad-hoc-Änderungen von Geschäftsprozessen dieser Art eine vorausgehende Kontrolle regulatorischer Datenschutzanforderungen unbrauchbar machen. Kittel stellt einen modellbasierten Ansatz zur Ad-hoc-Integration von Datenschutzkontrollen in Arbeitsabläufen vor, durch den die Abhängigkeiten zwischen Agilität und Compliance verringert werden sollen.

Ein weiteres Vorgehensmodell zur Vorbereitung auf datenschutzrechtliche Anforderungen wird von Wirmsperger, Buchholz und Wolff erarbeitet [BWW16]. Das Modell berücksichtigt rechtliche, technische, organisatorische und prozessuale Aspekte. Beginnend mit einer Vorprüfung und einer Umfeldanalyse zur Erfassung aller personenbezogenen Daten in Geschäftsprozessen solle der Status Quo auf Basis einer Fit-/Gap-Analyse erfasst sowie ein Risiko- und Maßnahmenplan erarbeitet werden.

Das von Chapman et. al. entwickelte CRISP-DM-Modell stellt die Entwicklung und Umsetzung spezifischer Data-Science-Projekte durch den Einsatz von Data Mining und künstlicher Intelligenz in den Mittelpunkt [Ca00]. Das Modell gilt als Standardvorgehen für die Ausführung von Data-Mining-Projekten und ist für diverse Projekte der künstlichen Intelligenz zur Sicherstellung des Datenschutzes anwendbar. Da das CRISP-DM-Modell in den Kontext des aktuellen Technologiestands eingeordnet ist, wird es als Rahmen für die vorliegende Untersuchung verwendet.

Während ein Großteil aktueller Untersuchungen den Einfluss regulatorischer Datenschutzvorgaben auf die Entwicklung intelligenter Technologien diskutieren, widmen sich einige wenige Quellen dem Unterstützungsgrad innovativer Technologien und deren Anwendungspotentialen zur Wahrung des Datenschutzes. Inwiefern Data Mining jedoch speziell bei der Verarbeitungsdokumentation und -aktualisierung in einem VVT unterstützt, wird in der Wissenschaft nicht vertieft betrachtet. Nach aktuellem Stand existiert kein Vorgehensmodell für diesen spezifischen Anwendungsfall.

3 Vorgehensmodell und methodische Unterstützung

3.1 Anforderungsanalyse

Die Anwendung einer Anforderungsanalyse hat unmittelbaren Einfluss auf die zielgerichtete Entwicklung des CRISP-DM-Modells. Zur vollständigen Ermittlung aller Anforderungen an CRISP-DM wird zunächst ein umfassender Anforderungskatalog entwickelt. Der Katalog differenziert zwischen technologischen Anforderungen, die primär die zu erbringenden Funktionalitäten, Mechanismen und Leistungen des Data Mining zur Gewährleistung der Konformität betreffen und Anforderungen von Seiten des Datenschutzrechts zur Pflege eines zentralen VVT. Letzteres orientiert sich an den Regularien der DSGVO aus Artikel 30.

3.2 CRISP-DM

CRISP-DM folgt einem iterativen Kreislauf mit insgesamt sechs Phasen, ohne dabei einen konkreten Endpunkt festzulegen. Stattdessen kann jede Phase und deren Iterationen, je nach Problemstellung mehrfach durchlaufen und ausdifferenziert werden. Jede Wiederholung des Gesamtprozesses bringt neue Fragestellungen hervor und kann zu einer Prozessoptimierung beitragen. Das Modell schreibt keine starre Sequenzierung der einzelnen Phasen vor. Rückkopplungen, die sich etwa aus unvorhergesehenen Problemfaktoren oder unzureichender Qualität eines Zwischenergebnisses ergeben, sind durchaus möglich und gewünscht. [Ca00]

Phase 1. Zur Erlangung vollständiger Kenntnis über die Geschäftsanforderungen und konkrete Aufgabenstellung hat eine präzise Erörterung der betriebswirtschaftlichen Problemstellung zu erfolgen [Ca00]. Dabei sind die zu erreichenden Zielkriterien festzulegen. Diese werden in Anforderungen an die Datenanalyse überführt, woraufhin ein konkreter Umsetzungsplan unter Berücksichtigung zeitlicher, personeller und sachlicher Ressourcen aufzusetzen ist [CL16].

Phase 2. Im nächsten Schritt werden relevante Datenbestände selektiert, deren Verarbeitung zur Erfüllung der zuvor bestimmten Ziele notwendig ist [CL16]. Es wird eine Datensammlung mit Beschreibung der typischen Eigenschaften der relevanten Daten angelegt, um ein generelles Verständnis über die selektierten Daten aufzubauen. Die Phase mündet letztlich in einer Bewertung der Datenqualität und -quantität [CL16].

Phase 3. Die Datenvorbereitung zielt auf die Auswahl der finalen Datenmenge ab, die in das Data-Mining-System integriert und entlang vordefinierter, anwendungsspezifischer Algorithmen analysiert werden soll [Ca00]. Es bedarf einer klaren Differenzierung zwischen irrelevanten und relevanten Daten. Das Ergebnis der Datenauswahl hängt von der jeweiligen Zielsetzung des Data-Science-Projektes ab. Ferner sind die Daten zu bereinigen, um eine Data-Mining-Verarbeitung zu ermöglichen. Diese Phase entscheidet darüber,

welche speziellen Merkmale und Charakteristiken die nachfolgende Modellbildung berücksichtigen soll [CL16].

Phase 4. Die Modellbildung nimmt sich der eigentlichen Datenanalyse an, indem ein Modell zum Umgang mit den selektierten Daten entwickelt wird [CL16]. Nach Auswahl und Parametrisierung einer passenden Modellierungstechnik wird ein Testmodell entwickelt, mit dessen Hilfe die Präzision und Qualität des Entwicklungsergebnisses geprüft und bewertet wird. Die Algorithmen der Modellbildung unterscheiden zwischen einem *Trainieren* und *Anwenden*, wobei das Modell entweder auf Basis des gewonnenen Wissens aus historischen Daten trainiert oder auf neue, bisher unbekannte Datensätze angewendet wird.

Phase 5. Zur Evaluation des Entwicklungsergebnisses wird die eingangs festgelegte Zielsetzung mit dem erarbeiteten Data-Mining-Verfahren abgeglichen. Für den Fall, dass die gewünschte Qualität des Modells zur Erfüllung der Zielkriterien nicht vollständig oder nur in Teilen erreicht wurde, muss CRISP-DM erneut durchlaufen werden [CL16].

Phase 6. Den Abschluss bildet die Planung und Umsetzung der Implementierung des Data Mining im Unternehmen. Das Modell kann je nach Anwendungsfall auf existierende oder auf neue, bislang unbekannte Datenbestände angewendet werden [Ca00].

4 Konzipierung des CRISP-DM-Modells

4.1 Anforderungsspezifizierung

Technologische Anforderungen. Tabelle 1 zeigt einen Überblick der technologischen Anforderungen an das Data-Mining-System.

Technologische Anforderungen
Zugriff auf den gesamten Datenpool des Unternehmens
Erschließen aller im Unternehmen verfügbaren, (un-)strukturierten Daten
Identifikation und Strukturierung aller verarbeiteten personenbezogenen Daten
Erfassung aller existierenden Verarbeitungsprozesse
Zentrale Steuerung der Pflege eines VVT
Vollständige Dokumentation aller Verarbeitungsprozesse im VVT
Gewährleistung kontinuierlicher Aktualität des VVT
Automatische Anpassung und Aktualisierung von Verarbeitungsprozessen
Erkennen von Trends, Veränderungen und datenschutzrechtlichen Anforderungen

Tab. 1: Technologische Anforderungen an das Data-Mining-System

Datenschutzrechtliche Anforderungen. Tabelle 2 zeigt einen Überblick der datenschutzrechtlichen Anforderungen an die Dokumentation von Verarbeitungstätigkeiten in

einem VVT gemäß Artikel 30 Absatz 1 DSGVO. Gleiches gilt für Auftragsverarbeiter unter Ausschluss der Beschreibung und Kategorisierung der Verarbeitungszwecke, der Beschreibung und Kategorisierung aller Datenempfänger im In- und Ausland sowie der Löschfristen der verschiedenen Datenkategorien (Art. 30 Abs. 2 DSGVO).

Datenschutzrechtliche Anforderungen
Name und Kontaktdaten des verantwortlichen Datenverarbeiters
Beschreibung und Kategorisierung der Verarbeitungszwecken
Beschreibung des Betroffenen und Kategorisierung der betroffenen Personen
Beschreibung und Kategorisierung der personenbezogenen Daten des Betroffenen
Beschreibung und Kategorisierung aller Datenempfänger im In- und Ausland
Beschreibung der Übermittlung in Drittländer oder internationale Organisationen und deren Benennung
Löschfristen der verschiedenen Datenkategorien
Dokumentation der technischen und organisatorischen Maßnahmen

Tab. 2: Datenschutzrechtliche Anforderungen

4.2 Grundlagen

Die erforderlichen Inhalte eines VVT ergeben sich aus der Analyse manuell gepflegter Verzeichnisse in der Praxis. Dies dient der nachgelagerten Lösungssuche, indem ermittelt wird, welche Strukturierungen und Klassifizierungen der relevanten Daten das Data-Mining-System zur vollumfänglichen Dokumentation zu berücksichtigen hat. Im Kontext der CRISP-DM-Entwicklung werden schließlich Regeln, Korrelationen und Muster zwischen Daten und deren Verarbeitungstätigkeiten abgeleitet. Eine beispielhafte Übersicht des Aufbaus eines VVT und der zu dokumentierenden Inhalte und Informationen ist in Tabelle 3 gegeben.

Themenbereiche eines VVT	Potentielle Inhalte
Dokumentation der Kontaktdaten des verantwortlichen Daten- oder Auftragsverarbeiters	Name, Funktion, E-Mail-Adresse, Telefonnummer und Anschrift des Verantwortlichen oder Auftragsverarbeiters
Dokumentation der Verarbeitungsprozesse	Bezeichnung, Beschreibung, Datenherkunft, Verwendetes IT-System
Dokumentation des Zwecks der Datenverarbeitung	Zweckkategorie, Zweckänderung, Zweck (Mit-)Bestimmung durch Dritte
Dokumentation des datenverarbeitenden Systems weitere	Name des datenverarbeitenden Systems

Tab. 3: Themenbereiche und potentielle Inhalte eines VVT

4.3 Konzeption

Phase 1. Zur Generierung eines exakten Verständnisses der Aufgaben- und Problemstellung werden im ersten Schritt die erwarteten Projektziele sowie -ergebnisse festgelegt. Im vorliegenden Kontext leiten sich diese aus den technologischen und datenschutzrechtlichen Anforderungen ab. Tabelle 4 zeigt die potentielle Zielsetzung des Data-Mining-Vorhabens.

Zielformulierung
Steigerung der Transparenz
Steigerung der Effizienz bei Anpassungen an Verarbeitungsprozesse
Verfolgbarkeit bei Prozessaktivitäten
Verfügbarkeit relevanter Informationen
Reduzierung der Arbeitsauslastung von Fachbereichen eines Unternehmens
Reduzierung des Abstimmungsaufwands zwischen den Fachbereichen
Standardisierung des Vorgehens
Vollständige Identifikation und Strukturierung personenbezogener Daten
Vollständige Erfassung aller Verarbeitungsprozesse
Unterstützung der Dokumentation der Verarbeitungsprozesse im VVT
Sicherstellung kontinuierlicher Aktualität des VVT

Tab. 4: Zielformulierung

Weiterhin ist eine ausführliche Risikoanalyse durchzuführen [CL16]. Ermittelte Risiken sind gemäß projektspezifischer Kriterien zu bewerten und individuell zu analysieren. Dies kann mittels einer Risiko-Matrix erfolgen, wobei die Eintrittswahrscheinlichkeit und Schadenhöhe für jedes Risiko geschätzt und in der Risiko-Matrix visualisiert werden. Zur Reduktion besonders schwerwiegender Risiken sind Gegenmaßnahmen festzulegen. Das Risikomanagement hat über die gesamte Projektdauer hinweg zu erfolgen.

Anhand der Zielsetzung werden Erfolgskriterien zur finalen Bewertung des Entwicklungsergebnisses spezifiziert [Ca00]. Das Vorhaben ist dann erfolgreich, wenn die Gesamtheit aller personenbezogenen Daten bekannt ist und diese entlang charakteristischer Merkmale strukturiert werden. Entsprechend sind Verarbeitungstätigkeiten automatisch zu erfassen, sodass Mitarbeitende bei Zentralisierung und Aktualisierung des VVT unterstützt werden. Weitere Erfolgskriterien betreffen die Reduzierung der Fehleranfälligkeit, des manuellen Arbeitsaufwands, der Komplexität sowie der Arbeitsauslastung innerhalb der Fachbereiche und -abteilungen.

Die Festlegung von Unternehmenszielen orientiert sich vor allem an der Firmenkultur und Vision eines Unternehmens, weshalb eine allgemein gültige Aussage nur bedingt möglich ist. Nichtsdestominder spiegeln sich zumeist einige Unternehmensziele wie Integrität, Verlässlichkeit und Vertrauenswürdigkeit in verschiedenen Unternehmen wider und ge-

hen demnach mit den Grundsätzen und Schutzziele der Datensicherheit und des Datenschutzes (Compliance) einher. Die strikte Einhaltung dieser drei Faktoren sind in der heutigen Zeit zur Wahrung der Wettbewerbsfähigkeit [He18] und Rechtmäßigkeit unabdingbar.

Die wichtigsten Fragestellungen zur Aufbereitung der ersten Phase sind nachfolgend zusammengefasst.

- Wie ist die Ausgangssituation und Problemstellung?
- Welche Ziele und Ergebnisse sollen durch Data Mining erreicht werden?
- Welche Risiken (finanziell, rechtlich, organisatorisch) können auftreten?
- Was sind die Erfolgskriterien und Unternehmensziele?
- Wie ist die aktuelle Unternehmenssituation?
- Welche Ressourcen sind zur Umsetzung des Vorhabens verfügbar?
- Welche Kosten sind für welchen Nutzen aufzubringen?
- Wurde ein Projektmanagementsystem etabliert und ein Projektplan aufgesetzt?

Phase 2. Zur Selektion der relevanten Datenbestände muss geklärt werden, was unter personenbezogenen Daten verstanden wird und wodurch sich diese kennzeichnen. Die DSGVO definiert personenbezogene Daten als „Informationen, die sich auf eine identifizierte oder identifizierbare natürliche Person beziehen“ (Art. 4 Abs. 1 DSGVO). Personen gelten dann als „identifizierbar“, wenn sie sich (in-)direkt eindeutig identifizieren lassen (Art. 4 Abs. 1 DSGVO). Es wird dann von personenbezogenen Daten gesprochen, wenn die erhobenen Daten einen direkten Bezug zu einer betroffenen Person hervorbringen.

Die Elemente und Kategorien der personenbezogenen Daten sind zu ermitteln. Dies erlaubt die Auswahl und Entwicklung eines passenden Data-Mining-Vorgehens. Datenelemente und -kategorien leiten sich aus den in Tabelle 3 dargelegten Informationen zur Dokumentation von Verarbeitungstätigkeiten ab. Die Erkenntnisse dienen der Entwicklung von Mustern und Regeln, die Data Mining zur vollumfänglichen Identifikation der personenbezogenen Daten, zur Ableitung und Dokumentation der resultierenden Verarbeitungstätigkeiten sowie dem Segmentieren der Elemente und Kategorien anzuwenden hat.

Verantwortliche mit weniger als 250 Beschäftigten sind gemäß Artikel 30 Absatz 5 DSGVO dazu verpflichtet, zu erheben, ob mit der Verarbeitung der personenbezogenen Daten besondere Risiken für die Rechte und Freiheiten für die Betroffenen einhergehen und inwieweit besondere Kategorien personenbezogener Daten gemäß Artikel 9 DSGVO verarbeitet werden. Weiterhin ist der Turnus der Datenverarbeitung zu bestimmen.

Phase 3. Die Datenvorbereitung gliedert sich in die Schritte Selektion und Integration, Säuberung, Reduktion und Transformation von Daten [Ca00]. Die Relevanz einer Daten-selektion und -integration resultiert aus den verschiedenen Datenbanken und Quellen, aus denen Daten potentiell entstammen. Nach erfolgter Datenselektion sind die Daten in einer konsistenten Datenbasis mit schlüssigen Datensätzen zu vereinheitlichen. Probleme, die

bei der Integration auftreten können, sind unter anderem Entitäten-Identifikationsprobleme, Redundanzen, Widersprüche oder Datenwertkonflikte [CL16].

Danach ist der Datenbestand manuell zu bereinigen. Es ist darauf zu achten, dass eingefügte Werte durch Bereinigung informationsneutral sind, ohne eine Verfälschung der vorhandenen Dateninformationen herbeizuführen. Neben fehlenden Daten stellen ebenso veräuschte Daten und Ausreißer oder inkonsistente und falsche Daten mögliche Problemstellung dar, die während des Säuberungsprozesses zu unterbinden sind [CL16].

Eine Reduktion der Daten ist dann notwendig, wenn ein Datensatz zur Ausführung des Data Mining zu groß ist. Einerseits kann die Komplexität eines Datensatzes mit Hilfe einer zeilen- oder spaltenweise Aggregation [CL16] verringert werden. Mehrere Daten werden also auf Basis von charakteristischen Attributen zusammengefasst. Ein konkreter Anwendungsfall ist etwa das Clustern von Datenelementen und -kategorien durch Anwendung der spaltenweisen Aggregation. Personenbezogene Daten können so von dem Rest des Datenbestandes abgespalten und kategorisiert werden. Eine zweite Lösung bietet die Dimensionsreduktion als Vorwärtsauswahl oder Rückwärtseliminierung [CL16], indem Stichproben einer repräsentativen Teilmenge der selektierten Daten durchgeführt werden. Eine Erfassung aller personenbezogener Daten kann etwa mittels der Vorauswahl erfolgen, indem alle Daten, die keinen direkten Personenbezug aufweisen, durch sukzessive Aufnahme neuer Anforderungen gelöscht werden.

Ziel der Datentransformation ist die Überführung der Daten in eine brauchbare Form, um in das Data-Mining-System integriert werden zu können. Verfahren zur Datentransformation sind etwa Codierungen, Zeichenketten (z.B. Umlaute), Maßeinheiten und Skalierungen, Kombinationen oder Separierungen von Attributen, Berechnungen abgeleiteter Werte, Aggregationen oder Datenglättungen (z.B. Regression) [CL16].

Phase 4. Die Modellbildung des Data Mining unterscheidet zwischen Potential- und Beschreibungsaufgaben. Potentialaufgaben umfassen die Datenklassifikation und das Ableiten von Prognosen, wohingegen Beschreibungsaufgaben der Segmentierung oder dem Aufstellen von Assoziationen zwischen Datensätzen dienen [CL16]. Bei der Klassifikation erfolgt eine Zuordnung eines Datenobjekts zu einer vordefinierten Klasse entlang charakteristischer Merkmale. Die Prognose hingegen zielt auf die Entwicklung eines Bewertungsmodells zur fortlaufenden Ermittlung stetiger Werte ab. Bisher unbekannte, numerische Merkmale werden auf Basis anderer Merkmale oder erlangter Erkenntnisse vorausgesagt und Abhängigkeiten zwischen diversen Variablen hergestellt. Im Rahmen der Segmentierung wird die Gesamtheit aller Daten in Teilmengen unterteilt und mehrere Datenobjekte mit gemeinsamen Merkmalen zu einer homogenen Gruppe zusammengeführt. Im Fokus der Assoziation steht die Ermittlung und Beschreibung von Mustern zwischen Datenobjekten, die in einer bestimmten Relation zueinanderstehen. Beispiele für Data-Mining-Verfahren sind Entscheidungsbäume, Cluster-Algorithmen oder Regressionen.

Für den vorliegenden Anwendungsfall muss primär eine ganzheitliche Erfassung und Strukturierung aller personenbezogenen Daten vorgenommen werden. Verschiedene Datenkategorien sind zur einheitlichen Dokumentation im VVT zu einem einzigen Datenelement zu reduzieren. Dafür eignen sich die Klassifikation und Segmentierung.

Zu Beginn erlaubt die Klassifikation eine Kategorisierung personenbezogener Daten gemäß charakteristischer Merkmale, durch die natürliche Personen eindeutig identifizierbar sind. Um eine solche Separierung zu erreichen, müssen dem Data-Mining-System die Merkmalanforderungen bekannt sein. Die Anforderungen ergeben sich vorrangig aus der Definition personenbezogener Daten des Artikel 4 DSGVO. Beispiele für charakteristische Merkmalanforderungen zur Datenklassifikation sind bspw. Name, Anschrift, E-Mail-Adresse oder Telefonnummer.

Überdies sind die als relevant klassifizierten Datenobjekte zu segmentieren und bestehende Datenkategorien bestimmten Datenelementen zuzuweisen. Eine Möglichkeit stellt das Rule-based Reasoning [Ch20] dar, indem Regeln entlang des Wenn-Dann-Sonst-Prinzips [FPZ95] erarbeitet werden. Ein Beispiel zur Segmentierung personenbezogener Daten kann etwa über die Regel „*Wenn die Datenkategorie die Angabe Name oder Anschrift enthält, dann sind diese personenbezogenen Daten dem Element persönliche Kontaktinformationen zuzuweisen*“ erfolgen. Dieses Schema muss für alle identifizierten Datenkategorien und -elemente umgesetzt werden – unter der Prämisse, dass die Möglichkeit einer Segmentierung besteht. Zur besseren Veranschaulichung ist die Vorgehensweise der Regelbildung in Abbildung 1 dargestellt.

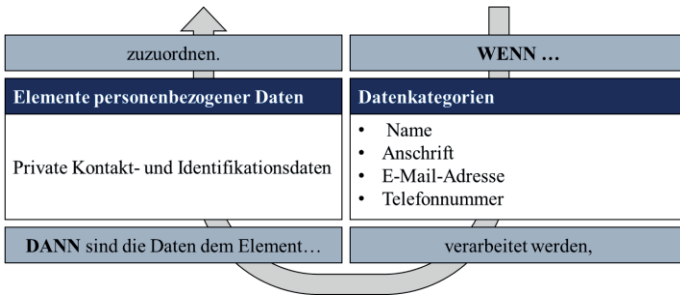


Abb. 1: Wenn-Dann-Regel zur Segmentierung personenbezogener Daten

Hinzu kommt die Notwendigkeit, auf Grundlage der identifizierten und segmentierten Daten resultierende Verarbeitungsprozesse zu erfassen, zu dokumentieren und zu aktualisieren. Hierfür kann auf die Assoziation und Prognose zurückgegriffen werden.

Mit Hilfe des Assoziationsverfahrens lassen sich Abhängigkeiten und Muster zwischen Daten und Verarbeitungstätigkeiten feststellen. Das stellt sicher, dass personenbezogene Daten, die nicht als solche kenntlich sind, identifiziert werden können. Eine Möglichkeit stellt dabei die semantische Interoperabilität [Gö10] zur Kollaboration zwischen diversen IT-Systemen mittels Klassifikationssystemen, Taxonomien oder Nomenklaturen dar. Die

semantische Interoperabilität eignet sich im Speziellen zur Erfassung und Dokumentation von Verarbeitungstätigkeiten auf Grundlage der verarbeiteten personenbezogenen Daten. Das Data-Mining-System wird dazu befähigt, Informationen mit den IT-Systemen der Organisation auszutauschen und so bspw. den Zweck einer Verarbeitung zu ermitteln. Mittels eines intelligenten und vernetzten Zusammenspiels zwischen IT-Systemen lassen sich die zentral im VVT festzuhaltenden Inhalte und Informationen detektieren und einheitlich dokumentieren. Dadurch wird ebenso eine Standardisierung der Dokumentation erreicht.

Als letzter Schritt unterstützt die Prognose dabei, Zusammenhänge zwischen bekannten und bisher unbekanntem Merkmalattributen herzustellen und Trendentwicklungen zu prognostizieren. Treten etwa Änderungen in den Angaben personenbezogener Daten auf, können diese analysiert und bei dokumentierten Verarbeitungstätigkeiten aktualisiert werden. Weiterhin können Risiken aufgrund von Datenlecks oder bei Nicht-Einhalten der geltenden Datenschutzvorgaben präventiv gemeldet werden.

Es lässt sich festhalten, dass eine Kombination der vier Data-Mining-Vorgehen essentiell ist, um den geschilderten Herausforderungen und Problemstellungen zu begegnen und die identifizierten Anforderungen zu erreichen.

Phase 5. Im Kontext der Evaluation werden die Analyseergebnisse geprüft. Ob die Umsetzung erfolgreich ist, ergibt sich aus einer Ermittlung des Erfüllungsgrads der initial spezifizierten Erfolgs- und Zielkriterien. Zentrale Fragestellung ist, ob der erwünschte betriebswirtschaftliche Nutzen durch das Entwicklungsergebnis erzielt wird [CL16].

Im weiteren Verlauf ist eine Analyse der auftretenden Fehler durchzuführen, woraus sich unter Umständen weitere Optimierungspotentiale ergeben [Ca00]. Tritt dieser Umstand auf, kann in eine der vorangegangenen Phasen zurückgekehrt und so das Data-Mining-Vorhaben sukzessive verbessert werden.

Phase 6. Den Abschluss bildet die praktische Implementierung des Data-Mining-Systems. Zur optimalen Einsatzvorbereitung wird ein im Detail ausgearbeitetes und strukturiertes Vorgehen zum künftigen Monitoring des Data Mining und der resultierenden Analyseergebnisse vorgesetzt. Des Weiteren muss eine ausreichende Motivation der Mitarbeitenden der Organisation, in der das System Anwendung finden soll sowie eine umfassende Unterstützung der durch das Data Mining betroffenen Mitarbeitenden (z.B. IT-Abteilung, Datenschutzbeauftragter, Fachabteilung etc.) gegeben sein, um das Scheitern des Projekts zu verhindern. Das System ist in den Regelbetrieb der Organisation zu überführen und in laufende Prozesse einzubetten.

5 Diskussion

Die ersten drei CRISP-DM-Phasen beanspruchen etwa 50 bis 70 Prozent des Arbeitsaufwands zur Entwicklung des Data Mining, wobei die einzelnen Phasen manuell vorzubereiten und umzusetzen sind [Wu20]. Ein direkter Vergleich des Status Quo zur Pflege

eines VVT und den Phasen des CRISP-DM impliziert, dass die initialen drei CRISP-Phasen gleichermaßen im manuellen Pflegeprozess eines VVT stattfinden. Unternehmen, die bereits ein VVT pflegen, haben die Schritte im Optimalfall durchlaufen. Auch wenn zur Pflege eines VVT kein Data-Mining-System etabliert werden soll, ist es sinnvoll, die Phasen gewissenhaft umzusetzen. Organisationen sollten nach Ausführung der initialen Phasen in Erwägung ziehen, ihre Ergebnisse in KI-Algorithmen und Regeln zu überführen und die manuellen Arbeitsaufwände auf ein Data-Mining-System zu verlagern. Unternehmen können auf ihrem bisherigen Arbeitsstand aufbauen, die Inhalte entsprechend dem dargestellten Vorgehen anpassen und letztlich in ein Data-Mining-Modell überführen. Jedoch muss das CRISP-DM-Modell nicht zwangsläufig in der Implementierung eines Data-Mining-Systems münden, auch wenn dies zu einem deutlichen Anstieg der Produktivität beiträgt. Stattdessen sehen sich Organisationen aufgrund der Rechenschaftspflicht ohnehin damit konfrontiert, Transparenz hinsichtlich der Datenverarbeitungen und -flüsse sicherzustellen. Um ebendiese Transparenz zu erreichen, haben Organisationen die initialen Phasen Business Understanding, Data Understanding und Data Preparation zur Erfassung, Dokumentation und Aktualisierung aller existierenden Verarbeitungsprozesse aufzubereiten. Folglich zieht das dargestellte Vorgehen kein Mehraufwand nach sich, sondern bietet Organisationen im Gegenteil die Möglichkeit, Synergien zu nutzen und in Zukunft bedarfsorientiert auf ihrem bisherigen Arbeitsstand aufzubauen, um eine technologische Unterstützung und Optimierung des Vorgehens zur Pflege eines VVT herbeizuführen.

Zwar kann Data Mining bei der vollständigen Erfassung und Kategorisierung personenbezogener Daten, dem Ableiten von Verarbeitungstätigkeiten sowie der Dokumentation und Aktualisierung im VVT unterstützen. Jedoch sind nachgelagert weiterhin manuelle Aufwände notwendig. Bspw. verantwortet das Rechtswesen einer Organisation die Zentralisierung von Verarbeitungstätigkeiten im VVT, die Überführung neuer datenschutzrechtlicher Vorgaben in konkrete Anforderungen an Data Mining oder die konstante Überwachung der Qualität der Analyseergebnisse.

Eine weitere Handlungsempfehlung betrifft die Erweiterung des CRISP-DM-Modells um die Monitoring-Phase. Neben einer fortwährenden Wartung des Systems sind auch die Ergebnisse der Datenanalysen durch Verantwortliche der Datenschutzorganisation (z.B. Datenschutzbeauftragter etc.) zu überwachen, da diese bestens mit den rechtlichen Grundlagen vertraut sind. Es ist sicherzustellen, dass alle Mitarbeitenden, die in Zukunft Berührungspunkte mit dem System haben, umfassend geschult werden.

Darüber hinaus ist ein Datensicherheitskonzept gemäß Artikel 32 DSGVO zu entwickeln, um softwareseitige Störungen und Systemausfällen vorzubeugen. Im Falle eines Absturzes wird etwa der Zugriff auf das zentrale Verzeichnis verweigert. Daher müssen in regelmäßigen Abständen automatische Backups des VVT durchgeführt werden und Mitarbeitende der Organisation dafür Sorge tragen, das System durch Schutzmaßnahmen (z.B. technisch-organisatorische-Maßnahmen) abzusichern.

Data Mining erlaubt eine ganzheitliche, transparente und zentrale Steuerung des VVT-Prozesses, sodass ein technologischer Einsatz zur Wahrung der Konformität geeignet ist.

Durch Einhalten der Vorgaben des Artikel 30 DSGVO wird das Haftungsrisiko wesentlich reduziert und der aktuell gelebte Prozess in der Praxis flexibilisiert und vereinfacht. Die ersten vier Phasen des Modells können als Basis für vergleichbare Aufgabenstellungen genutzt und spezifiziert werden.

Nichtsdestominder zeigt sich der Unterstützungsgrad des Data Mining erst durch Entwicklung entlang einer realen Aufgabenstellung und tatsächlichen Implementierung des Systems. Darüber hinaus ist die Modellentwicklung auf Annahmen und Theorien gestützt. Die CRISP-DM-Phasen wurden zwar detailliert ausgearbeitet, jedoch ohne die technologische Ebene vertieft zu betrachten. Dies ist unter anderem der Tatsache geschuldet, dass kein Testmodell unter realen Umständen entwickelt und zu Testzwecken implementiert wurden. Eine Aussage über die tatsächliche Um- und Einsetzbarkeit des Data-Mining-Systems kann somit nicht getroffen werden. Grundsätzlich ist die technische Sicht bei der Entwicklung des Vorgehensmodells unterrepräsentiert. Es kann vorkommen, dass Probleme, die in der Anwendung und Programmierung des Data-Mining-Systems auftreten, nicht vollständig erkannt und berücksichtigt wurden. Daher wird empfohlen, das System in der Praxis zunächst umfassend zu testen und mit erfolgreichem Abschluss der Testphase auf weitere Bereiche des Unternehmens auszuweiten. Ferner können in Zukunft vor- und nachgelagerte Teilprozesse oder weitere Anforderungen der DSGVO berücksichtigt und auf ähnliche Weise intelligent gesteuert und optimiert werden.

Es bleibt zu erwähnen, dass das CRISP-DM-Modell einem Standardmodell entspricht, welches der Standardisierung diverser Anwendungsfälle dient und damit nicht nur auf die Pflege eines VVT begrenzt ist. Stattdessen kann das CRISP-DM-Modell beliebig erweitert und ebenso auf andere Ausgangssituationen übertragen werden. Das Modell kann demnach genau wie VVT in der Praxis vielfältig ausfallen. Aus diesem Grund findet im Rahmen des vorliegenden Beitrags eine generische Darstellung und Entwicklung des Modells Anwendung. Auf die Beschreibung einzelner, konkreter Anwendungsfälle wird in diesem Zusammenhang bewusst verzichtet.

Als weiterer Forschungsbedarf kann das Process Mining und dessen Kombination mit Data Mining betrachtet werden. Process Mining vereint die Vorteile des Data Mining mit denen der Prozessmodellierung, sodass eine effiziente Überwachung und Erstellung von komplexen Echtzeitprozessen möglich ist [Re20]. Durch Standardisierung von Prozessen kann einerseits die Transparenz erhöht werden und damit Schwachstellen der aktuellen Prozessumsetzung effizient geprüft und bei Bedarf verbessert werden [NP19]. Andererseits werden durch Prozessautomatisierungen Redundanzen reduziert, Engpässe vermieden und damit einhergehend Kosten reduziert werden [NP19].

In einem nächsten Schritt kann das System um maschinelles Lernen erweitert werden. Diese Form der künstlichen Intelligenz ermöglicht die Entwicklung von Handlungsempfehlungen und Generierung von Maßnahmen anhand großer Datenbestände eines Prozesses [Re20]. Gemeinhin wird dieses Vorgehen als Predictive Process Mining bezeichnet. Anhand prozessbezogener Daten erkennt das System relevante Kausalitäten und erklärt diese. Während das System automatisch Trends und Muster

ableitet, werden die entwickelten Maßnahmen durch Mitarbeitende des Unternehmens bewertet und schließlich umgesetzt.

In einer vertieften Betrachtung sind die Potentiale einer Kombination des Process- und Data Mining zur Erreichung einer ganzheitlichen Standardisierung und Sicherstellung datenschutzrechtlicher Vorgaben zu untersuchen.

Literaturverzeichnis

- [BB18] Becker, M.; Buchkremer, R.: Implementierung einer Regulatory Technology Lösung bei Finanzinstituten unter Berücksichtigung agiler Vorgehensmodelle. In (Mikusz, M.; Volland, A.; Engstler, M., Hrsg.): Projektmanagement und Vorgehensmodelle 2018 - Der Einfluss der Digitalisierung auf Projektmanagementmethoden und Entwicklungsprozesse. Köllen Druck+Verlag GmbH, Bonn, S. 125-134, 2018.
- [BWW16] Buchholz, S.; Wirnsperger, P. J.; Wolff, D.: Zeitgemäßer Datenschutz in der datengetriebenen Wirtschaft - Effektive Umsetzung der EU-Datenschutz-Grundverordnung (DSGVO). Deloitte GmbH, 2016.
- [Ca00] Chapman, P. et.al.: CRISP-DM 1.0: Step-by-step data mining guide. The CRISP-DM consortium, 2000.
- [Ch20] Chowdhary, K.R.: Fundamentals of Artificial Intelligence. Springer India, 2020.
- [CL16] Cleve, J.; Lämmel, U.: Data Mining. De Gruyter Oldenbourg, Berlin, Boston, 2016.
- [FPZ95] Frye, D.; Palfai, T.; Zelazo, D.P.: Theory of mind and rule-based reasoning. In (Cognitive Development, Hrsg.): Cognitive Development, Volume 10, Issue 4. S. 483-527, 1995.
- [Gö10] Gödert, W.: Semantische Wissensrepräsentation und Interoperabilität. In (Deutsche Gesellschaft für Informationswissenschaft und Informationspraxis e. V., Hrsg.): Information - Wissenschaft & Praxis, 61. Jahrgang, Nr. 1. S. 5-28, 2010.
- [Ha16] Hackett, D.: Big Data in Life Insurance, <https://www.mlc.com.au/content/dam/mlc/documents/pdf/media-centre/big-data-report.pdf>, Stand: 14.05.2021.
- [He18] Hellmann, R.: IT-Sicherheit: Eine Einführung. De Gruyter Oldenbourg, Berlin, Boston, 2018.
- [Ki13] Kittel, K.: Agilität von Geschäftsprozessen trotz Compliance. In (Wirtschaftsinformatik Proceedings, Hrsg.): Wirtschaftsinformatik Proceedings 2013. S. 967-981, 2013.
- [KRS20] Köhler, W.; Schultz, C.; Rasche, C.: Das 100% Problem im Datenschutz. G-Forum Konferenz, Karlsruhe, 2020.
- [NP19] Peters, R.; Nauroth, M.: Process-Mining Geschäftsprozesse: smart, schnell und einfach. Springer Gabler, Wiesbaden, 2019.
- [Re20] Reinkemeyer, L.: Process Mining in Action Principles, Use Cases and Outlook. Springer Nature Switzerland AG, 2020.

- [Wu20] Wuttke, L.: Datasolut. Von CRISP-DM: Grundlagen, Ziele und die 6 Phasen des Data Mining Prozess, <https://datasolut.com/crisp-dm-standard/>, Stand: 14.05.2021.

ExPD: Semi-automated Web Extraction of Personal Data

Alexander Böhner,¹ Dominik Herrmann²

Abstract: According to the GDPR, data subjects can issue a Subject Access Request (SAR) to obtain all personal data concerning them from a data controller. Data controllers are interested to automate the handling of SARs, which is challenging for legacy applications and services that lack suitable export functionality. This challenge is addressed by the ExPD prototype, a tool that automates the task of collecting personal data from the pages of web applications. The ExPD operator specifies extraction rules interactively in the browser using a small set of pages. After extraction, a tool-assisted refinement stage allows to fine-tune the exported data. Besides automating the processing of SARs, ExPD may also be useful for citizens who want to obtain an overview of their public data footprint on sites that share user contributions, which is demonstrated in a case study.

Keywords: Subject Access Requests; Crawler; Case Study

1 Introduction

Article 15 of the General Data Protection Regulation (GDPR) of the European Union grants data subjects the right to obtain all personal data concerning them from a data controller. In this paper, we focus on data controllers that run web-based services and applications. This includes, for instance, Software-as-a-Service providers but also employers who provide internal web applications that are used by their employees. While large providers have started to offer self-service portals (e. g., Google Takeout), many small providers process Subject Access Requests (SARs) manually by compiling a human-readable report with data collected from various applications.

Manual collection is time-consuming and error-prone, resulting in dissatisfactory responses [AD18, HL16, KLH20, MAve18, Ur19]. When services and legacy tools lack dedicated export functionalities, some data controllers resort to sharing screenshots with the requested information [HL16]. In at least one documented case a controller even responded with personal information of the wrong subject [KLH20].

According to a 2020 Gartner report, there is increasing interest in the automation of SARs [Ga20]. An early attempt from 2018 used Robotic Process Automation (RPA), which is a sophisticated version of recording and replaying manual interactions with software [Fi18].

¹ University of Bamberg, Privacy and Security in Information Systems Group, 96045 Bamberg, Germany
alexander.boehner@uni-bamberg.de

² University of Bamberg, Privacy and Security in Information Systems Group, 96045 Bamberg, Germany
dh.psi@uni-bamberg.de

There are, however, no public reports about the used techniques, their effectiveness, and their limitations. What is a feasible approach and what are its limitations? To address this question, we present the *ExPD* prototype, a lightweight tool for semi-automated extraction of personal data from web applications. *ExPD* runs within the browser and records to what parts of an application a human operator navigates and what parts of the pages the operator marks for extraction. From that, *ExPD* generates XPath rules and a graph model. The model is used in subsequent runs of *ExPD* to automatically navigate through the whole web application, extracting only the parts on the relevant pages that concern a given (parametrized) data subject (identified by, e. g., a user ID).

*ExPD*s targets both *data controllers* and *data subjects*: *ExPD* enables data controllers to extract personal data without knowledge about programming or databases. Because to this day data controllers reportedly fail to provide satisfactory responses to SARs, *ExPD* enables data subjects to collect at least those pieces of personal data that are exposed on the web interface of an application. As *ExPD* rules are typically agnostic of a particular user identity, data subjects could share *ExPD* rule sets for popular services among themselves.

Outline Section 2 presents related work. After summarizing the requirements in Sect. 3, we explain our approach in Sect 4. After that, we present a case study in Sect. 5 and discuss our approach in Sect. 6, before we conclude in Section 7.

2 Related Work

There are various approaches for data extraction from web pages. *Wrapper induction* systems [LRNS02, Fu14, Lo18] generate programs that extract data from web sources. These approaches are not suitable for our scenario, since they do not traverse an application but are meant to work with a predefined set of pages. The set of pages that contain personal data may change when content is added.

Record and Replay systems record browser interactions and produce scripts for replaying the recorded interaction [AC11, Ch15, CMB18, Li09, Ne15]. *ExPD* is similar to existing Programming by Demonstration (PBD) approaches that are designed for end-users and do not require web or programming knowledge. Like *ExPD*, Vegemite [Li09], Webcombine [Ch15], und Rousillon [CMB18] can *expand* the recorded interactions. Expansion improves the usability for operators because it decreases the number of interactions that have to be recorded.

Handling SARs requires the precise extraction of the personal data of a particular subject, which is not supported by the existing tools. We did not find ready-to-run code to extend one of the existing tools. Thus, we resorted to building an *ExPD* proof-of-concept application from scratch, allowing us to tailor it to the specific set of requirements.

3 Requirements

ExPD is supposed to meet the following three requirements:

R1: Precise Extraction ExPD is required to extract personal data of one specific data subject only – even if a page or a relevant fragment contains data of other subjects. ExPD must enable operators to precisely specify the desired data before collection as well as allow them to review and sanitize the data after extraction.

R2: Efficient Usage Specifying the desired data should be as effortless as possible and not require programming skills. Data spread over reoccurring structures or multiple pages (pagination) should be discovered and collected automatically. It should also be possible to parameterize the data subject so that a given specification can be used to handle SARs for various subjects.

R3: Simple Deployment Setup of ExPD should be easy to encourage adoption by data controllers, researchers, and data subjects. It should run as an application in a desktop environment without depending on third-party services or infrastructure.

4 Approach

ExPD uses three phases to extract personal data from web applications. Firstly, the *learning phase* builds a graph model based on the operator's input. Secondly, the *extraction phase* uses this model to automate both the interaction and extraction in the application. Thirdly, the *refinement phase* presents extracted elements to the operator, who decides what data to export from these elements. In the following, we describe each of the three phases separately. Before that, we clarify the basic concepts of how we model a web application using an example application.

4.1 Example application

We explain our approach with the following example web application, a web application where users can share comments (Fig. 1). The application provides:

1. a list of comments made by different users,
2. detailed descriptions of particular comments,
3. a more accurate description of a comment's date of origin as tooltip,
4. and a listing of the latest logins of the current user (view not included in Fig. 1).

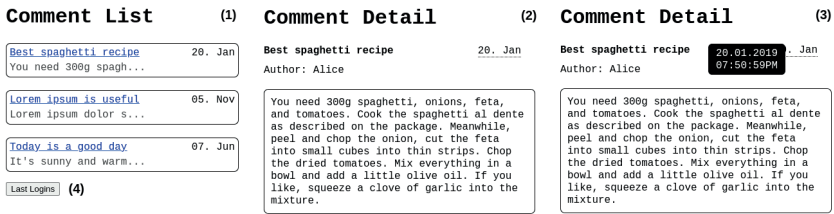


Fig. 1: Comment-sharing web application.

The application displays a comment's description after clicking on the comment in the list. In a comment's description view, hovering with the cursor above its year of origin will show the exact date of creation in a custom-styled tooltip.

In our example application, we want to extract the personal data of the data subject that is a certain user in the application. Personal data is present in (2) and (3) of the application. To screen these pieces of data, our system has to navigate from (1) over (2) to (3). Then, our system has to locate the personal data on the webpage and eventually persist it.

To explain our approach, we first define our terminology with the example application: In the application, (1), (2), and (3) each result from an interaction with the browser, namely, entering a URL, clicking a button, and hovering to reveal a tooltip. Each interaction with the browser leads to an *application state*. We define an application state as the result of an interaction with the application. For instance, a click on a comment in (1) is an interaction that results in the application state shown in (2). A similar interaction can be done with each comment in the list, each resulting in a separate application state. However, in all these states, the application describes a comment in more detail. The states, therefore, share the same concept in the domain of the application. Additionally, the states resulted from similar interactions, i. e., clicking a comment in the list. Because the states are similar regarding the interactions and their domain concept, we introduce the *statetype*. A statetype models similar states. The statetype is the central concept of modeling an application in our system. We assume that states of the same statetype provide

- similar interactions,
- similar patterns in the DOM,
- representation of the webpage, and
- similar types of data.

Furthermore, the notion of statetypes allows us to formalise the interactions between statetypes as *interaction rules*. An interaction rule consists of both an action to take and HTML elements in the webpage, located by XPath. For instance, the list of comments in (1) is formalized as an XPath expression defining the comment elements and the action "click".

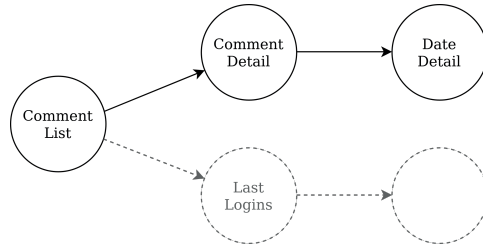


Fig. 2: Graph model of the Comment-sharing web application. Each node represents a statetype.

Clicking each of those elements results in a state of the same type as in (2). Therefore, applying an interaction rule of a statetype results in states of equal statetype.

In contrast to interaction rules, *extraction rules* define elements that contain personal data. Extraction rules use XPath to identify elements in the DOM of an application state, as well. For instance, in (2), an extraction rule would define the element containing the comment’s author. The same extraction rule applies to all other application states that share the same statetype as (2) because each of them describes a comment with its author. Extraction rules additionally may specify another element that enables *conditional extractions*. Such rule only extracts if the conditional element contains given keywords that identify the data subject. For instance, in (2), an extraction rule would define the comment’s full text only to be extracted if the “author” element contains the keyword “Alice”.

Statetypes, interaction rules, and extraction rules are used to build a graph model of the web application. In this model, statetypes are nodes connected by interaction rules as edges. Together, they model possible interactions in the web application. For the example above, Fig. 2 illustrates a possible graph. Here, the statetypes map to the sample states in Fig. 1. At each node, extraction rules may be specified (not shown in Fig. 2).

4.2 Learning Phase

To construct the graph, ExPD captures the operator’s input in an instrumented browser. At the beginning of the learning phase, ExPD opens a given URL for the desired web application. ExPD executes a JavaScript script in the instrumented browser to display a control overlay for the operator’s input (Fig. 3). With this overlay, the operator may specify interaction and extraction rules for the start state in the graph model. To define a rule, the operator selects page elements by clicking the highlighted bounding rectangles. For these elements, the corresponding action can be defined in the control overlay. For instance, the operator may select multiple comment elements in the list with the action “click” to crawl as a list. The resulting interaction rule is shown in Figure 4.

In this rule, the XPath expression defines all “li” elements in the comment list. The expression captures all “li” (comment) elements, because no index for these in the correspondig DOM

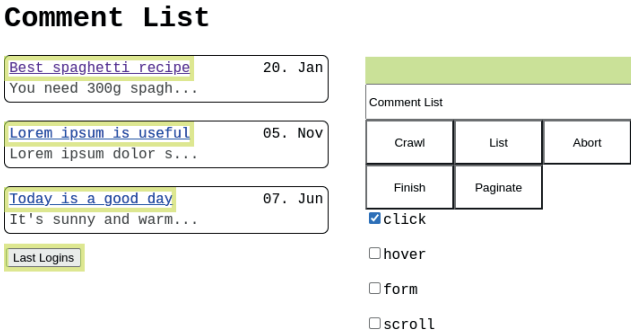


Fig. 3: User interface for operating the ExPD prototype. The control overlay and bounding boxes of elements are used for the definition of rules.

```
'action': 'click'
expression': 'body[1]/div[5]/ul[2]/li/div[1]'
```

Fig. 4: Interaction rule for clicking on items in a list. The “li” segment applies to all “li” elements in that level of hierarchy.

hierarchy is given. ExPD appends the rule to the start state in the graph model. After the operator has specified the rules, clicking the “Finish” button in the overlay completes the processing of the current state. ExPD then picks an interaction rule and executes it in the browser. The application reaches a new state exemplary for a different statetype than the start state. ExPD appends a new statetype in the graph with the just executed interaction rule as the edge to the start state. As with the start state, the operator is asked to specify interaction rules which eventually form an edge to a future statetype. This way, ExPD builds the graph statetype by statetype until there are no interaction rules left to execute. If an interaction rule of a different statetype than the current should be executed, ExPD first navigates to a state of the corresponding statetype. For this, ExPD calculates the shortest path in the graph model and executes the resulting interactions accordingly. To avoid navigating through the application for each new statetype, ExPD executes interaction rules in a depth-first manner.

In the process of defining rules for statetypes, the operator may also define extraction rules with the overlay controls. In contrast to interaction rules, the operator may specify a conditional element for extraction rules. The keywords used in conditional extractions are specified in a configuration file.

ExPD also supports interaction rules that are executed before reaching the start state. The rules are defined by the operator before the regular learning phase. They are called *one-time rules* because they are executed only once in the extraction phase, before the actual extraction. With these rules, the operator may log in as a specific user, click away cookie banners, or do any other task required before the extraction.

4.3 Extraction Phase

Once all rules have been added, the operator starts the extraction. ExPD uses the learned application model in the extraction phase to automatically extract personal data. First, the operator has to define an URL as the starting point in the configuration file. The application state resulting from opening the URL has to be of the same statetype as the start state in the graph. In addition to the URL, keywords for conditional extractions have to be defined in the configuration file. Because both the URL and keywords can be parametrized, the same learned model can be used for extractions with different data subjects. For instance, ExPD could start at different user profiles (one for each data subject) in the first state and perform conditional extractions only if that user's name is present.

ExPD begins the extraction at the provided URL. ExPD searches for possible interactions of each interaction rule in the corresponding statetype. For instance, all comment links in the list are possible elements to click on according to the rule in Fig. 4. ExPD appends the current application state, including the possible interactions to the corresponding statetype in the model. Analogous to the learning phase, ExPD picks an interaction to execute and explores new application states. At each state, possible interactions are determined and appended to the graph. ExPD terminates if it executed all possible interactions at least once. Because each interaction rule may result in multiple possible interactions at each state, ExPD traverses many more states than the operator used for learning the model. Still, ExPD does not leave the application model, because each visited state is linked to a statetype.

While ExPD navigates through application states, extraction rules are applied at each visited state. Extraction rules are looked up in the corresponding nodes of the statetypes. ExPD saves extracted elements that match the XPath expressions in the extraction rules at the application state in the graph.

4.4 Refinement Phase

After the extraction phase, the operator can review the extracted elements and determine the final personal data export. The operator can browse the extraction rules of each statetypes and view the extracted data, as shown in Fig. 5. Here, the operator may select text in the extracted elements to form refinement patterns. Refinement patterns specify either an HTML element's attribute or the innerText. The patterns can apply to all other data of the same rule. Therefore, the operator may specify patterns and then apply these to all extracted pieces of data to save time. Additionally, the operator can exclude elements to avoid exporting, e. g., business secrets, or other data subjects' data. Once the operator has decided on the personal data to export, ExPD prompts for additional data such as the purposes of the processing, or the categories of personal data. Finally, the operator can export the data in PDF, CSV, or JSON format.

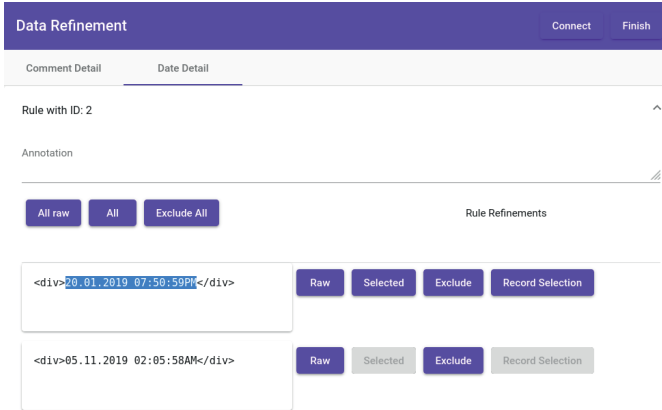


Fig. 5: Refinement interface of the prototype for selecting personal data in extracted elements.

4.5 Technical Details

The ExPD prototype³ is implemented in two modules. ExPD’s core is a Python application that instruments the browser with Selenium WebDriver⁴. We use Google Chrome as the instrumented browser. The refinement UI is developed as a separate angular application. Both modules use the Neo4J Graph database⁵ for managing the application graph model.

5 Case Study

In this section we demonstrate the ramifications of using ExPD in practice using the question and answers platform *stackoverflow.com* as an example. This platform accepts SARs via a web form. For the case study, however, we are interested in the use case of a private individual who wishes to obtain their publicly accessible data footprint. We will collect selected personal data with ExPD. To obtain a ground-truth baseline, we will use the public API of the platform.

5.1 Extracted Personal Data

We determined our data subject by picking a random user from the top 0.01% users sorted by highest reputation. Our data subject has posted more than 90 questions, 9000 answers, and 25,000 comments.

³ <https://github.com/UBA-PSI/ExPD>

⁴ <https://www.selenium.dev/>

⁵ <https://github.com/neo4j/neo4j>

The data subject is assumed to be interested in all their answers, questions, and comments on the platform. To allow for an unambiguous comparison of ExPD and API, we will only extract the times of creation, last activity, and last edit associated with these entities.

5.2 Extraction with API

One co-author, an experienced software developer, familiarized himself with the Stackoverflow API, obtained an API key, and implemented a Python script that complied to the prescribed rate limit. The whole process took less than two hours. The execution time of the script was about 75 minutes, resulting in 38,739 extracted timestamps.

5.3 Extraction with ExPD

Before recording rules with ExPD, one co-author familiarized himself with the structure of Stackoverflow and the location of the desired timestamp. He decided to take the *profile page* of the data subject as the start-state by putting its URL into ExPD's configuration file. Additionally, he specified the username of the data subject as a keyword in the configuration file. After that, he launched ExPD's learning phase, which opens an instrumented browser. Then, the operator subsequently enters actions in the provided overlay to gather rules for the automatic extraction.

In total, the operator defined eight interaction rules and eight extraction rules. The training required six minutes of manual work, then ExPD started to extract the data automatically (extraction phase), which took 91 minutes. In the subsequent refinement phase the operator worked through the statetypes and extraction rules. After selecting the desired text for each rule, he obtained a CSV file with all exported data. In this case, refinement took three minutes. ExPD extracted 29,401 date values in total.

5.4 Comparison

First of all, we note that ExPD extracted a strict subset of the ground-truth data, i. e., ExPD did not extract data pertaining to other users. However, ExPD only extracted 75.8% of the timestamps contained in the ground-truth. We investigated the missing values and found that we failed to define an extraction rule for the edit-date of a user's answers. None of the pages that the operator encountered before ExPD started the automated extraction contained an edit-date.

To address this issue, which is a central issue in any programming-by-example tool, we plan to give the operator additional control to increase page coverage in the learning phase. Apart from that, ExPD managed to extract all the remaining dates without manual intervention or

programming on the part of the operator. While the API script was faster, the time of manual interaction with ExPD was shorter than the time needed to implement the API script.

6 Discussion

The ExPD prototype is capable of extracting personal data from web applications without programming experience. It takes less human effort than a manual extraction, avoids extracting critical data, and performs all tasks required from crawling to PDF exporting to ease its integration in a process. Because of this, ExPD is a viable option for extracting personal data in a subject access process. For private individuals, ExPD encourages them to do their own research. If an individual is interested in its personal data on a website, ExPD lessens manual labor that this individual might not be willing to do. Overall, ExPD intends to make personal data more accessible for private individuals and support data controllers to comply with the GDPR.

Ethical Considerations With ExPD, the personal data of any data subject in web applications may be extracted. Therefore, ExPD could also be used to gather personal data without legitimate interest or the data subject's consent. However, especially when using an API, extracting personal data is feasible without ExPD, nevertheless. Instead, ExPD at least grants private individuals the possibility to access their public personal data with reasonable effort. Individuals can use this data to understand what others can find out about them.

Limitations Our approach has several limitations.

The web application may not act as defined in the stypetype model because of several reasons. Firstly, technical issues like e. g., software errors can cause a state to differ from an expected state. Secondly, the assumption that states of certain stypetype share the same interaction options, structure, and data may not always apply. The assumption does not hold if, e. g., a state only includes a button if certain data is present. Thirdly, the application may change during extraction. Especially, if the operator is not in control of the application's deployment, A/B testing may interfere in the extraction. To mitigate these, ExPD checks if elements include similar text before executing the interaction. To make the extraction less vulnerable to these limitations, ExPD could adopt more sophisticated element localisations with the XPath language.

Additionally, ExPD is limited by the operator's knowledge about the web application. ExPD will not extract personal data if the operator does not know where to find personal data in the application. To avoid missing personal data in this way, ExPD could autonomously crawl the application to find applications states containing similar data to already extracted data. ExPD could then point the operator to these application states.

Future Work We intend to make the applications' graph models portable and ultimately distributable. Distributed models can be learned, distributed publicly, and eventually used or improved by a community of users. Consequently, operators can save time using these models and are less prone to make errors during learning.

Additionally, we will reduce missed statetypes in the learning phase by providing suggestions for states that contain potential desired data.

7 Conclusion

ExPD instruments a web browser to extract personal data from web applications. Its architecture is more lightweight than sophisticated Robot Process Automation (RPA) or Record-Replay systems presented in the literature. Thus, ExPD can be set up on desktop machines with little effort.

Operators do not need programming skills to extract data with ExPD. They specify interaction and extraction rules by browsing the web application and labelling the relevant parts of the page via a control panel that floats on top of websites. Rules can be conditioned to ensure that only personal data that is relevant to a particular user is extracted. Using parametrization, a given rule set can be reused for Subject Access Requests (SARs) of different users.

As illustrated in the StackExchange case study, ExPD can also be used by *data subjects*, e. g., when a controller fails to provide a satisfactory response to a SAR. In such cases, ExPD allows data subjects to get an overview of their *public* data footprint on the respective site, which may encourage them to review and amend their profile on that site.

Bibliography

- [AC11] Andrica, Silviu; Candea, George: WaRR: A tool for high-fidelity web application record and replay. In: DSN 2011, Proceedings of the 2011 IEEE/IFIP International Conference on Dependable Systems and Networks. IEEE Compute Society, pp. 403–410, 2011.
- [AD18] Ausloos, Jef; Dewitte, Pierre: Shattering One-Way Mirrors. Data Subject Access Rights in Practice. Data Subject Access Rights in Practice. International Data Privacy Law, 8(1):4–28, 2018.
- [Ch15] Chasins, Sarah; Barman, Shaon; Bodík, Rastislav; Gulwani, Sumit: Browser Record and Replay as a Building Block for End-User Web Automation Tools. In (Gangemi, Aldo; Leonardi, Stefano; Panconesi, Alessandro, eds): WWW 2015, Proceedings of the 24th International Conference on World Wide Web Companion Volume. ACM, pp. 179–182, 2015.
- [CMB18] Chasins, Sarah E.; Mueller, Maria; Bodík, Rastislav; Rousillon: Scraping Distributed Hierarchical Web Data. In (Baudisch, Patrick; Schmidt, Albrecht; Wilson, Andy, eds): UIST 2018, Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology. ACM, pp. 963–975, 2018.

- [Fi18] Fieldfisher: , A novel approach to data subject requests. <https://www.fieldfisher.com/en-ie/locations/ireland/ireland-blog/a-novel-approach-to-data-subject-requests>, 2018.
- [Fu14] Furche, Tim; Gottlob, Georg; Grasso, Giovanni; Guo, Xiaonan; Orsi, Giorgio; Schallhart, Christian; Wang, Cheng: DIADEM: Thousands of Websites to a Single Database. *Proc. VLDB Endow.*, 7(14):1845–1856, 2014.
- [Ga20] Gartner: , Gartner Identifies the Legal and Compliance Technologies to Focus on Post COVID-19. <https://www.gartner.com/en/newsroom/press-releases/2020-10-05-gartner-identifies-the-legal-and-compliance-technologies-to-focus-on-post-covid-19>, 2020.
- [HL16] Herrmann, Dominik; Lindemann, Jens: Obtaining personal data and asking for erasure: do app vendors and website owners honour your privacy rights? In (Meier, Michael; Reinhardt, Delphine; Wendzel, Steffen, eds): *Sicherheit 2016: Sicherheit, Schutz und Zuverlässigkeit, Beiträge der 8. Jahrestagung des Fachbereichs Sicherheit der Gesellschaft für Informatik e.V. (GI)*. volume P-256 of LNI. GI, pp. 149–160, 2016.
- [KLH20] Kröger, Jacob Leon; Lindemann, Jens; Herrmann, Dominik: How do app vendors respond to subject access requests? A longitudinal privacy study on iOS and Android Apps. In (Volkamer, Melanie; Wressnegger, Christian, eds): *ARES 2020, Proceedings of the 15th International Conference on Availability, Reliability and Security*. ACM, pp. 10:1–10:10, 2020.
- [Li09] Lin, James; Wong, Jeffrey; Nichols, Jeffrey; Cypher, Allen; Lau, Tessa A.: End-user Programming of Mashups with Vegemite. In (Conati, Cristina; Bauer, Mathias; Oliver, Nuria; Weld, Daniel S., eds): *IUI 2009, Proceedings of the 14th International Conference on Intelligent User Interfaces*. ACM, pp. 97–106, 2009.
- [Lo18] Lockard, Colin; Dong, Xin Luna; Shiralkar, Prashant; Einolghozati, Arash: CERES: Distantly Supervised Relation Extraction from the Semi-Structured Web. *Proc. VLDB Endow.*, 11(10):1084–1096, 2018.
- [LRNS02] Laender, Alberto H. F.; Ribeiro-Neto, Berthier A.; Silva, Altigran Soares da: DEByE – Data Extraction By Example. *Data Knowl. Eng.*, 40(2):121–154, 2002.
- [MAvE18] Mahieu, René L. P.; Asghari, Hadi; van Eeten, Michel: Collectively exercising the right of access: individual effort, societal effect. *Internet Policy Review*, 7(3), 2018.
- [Ne15] Netravali, Ravi; Sivaraman, Anirudh; Das, Somak; Goyal, Ameesh; Winstein, Keith; Mickens, James; Balakrishnan, Hari; Mahimahi: Accurate Record-and-Replay for HTTP. In (Lu, Shan; Riedel, Erik, eds): *USENIX ATC 2015, USENIX Annual Technical Conference*. USENIX Association, pp. 417–429, 2015.
- [Ur19] Urban, Tobias; Tatang, Dennis; Degeling, Martin; Holz, Thorsten; Pohlmann, Norbert: A Study on Subject Data Access in Online Advertising After the GDPR. In (Pérez-Solà, Cristina; Navarro-Arribas, Guillermo; Biryukov, Alex; García-Alfaro, Joaquín, eds): *DPM 2019 and CBT 2019, Proceedings of the ESORICS 2019 International Workshops on Data Privacy Management, Cryptocurrencies and Blockchain Technology*. volume 11737 of *Lecture Notes in Computer Science*. Springer, pp. 61–79, 2019.

Technical & Data Protection Aspects of a Smart Digital Control Center for Smart Cities

Jan-Philipp Stroscher¹, Marius Schnaubelt², Alejandro Sanchez Guinea³, Stefan Fabian², Julius von Willich³, Yasin Alhamwy⁴, Maximilian Bauer⁵, Oskar von Stryk², Max Mühlhäuser³, Kurt Geihs⁶, Uwe Klingauf⁵, Gerrit Hornung¹

Abstract: The advances of smart city infrastructure spark an increasing interest in digital twins for cities using the plethora of new sensors that will be deployed across the city to monitor the urban environment. In normal operation, the digital twin provides an overview of the current state of the city. During disaster scenarios, the digital twin of the city will be able to provide the emergency services with an overview of the situation, increasing their effectiveness and supporting them in planning their course of action. This calls for a discussion about the data protection laws as well as data protection by design approaches and secure processing techniques in accordance with current European law. Our paper will discuss the related data protection laws and describe the technical overview of a system with both, permanent stationary sensors and mobile on-demand agents.

Keywords: Situation Control Center; Smart Cities; Data Protection Laws; Smart Street Lamps; Rescue Robotics; Emergency Response

1 Introduction & Motivation

Future smart cities are dependant on sensors that are distributed across the city to maintain their digital twin. Smart street lamps have been proposed as the ideal platform [Mü20] to introduce sensors into a smart city since street lamps are an essential facility of a city's environment and are already densely deployed in the urban landscape.

In normal operation, i.e., during normal days the smart street lamps continuously capture data and relay the required data where needed. The smart street lamps have a wide array of sensors such as laser scanners (LiDARs), cameras, temperature sensors, and other environmental sensors, to capture data from the surrounding environment [Mü20]. This poses multiple questions regarding the data protection laws, such as: What sensors can be used? What data can be captured and processed during normal operation?

¹ {jan-philipp.stroscher, gerrit.hornung}@uni-kassel.de – University of Kassel, Institute of Economic Law

² {fabian, schnaubelt, stryk}@sim.tu-darmstadt.de – Technical University of Darmstadt, Simulation, Systems Optimization and Robotics Group (SIM)

³ {max, sanchez, willich}@tk.tu-darmstadt.de – Technical University of Darmstadt, Telecooperation Lab (TK)

⁴ alhamwy@vs.uni-kassel.de – University of Kassel, Distributed Systems

⁵ {bauer, klingauf}@fsr.tu-darmstadt.de – Technical University of Darmstadt, Institute of flight systems and automatic control (FSR)

⁶ geihs@uni-kassel.de – University of Kassel, Scientific Center for Information System Design (ITeG)

During a state of emergency, be it natural or man-made disasters, getting an overview of the current situation is of paramount importance for the emergency services, helping them plan their course of action [NKP15]. A Smart Control Center (SCC) acts as a hub for the information during an emergency, combining, processing, and visualizing the sensor information provided by all the agents. The SCC should encompass the following characteristics: human-centered, comprehensive, proactive, coordinated, and self-healing as defined by the vision presented in [ZLB10].

Let us assume that due to a natural or man-made disaster (be it an earthquake, gas explosion, fire, etc.) a shopping mall has collapsed. Some people will be trapped and others will be fleeing the area. The smart street lamps will be able to gather information from the surrounding area which will provide an overview of the current situation from their perspective area, which in turn can help save many lives. However, the smart street lamps are stationary and, therefore, will have blind spots that require exploration to get a near-complete overview.

Mobile search and rescue robots are able to reach and explore these blind spots and provide a more detailed and maneuverable overview of the situation [CM03; Su19]. Unmanned ground vehicles (UGVs) and unmanned aerial vehicles (UAVs) can assist the human operators with performing complex tasks from a safe distance. Additionally, UAVs can provide a bird's eye view of the scene, e.g., as a rough situation overview or locally, to aid the UGVs in their manipulation tasks. This leads to further questions regarding the data protection laws. The type of data (anonymous data, personal data, and/or special categories of personal data) has a significant impact on the legal requirements and thus on the concrete design of the systems used. In addition to the basic applicability of the General Data Protection Regulation (GDPR), it is also a matter of fulfilling the other obligations under data protection law. Data protection should already be comprehensively considered during the technical development of the systems.

In addition to the question of what can be permissible under data protection law, data protection-friendly technology design should also be used in order to comply with the right to informational self-determination on which data protection is based on (BVerfGE 65, 1, 41 et seq. (census ruling)) to the greatest possible extent in order to achieve broad acceptance among the population [KN19] and preventing malicious use [Wh21] of such sensor data. Concretely, the following questions arise: What changes from the normal operation and what can now be processed?

2 Technical Aspects

In this section, we shortly introduce the technical aspects of the smart control center.

2.1 Smart street lamps

The vision of smart cities revolves around the idea of having the city providing its citizens a wide range of services that improve their quality of life, in terms of mobility, health, resource management, and overall experience [A118; Ri18]. These sensors can also assist first responders in detecting and handling disasters. To achieve this vision, street lamps as a platform (termed SLaaP) has been proposed, allowing to provide novel services to citizens through an innovative true city infrastructure [Mü20].

2.1.1 Street lamps in perspective

Street lamps are an essential and pervasive facility of city environments. Three of their natural characteristics have been recognized as key aspects for the development of smart digital city capabilities [Mü20]. First, street lamps are in general connected to existing electric power lines, which can become the basis for a next-generation digital infrastructure that incorporates various computing, networking, and Internet-of-Things (IoT) components [AIM10]. The second natural characteristic of street lamps refers to how densely deployed they are in the urban landscape. This makes them already ubiquitous in the life of the cities' inhabitants and thus the perfect starting point towards establishing an urban digital infrastructure that is easily accessible and scalable. The third characteristic is that street lamps are at large publicly owned, which allows to establish broad and comprehensive control mechanisms and regulations at the city level.

2.1.2 Key aspects of smart street lamps

In order to use street lamps as the basis of a citywide infrastructure that can enable a smart control center as proposed, they should be augmented both in hardware and software in a selective and strategical manner across the city to ensure effective and efficient coverage. In terms of hardware, there are four aspects that have been identified as essential for this purpose [Mü20].

- *Sensors* and *actuators* should be installed to allow for data acquisition and interaction. Specifically, for a smart control center, it is necessary to obtain accurate spatio-temporal data, for which range active sensors such as LiDAR and radar are ideal, as they allow to create 3D representations of the scanned environments with high precision. Concerning interaction, next-generation laser based projectors installed on the street lamps can provide situation-aware and personalized information to citizens.
- Concerning *computing* and *storage* resources, 4D visualizations such as the one proposed by our smart control center require nearby computational resources such as fog/edge computing, edgeClouds, or cloudlets [Sa17]. These resources can be

provided by street lamps with their unique characteristics detailed above. Single-board computers suitable for executing complex neural network based machine learning and computer vision approaches directly on the street lamps are required to enable privacy-preserving processing of the 4D data, preserving privacy-critical information in the lamps, while transferring the rest of the data to the cloud.

- In terms of *communications* and *networking*, as the focus of our smart control center relates to critical situations, it is essential to augment the lamps as to guarantee connectivity to the Internet, to nearby users and devices, as well as to other street lamps. This can be achieved by taking advantage of the natural proximity and density of street lamps, which opens up the possibility of establishing a true decentralized mesh network. To this end, a minimum viable set of street lamps would be connected to a fiber network, acting as the gateway nodes for the street lamp network. Then, the gateway nodes would relay the traffic to nearby lamps via WiFi or millimetre wave (mmWave) communication, taking advantage of the spatial distribution of street lamps.
- Smart street lamps can be of help for *energy management* by providing charging points for mobile agents.

2.2 Mobile Agents

Mobile rescue robots in the air, called UAVs, and on the ground, called UGVs, can temporarily provide a more detailed and movable overview of the situation in local areas of interest than provided by the stationary smart street lamps. The mobile agents can be tele-operated by human operators to perform tasks from a safe distance in hostile environments. Using complex software, they can also perform tasks autonomously with minimal or – in the case of connection loss – no supervision, and assist the human operators in performing complex tasks with smart assistance functions. Furthermore, teams of agents could collaborate to perform tasks that are impossible or too difficult for a single agent. Table 1 shows an overview of the sensor types that are presented in the following section and for what purposes they may be aggregated. The listed sensor types cover the types usually employed in a rescue robot such as the robot built by and used in the SIM research group [Sc21b]. UAVs supporting the ground robots from air can provide detailed camera images in a bird's eye view to facilitate navigation or manipulation tasks.

- *LiDARs* are visual distance sensors that provide a cloud of surface points in the environment by emitting laser impulses and measuring the time for the reflected light to return to the sensor. These surface points are aggregated to collect information about the surface of the robot's environment. Using Simultaneous Localization and Mapping (SLAM) techniques, the robot can localize itself in the aggregated pointcloud data and match new pointclouds to build a consistent map. Based on this localization of the incoming pointclouds multiple environment representations are

<i>Sensor</i>	<i>Usage</i>	<i>Aggregation</i>
LiDAR (Pointcloud)	On Robot, Streamed	Low-resolution static 3D model
Radar, Sonar	On Robot, Streamed	Low-resolution static 2D model, Vital signs
Cameras (Color, Thermal)	On Robot, Streamed	Optional
Point Measurements	On Robot, Streamed	Single hypothesis (e.g. GNSS position, Radiation source localization)

Tab. 1: Overview of the sensor data types present used by the mobile agents. On Robot means that the data is processed on robot and aggregated into a model, streamed means that the the data is streamed to operator station for visualization.

built including but not limited to an elevation map and an occupancy grid. Elevation maps are a 2D grid representation of the ground surface that is often employed for UGV path planning where each cell represents the maximum elevation at this cell above a previously set zero-plane. Occupancy grids are used for collision avoidance and divide the world into 3D voxels (cubes with fixed size) that are either occupied, free or unknown. The pointclouds can also be colored using camera information and aggregated into meshes.

- Visual sensors such as LiDAR and cameras provide only inadequate and unreliable measurement values under the influence of aerosols such as smoke, fire, swirling dust or fog. To ensure robust localization and navigation under these conditions, *radar* and *sonar* sensors are deployed as these are not significantly affected by these aerosols. Furthermore, radars can be used to perform vital sign detection, e.g., to estimate the vital parameters of human beings.
- *Camera data* (e.g., color, omnidirectional or thermal images) can improve the situation awareness when combined with the 3D information, e.g., using the omnidirectional images, the 3D mesh can be colorized. By using thermal data, heat spots can be detected or color images can be processed to add semantic information to the world model.
- The robot is also equipped with various sensors that measure environmental factors at the location of the robot, which are called *point measurements*. This can be, for example, the radio signal strength at the robot's antennas, the concentration of gases such as CO₂, Global Navigation Satellite System (GNSS) location information, and radiation measurements.

Due to the nature of the sensors used on mobile agents, they should only be used in the event of a disaster to limit the processing of personal data to what is necessary for the purpose of crisis handling.

2.3 Visualization

An operation center with access to the information gathered by street lamps and mobile agents enables its users to get a quick and detailed overview of the disaster at hand. This knowledge in turn can not only be used to plan rescue operations but also identify areas in need of exploration by either the mobile agents or the rescue personnel.

With the immense volume of data generated by the plethora of sensors deployed in such a potential emergency scenario, pre-processing – using fog or edge computing – becomes a necessity in order to keep the data manageable. A possible solution for this problem is an abstract, model-based approach breaking down sensor data to a sensible minimum before transmitting it away from the sensor node, in order to enforce privacy. Model-based here meaning transmitting, e.g., *human, adult* instead of a pointcloud and using this data to drive a rendered 3D representation.

This data can then be used to drive a model-based immersive and interactive virtual representation offering high-fidelity to the user while autonomously managing the level of detail presented as demonstrated by Prandi et al. [Pr14]. Such an extensive virtual scene offers the operator an overview of the whole disaster while more detailed or even raw data can then be requested as required.

With the spatial nature of city data, modern technologies such as Augmented Reality (AR) and Virtual Reality (VR) with their inherent 3D representation offer the ideal visualization environment [Mo18]. The 6-Degrees of Freedom (DoF) controllers usually employed alongside current VR Head Mounted Displays (HMDs) also offer an intuitive interaction method for the spatial data at hand [Ja17]. The same holds true for mid-air interaction methods featured by current AR headsets [Zh18].

Additionally, information could be shared through secure communication channels with authorized rescue personnel equipped with AR devices enabling them to better react to the situation at hand [SMS21] while maintaining security of processing in accordance to Art. 32 GDPR. Such support can range from providing expert knowledge [Gu12] to supporting their decision-making in, e.g., triage [Mi12] and enhance their situational awareness [Al14].

Finally, local storage in the operation center enables the operator to examine the timeline of the catastrophe. This would also allow conducting less urgent tasks after the scenario is concluded, such as prosecuting onlookers. Additionally, this central storage approach also allows for easier deletion of the data within legal deadlines once the incident has concluded in compliance with the principle of storage limitation Art. 6 (1) (e) GDPR. Here, suitable visualization techniques such as the space-time cube presented by Filho et al. [FSN20] can aid the operator in understanding events in their temporal context or even highlight correlations not readily apparent with current visualization techniques.

3 Data Protection Aspects

The use of the various sensors described above generates a large amount of data, which always includes personal data. In order to be able to operate such systems in a legally secure manner, the first step is to comply with the data protection requirements, in particular, those of the GDPR. This includes the implementation of data protection requirements through technical design (Art. 25 (1) GDPR) [Ho11]. In the second step, consideration should also be given to further measures for the data protection-friendly and data-secure design of such systems which under certain circumstances may go beyond mere permissibility to strengthen acceptance among the public. Furthermore, the different operation conditions (non-crisis/crisis) also lead to questions of differentiation in terms of data protection law with regard to the existing legal bases for processing personal data [HS21a; HS21b].

For the assessment of the applicability of the GDPR, with the accompanying follow-up questions, the “processing” of “personal data” within the meaning of Art. 4 No. 1 GDPR must first be determined in accordance with Art. 2 (1) GDPR. Processing in the sense of Art. 4 No. 2 GDPR is given both in the collection of data by means of the sensors installed in the smart street lamps, as well as by the UAVs and UGVs, since various data is automatically collected, processed and transmitted to the Control Center in these systems (see Section 2).

Then there is the question of the personal reference of the data within the meaning of Art. 4 No. 1 GDPR. According to the definition pursuant to Art. 4 No. 1 GDPR, “personal data means any information relating to an identified or identifiable natural person [...]”. In addition, the GDPR considers an identifiable natural person as “one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person”. Recital 26 pp. 1 to 4 to the GDPR provides further guidance on how the European legislator believes the provision should be made. However, these again include terms that are subject to interpretation, such as “according to general discretion”, “probable”, “objective factors” or the technological standard to be used as a basis.

Due to the terms requiring interpretation, it is disputed how exactly the allocation is to be determined in terms of identifiability [BE15; Ec21]. This discussion has been controversial for years and has been repeatedly decided by the highest court for various constellations [see, for example, ECJ v. 19.10.2016 - C-582/14, CR 2016, 791 m. Nink - Breyer (dynamic IP addresses), BGH v. 16.5.2017 - VI ZR 135/13, CR 2017, 662 m. Note Keppeler = NJW 2017, 2416]. The ECJ’s decision in Case C-582/14 was decisive in this regard: In this decision, the ECJ pointed out that recital 26 requires controllers to take into account the means that they themselves or another person are generally likely to use. On the other hand, it is not required that the necessary information or means are directly available to the responsible body or are used in a specific individual case. Rather, it is sufficient if the responsible entity can legally and reasonably access this data. The access does not have to have actually taken place. Personal data, on the other hand, must always be denied if identification is not practically

feasible. Such cases can arise, for example, from a disproportionate effort in terms of time, costs and manpower. In such cases, the risk of identification is so low that it is negligible [ECJ, NVwZ 2017, 213 para. 39 et seq.; in summary: Karg, in: Simitis/Hornung/Spiecker gen. Döhmman, Datenschutzrecht, 2019, Art. 4 no. 1 marginal no. 61 with further references]. After the very abstract clarification of the question of the personal reference of data, the core of the future dispute will move to a lower level of abstraction by asking the question of the use of the means by the responsible party.

Based on this, the following applies to sensor data at hand: LiDARs do not in themselves allow for the identifiability of individuals, so that the scope of the GDPR is not opened for this type of data due to the lack of personal reference. The same applies to radar, sonar and point measurements. In the case of cameras (as per the specific design selected here), identifiability is possible on the basis of the real image, so that a reference to persons is given [Sc21a, ECJ, 11.12.2019 – C-708/18, recital 34]. When using thermal imaging cameras, on the other hand, it depends on the concrete image reproduction, i.e., the question of how detailed the imaging is. Therefore, as an interim conclusion, the GDPR initially only applies to the data generated by the cameras.

When a large amount of data is combined, it is precisely the combination of these different data that creates a reference to a person. Therefore, in particular for LiDAR, radar, and sonar data, it must be answered to what extent a combination with other data can lead to identifiability. In this constellation, special consideration must be given to the fact that the data is merged into a database that is operated by a public entity. Therefore, in addition to the restrictions under data protection law, questions about the informational separation of powers (separation of the processed data according to the task assignment/purpose fulfilment of different or also the same authorities) [La10, p. 337 et seq.] must also be taken into account, but this is not the focus of this paper. Along with this, the purpose limitation principle acquires greater weight since it has implications for the separation of informational powers.

Based on the criteria established by the ECJ to be taken into account when assessing the purpose of a person, the fundamental question arises as to what measures are reasonably taken by a public body. In principle, public authorities act on the basis of powers granted to them by law or otherwise. Within this allocation of tasks, the need for data processing is also recognized in principle by the GDPR, as can be seen from Art. 6(1)(e) of the GDPR. Here, the corresponding national task allocation norm or the corresponding norm in the Federal Data Protection Act (BDSG) or the state data protection laws must also be consulted. Therefore, it depends on which specific authority collects this sensor data and what other data is already available to this authority. This leads to the conclusion that this question must be determined in each case for the specific application scenario. With reference to the sensor data described here, for example, identifiability can be achieved by combining the recorded LiDAR data with real images of the UAVs or UGVs, thus enabling identification of the corresponding person by means of motion tracking. Furthermore, it must be taken into account that such a reference to a person can also only occur “creepingly” in the course

of time. This should also be taken into account in the technical design in such a way that technical measures to minimise this danger should be considered [HW19].

These possibilities show that in the vast majority of cases personal data is processed. This means that the requirements of the GDPR must be met in their entirety. This applies in particular to the principles of processing in Art. 5 of the GDPR and, in particular, the lawfulness of processing. Thus, as far as data is used for disaster preparedness and response, a legal basis from Art. 6 of the GDPR is required. Since area-wide consent is unrealistic in the smart city, legal bases in national law will be required for state operators of the systems (Art. 6(1)(1)(e), (2), (3) GDPR). With regard to the question of the relevant legal basis, a distinction must be made between “normal operation”, i.e., such operation that does not involve any special incidents and “disaster operation”. Data processing primarily by UAVs and UGVs in the event of a disaster will *prima facie* be possible within the framework of the applicable security law. For this, the corresponding legal requirements must be determined for the respective concrete operational scenarios. The specific reason for the processing results in restrictions regarding the collection (such as the link to a relevant risk) but also the further processing for other purposes (principle of purpose limitation). Therefore, new legal bases are required for the permanent use of smart street lamps (as part of informational disaster preparedness), insofar as, as explained above, a merging of data is intended, since in these cases personal data will regularly be available. This also applies in particular against the background that the data obtained is also to be used further for damage analysis and prevention [HS21b].

In addition, the requirements of Art. 25, 32 GDPR must be complied with [see: ENISA, Privacy and Data Protection by Design - from Policy to Engineering]. At this point, it is possible to comprehensively take into account the protection of the data subjects' informational self-determination from the outset when designing the systems. Article 25 of the GDPR, through its reference to Article 5 of the GDPR, but also the entire GDPR, intends a comprehensive implementation of the requirements through an appropriate “design” whereby organisational processes are also addressed in addition to the technical design [Hansen, in: Simitis/Hornung/Spiecker gen. Döhmman, *Datenschutzrecht*, 1st ed. 2019, Art. 25 recital 15 et seq.]. To this end, data protection requirements are to be included in the development at an early stage by considering appropriate measures. However, these measures must explicitly still be fulfilled during the processing itself [BG17, Martini, in: Paal/Pauly, 3rd ed. 2021, DS-GVO Art. 32 recital 1b et seq., Hansen, in: Simitis/Hornung/Spiecker gen. Döhmman, *Datenschutzrecht*, 1st ed. 2019, Art. 25 recital 19].

In contrast to Art. 25 GDPR, Art. 32 GDPR addresses aspects of security of processing and thus concretizes the data protection principle of “integrity and confidentiality” from Art. 5(1)(f) GDPR (although this could also be considered to be encompassed by Art. 25 GDPR). This includes measures that serve in particular to protect personal data and processing systems both from attacks and from a failure or impairment of the reliability of the systems [Hansen, in: Simitis/Hornung/Spiecker gen. Döhmman, *Datenschutzrecht*, 1st ed. 2019, Art. 32 recital 1 et seq.].

First of all, it should be noted that the most data protection-friendly systems are those that only collect data, that do not originally contain any reference to individuals, since in this respect there are also no risks to the informational self-determination of individuals. Therefore, an area analysis with LiDAR sensors is to be welcomed from a data protection point of view. The same applies to the survey using radar and sonar sensors.

However, since as shown, many systems cannot work without personal data, data protection within the meaning of Art. 25, 32 DSGVO must be considered comprehensively. Measures can be implemented at various points within the processing procedures (collection, storage, transmission, etc.) of the data. In the following, individual measures corresponding to the processing steps will be presented and explained by way of example:

- *Data collection:* In the context of data collection by the smart street lamps, UAVs, and UGVs, the data should be stored in a ring buffer⁷ containing a limited time window to ensure data-protection in normal operation but also enabling the reconstruction of events in case of an incident.
- *Data transmission:* Since a transmission of data to the control center is necessary for the presentation and further coordination of operations, the following measures can be taken to secure the transmission in terms of data security according to Art. 32 GDPR:
 - Minimizing the amount of data transmitted to the control center by applying edge computing where-ever possible. Sending abstract model-based data instead of raw data also reduces the requirements regarding the mobile connection quality.
 - Anonymize the data before transmission, e.g., by pixelating video streams using onboard computing for teleoperation using detected objects. By this operation, biometric data can be waived according to Art. 9 DSGVO as well.
- Apply *access restrictions* with appropriate technical protection (e.g., role concepts) in the control center to reduce the information displayed to each role to a minimum.

4 Conclusion & Outlook

In this paper we presented the legal and technical frame for a potential Smart Control Center in future smart cities as well as its potential benefits for crisis management. We elaborated possible, privacy compliant approaches for data collection, both in crisis scenarios as well as normal smart city operation. To comply with law, a smart city SCCs needs to be designed from the ground up with data protection by design and security of processing. Considering the data protection by design approach in the conceptualisation of the SCCs could lead to broad public acceptance in the use of the proposed infrastructure. Finally, an automated, GDPR compliant, data deletion mechanism should be employed.

⁷ A ring buffer is a storage structure that overwrites old data as new data is received.

In future work, the subject of data storage will be considered in more detail. Simulations will likely be the main method of emergency response training, however, real world data from previous emergency responses is still beneficial for the comprehensive training of rescue personnel. Especially, if and when the purpose merits, unanonymized storage should be discussed separately.

5 Acknowledgments

This work has been funded by the LOEWE initiative (Hesse, Germany) within the emergenCITY center.

References

- [AIM10] Atzori, L.; Iera, A.; Morabito, G.: The internet of things: A survey. *Computer networks* 54/15, pp. 2787–2805, 2010.
- [Al14] Albert, A.; Hollowell, M. R.; Kleiner, B.; Chen, A.; Golparvar-Fard, M.: Enhancing Construction Hazard Recognition with High-Fidelity Augmented Virtuality. *Journal of Construction Engineering and Management* 140/7, 2014.
- [Al18] Alfa, A. S.; Maharaj, B. T.; Ghazaleh, H. A.; Awoyemi, B.: The role of 5G and IoT in smart cities. In: *Handbook of smart cities*. Springer, pp. 31–54, 2018.
- [BE15] Brink, S.; Eckhardt, J.: Wann ist ein Datum ein personenbezogenes Datum. *Anwendungsbereich des Datenschutzrechts, ZD 205/*, p. 212, 2015.
- [BG17] Baumgartner, U.; Gausling, T.: Datenschutz durch Technikgestaltung und datenschutzfreundliche Voreinstellungen. Was Unternehmen jetzt nach der DS-GVO beachten müssen, *ZD 308/*, 2017.
- [CM03] Casper, J.; Murphy, R.: Human-robot interactions during the robot-assisted urban search and rescue response at the World Trade Center. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 33/3, pp. 367–385, 2003.
- [Ec21] Eckhardt, J.: Wann ist ein IoT-Gerät datenschutzrelevant? *Datenschutz und Datensicherheit-DuD* 45/2, pp. 107–113, 2021.
- [FSN20] Filho, J. A. W.; Stuerzlinger, W.; Nedel, L.: Evaluating an Immersive Space-Time Cube Geovisualization for Intuitive Trajectory Data Exploration. *IEEE Transactions on Visualization and Computer Graphics* 26/1, pp. 514–524, Jan. 2020.
- [Gu12] Gurevich, P.; Lanir, J.; Cohen, B.; Stone, R.: TeleAdvisor: A versatile augmented reality tool for remote assistance. In: *Conference on Human Factors in Computing Systems - Proceedings*. ACM Press, New York, New York, USA, pp. 619–622, 2012.
- [Ho11] Hornung, G.: Datenschutz durch Technik in Europa. Die Reform der Richtlinie als Chance für ein modernes Datenschutzrecht, *ZD/*, pp. 51–56, 2011.
- [HS21a] Hornung, G.; Stroscher, J.: i.E. Datenschutz in der Katastrophe, Teil 1: Anwendbarkeit, Systematik und Kompetenzfragen. *GSZ/*, to appear, 2021.
- [HS21b] Hornung, G.; Stroscher, J.: i.E. Datenschutz in der Katastrophe, Teil 2: Zulässigkeitstatbestände und Vorgaben für die Datenverarbeitung zur Katastrophenvorsorge und -bekämpfung. *GSZ/*, to appear, 2021.
- [HW19] Hornung, G.; Wagner, B.: Der schleichende Personenbezug. *Computer und Recht* 35/9, pp. 565–574, 2019.

- [Ja17] Jamei, E.; Mortimer, M.; Seyedmahmoudian, M.; Horan, B.; Stojcevski, A.: Investigating the Role of Virtual Reality in Planning for Sustainable Smart Cities. *Sustainability* 9/11, p. 2006, Nov. 2017.
- [KN19] Khan, M. N. H.; Neustaedter, C.: An Exploratory Study of the Use of Drones for Assisting Firefighters During Emergency Situations. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI '19, Association for Computing Machinery, Glasgow, Scotland UK, pp. 1–14, 2019.
- [La10] Laue, P.: *Vorgangsbearbeitungssysteme in der öffentlichen Verwaltung: rechtliche Rahmenbedingungen und Gestaltungsanforderungen*. kassel university press GmbH, 2010.
- [Mi12] Mizumoto, T.; Imazu, S.; Sun, W.; Shibata, N.; Yasumoto, K.: Emergency medical support system for visualizing locations and vital signs of patients in Mass Casualty Incident. In: *2012 IEEE Int. Conf. on Pervasive Computing and Communications Workshops*. IEEE, pp. 740–745, Mar. 2012.
- [Mo18] Moloney, J.; Spehar, B.; Globa, A.; Wang, R.: The affordance of virtual reality to enable the sensory representation of multi-dimensional data for immersive analytics: from experience to insight. *Journal of Big Data* 5/1, Dec. 2018.
- [Mü20] Mühlhäuser, M.; Meurisch, C.; Stein, M.; Daubert, J.; Von Willich, J.; Riemann, J.; Wang, L.: Street Lamps as a Platform. *Commun. ACM* 63/6, pp. 75–83, May 2020.
- [NKP15] Ntuen, C. A.; Kim, G.-M.; Park, E. H.: *Designing a fire incident response information support system.*, 2015.
- [Pr14] Prandi, F.; Soave, M.; Devigili, F.; Andreolli, M.; De Amicis, R.: Services Oriented Smart City Platform Based On 3d City Model Visualization. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences II-4/*, pp. 59–64, Apr. 2014.
- [Ri18] Rivano, H.; Augé-Blum, I.; Bechkit, W.; Boussetta, K.; Fiore, M.; Stanica, R.; Valois, F.: Wireless Access Networks for Smart Cities. In: *Smart Technologies: Breakthroughs in Research and Practice*. IGI Global, pp. 476–507, 2018.
- [Sa17] Satyanarayanan, M.: The emergence of edge computing. *Computer* 50/1, pp. 30–39, 2017.
- [Sc21a] Schindler, S.: i.E. Biometrische Videoüberwachung - Zur Zulässigkeit biometrischer Gesichtserkennung in Verbindung mit Videoüberwachung zur Bekämpfung von Straftaten./, to appear, 2021.
- [Sc21b] Schnaubelt, M.; Ullrich, T.; Torchalla, M.; Diegelmann, J.; Hoffmann, M.; von Stryk, O.: Entwicklung eines autonomiefokussierten hochmobilen Bodenrobotersystems für den Katastrophenschutz. In: *Digital-Fachtagung VDI-MECHATRONIK 2021*. Pp. 20–25, Mar. 2021.
- [SMS21] Schlosser, P. D.; Matthews, B.; Sanderson, P. M.: Head-Worn Displays for Healthcare and Industry Workers: A Review of Applications and Design. *Int. Journal of Human-Computer Studies*, p. 102628, Mar. 2021.
- [Su19] Surmann, H.; Worst, R.; Buschmann, T.; Leinweber, A.; Schmitz, A.; Senkowski, G.; Goddemeier, N.: Integration of UAVs in Urban Search and Rescue Missions. In: *2019 IEEE Int. Symposium on Safety, Security, and Rescue Robotics (SSRR)*. Pp. 203–209, 2019.
- [Wh21] Whitney, C. D.; Naval, T.; Quepons, E.; Singh, S.; Rick, S. R.; Irani, L.: HCI Tactics for Politics from Below: Meeting the Challenges of Smart Cities. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI '21, Association for Computing Machinery, Yokohama, Japan, 2021.
- [Zh18] Zhang, L.; Chen, S.; Dong, H.; El Saddik, A.: Visualizing Toronto City Data with HoloLens: Using Augmented Reality for a City Model. *IEEE Consumer Electronics Magazine* 7/3, pp. 73–80, May 2018.
- [ZLB10] Zhang, P.; Li, F.; Bhatt, N.: Next-generation monitoring, analysis, and control for the future smart control center. *IEEE Transactions on Smart Grid* 1/2, pp. 186–192, 2010.

Adversarial Examples zum Selbstdatenschutz? Der Fall biometrischer Gesichtserkennung im öffentlichen Raum

Niklas Bunzel,¹ Neele Zander,² Robert Landwirth,³ Ann-Katrin Riedel¹

Abstract: Videoüberwachung im öffentlichen Raum kann zur Aufklärung von Verbrechen, Identifikation von Personen und Erkennung von Waffen oder von Gruppenbildung genutzt werden. Durch die Unterstützung von künstlicher Intelligenz (KI) können große Datenmengen automatisch und effizient erhoben und ausgewertet werden. KI-gestützte Videoüberwachung im öffentlichen Raum, insbesondere die Gesichtserkennung, intensivieren Grundrechtseingriffe und erhöhen datenschutzrechtliche Risiken. Der im April 2021 veröffentlichte Gesetzesentwurf der EU-Kommission zur Regulierung von KI-Anwendungen erkennt zwar das Risiko der biometrischen Gesichtserkennung mittels Videoüberwachung im öffentlichen Raum, sieht jedoch von einem allumfassenden Verbot ab. Dies stellt ein Problem für die Privatheit der betroffenen Bürger*innen dar. Im Rahmen des Papers diskutieren wir Adversarial Examples als Möglichkeit des Selbstdatenschutzes darauf zu reagieren.

Keywords: Künstliche Intelligenz; Videoüberwachung; Regulierung von KI-Anwendungen; hochrisikante Anwendung; Selbstdatenschutz; Adversarial Examples

1 Einleitung

In vielen deutschen Großstädten gibt es kameraüberwachte öffentliche Plätze. Dabei wird zur Erhöhung des Sicherheitsgefühls, zur Kriminalitätsprävention und der Aufklärung von Straftaten auf Videoüberwachung gesetzt. Allerdings ist Videoüberwachung kein neues Phänomen. In Deutschland wird sie schon seit den 50er Jahren eingesetzt [Ka08]. Oft findet eine Live-Überwachung statt, bei der das Kamerabild direkt in eine Schaltzentrale geleitet und von Fachpersonal begutachtet wird. Weltweit gibt es bereits über 770 Millionen Kameras [LP19], alleine in Berlin sind es über 17.000 installierte Kameras [Bi20]. Da die Zahl an Kameras steigt, ist die Live-Überwachung vor allem in personeller Hinsicht eine Herausforderung. Um damit umzugehen wird vermehrt auf eine automatische Auswertung mittels Algorithmen gesetzt, beispielsweise zur Personenerkennung. Einzelne Personen können am Gesicht oder auch am Gang erkannt werden. Dabei werden für die Gesichtserfassung in aktuellen Forschungsarbeiten Erfolgsraten von bis zu 93% auf der Basis von Bilddaten erzielt, welche nicht unter Laborbedingungen entstanden sind [Li21]. Aufbauend auf der Gesichtserkennung können auch Personen gezählt werden, beispielsweise

¹ Fraunhofer SIT, Rheinstraße 75, 64295 Darmstadt, Deutschland vorname.nachname@sit.fraunhofer.de

² Johann Wolfgang Goethe-Universität Frankfurt, Theodor-W.-Adorno-Platz 1, 60323 Frankfurt am Main neelezander@gmx.de

³ Technische Universität Darmstadt, Karolinenplatz 5, 64289 Darmstadt robert.landwirth@tu-darmstadt.de

bei Demonstrationen. Beispiele für die Gesichtsdetektion, also das Erkennen von Gesichtern, sind in Abbildung 1 zu sehen. Dabei werden selbst teilweise verdeckte Gesichter, aber auch vollständig maskierte Gesichter oder auch Personen im Profil erkannt.

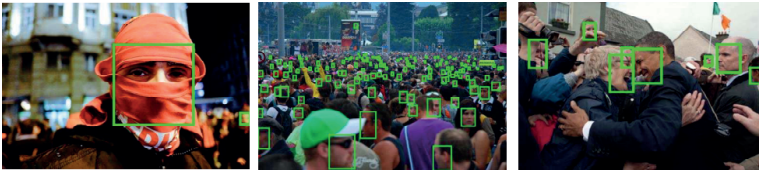


Abb. 1: Beispiele aus dem WIDER FACE Benchmark [Ya16]

Viele kommerzielle Gesichtserfassungssysteme, die vor der COVID-19 Pandemie entwickelt wurden, haben Schwierigkeiten beim Erfassen von Gesichtern mit Mund-Nasen-Schutz. Dies spiegelte sich in Fehlerraten von bis zu 50% bei maskierten Gesichtern wider. Systeme, die nach 2020 entwickelt oder angepasst wurden, haben selbst für maskierte Gesichter eine sehr hohe Erkennungsrate von bis zu 97% [NGH21].

Im April 2021 stellte die EU-Kommission ihren Gesetzesentwurf zur Regulierung von Künstlicher Intelligenz vor, bei der die Risiken von KI-Anwendungen für Bürger betrachtet werden. Hierbei liegt der Fokus eigener Aussage nach auf der konkreten Anwendung und Ausgestaltung des jeweilig eingesetzten KI-Systems, nicht von KI-Technologie im Allgemeinen. Je höher das zu erwartende Risiko durch die konkrete Nutzung und Ausgestaltung der KI ausfällt, desto strengeren Regelungen soll die KI-Anwendung unterworfen sein. Der Entwurf unterscheidet das Gefahrenpotential von KI-Anwendungen durch vier verschiedene Kategorien: minimales Risiko, limitiertes Risiko, hohes Risiko und unakzeptables Risiko. Ein unakzeptables Risiko führt zu einem gesetzlichen Verbot der Anwendung [EC21].

Die biometrische Gesichtserkennung⁴ in Echtzeit an öffentlichen Orten wird als unakzeptables Risiko eingestuft. So soll die generalisierte Überwachung der Bevölkerung und das *social scoring* zur Bewertung des sozialen Verhaltens, wie es in China Anwendung findet, nicht erlaubt sein [Re21]. Allerdings enthält der Entwurf diverse Ausnahmefälle, in welchen die biometrische Gesichtserkennung im öffentlichen Raum durch KI-Videotechnik als „hochriskante Anwendung“ (hohes Risiko) eingestuft wird und zu geografisch sowie zeitlich definierten Ausnahmefällen zugelassen sein soll. Hierzu zählen die Terrorabwehr, die Personenverfolgung von Straftätern besonders schwerer Verbrechen oder die Suche nach vermissten Kindern [Fa21].

⁴ Die biometrische Gesichtserkennung dient dazu, charakteristische Merkmale des Gesichts auszuwerten und die Zuordnung zu einer Person zu ermöglichen. Anhand eines digitalen Bildes erfolgt ein Abgleich der Charakteristika mit biometrischen Referenzen durch eine Analysesoftware. Hierbei werden hauptsächlich Charakteristika analysiert, die sich nicht durch die Mimik verändern lassen, wie etwa Seitenpartien des Mundes oder der Augenabstand [BSI08]. Die biometrischen Charakteristika werden in digitale Muster (Templates) übersetzt und miteinander verglichen. Die einfachste Form der Gesichtserkennung ist die Verifikation im „Eins-zu-eins-Vergleich“ zweier Templates. Zudem kann eine Identifikation anhand eines Abgleichs von Gesichtsbildern aus Datenbanken im „Eins-zu-n-Vergleich“ erfolgen. Die intelligente Videoüberwachung ermöglicht diesen Abgleich bei laufender Beobachtung mittels vorher festgelegter Algorithmen [Ra17].

Kritiker*innen mahnen an, dass im Ausnahmefall der Hochrisikoanwendung der Gesetzesentwurf lediglich Auflagen für die biometrische Gesichtserkennung und nicht ihr gänzlich Verbot vorsieht. Dies ermöglicht unter Umständen einen Massendatenabgleich des öffentlichen Raumes. Die Echtzeit-Identifizierung eines Menschen im öffentlichen Raum durch KI-basierte Gesichtserkennung bleibt daher, wenn auch auf Ausnahmen beschränkt, zunächst möglich [Fa21]. Allerdings ist Terrorabwehr (als Zwecksetzung) ein schwer bestimmbares Einsatzfeld, da „Terror“ eine unbestimmte Bedrohung zu einem unbestimmten Zeitpunkt markiert (siehe bpsw. [DR16, S.16], allgemein vgl. [Mo14, S.218]). Auch aus rechtlicher Sicht erfordert die Ausnahmeregelung zum Zweck der Terrorabwehr eine kritische Betrachtung, da hierbei zur Rechtfertigung eines Einsatzes dieser Technologie die Reduzierung auf den extremsten Fall erfolgt.

Videoüberwachung an öffentlichen Plätzen ist mit einer prinzipiellen Unmöglichkeit des Selbstschutzes verknüpft, wenn man weiterhin ‚normal‘ am öffentlichen Leben teilnehmen möchte. Der folgende Beitrag thematisiert vor diesem Hintergrund *Adversarial Examples*⁵ als neue Möglichkeit eines persönlich erwirkten Selbstschutzes. Zunächst erfolgt eine Betrachtung der rechtlichen Risiken durch automatisierte, biometrische Gesichtserkennung im öffentlichen Raum und ihrer Einstufung durch die EU-Kommission. Anschließend wird auf technische Möglichkeiten zur Provokation einer Fehlklassifikation durch die KI-Videotechnik eingegangen. Wir schließen den Aufsatz mit einem Resümee.

2 Rechtliche Betrachtung

Der Einsatz biometrischer Gesichtserkennungssoftwares mittels KI-Videotechnik im öffentlichen Raum wirft aus rechtlicher Perspektive die Frage auf, welchen Risiken Bürger*innen ausgesetzt werden und welche Möglichkeiten es gibt, diese zu minimieren. Videoüberwachung im öffentlichen Raum dient der Steigerung und dem Erhalt der öffentlichen Sicherheit. Eingesetzt wird sie zur Strafverfolgung und Gefahrenprävention in der Stadtüberwachung und dient unter anderem der Beweissicherung und Aufklärung. Dafür bietet die intelligente Videoüberwachung zusätzliches Unterstützungspotenzial durch Technologien, wie der automatisierten, biometrischen Gesichtserkennung. Sie erhöht die Auswertungsmöglichkeiten, indem eine automatisierte Erkennung und der automatisierte Abgleich mit Gesichtsbilddatenbanken ermöglicht werden [HS17]. Das Identifizieren einzelner Personen wird umfangreicher und schneller möglich und gewährleistet somit eine zuverlässigere Überwachung. Durch den Zusammenschluss von *smart cameras* zu komplexen Überwachungssystemen wird dabei die automatisierte Verfolgung von Personen flächenmäßig ausgedehnt. Videoüberwachung birgt Risiken, vor allem in Bezug auf Grundrechte und die personenbezogenen Daten der Betroffenen und steht somit im Spannungsverhältnis zu Privatheit im öffentlichen Raum. Schon die gesteigerte Möglichkeit der qualitativen und quantitativen Datenerhebung, sowie die Kombinationsmöglichkeit einzelner biometrischer

⁵ Ein Adversarial Example ist ein Bild welches, durch das Hinzufügen eines künstlich generierten Musters, in einem algorithmischen Klassifikator eine Fehlklassifikation hervorrufen soll.

Datensätze durch einen intelligenten Rückschluss auf ihren Informationsgehalt, öffnet die Tür für eine sich stetig ausweitende Datensammlung zur Aufrechterhaltung der öffentlichen Sicherheit. Hiermit einher geht ein Kontrollverlust, da menschliche Entscheidungsfindung durch letzten Endes intransparente Entscheidungsfindung einer Maschine ersetzt wird.

2.1 Chilling Effects

Der Rückgang überwachungsfreier Räume durch die biometrische Erfassung und Wiedererkennung mindert die Anonymität der Betroffenen (und damit auch den Schutz ihrer Privatheit). Verschärfend kommen hinzu, dass die Überwachung Einzelner in der Öffentlichkeit meist anlasslos erfolgt und eine unbestimmte Vielzahl von unverdächtigen Bürger*innen betrifft. Für Betroffene ist schwer nachvollziehbar, was und wie genau überwacht wird und welche Analysen auf die erhobenen Daten angewandt werden. So bestehen Unsicherheiten darüber, welche biometrischen Charakteristika von einem Analysetool erfasst werden und ob und wie man sich dagegen schützen kann. Ein damit verbundenes Gefühl des Überwachtwerdens und der Einschüchterung kann als Nebenfolge zu *chilling effects* führen, also zu Verhaltensanpassungen im öffentlichen Raum und zu eingeschränkter Grundrechtsausübung⁶. Vor allem betroffen sind dabei Freiheitsgrundrechte wie beispielsweise die Versammlungsfreiheit aus Art. 8 I GG.

2.2 False-Positive-Alarm

Dabei funktioniert auch die KI-gestützte Gesichtserkennung nicht vollkommen zuverlässig. Ein Problem ist die fälschliche Klassifikation eines Falls in der Zielklasse. So können beispielsweise unverdächtige Bürger*innen als gesuchte Straftäter*innen klassifiziert werden [Hi21]. Eine hohe *false-positive*-Rate kann zu hohen Kosten durch unbegründete Polizeieinsätze führen und tiefgreifende Eingriffe in die Rechte der Betroffenen zur Folge haben, indem diese ungerechtfertigter Weise in den Operator*innenfokus gelangen. Eine Fehlklassifikation kann dabei durch ungenaue und qualitativ minderwertige Trainingsdaten und Referenzbilder hervorgerufen werden, aber auch durch Scheinkorrelationen im Datensatz bedingt sein⁷. Hinzu kommt die Beeinflussung der Klassifikation durch äußere Faktoren. So spielt die Kameraposition, die Beleuchtung des überwachten Bereiches und der Lichteinfall auf das zu erkennende Gesicht eine Rolle [SS18]. Auch ein Fleck auf der Linse, Rauschen, oder sonstige Umwelteinflüsse können die Klassifikation beeinflussen und so zu Fehlklassifikationen führen.

⁶ BVerfG v. 11.03.2008, 1 BvR 2074/05, Rn. 78 - Verweis auf BVerfG 100, 313, 376, 392. Vergleichend wird in ähnlicher Argumentation aus sozialphilosophischer Sicht darauf hingewiesen, dass eine großflächig eingesetzte Videoüberwachung durch den Staat zu Geringschätzung eigener Privatheit (und damit Autonomie) führen kann [Rö01, S.39].

⁷ Bspw. weist der Datensatz eine Verbindung von Hautfarbe und Straffälligkeit auf, die empirisch repräsentativ ist, aber nicht in einem Kausalitätsverhältnis steht. Siehe unten.

2.3 Bias

Intelligente Gesichtserkennungssoftware birgt somit Risiken für den Gleichheitsgrundsatz oder das Diskriminierungsverbot aus Art.3 Abs.1 und Abs.3 GG, wenn die zugrundeliegenden Videodaten, aufgrund unausgewogener Trainingsdaten, fehlerhaft analysiert werden. Dies führt zu Verzerrungen der Analyseergebnisse. Hierdurch kann es zu Vergleichsgruppenbildung aufgrund sachlich nicht-gerechtfertigter Merkmale kommen und Ungleichbehandlung auslösen [De18]. So können beispielsweise männliche gegenüber weiblichen Gesichtszügen eher zu einer Klassifikation als "potentiell gefährlich" führen. Weiterhin können einzelne Hautfarben gegenüber anderen benachteiligt werden. In den USA kam es beispielsweise immer wieder zu Falscherkennungen wegen fehlerhaften KI-Fotoabgleichen [An20; Hi20; Hi21]. Durch die Evaluation von Gesichtserkennungssoftware konnte empirisch nachgewiesen werden, dass *People of Color* gegenüber anderen Personengruppen einer höheren Fehlerrate ausgesetzt und fälschlich als Straftäter*innen identifiziert wurden [St21]. Diese Fehlerraten führen zu rassebezogenen Ungleichbehandlungen durch die Gesichtserkennung. Zu solchen Verzerrungen, dem sogenannten Bias, kann es neben unausgewogener Trainingsdatensätze, durch den Trainingsprozess selbst oder durch Fehler im Analysemodell kommen.

2.4 Recht der informationellen Selbstbestimmung

Die öffentliche Videoüberwachung führt zu tiefgreifenden Beeinträchtigungen des Rechts auf informationelle Selbstbestimmung aus Art.2 Abs.1 iVm. Art.1 Abs.1 GG. Dieses gewährleistet jedem Einzelnen selbst zu entscheiden, wann und welche persönlichen Daten und Lebensumstände offenbart werden und gibt der einzelnen Person somit die Kontrolle über ihre Daten und deren Verarbeitung⁸. Kern des Schutzes des informationellen Selbstbestimmungsrechts ist der Schutz vor Datenmissbrauch, einer totalen Registrierung und anschließenden Katalogisierung der Betroffenen durch moderne Informationstechniken. Vom Schutz umfasst sind auch öffentlich zugängliche personenbezogene Daten, wie biometrische Gesichtsmerkmale [Rü21]. Charakteristisch für Videoüberwachung ist, dass zunächst meist mehr Daten erhoben werden als für die Erreichung des verfolgten Zweckes erforderlich ist. Zudem ist die Maßnahme schwer nachvollziehbar oder der Vorgang läuft verdeckt ab.

Intelligente Gesichtserkennungstechnologie als automatisierte Datenverarbeitungsform steigert die Eingriffsintensität der öffentlichen Videoüberwachung erheblich, sodass in diesen Fällen laut Bundesverfassungsgericht prinzipiell von einem Eingriff in das informationelle Selbstbestimmungsrecht bereits ab Datenerhebung ausgegangen wird⁹. Automatisierte

⁸ BVerfG v. 15.12.1983, 1 BvR 209/83, Rn.1.

⁹ BVerfG v. 18. 12. 2018, 1 BvR 142/15, Rn.43.

Datenverarbeitungsverfahren können enorme Mengen an Datensätzen erheben, verarbeiten und somit auch, teils unkontrolliert, analysieren¹⁰. Die Extrahierung biometrischer Gesichtsdaten kann dabei ein direktes Identifikationsmerkmal liefern [Co17]. Aus den biometrischen Mustern der Gesichtserkennung könnte das ungefähre Alter in Echtzeit bestimmt werden, woran sich eine Verknüpfung mit weiteren Datenbeständen anschließen kann [He14]. Die biometrische Gesichtserkennung ermöglicht es so, dass auch im öffentlichen Raum vermehrt sensible Daten erhoben und verarbeitet werden können, die die Betroffenen bei Kenntnis der stattfindenden Überwachung und Analyse nicht preisgegeben hätten. Die öffentliche Videoüberwachung mittels Gesichtserkennungssoftware ist daher als zusätzlicher Grundrechtseingriff einzuordnen.

2.5 Datenschutz

Die Vielzahl unverdächtiger Bürger*innen dürfte somit ein nachvollziehbares Interesse daran haben, den Schutz ihrer Privatheit zu wahren, Möglichkeiten der informationellen Selbstbestimmung zu erhalten und sich drohenden Ungleichbehandlungen durch Gesichtserkennungssoftware eigenständig zu entziehen.

Unter Selbstschutz ist die Gesamtheit aller Maßnahmen technischer und organisatorischer Art zu verstehen, die die einzelnen Bürger*innen zum Schutz ihrer Datenschutzgrundrechte ergreifen können [Ka14]. Sie helfen ihnen eine aktivere Position in der Gewährleistung ihrer informationellen Selbstbestimmung einnehmen zu können, unabhängig vom datenschutzrechtlich garantierten Schutz und seiner tatsächlichen Umsetzung durch Dritte.

Die vorgebrachten Risiken automatisierter biometrischer Gesichtserkennung sind auch aus datenschutzrechtlicher Perspektive zu betrachten, dadurch rückt die gegenwärtige Wichtigkeit von Selbstschutztechniken bei biometrischer Gesichtserkennung mit Blick auf den Gesetzesentwurf der EU-Kommission ins Zentrum. Videoüberwachung stellt einen datenschutzrelevanten Vorgang dar¹¹. Durch die Technik der intelligenten Videoüberwachung werden die Überwachungs- und Wahrnehmungsmöglichkeiten und die Möglichkeiten der Datenerhebung sowie der Verarbeitung ausgeweitet. Die Videoüberwachung erfolgt mit hochauflösenden Kameras. Die hohe Bildqualität, starke Zoomfunktion und Sensortechnologie ermöglichen dabei eine reliable Erkennung biometrischer Gesichtsdaten in vielen Einsatzkontexten. Ein datenschutzrelevanter Vorgang ist durch die bloße Erfassung biometrischer Gesichtsmarkmale anzunehmen, wenn die Erhebung die Daten verfügbar macht und somit die Basis für anschließende Verarbeitungsschritte bildet.

Eine Möglichkeit Selbstschutz auszuüben ist, videoüberwachte öffentliche Räume zu meiden oder das Gesicht zu verdecken. Dies ist im Alltag allerdings wenig praktikabel. Somit ist Videoüberwachung eine Herausforderung für den Selbstschutz.

¹⁰ BVerfG v. 18. 12. 2018, 1 BvR 142/15, Rn.37.

¹¹ BVerfG, v. 23. 02. 2007, 1 BvR 2368/06, Rn.38.

Dabei wird Selbstschutz vor allem dort wichtig, wo die Verarbeitung besonderer Kategorien personenbezogener Daten stattfindet und mangels Transparenz, Erklärbarkeit und Nachvollziehbarkeit die Durchsetzung der Betroffenenrechte, wie Löschungs- und Berichtigungsansprüche oder Auskunftsrechte¹² erschwert werden. Durch intelligente Videoüberwachung stellt sich nun die zusätzliche Frage, wann, neben der Beobachtung, auch eine Analyse stattfindet und welche Daten von dieser Analyse umfasst sind. Dies führt dazu, dass Betroffene nicht nur mit den Fragen „Wird gefilmt und wann wird gefilmt?“, „Welcher Bereich wird gefilmt?“, sondern auch mit der Frage nach „Welche Analysen werden auf dem Filmmaterial ausgeführt?“ konfrontiert sind.

Die Verarbeitung von sensiblen Daten zur eindeutigen Identifikation einer natürlichen Person ist nur unter der Berücksichtigung der erhöhten Zulässigkeitsvoraussetzungen des Art. 9 DSGVO möglich und steht unter besonderem Schutz. Unter „sensible Daten“ fallen nach Art. 9 I DSGVO auch biometrische Daten, also körperliche Merkmale, die unmittelbar personengebunden und schwer veränderbar sind. Die DSGVO nennt hierbei in Artikel 4 Nr.14 konkret Gesichtsbilder, die durch spezielle technische Verfahren die Identifizierung eines Menschen ermöglichen [Er21]. Zu diesen technischen Verfahren gehört auch der Einsatz von Gesichtserkennungssoftware [Er21]. Um das informationelle Selbstbestimmungsrecht zu gewährleisten, hat der europäische Gesetzgeber unter anderem durch Regelungen zu technisch-organisatorischen Maßnahmen ein Augenmerk auf die tatsächliche Umsetzung des Schutzes dieser Daten gelegt und das besondere Risiko datenschutzrelevanter Verarbeitungen hervorgehoben¹³. So können in Hard- und Software der *smart camera* datenschutzfreundliche Techniken durch Privacy-Filter zur Unkenntlichmachung einzelner Bildbereiche implementiert werden [Br18a]. Auch die geforderte Folgeabschätzung für risikoreiche Technologien kann datenbezogene Risiken sichtbar machen und einen sensiblen Umgang befördern¹⁴.

Im Fall Videoüberwachung sind Bürger*innen also stets auf gesetzliche Regelungen angewiesen und von deren tatsächlichen Umsetzung durch die Verantwortlichen abhängig. Hinzu kommt, dass klassische aktive Maßnahmen des Selbstschutzes, wie die Angabe eines Pseudonyms, die Nutzung von datenschutzfreundlicher Technik oder eine Verschlüsselung bei Gesichtserkennung durch Videoüberwachung von Bürger*innen nicht eingesetzt werden können. Auch Maßnahmen wie Wegdrehen oder Verdecken des Gesichts sind durch das derzeitige Niveau der Erkennung oft nicht mehr effektiv [CS17; LC21; Li21; NGH21]. Gerade der intensive Eingriff durch den unmittelbaren Personenbezug bei biometrischen Gesichtsbildern und die schwierige Veränderbarkeit körperlicher Charakteristika bedrohen dabei die Privatheit von Personen. Somit steht die Gesichtserkennung und -detektion als Technik im Spannungsverhältnis zwischen dem politischen Schutzziel der Steigerung des Sicherheitsgefühls und der objektive Sicherheit von Bürger*innen und der dadurch eröffneten Bedrohung von Freiheits- und Datenschutzrechten. Da Bürger*innen im Alltag wenige

¹² vgl. Art 15ff DSGVO.

¹³ vgl. Art. 25 I DSGVO.

¹⁴ vgl. Art. 35 DSGVO.

Möglichkeiten haben auf die Überwachungssituationen zu reagieren, kann aus datenschutzrechtlicher und grundgesetzlicher Perspektive attestiert werden, dass neue Möglichkeiten des Selbst Datenschutzes diskutiert werden sollten.

3 Technische Möglichkeiten zum Selbstschutz

Im Bereich der Videoüberwachung kann ein Ziel sein, Gesichter zu detektieren oder wiederzuerkennen. Algorithmen zur Gesichtsdetektion stellen fest, ob auf einem Bild ein Gesicht dargestellt ist und markieren es mit einer *Bounding Box* wenn möglich. Eine Gesichtserkennung ist das Zuordnen eines Gesichts zu einer Person, wie im eingangs besprochenen Beispiel. Angriffe auf die Gesichtsdetektion oder -identifikation könnten als Mittel zum Selbstschutz verwendet werden. Bei einem Angriff auf die Gesichtsdetektion wird verhindert, dass das neuronale Netz überhaupt ein Gesicht erfasst. Dann kann eine Person über eine Gesichtserkennung auch nicht mehr wiedererkannt und zugeordnet werden. Damit ist ein Angriff auf die Gesichtsdetektion auch implizit ein Angriff auf die Gesichtserkennung [BSI08].

3.1 Adversarial Examples

Eine Angriffsmöglichkeit bieten speziell generierte Bilder, sogenannte Adversarial Examples, welche eine Fehlklassifikation im Zielsystem provozieren sollen. Abbildung 2 stellt ein solches Adversarial Example dar. Dieses Beispiel demonstriert einen Angriff auf die Gesichtserkennung im digitalen Raum. Links ist ein Bild von John Howard zu sehen, welcher durch ein neuronales Netz korrekt klassifiziert wird. Diesem Foto wird ein speziell generiertes Rauschmuster hinzugefügt. Das Rauschmuster sehen wir in verstärkter Intensität in der Mitte. Das Ergebnis der Addition der beiden Bilder ist rechts zu sehen. Wir erkennen auf diesem Foto noch eindeutig John Howard, aber das KI-System ist sich sicher, dass das Foto Saddam Hussein zeigt. Adversarial Examples gibt es als *targeted attacks* – also als Angriffe die eine bestimmte Zielklasse provozieren sollen (hier: das Erkennen einer Person als „Sadam Hussein“) oder als *untargeted attacks*, deren Ziel lediglich ist, eine beliebige Fehlklassifikation herbeizuführen (Also die Person auf dem Bild spezifisch nicht als „John Howard“ zu erkennen).

Adversarial Examples existieren nicht nur im digitalen Raum, sondern können auch in der „realen Welt“ Anwendung finden. Beispielsweise haben Eykholt et al. in [Ey18] die Klassifikation von Verkehrsschildern durch gezieltes Anbringen von Aufklebern in die Irre geführt. Ein weiterer Angriff auf Bildklassifikation ist Brown et al.’s Adversarial Patch [Br18b]. Hierbei wird ein Sticker erzeugt, der die Aufmerksamkeit des neuronalen Netzes auf sich zieht und somit die korrekte Klassifikation des primären Objekts verhindert. Diese Art von Angriff kann in der physischen Welt auch bei der Gesichtsdetektion eingesetzt werden.

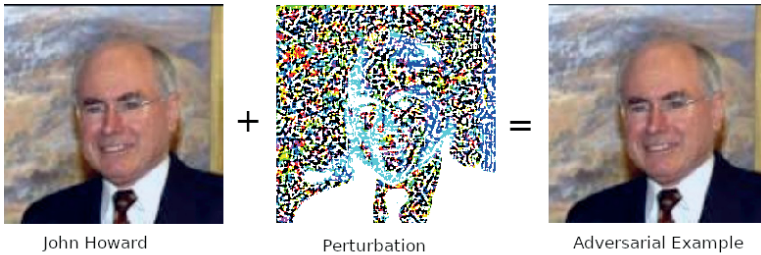


Abb. 2: Beispiel für einen Angriff durch Adversarial Examples im digitalen Raum. Neuronales Netz erkennt John Howard als Saddam Hussein.

3.2 Gesichtsdetektion

Ein Beispiel für eine Gesichtsdetektion ist in Abbildung 3 dargestellt. Die Sticker haben ein QR-Code ähnliches Muster und lassen das neuronale Netz MTCNN [Zh16] kein Gesicht mehr im Bild erkennen. Dieses neuronale Netz hat bisher sehr gute Ergebnisse in dem *WIDER FACE* Gesichtsdetektion-Benchmark [Ya16] erzielt. MTCNN ist aufgrund seiner Architektur robust gegenüber einfachen Angriffen, wie dem aus Abbildung 2 [Ka19].



Abb. 3: Forscher mit Angriffs-Stickern wird nicht von MTCNN detektiert [Ka19].

Weitere Angriffe auf die Gesichtsdetektion des bekannten MTCNN sind in [Ot20; Ts20] exploriert worden. In [Ts20] wurden diverse Platzierungen der Angriffs-Patches und der Abstand zur Kamera betrachtet. In [Ot20] wurden unter anderem künstlich generierte Brillengestelle mit einer Perturbation versehen, um die Erkennung durch das MTCNN zu verhindern.

Allerdings sind Angriffe auf die Gesichtsdetektion schwer generalisierbar. So werden

viele Angriffe in Interaktion mit einer bestimmten Detektionssoftware entwickelt und funktionieren auch nur für diese. Ein Beispiel dafür ist ein Versuch von McBrown [Ma19], in dem die Erkennung dadurch abgelenkt werden soll, dass die abstrahierten Features eines Gesichts (den *histogram of oriented gradients*-Features oder HOG-Features) auf eine Maske gedruckt werden, siehe Abbildung 4 Mitte. Dieser Angriff funktioniert nur für ein klassisches KI-System, welches mittels einer linearen *support vector machine* (SVM) auf eben diese HOG-Features trainiert wurde. Ein neuronales Netz oder eine SVM, die auf andere Features trainiert wurden, erkannten das Gesicht trotz der Maske, wie in Abbildung 5 zu sehen ist. Weiterhin haben wir durch digitale Reproduktion des Angriffs festgestellt, dass schon die Maske selbst, die Detektion des Gesichts verhindert, da dafür notwendige Kerneigenschaften des Gesichts verdeckt werden. Der Angriff kann dabei zum Erkennen eines „anderen Gesichts“ auf der Maske führen. Das hier besprochene Beispiel ist in Abbildungen 4 und 5 dargestellt.



Abb. 4: HOG-Klassifizierer. Links: Im Originalbild wird das Gesicht erfasst, Mitte: Muster auf der Maske wird als Gesicht erfasst, Rechts: Maske ohne Muster – es wird kein Gesicht erfasst.

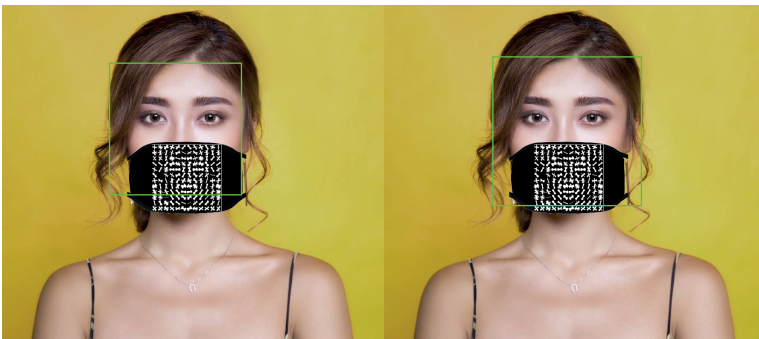


Abb. 5: Links „convolutional neural network“(CNN) und rechts Haar-Klassifizierer (SVM) erfassen das durch die Maske verdeckte Gesicht.

Eine ähnliche Art von Angriff ist das CV Dazzle-Konzept [Ha20]. Dabei werden Merkmale für die Gesichtserfassung wie Konturen und Symmetrie durch Frisuren und Make-Up unterbrochen oder verdeckt. CV Dazzle zielt dabei auf die Viola-Jones Gesichtserfassung und ist nicht zum Schutz gegen KI-Systeme die auf CNNs basieren geeignet [Ha20]; siehe Abbildung 6.

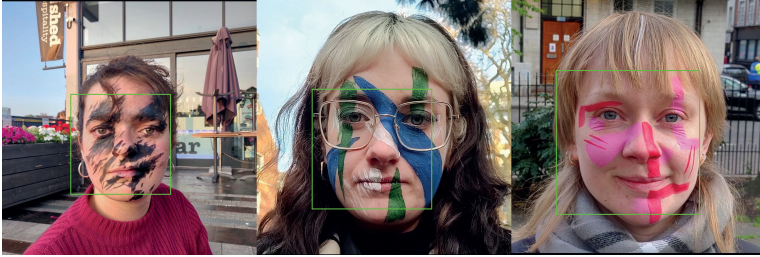


Abb. 6: Make-Up nach dem CV Dazzle-Konzept gegen Instagram- und Snapchat-Gesichtsdetektion mit CNN detektiert [Th21].

3.3 Gesichtserkennung

Ein Beispiel für einen Angriff auf die Gesichtserkennung findet sich in Sharfi et al. 2019 [Sh19]. Dort generieren die Autor*innen ein Brillengestell, das dazu verwendet werden kann, nicht identifiziert oder als eine bestimmte andere Person erkannt zu werden. Im letztgenannten Fall werden die relevanten Gesichtsmerkmale der Zielperson, als die man erkannt werden möchte, in einer für den Menschen nicht erkennbaren Form, auf die Brille übertragen. Dieser Prozess geschieht in Abhängigkeit der Gesichtsmerkmale der intendierten Träger*in der Brille. Dadurch wird die Forscherin in Abbildung 7 als Brad Pitt erkannt. Dieser Angriff wurde erfolgreich auf verschiedene neuronale Netze zur Gesichtserkennung angewendet [Sh19]. Mit einer solchen Brille könnte man beispielsweise automatisierte Passkontrollen, wie sie an Flughäfen zu finden sind, umgehen. Dabei wurden in der Evaluation Erfolgsraten von über 80% erreicht, wenn dieser Angriff als *untargeted attack* eingesetzt wird. Bei *targeted attacks* variieren die Erfolgchancen stark in Abhängigkeit des Trägers und der Zielperson.



Abb. 7: Forscherin mit der generierten Brille wird als Brad Pitt erkannt [Sh19].

3.4 Adversarial Examples abwehren

Allerdings stellen auch Adversarial Examples kein absolutes Problem für Detektions- und Erkennungsalgorithmen dar. Forscher*innen arbeiten an Maßnahmen, um die Robustheit von neuronalen Netzen zu verbessern. So werden bei *adversarial training* [Ma18] Adversarial Examples mit in den Trainingsprozess integriert. Dadurch werden neuronale Netze robust gegen Angriffe, mit denen sie trainiert wurden, jedoch nicht gegen neue Angriffe.

Preprocessing-Verfahren wie etwa JPEG-Kompression oder eine Glättung versuchen den Angriffs-Charakter des Bildes vor der Klassifikation zu entfernen. So können Angriffe wie in Abbildung 2 verhindert werden, wenn das Rauschmuster durch eine vorgestellte JPEG-Kompression entfernt wird ¹⁵. Eine weitere Abwehrmaßnahme kann bereits eine einfache Rotation des Eingangsbildes sein [Gu18]. Problematisch ist, dass bisher für jede *Preprocessing*-Maßnahme wieder ein neuer erfolgreicher Angriff gefunden werden konnte [ACW18; At18].

Letztlich gibt die es Möglichkeit die Struktur des neuronalen Netzes zu modifizieren, um seine Robustheit zu erhöhen, beispielsweise durch das Ändern seiner Aktivierungsfunktion [XZZ20]. Eine weitere Möglichkeit ist es die Adversarial Examples mittels eines Detektors zu entdecken und den Angriff zu umgehen [RKH19; XEQ18].

Resümierend können wir schließen, dass aufgrund des Blackboxcharakters der neuronalen Netze, die zur Klassifikation von Gesichtern eingesetzt werden, Angriffe nur in direkter Interaktion mit den verwendeten Algorithmen entwickelt werden können. Da Angriffe allerdings genau die Aspekte der Algorithmen ausnutzen, die auch zu einer legitimen Klassifikation führen, können auf der Gegenseite Schutzmaßnahmen ebenfalls schlecht generalisiert werden und setzen im Regelfall die Kenntnis spezifischer Angriffsbilder (adversarial training) oder zumindest Angriffsmuster (*Preprocessing*-Transformationen) voraus. So zeigt die momentane Forschungslandschaft das Wechselspiel von Angriff und Schutz von Gesichtsdetektionsalgorithmen als „Katz und Maus Spiel“: Angriffe können nur unter Kenntnis der Klassifikation entwickelt werden und Schutz vor einem Angriff nur unter Kenntnis des Angriffs.

4 Fazit

Der aktuelle Gesetzesentwurf der EU-Kommission macht deutlich, dass ein allumfassendes Verbot von intelligenten Gesichtserkennungstechnologien in der öffentlichen Videoüberwachung zur Abwendung von Massendatenabgleichen und allumfänglicher Überwachung,

¹⁵ Bei einfachen Angriffen wird ein hochfrequentes Rauschmuster erzeugt, welches den Klassifikator in die Irre führt und gleichzeitig für den Menschen kaum wahrnehmbar ist. Bei der JPEG-Kompression werden für den Menschen nicht sichtbare hochfrequente Anteile im Bild entfernt. Damit ist die JPEG-Kompression ein wirksames Mittel gegen einfache Angriffe wie in Abbildung 2 gezeigt.

wie es von Kritikern gefordert wird, nicht dem derzeitigen Willen des europäischen Gesetzgebers entspricht. Aus Sicht von Sicherheitsbehörden ist dies mit Blick auf das enorme Unterstützungspotenzial von automatisierter biometrischer Gesichtserkennung zur Personenverfolgung oder Terrorabwehr nachvollziehbar. Dadurch erfolgt regelmäßig die Begrenzung des Privatheitsschutzes zu übergeordneten Zwecken, wie Sicherheitsinteressen. Die intensivierten und oft anlasslosen Eingriffe in das informationelle Selbstbestimmungsrecht, drohende Einschüchterungseffekte durch Gesichtserkennung, Diskriminierungspotenzial durch Bias und die umfangreiche Verarbeitung und Analyse personenbezogener Daten, stellen, trotz eng beschränkter Einsatzszenarien für die Echtzeit-Gesichtserkennung durch die offen gebliebenen Möglichkeiten der Nutzung, ein enormes Risiko dar. Hinzu kommt, dass meist unverdächtige Bürger*innen zu den Betroffenen zählen. Dadurch wird die Gewährleistung des informationellen Selbstbestimmungsrechtes für Bürger*innen wichtig. Die Besonderheit des Szenarios der biometrischen Gesichtserkennung durch Videoüberwachungsmaßnahmen im öffentlichen Raum rückt die Frage nach Selbstschutzmaßnahmen in den Vordergrund. Datenschutzfreundliche Voreinstellungen können Betroffene nicht vornehmen. Hinzu kommt die Beständigkeit und hohe Individualität personenbezogener Gesichtsdaten. Eine Reaktionsmöglichkeit bieten nur komplexe Selbstschutztechniken – einfaches Wegdrehen, maskieren oder Schminken sind dabei kein ausreichendes Mittel mehr, um einen effektiven Selbstschutz zu garantieren, somit sollten Wege aufgezeigt werden, wie Bürger Selbstschutz effektiv verfolgen können. Eine gesetzgeberische Förderung von Selbstschutztechniken würde dem staatlichen Überwachungsinteresse öffentlicher Plätze entgegenstehen. Die Selbstvornahme durch adversarial examples kann Privatheitsgefahren entgegenwirken und den Schutz für den Einzelnen absoluter ausgestalten.

Im Rahmen des vorliegenden Papers haben wir unterschiedliche Techniken vorgestellt, die einen solchen Selbstschutz durch adversarial examples ermöglichen könnten. So kann das Drucken von Störmustern auf einer Maske oder das Herstellen einer Brille mit einem bestimmten Farbmuster potentiell sowohl die Gesichtserkennung, als auch die Gesichtsdetektion erschweren oder sogar unmöglich machen. Während Adversarial Examples unter Laborbedingungen Erfolgsraten bei der Verhinderung der Gesichtserkennung von über 80% versprechen [Sh19], ist die Frage offen, ob diese Schutztechniken generalisiert Einsatz finden können. Bisherige Angriffe auf Erkennungsalgorithmen werden stets in Interaktion mit einer bestimmten Software entwickelt — die Frage nach der Übertragbarkeit der Schutzfunktion auf weitere Erkennungsalgorithmen bleibt dabei offen. Auch wie gut solche Schutztechniken in der tatsächlichen Anwendung im öffentlichen Raum funktionieren, wurde noch nicht evaluiert. Dennoch denken wir, dass als Adversarial Examples gewonnene, tragbare Schutzmuster, zukünftig eine Möglichkeit zum Selbstschutz von Bürger*innen bieten können, die wissenschaftlich und gesellschaftlich nicht unberücksichtigt bleiben sollte.

Danksagung

Diese Forschungsarbeit wurde vom Bundesministerium für Bildung und Forschung (BMBF) und vom Hessischen Ministerium für Wissenschaft und Kunst (HMWK) im Rahmen ihrer gemeinsamen Förderung für das Nationale Forschungszentrum für angewandte Cybersicherheit ATHENE unterstützt.

Literatur

- [ACW18] Athalye, A.; Carlini, N.; Wagner, D.: Obfuscated Gradients Give a False Sense of Security: Circumventing Defenses to Adversarial Examples. In (Dy, J.; Krause, A., Hrsg.): Proceedings of the 35th International Conference on Machine Learning. Bd. 80. Proceedings of Machine Learning Research, PMLR, Stockholmsmässan, Stockholm Sweden, S. 274–283, Okt. 2018.
- [An20] Anderson, E.: Controversial Detroit facial recognition got him arrested for a crime he didn't commit, <https://eu.freep.com/story/news/local/michigan/detroit/2020/07/10/facial-recognition-detroit-michael-oliver-robert-williams/5392166002/>, Stand: 05.05.2021, 2020.
- [At18] Athalye, A.; Engstrom, L.; Ilyas, A.; Kwok, K.: Synthesizing Robust Adversarial Examples. In: Proceedings of the 35th International Conference on Machine Learning. 2018.
- [Bi20] Bischoff, P.: Surveillance camera statistics: which cities have the most CCTV cameras?, <https://www.comparitech.com/vpn-privacy/the-worlds-most-surveilled-cities/>, Stand: 10.05.2021, 2020.
- [Br18a] Bretthauer, S.: Intelligente Videoüberwachung - Fluch und Segen, <https://www.edvgt.de/wp-content/uploads/2018/08/Vortrag-Bretthauer.pdf>, Stand: 12.05.2021, 2018.
- [Br18b] Brown, T. B.; Mané, D.; Roy, A.; Abadi, M.; Gilmer, J.: Adversarial Patch, 2018, arXiv: 1712.09665 [cs.CV].
- [BSI08] Bundesamt für Sicherheit in der Informationstechnologie: Gesichtserkennung, https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Biometrie/Gesichtserkennung_pdf.pdf?__blob=publicationFile&v=1, Stand: 23.04.2021, 2008.
- [Co17] Conrad, C. S.: Künstliche Intelligenz - Die Risiken für den Datenschutz. Datenschutz und Datensicherheit 41/12, S. 740–744, 2017.
- [CS17] Chowdhury, S.; Sil, J.: FACE RECOGNITION from NON-FRONTAL IMAGES Using DEEP NEURAL NETWORK. In: 2017 Ninth International Conference on Advances in Pattern Recognition (ICAPR). S. 1–6, 2017.

- [De18] Desoi, M.: *Intelligente Videoüberwachung - Rechtliche Bewertung und rechts-gemäße Gestaltung*. Springer Vieweg, Wiesbaden, 2018.
- [DR16] Daase, C.; Rühlig, T.N.: *Der Wandel der Sicherheitskultur nach 9/11*. In (Fischer Susanne und Masala, C., Hrsg.): *Innere Sicherheit nach 9/11*. Springer, S. 13–33, 2016.
- [EC21] Commission, E.: *New rules for Artificial Intelligence – Questions and Answers*, https://ec.europa.eu/commission/presscorner/detail/en/QANDA_21_1683#1, Stand: 12.05.2021, 2021.
- [Er21] Ernst, S.: *DS-GVO Art. 4 Rn. 100*. In (Paal, B.; Pauly, D., Hrsg.): *Datenschutz-Grundverordnung Bundesdatenschutzgesetz: DS-GVO BDSG. 3.*, C.H.Beck, 2021.
- [Ey18] Eykholt, K.; Evtimov, I.; Fernandes, E.; Li, B.; Rahmati, A.; Xiao, C.; Prakash, A.; Kohno, T.; Song, D.: *Robust Physical-World Attacks on Deep Learning Visual Classification*. In: *Computer Vision and Pattern Recognition (CVPR)*. 2018.
- [Fa21] Fanta, A.: *EU verbietet automatisierte Gesichtserkennung an öffentlichen Orten – „mit wenigen Ausnahmen“*, <https://netzpolitik.org/2021/kuenstliche-intelligenz-eu-verbietet-automatisierte-gesichtserkennung-an-oeffentlichen-orten-mit-wenigen-ausnahmen/>, Stand: 12.05.2021, 2021.
- [Gu18] Guo, C.; Rana, M.; Cisse, M.; van der Maaten, L.: *Countering Adversarial Images using Input Transformations*. In: *International Conference on Learning Representations*. 2018.
- [Ha20] Harvey, A.: *Computer Vision Dazzle Camouflage*, <https://cvdazzle.com/>, Stand: 23.04.2021, 2020.
- [He14] Held, C.: *Intelligente Videoüberwachung - Verfassungsrechtliche Vorgaben für den polizeilichen Einsatz*. Duncker & Humblot GmbH, 2014.
- [Hi20] Hill, K.: *Wrongfully Accused by an Algorithm*, <https://www.nytimes.com/2020/06/24/technology/facial-recognition-arrest.html>, Stand: 05.05.2021, 2020.
- [Hi21] Hill, K.: *Another Arrest, and Jail Time, Due to a Bad Facial Recognition Match*, <https://www.nytimes.com/2020/12/29/technology/facial-recognition-misidentify-jail.html>, Stand: 03.05.2021, 2021.
- [HS17] Hornung, G.; Schindler, S.: *Das biometrische Auge der Polizei. Rechtsfragen des Einsatzes von Videoüberwachung mit biometrischer Gesichtserkennung*. *Zeitschrift für Datenschutz*, S. 203–209, 2017.
- [Ka08] Kammerer, D.: *Die Anfänge von Videoüberwachung in Deutschland*. *Kriminologisches Journal* 2008/4, S. 257–268, 2008.
- [Ka14] Karaboga, M.; Masur, P.K.; Matzner, T.; Mothes, C.; Nebel, M.; Ochs, C.; Schütz, P.; Fhom, H.S.: *White Paper Selbstschutz*, 2014.

- [Ka19] Kaziakhmedov, E.; Kireev, K.; Melnikov, G.; Pautov, M.; Petiushko, A.: Real-world Attack on MTCNN Face Detection System. 2019 International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON)/, S. 0422–0427, Okt. 2019.
- [LC21] Liu, Y.; Chen, J.: Multi-factor joint normalisation for face recognition in the wild. *IET Computer Vision* n/a/n/a, 2021, eprint: <https://ietresearch.onlinelibrary.wiley.com/doi/pdf/10.1049/cvi2.12025>.
- [Li21] Liu, Y.; Wang, F.; Sun, B.; Li, H.: MogFace: Rethinking Scale Augmentation on the Face Detector, 2021, arXiv: 2103.11139 [cs.CV].
- [LP19] Lin, L.; Purnell, N.: A World With a Billion Cameras Watching You Is Just Around the Corner, <https://www.wsj.com/articles/a-billion-surveillance-cameras-forecast-to-be-watching-within-two-years-11575565402>, Stand: 10.05.2021, 2019.
- [Ma18] Madry, A.; Makelov, A.; Schmidt, L.; Tsipras, D.; Vladu, A.: Towards Deep Learning Models Resistant to Adversarial Attacks. In: *International Conference on Learning Representations*. 2018.
- [Ma19] MacDonald, B.: Fooling Facial Detection with Fashion, <https://towardsdatascience.com/fooling-facial-detection-with-fashion-d668ed919eb>, Stand: 23.04.2021, 2019.
- [Mo14] Molotch, H.: *Against Security: How We Go Wrong at Airports, Subways, and Other Sites of Ambiguous Danger-Updated Edition*. Princeton University Press, 2014.
- [NGH21] Ngan, M.; Grother, P.; Hanaoka, K.: Ongoing Face Recognition Vendor Test (FRVT) Part 6B: Face recognition accuracy with face masks using post-COVID-19 algorithms, https://pages.nist.gov/frvt/reports/facemask/frvt_facemask_report.pdf, 2021.
- [Ot20] Ottaviano, A.: *MTCNN_adversarial_attack*, https://github.com/alex96295/MTCNN_adversarial_attack, Stand: 03.05.2021, 2020.
- [Ra17] Rath, C.: Was Gesichtserkennung kann. *Deutsche Richterzeitung* 2017/1, S. 8–10, 2017.
- [Re21] Rettinger, L.: The Human Rights Implications of China’s Social Credit System. *Journal of High Technology Law* 21/1, S. 7, 2021.
- [RKH19] Roth, K.; Kilcher, Y.; Hofmann, T.: The Odds are Odd: A Statistical Test for Detecting Adversarial Examples. In (Chaudhuri, K.; Salakhutdinov, R., Hrsg.): *Proceedings of the 36th International Conference on Machine Learning*. Bd. 97. *Proceedings of Machine Learning Research*, PMLR, S. 5498–5507, Sep. 2019.
- [Rö01] Rössler, B.: *Der Wert des Privaten*. Suhrkamp Verlag AG, 2001, ISBN: 3518291300.

- [Rü21] Rückert, C.: Mit künstlicher Intelligenz auf Verbrecherjagd - Einsatz von Gesichtserkennungstechnologie zur Aufklärung der „Kapitolverbrechen“, <https://verfassungsblog.de/ki-verbrecherjagd/>, Stand: 12.05.2021, 2021.
- [Sh19] Sharif, M.; Bhagavatula, S.; Bauer, L.; Reiter, M. K.: A General Framework for Adversarial Examples with Objectives. *ACM Trans. Priv. Secur.* 22/3, Juni 2019.
- [SS18] Schindler, S.; Salzmann, M.: Polizeiliche Gesichtserkennung in Deutschland. *ZD-Aktuell*/, S. 06344, 2018.
- [St21] Strong, J.: Klage in Amerika: Gesichtserkennung wird zur Bürgerrechtsfrage, https://www.heise.de/hintergrund/Klage-in-Amerika-Gesichtserkennung-wird-zur-Buergerrechtsfrage-6018527.html?wt_mc=sm.red.ho.6018527.posting.link.link&fbclid=IwAR20cwh4wS9yNFKmjXojeTnqPbWRcYDyPFUrT4JF74Qqu1RGb_8_SUCgy5E, Stand: 12.05.2021, 2021.
- [Th21] The, D. C.: The Dazzle Club, <https://www.instagram.com/thedazzleclub/>, Stand: 07.05.2021, 2021.
- [Ts20] Tsai, A.: Adversarial-attack-on-mtcnn, <https://github.com/yahi61006/adversarial-attack-on-mtcnn>, Stand: 03.05.2021, 2020.
- [XEQ18] Xu, W.; Evans, D.; Qi, Y.: Feature Squeezing: Detecting Adversarial Examples in Deep Neural Networks. *Proceedings 2018 Network and Distributed System Security Symposium*/, 2018.
- [XZZ20] Xiao, C.; Zhong, P.; Zheng, C.: Enhancing Adversarial Defense by k-Winners-Take-All. In: *International Conference on Learning Representations*. 2020.
- [Ya16] Yang, S.; Luo, P.; Loy, C.-C.; Tang, X.: WIDER FACE: A Face Detection Benchmark. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. S. 5525–5533, Juni 2016.
- [Zh16] Zhang, K.; Zhang, Z.; Li, Z.; Qiao, Y.: Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks. *IEEE Signal Processing Letters* 23/10, S. 1499–1503, 2016.

Technische Herausforderungen bei der Umsetzung von Uploadfiltern

Erkennen von 15 Sekunden Video oder Audio

Martin Steinebach ¹

Abstract: Uploadfilter sind derzeit Gegenstand der öffentlichen Diskussion. Dabei wird in erster Linie über ihren erwünschten Einsatz und ihre Wirkung gesprochen, aber kaum über ihre technischen Hintergründe. Wir beleuchten daher die Verfahren, die Uploadfilter ermöglichen und vergleichen rechtliche Vorgaben mit technischen Eigenschaften. Insbesondere betrachten wir die Vorgabe, dass Ausschnitte von Audio und Video von mehr als 15 Sekunden Länge erkannt werden sollen und zeigen auf, welchen Interpretationsspielraum diese Vorgabe mit sich bringt. Denn für die Performanz der Verfahren ist es von großer Bedeutung, ob die 15 Sekunden zusammenhängend sein müssen oder nicht.

Keywords: Urheberrecht; Robuste Hashverfahren

1 Motivation

Unter dem Begriff "Uploadfilter" werden Systeme verstanden, die digitale Werke beim Hochladen auf die Plattform eines Onlinedienstes untersuchen und basierend auf dem Untersuchungsergebnis eine Entscheidung über die nachfolgende Verfahrensweise mit diesem Inhalt treffen. Lösungen wie ContentID des Onlinedienstes YouTube sind sicher die bekanntesten Vertreter von Uploadfiltern, die bereits lange in der Praxis verwendet werden. Ihre Aufgabe ist es, urheberrechtlich geschütztes Material zu erkennen, welches auf einer Internetplattform angeboten werden soll. Dazu werden die zu filternden Materialien entweder im Vorhinein in einer geeigneten Form in einer Datenbank hinterlegt oder bei einer Urheberrechtsverletzung auf der Plattform im Nachhinein der Datenbank hinzugefügt. Solche Lösungen regieren nur auf inhaltlich identische Kopien, ein Konzept von Ähnlichkeit wird hier nicht verfolgt: Ein geschütztes Musikstück wird erkannt, nicht aber ähnlich klingende Werke oder andere Werke des gleichen Musikers.

Von der Aufgabe her ähnlich sind andere Filterverfahren wie beispielsweise das Erkennen und Blockieren von erotischen Inhalten in einem Firmennetz oder in sozialen Medien oder das Unterdrücken von Formulierungen mit Beschimpfungen in einem Forum.

¹ Fraunhofer SIT, MSF, Rheinstrasse 75, 64295 Darmstadt, Deutschland martin.steinebach@sit.fraunhofer.de

Technisch ähnlich zu Uploadfiltern sind Methoden zum Monitoring von Werbevideos, die automatisiert Fernsehkanäle überwachen, um zu prüfen, ob bezahlte Werbeslots auch tatsächlich mit den Werbeinhalten bestückt wurden oder Verfahren zum automatisierten Erfassen von Musikstücken, die bei einem Radiosender gespielt wurden, um Abrechnungen oder Hitlisten zu erstellen.

1.1 Rechtliche Vorgaben

In dieser Arbeit soll der öffentlich bekannte Stand der Technik mit den aktuellen rechtlichen Vorgaben [RS20] verglichen werden. Grundlage dafür ist § 10 UrhDaG-E. Hier werden die Grenzen für "geringfügige Nutzung geregelt"²: Erlaubt sind 15 Sekunden Audio oder Video, Bilder und Grafiken mit einer Dateigröße bis zu 125kB und Texte bis zu 160 Zeichen.

Dabei soll der Fokus auf Audio und Video liegen. Die erlaubte Spieldauer von bis zu 15 Sekunden bedeutet im Umkehrschluss, dass ein Uploadfilter erkennen soll, dass eine Verwendung von mehr als 15 Sekunden vorliegt. Nicht klar geregelt ist allerdings, ob die Spieldauer zusammenhängend sein muss oder auch beispielsweise vier unterschiedliche Ausschnitte aus dem gleichen Werk zu je vier Sekunden bereits erkannt werden müssen. Dies führt zu sehr unterschiedlichen Einschätzungen hinsichtlich der technischen Umsetzbarkeit abhängig von der Interpretation dieser Vorgabe. Diese wird im Abschnitt "Herausforderungen" ausführlich diskutiert.

Beachtet werden muss, dass nach § 7 UrhDaG (1) nur ein *bestmöglicher* Ansatz gefordert ist. Auch deshalb ist eine kritische Betrachtung des Machbaren wichtig, um "bestmöglich" einschätzen zu können.

1.2 Inhalte

Ein wichtiger Aspekt bei der Umsetzung von Uploadfiltern ist die Frage, wie viele Inhalte erkannt werden müssen bzw. wie groß die Menge ist, die potentiell erkannt werden soll. Prinzipiell ist die Obergrenze hier die Summe aller Inhalte, die bekannt sind und auf denen Rechte angemeldet werden können. Als Näherung soll hier ein Blick auf die Sammlungen geworfen werden, die von großen Anbietern verfügbar gemacht werden.

Im Audibereich bietet amazon music 40 Millionen Musikstücke an, Spotify 35 Millionen und Apple Music 45 Millionen³. Es sind also zumindest 45 Millionen Musikstücke potentiell zu erkennen, wenn Apple die Sammlungen von amazon und Spotify vollständig abdeckt.

Filme sind weniger präzise abzuschätzen, da hier die Anbieter oft das Angebot wechseln. Bei amazon prime video stehen 20.000 Filme zur Auswahl, bei netflix 2.800 und bei google play

² Gesetzentwurf der Bundesregierung: Entwurf eines Gesetzes zur Anpassung des Urheberrechts an die Erfordernisse des digitalen Binnenmarktes

³ <https://www.telegraph.co.uk/technology/0/best-music-streaming-services-apple-music-spotify-amazon-music>

13.000⁴. Hier ist allerdings die Summe aller Filme, die potentiell erkannt werden müssen, deutlich höher. Die Filmdatenbank IMDb kennt beispielsweise gut 7,5 Millionen Filme⁵. Hinzu kommen noch Serien, die ebenfalls eine große Menge von Inhalten produzieren, sowie beispielsweise Fernsehsendungen, Musikvideos und Dokumentationen.

2 Technik

Uploadfilter sind eine verkürzte Bezeichnung für komplexe Systeme, in denen das eigentliche Filtern nur eine von vielen Komponenten ist. Entsprechende Lösungen müssen ja neben der Analyse beispielsweise auch die Umsetzungen der Reaktionen auf das Ergebnis der Analyse bereitstellen. Das Wiedererkennen bzw. Re-Identifizieren von Inhalten geschieht entweder über sogenannte robuste Hashverfahren oder über Verfahren zur Extraktion von Merkmalen (“Features”). Entsprechende Verfahren erzeugen eine kompakte Darstellung des Inhalts und speichern diese in einer Datenbank. Zum Prüfen wird dann vom Medium mit derselben Methode ein weiterer Hash oder ein Merkmalsvektor errechnet und mit der Datenbank verglichen. Diese Methoden sind schnell und einfach zu berechnen, weiterhin weisen sie niedrige Fehlerraten auf. Entsprechende Verfahren sind in erster Linie für Multimedia Daten bekannt und werden für Bild, Video und Ton in zahlreichen Anwendungen eingesetzt. Aber auch für Texte sind robuste Hashverfahren bekannt. Ein Uploadfilter besteht aus einer Reihe von Komponenten. Abstrakt müssen dabei zumindest die folgenden vorhanden sein:

Referenzdatenbank: Anhand dieser Referenzdaten entscheidet der Uploadfilter, wie seine Reaktion ausfallen soll. Üblich ist hier eine Liste von Werken, in der die Fälle gespeichert sind, auf die der Uploadfilter mit weiteren Maßnahmen reagieren soll.

Entscheidungsverfahren Im Kern des Uploadfilters muss immer die Frage geklärt werden, ob ein eingehendes Datum (also beispielsweise ein Bild, ein Text oder ein Video) in einer Referenzdatenbank hinterlegt ist. Parametrierbar ist hierbei, wie tolerant das Verfahren gegen Abweichungen zwischen dem vorliegenden Datum und der Referenz ist.

Reaktionsmechanismus Auf Basis der Entscheidung des Einordnungsverfahrens muss der Uploadfilter in der Lage sein, eine abhängige Handlung durchzuführen. Diese Reaktion kann in Abhängigkeit vom Einsatzszenario wieder vielfältig ausfallen. Zur Vermeidung von Urheberrechtsstreitigkeiten wird ein Inhalt vielleicht blockiert oder dem Rechteinhaber und dem Uploader zur Klärung übergeben.

3 Erkennen von Inhalten

Es gibt unterschiedliche Ansätze, die sich mit der Einordnung von Inhalten befassen, also im Kontext eines Uploadfilters die Aufgabe übernehmen, über einen eingehenden Inhalt

⁴ <https://streaming-geraete.de/filme-netflix-amazon-disney>

⁵ https://de.wikipedia.org/wiki/Internet_Movie_Database

eine Entscheidung zu treffen. Dabei unterscheiden sich die Verfahren für verschiedene Inhalte deutlich. Eine Lösung für Bilder ist nicht ohne weiteres auf Texte übertragbar. Generell geht es um das Wiedererkennen von Inhalten. Es wird angenommen, dass ein Inhalt bereits bekannt ist, z.B. aus einer früheren Untersuchung. Er soll re-identifiziert und mit einigen in einer Datenbank gespeicherten Informationen darüber abgeglichen werden. Diese kann beispielsweise durch kryptographische (üblich bei Texten) oder robuste Hashes (üblich bei Medien wie Audio, Video und Bild) geschehen. Es muss betont werden, dass es bei dieser Aufgabe darum geht, die Inhalte selbst zu identifizieren und nicht darum, weitere Informationen daraus zu gewinnen. So zählt die Aufgabe, in einem Bild eine abgebildete Person zu erkennen, wie z.B. in [Go14] besprochen, nicht zur hier diskutierten Re-Identifikation.

In den folgenden Abschnitten gehen wir auf eine Reihe von Methoden ein, mit denen Wiedererkennen und Klassifizierung umgesetzt werden können. Diese sind eher beispielhaft zu sehen. Ziel ist es zu zeigen, wie unterschiedlich eine Einordnung erfolgen kann.

3.1 Kryptographische Hash-Funktionen

Kryptographische Hash-Funktionen (siehe z.B. [Ka96]) sind ein Primitiv der Sicherheitsprotokolle mit vielen Anwendungen, die in der IT-Sicherheit schon sehr lange bekannt sind [Da87]. Sie berechnen Hash-Werte fester Länge aus Informationen beliebiger Länge. Sie müssen eine Reihe von Anforderungen erfüllen, unter anderem Effizienz und Kollisionsresistenz und die Eigenschaften einer Einwegfunktion aufweisen. Diese Eigenschaften führen dazu, dass kryptographische Hash-Funktionen nur dazu geeignet sind, identische Kopien eines Werkes zu erkennen. Sobald auch nur minimale Änderungen an der Datei auftreten, die die Informationen speichert, ist der Hash ein vollständig anderer. Dazu genügt es, die Datei mit einem verlustbehafteten Kompressions-Algorithmus wie JPEG für Bilder oder h.264 für Videos zu speichern.

3.2 Robuste Hash-Funktionen

Es sind mehrere robuste oder wahrnehmungsbezogene Hashes für verschiedene Medientypen bekannt, die unterschiedliche Robustheitsgrade bieten. Da es zu viele Algorithmen gibt, um sie hier zu erwähnen, empfehlen wir Erhebungen wie die von Haouzia et al. [HN08] oder Neemila und Singh [NS14]. Es existieren auch Methoden für Audio [HKO01]- und Videostreams [OKH01] sowie für Textdaten [St13]. Auch kommerzielle Lösungen wie Shazam sind bekannt, die auf entsprechende Verfahren setzen⁶.

Robuste Hash-Funktionen extrahieren wahrnehmungsrelevante Merkmale aus Multimedia-Inhalten zu Identifikationszwecken. Sie müssen eine Reihe von Anforderungen erfüllen. Die wichtigsten sind:

⁶ <https://www.heise.de/ct/artikel/Wie-Shazam-Songs-erkennt-4192471.html>

- Unterscheidung: Wahrnehmbar unterschiedliche Stücke von Mediendaten sollen unterschiedliche Hash-Werte haben.
- Robustheit: Die robusten Hash-Werte sollen eine gewisse Wahrnehmungsinvarianz aufweisen, d.h. zwei Mediendaten, die für einen durchschnittlichen Zuschauer/Zuhörer hinsichtlich seiner Wahrnehmung ähnlich sind, sollen auch ähnlich sein.
- Sicherheit: Die Merkmale müssen Angriffe überstehen, die direkt auf die Merkmals-Extraktion und nachfolgende Verarbeitungsschritte abzielen. Ähnlich wie bei kryptographischen Hash-Funktionen müssen die robusten Hash-Werte gleichmäßig auf alle möglichen Mediendaten verteilt und paarweise statistisch unabhängig für zwei Mediendaten sein, die sich in der Wahrnehmung unterscheiden.

Die Zuverlässigkeit robuster Hashverfahren bei der Wiedererkennung ist hoch. So weist beispielsweise das Verfahren für Bilder aus [SLY12] in dem dort durchgeführten Test eine Falsch-Positiv-Rate von 0% und eine Falsch-Negativ-Rate von 0.2% auf. Ein Verfahren für Videodaten [LFS20] kommt bei einer ausgeglichenen Optimierung auf beide Raten auf eine Falsch-Positiv-Rate von 1% bei einer Falsch-Negativ-Rate von 5%. Bei dem Audioverfahren AudioID [AI01] wird eine Erkennungsleistung von über 95% auch nach starken Veränderungen angegeben, teilweise werden auch 100% Erkennung erreicht. Das liegt an den Merkmalen, die zur Berechnung der Hashbits verwendet werden, und an dem Umstand, dass die Hashes nicht identisch sein müssen, um zwei Werke als gleich anzusehen. Hier wird mit Distanzen gearbeitet, häufig mit der Hamming-Distanz, die misst, wie viele Bits zweier Hashes sich unterscheiden. Die Distanz, die erlaubt ist, um zwei Werke als gleich anzusehen, ist dann der Schwellwert des Verfahrens. Dieser wird in absoluten Werten (x Bits von n Bits) oder in Prozent ($x\%$ von n Bits) angegeben.

Um einen besseren Eindruck eines Hashverfahrens zu haben, stellen wir hier noch eine vereinfachte Variante des Hashes aus [AI01] vor. Der Hash wird gebildet, indem 16 Frequenzbänder aus einem Abschnitt des Audios gebildet und ihre Veränderung über 16 Zeitabschnitte betrachtet werden. Ein Hashbit wird gesetzt, wenn die Summe der Energie im Frequenzband F zum Zeitpunkt t und im drüber liegenden Frequenzband $F + 1$ zum Zeitpunkt $t + 1$ größer ist als die Summe der Energie im Frequenzband $F + 1$ zum Zeitpunkt t und im darunter liegenden Frequenzband F zum Zeitpunkt $t + 1$. So wird der 16×16 Hash aus Abbildung 2 (links) für die drei Sekunden Audio aus Abbildung 1 gebildet. Der Hash daneben ist aus dem gleichen Audio nach einem Hochpassfilter, der das Audio unter 80Hz abschneidet, gebildet. Er unterscheidet sich nur an sechs Stellen vom ersten Hash, ist also robust gegen die Filterung.

Um eine längere Passage zu hashen, werden mehrere Hashes in einer überlappenden Sequenz gebildet. Abbildung 3 zeigt eine solche Sequenz. Hier sind 50 Zeilen dargestellt, die zusammen 10 Sekunden Audio darstellen. Die einzelnen Zeilen stehen dabei jeweils für einen Ausschnitt von 0.2 Sekunden, die 256 Spalten entsprechend den Hashes der

Matrixdarstellung aus Abbildung 2, nur dass diese nun als Folge hintereinander aufgeführt werden und nicht mehr in der ursprünglichen 16x16 Matrix.

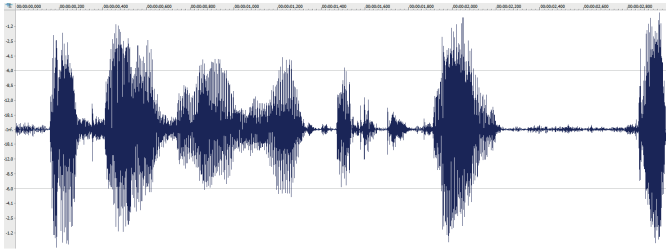


Abb. 1: Amplituden Plot von drei Sekunden Audio. Die X-Achse stellt den Zeitverlauf dar, die Y-Achse die Energie des Audiosignals.

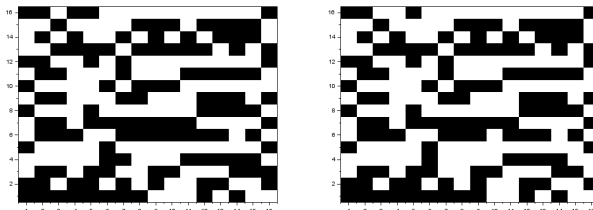


Abb. 2: Links: Hash des Originals, Rechts: Hash nach Hochpassfilter bei 80Hz. Die Hammingdistanz der Hashes liegt bei 6.

4 Herausforderung

Wie gut sind nun die rechtlichen Vorgaben technisch mit dem bekannten Stand der Technik umzusetzen? In diesem Abschnitt sollen die öffentlich verfügbaren Kennzahlen, die wissenschaftlichen Publikationen entstammen, mit den Anforderungen verglichen werden. Von in Unternehmen eingesetzten Verfahren sind entsprechende Daten nicht bekannt, allerdings werden Erfahrungen von Anwendern auf Foren und Blogs verbreitet. Hier wird davon berichtet⁷, dass ContentID auf Ausschnitte ab fünf Sekunden Spieldauer reagiert. Sollte das die Grenze sein, unter der eine Erkennung nicht möglich ist, so wären die erlaubten 15 Sekunden nur dann eine technisch umsetzbare Grenze, wenn die Nutzer die verwendeten Inhalte auf maximal drei Ausschnitte aufteilen würden. Würden 16 Sekunden eines Videos in vier Teile zu je 4 Sekunden geteilt werden, würde die Erkennung scheitern. Um die Herausforderungen zu diskutieren, die sich durch die Uploadfilter und die Vorgaben ergeben, verwenden wir ein abstraktes Modell eines robusten Hashverfahrens, wie weiter oben beschrieben. Wir fokussieren uns dabei der Einfachheit halber auf Audiohashes. Die Herausforderungen, auf die eingegangen werden soll, sind folgende:

⁷ <https://www.eff.org/de/wp/unfiltered-how-youtubes-content-id-discourages-fair-use-and-dictates-what-we-see-online>



Abb. 3: Plot von 50 Hashes aus einem Audio von 10 Sekunden Spieldauer

- Datenmengen: Wie groß werden die Datenbanken der Hashes, wenn eine gegebene Anzahl an Werken erkannt werden soll?
- Erkennungsleistung: Wie kann das Verhältnis zwischen zu erkennender Spieldauer abhängig von der erlaubten Fragmentierung zu der Wahrscheinlichkeit einer Erkennung durch das Verfahren abgeschätzt werden?
- Suchverfahren: Welche Möglichkeiten zu einer effizienten Suche nach übereinstimmenden Stellen können umgesetzt werden und welche Auswirkungen haben diese?

4.1 Datenmengen

Robuste Hashverfahren erfordern eine im Vergleich zu den bekannteren kryptographischen Hashverfahren einen größeren Speicherplatz und sind linear abhängig von der Spieldauer von Audio und Video. Das führt bei umfangreichen Sammlungen zu großen Datenbanken. Verwendet man den bekannten robusten Audiohash von Haitzma et al. für alle Musikstücke, die Apple Music anbietet, so gilt Folgendes: In ihrer Veröffentlichung [HKO01] beschreiben die Autoren ein Verfahren, welches in Schritten von 12,5 Millisekunden einen Hash von 32 Bit erzeugt. Pro Minute entspricht dies 4800 Hashes. Nimmt man eine Spieldauer von durchschnittlich 3 Minuten pro Musikstück an, so ergeben sich bei der genannten Sammlung 135 Millionen Minuten Spieldauer. Die ergibt 648 Milliarden Hashes zu 32 Bit, was 2,35 Terabyte Speicherplatz entspricht. Auch andere Verfahren im Audiobereich, die in der Literatur benannt werden, führen zu vergleichbaren Datenbankgrößen. In Tabelle 1 sind die Angaben von Özer et al. [Oz05] und Allamanche et al. [Al01] eingearbeitet.

Tab. 1: Beispiele für die Datenbankgrößen verschiedener Verfahren bei 45 Millionen Musikstücken.

	Hash/s	Bit per Hash	DBS (TB)
Haitsma	80	32	2.00497352
Özer	78	16	0.97742459
Allamanche	43	128	4.31069306

Die Menge an Stücken hat auch eine direkte Auswirkung auf die Erkennungsleistung, die in der Literatur nur selten ausführlich diskutiert wird. Eine Ausnahme ist hier [A101]. Dort wird beschrieben, dass eine Erweiterung der Referenzdatenbank von 1.000 auf 15.000 Titel zur Folge hatte, dass die Anzahl der Merkmale, die als Hash verwendet wurden, viermal so groß sein musste, um eine vergleichbare Erkennung zu erreichen. Ohne diese Erweiterung wäre die Verwechslungsgefahr durch den geringen Merkmalsraum zu groß. Im nächsten Schritt⁸ wurde die Referenzdatenbank von 15.000 auf 90.000 Titel erweitert, ohne die Merkmale weiter anzupassen. Hier bleibt die Erkennungsleistung vergleichbar. Nicht abgeleitet werden kann allerdings, wie sich eine Datenbank von 45 Millionen Stücken auf die Erkennung oder die notwendigen Merkmale auswirkt. Hier kann eine Häufung von Fehlerkennungen auftreten, die eventuell weitere Verbesserung am Verfahren erfordert, was wiederum zu größeren Datenbanken führen kann. Das Verfahren von Haitsma et al. gibt eine Datenbank von 10.000 Stücken an [HK02]. Auch hier ist nicht abschätzbar, wie sich eine deutlich größere Menge von Musikstücken auf die Erkennung auswirkt.

4.2 Erkennungsleistung

Die Frage, wie groß die Herausforderung ist, eine Nutzung eines fremden Inhalts von über 15 Sekunden Spieldauer zu erkennen, ist abhängig davon, wie diese verwendet werden. Liegen beispielsweise die 15 Sekunden am Stück vor, so ist eine Erkennung möglich, wenn man die Angaben in der Literatur betrachtet: So wird in [A101] von Testsequenzen von 20 Sekunden gesprochen. In [HK02] werden nur "kurze Ausschnitte" erwähnt. Das Verfahren selbst basiert auf Hashes von 3 Sekunden Länge, es kann aber nicht direkt gefolgert werden, dass dies die Länge der Ausschnitte war, da zur Erkennung eventuell mehrere Hashes notwendig sind. In diesem Kontext sind sowohl die falsch-negativen, also nicht erkannte Werke, als auch die falsch-positiven, also Werke, die angezeigt werden, obwohl sie nicht verwendet wurden, von Bedeutung. In der Literatur wird oft nur von der Erkennungsleistung im Sinne der richtig-positiven gesprochen, was dementsprechend erlaubt, auf die falsch-negativen zu schließen.

⁸ http://virtualgoods.tu-ilmeneau.de/2003/Robust_Audio.ppt

4.2.1 Falsch-Positive

Das führt aber auch dazu, dass ein wichtiges Problem im Betrieb eines Uploadfilters schwer abschätzbar ist: Wie oft kommt es vor, dass ein Uploadfilter auf ein hochgeladenes Medium fälschlicher Weise reagiert? Theoretisch müsste diese Wahrscheinlichkeit verschwindend gering sein. Dies wird auch in [HK02] entsprechend dargestellt, wo eine falsch-positiv Rate in der Größenordnung von 10^{-20} errechnet wird. Nehmen wir die Verfahren aus Tabelle 1 und gehen von nur 3 Sekunden Audiolänge aus. Das Verfahren von Haitzma würde dann $3 * 60 * 32 = 5760$ Bit verwenden. Das bedeutet, wir haben 2^{5760} mögliche Hashsequenzen. Auch wenn nur 75% der Bits übereinstimmen müssen, was hier als Erkennungsgrenze erwähnt wird, so ist die Anzahl der möglichen Kombinationen noch immer extrem hoch. Bei kryptographischen Hashes bewegt man sich in deutlich niedrigeren Bereichen, gebräuchlich sind 160 bis 512 Bits. Trotzdem treten bei kryptographischen Hashfunktionen nur sehr selten Kollisionen auf, während bei den robusten Hashverfahren immer wieder falsch-positive entdeckt werden. Ein Bildhashverfahren auf Basis von Helligkeitsblöcken z.B. hat durchaus falsch-positiv-Raten im Prozentbereich bei einer Größe von 256 Bit und muss für eine bessere Leistung besonders optimiert werden [SLY12].

Der Grund dafür ist ein grundsätzlicher: Wie bereits oben beschrieben, arbeiten robuste Hashverfahren anders wie kryptographische. Da sie leichte Änderungen akzeptieren, führen sehr ähnliche Eingaben auch zu ähnlichen Hashes. Bei kryptographischen Hashfunktionen ist dies anders, hier führen schon kleine Veränderung zu maximalen Unterschieden im Hash. Das führt dazu, dass robuste Hashverfahren zwar in der Theorie durch ihre Größe auf einen großen Zahlenraum abbilden, in der Praxis dieser Zahlenraum aber nicht ausgenutzt wird, da die Eingaben nicht zufällig generiert sind, sondern Werke darstellen, die sich oft in gewissen Maßen ähnlich sind. Ein Musikgenre wird unter anderem durch Ähnlichkeiten in Rhythmus, Geschwindigkeit und Tonalität definiert. In Genres wie HipHop mit dem bekannten Amen-Break oder in House mit dem Grundrhythmus von vier Schlägen mit dem Roland 909 Drumcomputer treten solche Ähnlichkeiten ganz deutlich hervor. Dadurch kommt es zu häufigeren Kollisionen zwischen zwei Hashes. Wie hoch diese Wahrscheinlichkeit genau ist, kann nur durch Evaluierungen festgestellt werden. Dass diese Fragestellung noch immer aktuell ist, zeigt beispielsweise [SBL20], wo falsch-positiv diskutiert werden, allerdings keine falsch-negative.

Die Wahrscheinlichkeit, dass sich kurze Passagen von beispielsweise einer Sekunde Spieldauer in zwei Stücken ähneln, ist deutlich größer als bei 15 Sekunden. Schon durch wenige Takte, in denen ein Rhythmus ohne Begleitung spielt, kann hier ein Alarm ausgelöst werden. Solche Stellen treten in manchen Genres häufig auf. Das gleiche gilt aber auch für klassische Musik: Für einen Hash kann ein lang gehaltener Ton einer Tonart schnell zu Verwechslungen führen. Das bedeutet, dass eine zu erkennende Stückelung in kurze Passagen durch einen Uploadfilter zu vermehrten falsch-positiven Treffern führen kann.

4.2.2 Falsch-Negative

Da die primäre Aufgabe eines Uploadfilters liegt darin, einmal in der Datenbank hinterlegte Werke wiederzuerkennen. Liegen Teile eines in der Datenbank bekannten Werkes in einem Upload vor, sollte er diese also erkennen. Gelingt ihm dies nicht, liegt eine falsch-negative Entscheidung vor. In der Praxis können Algorithmen Werke desto besser erkennen, je mehr zusammenhängendes Material ihnen zur Prüfung vorliegt. Das liegt daran, dass nicht jede Sekunde eines Werkes gleich gut erkannt werden kann. Teile sind vielleicht durch eine Nachbearbeitung besonders gestört, andere haben schon grundlegend die Eigenschaft, nur schlecht für den Hash geeignet zu sein: Der Algorithmus aus [HK02] basiert beispielsweise darauf, Änderungen in benachbarten Frequenzen und Zeitabschnitten abzubilden. Liegen aber keine Änderungen vor, ist die Energie in den betreffenden Frequenzen über längere Zeit statisch, so entsteht ein schwacher Hashwert, der leicht gestört werden kann. Solche Stellen sind üblicher Weise nur kurz und führen dann nicht zu Problemen, wenn ausreichend besser geeignete Stellen zur Verfügung stehen.

Unter anderem deshalb ist es für einen robusten Hash deutlich einfacher, eine robuste Wiedererkennung zu erreichen, wenn 15 Sekunden Material am Stück vorliegen, als wenn der Hash mehrere Ausschnitte von wenigen Sekunden erkennen soll. Würden beispielsweise 20 Sekunden verwendet, die aber in 10 Stücke zu je zwei Sekunden aufgeteilt sind, und sind dann drei Stücke für eine Erkennung ungeeignet, dann werden nur noch 14 Sekunden erkannt und der Uploadfilter gibt eine falsch-negative Entscheidung aus. Eine vergleichsweise umfangreiche Betrachtung der Fehlerraten wird in [Oz05] erstellt. Hier liegen die Erkennungsraten bei ca. 99%, die Fehlalarme bei ca. 1%. Relativiert werden muss die Evaluierung leider durch die geringe Anzahl von nur 1550 Teststücken.

Auch wenn der Fokus der Betrachtungen hier auf Audio liegt, gibt es die gleichen Überlegungen und Eigenschaften auch bei Hashes für Videodaten. So werden in [LM12] Analysen der Abhängigkeit für falsch-positive und falsch-negative für mehrere Hashverfahren durchgeführt. Selbst bei schwachen Veränderungen gilt für alle Verfahren: Eine niedrige Wahrscheinlichkeit, ein Video nicht zu erkennen, wird immer mit einer hohen Chance erkaufte, einen Fehlalarm auszulösen. So liegt bei einer falsch-negativen Einschätzung von ca. 10% die Fehlalarmrate für den LRTA (low-rank tensor approximations)- Algorithmus schon bei 20%.

4.3 Suchverfahren

Ein robuster Hash erfordert andere Suchstrategien wie ein kryptographischer Hash oder auch allgemein ein Datenbankeintrag. Da der robuste Hash nur ähnlich zu einem anderen sein muss und diese Ähnlichkeit durch die Hamming Distanz berechnet wird, sind bei der Suche zahlreiche Operationen notwendig. In der Praxis wird der Hash des hochgeladenen Datums errechnet und dann mit allen Hashes in der Datenbank verglichen. Dieser Vergleich

zählt die Bit-Positionen, an denen Such-Hash und Datenbankeintrag nicht übereinstimmen. Die Anzahl dieser Fehlpositionen ist die Hamming Distanz. Bei der großen Menge von Hashes wie in Tabelle 1 dargestellt, ist eine Suche rechenaufwändig. Dementsprechend müssen Optimierungen erfolgen, wie bereits in [AI01] dargestellt. Hier werden besonders prägnante Passagen der Such-Hashes identifiziert und zuerst nur nach diesen gesucht, was nur einen Verlust bei der Erkennung im Bereich eines Prozentes mit sich bringt. Auch in [WSY14] werden Methoden zum schnellen Suchen vorgestellt und auch hier ist der Verlust an erkannten Bildern nur gering im Vergleich zum Gewinn an Geschwindigkeit.

Wie sich die zu erwartenden großen Datenbanken auf die Leistung der Verfahren auswirken, lässt sich schwer abschätzen. Offensichtlich ist allerdings, dass ein Kompromiss zwischen Erkennungsleistung und Performanz notwendig ist. Ein praktikabler Uploadfilter wird bei der Suche eine Optimierung einsetzen müssen, um mit vertretbarem Aufwand Werke zu suchen. Diese werden deutlich stärker ausfallen müssen als in der Literatur diskutiert wird, da die Datenbanken in der Praxis um mehrere Größenordnungen umfangreicher ausfallen als die bei den Evaluierungen eingesetzten.

5 Diskussion

Der vorhergehende Abschnitt zeigt, dass eine Umsetzung eines Uploadfilters eine Reihe signifikanter Herausforderungen mit sich bringt, die teilweise wissenschaftlich nur schwer abgeschätzt werden können. Dies gilt vor allem für die Frage, wie die Erkennung von mehr als 15 Sekunden Material gehandhabt werden soll. Sind die 15 Sekunden am Stück zu erkennen, kann dies mit dem Stand der Technik gut abgebildet werden, wenn robuste Hashverfahren oder Inhaltsmerkmale eingesetzt. Die Fehlerraten sollten hier akzeptabel sein. Kryptographische Hashverfahren sind für den Einsatz in Uploadfilter bei Medien nicht akzeptabel. Mit ihnen könnten zwar die effizienten Ansätze umgesetzt werden, sowohl was Berechnungs- als auch Suchaufwand betrifft, allerdings ist bei ihnen ein Umgehen trivial.

Bei einer Verteilung der 15 Sekunden auf mehrere kürzere Stücke ist mit einem signifikanten Anstieg an Fehlern zu rechnen. Hier muss dann entschieden werden, ob fehlerhafte Ablehnungen oder Fehlalarme die größere Problematik mit sich bringen. Darauf basierend müssen die zur Erkennung eingesetzten Algorithmen optimiert werden.

Aber auch die schiere Menge an Daten, die durchsucht werden muss, wird in der Praxis eine Herausforderung sein. Dass dies technisch möglich ist, zeigen Lösungen wie ContentID von Google. Dahinter steht allerdings ein Aufwand, der für Plattformbetreiber erheblich sein kann. Dies gilt insbesondere, wenn die Rechteinhaber all ihre Werke in die Referenzdatenbanken des Uploadfilters einstellen können, um deren Verwendung erkennen zu können. Um die Uploadfilter in die Praxis einzuführen empfiehlt es sich, die Datenbanken schrittweise zu befüllen und hier mit Grenzen zu arbeiten, die beispielsweise auf der Verbreitung der Werke basieren, wenn davon ausgegangen werden kann, dass bekannte und aktuelle Werke häufiger verwendet werden. So kann die Infrastruktur kontinuierlich wachsen. Auch eine Regelung,

die die 15 Sekunden anfangs nicht segmentiert und im Laufe der Zeit die Erkennung immer kürzere Segmente fordert, würde einer realistischen Entwicklung der Technologie entgegenkommen.

Bei der Suche nach geeigneten Technologien werden auch immer wieder Anbieter der Eingangs erwähnten verwandten Anwendungen wie Werbemonitoring oder das Erkennen von Musikstücken im Radio. Diese Technologien haben die gleichen Wurzeln, allerdings eine deutlich abweichende Ausgangslage. Prüft beispielsweise ein Unternehmen, ob seine Werbespots im Fernsehen gezeigt werden, dann ist die Referenzliste nur sehr kurz. Wahrscheinlich hat das System höchstens eine zweistellige Anzahl von Inhalten, die es erkennen muss. Gleichzeitig ist ein Werbespot meist 30 Sekunden lang und nicht unterbrochen oder von anderen Inhalten überlagert. Das bedeutet, dass eine in diesem Umfeld erfolgreiche Technologie nicht zwingend auch für einen Uploadfilter geeignet sein muss.

Förderhinweis

Diese Forschungsarbeit wurde vom Bundesministerium für Bildung und Forschung (BMBF) und vom Hessischen Ministerium für Wissenschaft und Kunst (HMWK) im Rahmen ihrer gemeinsamen Förderung für das Nationale Forschungszentrum für angewandte Cybersicherheit ATHENE unterstützt.

Literaturverzeichnis

- [Al01] Allamanche, Eric; Herre, Juergen; Hellmuth, Oliver; Froeba, Bernhard; Kastner, Throsten; Cremer, Markus: Content-based Identification of Audio Material Using MPEG-7 Low Level Description. In: ISMIR. 2001.
- [Da87] Damgård, Ivan Bjerre: Collision free hash functions and public key signature schemes. In: Workshop on the Theory and Application of of Cryptographic Techniques. Springer, S. 203–216, 1987.
- [Go14] Gong, Shaogang; Cristani, Marco; Loy, Chen Change; Hospedales, Timothy M: The re-identification challenge. In: Person re-identification, S. 1–20. Springer, 2014.
- [HK02] Haitsma, J.; Kalker, T.: A Highly Robust Audio Fingerprinting System. In: ISMIR. 2002.
- [HKO01] Haitsma, Jaap; Kalker, Ton; Oostveen, Job: Robust audio hashing for content identification. In: International Workshop on Content-Based Multimedia Indexing. 2001.
- [HN08] Haouzia, Adil; Noumeir, Rita: Methods for image authentication: a survey. *Multimedia tools and applications*, 39(1):1–46, 2008.
- [Ka96] Katz, Jonathan; Menezes, Alfred J; Van Oorschot, Paul C; Vanstone, Scott A: *Handbook of applied cryptography*. CRC press, 1996.
- [LFS20] Liu, Huajian; Fach, Sebastian; Steinebach, Martin: Motion vector based robust video hash. *Electronic Imaging*, 2020(4):218–1, 2020.

-
- [LM12] Li, Mu; Monga, Vishal: Robust video hashing via multilinear subspace projections. *IEEE transactions on image processing*, 21(10):4397–4409, 2012.
- [NS14] Neelima, Arambam; Singh, Kh Manglem: A short survey on perceptual hash function. *ADBU Journal of Engineering technology*, 1, 2014.
- [OKH01] Oostveen, Job C; Kalker, Ton; Haitsma, Jaap: Visual hashing of digital video: applications and techniques. In: *Applications of digital image processing XXIV*. Jgg. 4472. International Society for Optics and Photonics, S. 121–131, 2001.
- [Oz05] Ozer, Hamza; Sankur, Bulent; Memon, Nasir; Anarim, Emin: Perceptual audio hashing functions. *EURASIP Journal on Advances in Signal Processing*, 2005(12):1–14, 2005.
- [RS20] Raue, Benjamin; Steinebach, Martin: Uploadfilter - Funktionsweisen, Einsatzmöglichkeiten und Parametrisierung. *Zeitschrift für Urheber- und Medienrecht*, 64(5):355–364, 2020.
- [SBL20] Son, Heui-Su; Byun, Sung-Woo; Lee, Seok-Pil: A Robust Audio Fingerprinting Using a New Hashing Method. *IEEE Access*, 8:172343–172351, 2020.
- [SLY12] Steinebach, Martin; Liu, Huajian; Yannikos, York: Forbild: Efficient robust image hashing. In: *Media Watermarking, Security, and Forensics 2012*. Jgg. 8303. International Society for Optics and Photonics, S. 83030O, 2012.
- [St13] Steinebach, Martin; Klöckner, Peter; Reimers, Nils; Wienand, Dominik; Wolf, Patrick: Robust hash algorithms for text. In: *IFIP International Conference on Communications and Multimedia Security*. Springer, S. 135–144, 2013.
- [WSY14] Winter, Christian; Steinebach, Martin; Yannikos, York: Fast indexing strategies for robust image hashes. *Digital Investigation*, 11:S27–S35, 2014.

13. Workshop KI-basiertes
Management und Optimierung
komplexer Systeme (MOC 2021)

Towards Predictive Maintenance as a Service in the Smart Housing Industry

Maximilian Lowin¹ and Cristina Mihale-Wilson²

Abstract: Maintenance is a significant cost driver in many industries with tangible assets. Aiming to predict damages before they occur, this paper focuses on predictive maintenance (PdM) for smart buildings and apartments – a multi-billion-dollar market with substantial cost savings potential. Based on stakeholder groups' heterogeneity within the smart housing industry, PdM cannot be a one-fits-all solution. To be effective, practitioners can enrich PdM with Artificial Intelligence (AI). However, to match very heterogeneous environments and the various needs of the stakeholders, PdM must be modular and flexible. Motivated by the challenges and peculiarities for implementing Predictive Maintenance as a Service (PdMaaS) in the smart housing industry, we provide a concept to support managers to overview and optimize complex PdM needs in complex and heterogeneous environments.

Keywords: smart services, predictive maintenance, smart housing industry, visualization.

1 Motivation

Maintenance is a significant cost driver in almost all industries with tangible assets [WY07]. Ineffective maintenance in industrial settings represents a yearly loss of more than \$60 billion [Mo02]. Such losses occur because traditional maintenance methods such as **run-to-failure** management and **preventive maintenance** are not optimal. In the case of **run-to-failure** management, for instance, machines only get maintained if they break. Run-to-failure results in low machine availability due to long downtimes until artisans correct machine failures and higher overtime labor costs [Mo02]. On the contrary, in the case of **preventive maintenance**, a machine gets maintained on a time-driven schedule, e.g., based on elapsed time or hours of operation. Such time-driven approaches in maintenance face two significant problems: First, companies replace parts that are not yet broken and would have lasted longer. As a result, preventive maintenance creates regular but sometimes unnecessary repairs. Second, parts may break before their scheduled maintenance. In this case, preventive maintenance generates a similar cost as run-to-failure approaches. Flexible and intelligent forward-looking maintenance promises enormous potential savings. Hence, scholars and practitioners are continually seeking to develop more efficient forms of maintenance. One promising concept is **Predictive Maintenance**

¹ Goethe University Frankfurt, Professur für Wirtschaftsinformatik und Informationsmanagement, Theodor-W.-Adorno-Platz 4, 60629 Frankfurt am Main, lowin@wiwi.uni-frankfurt.de

² Goethe University Frankfurt, Professur für Wirtschaftsinformatik und Informationsmanagement, Theodor-W.-Adorno-Platz 4, 60629 Frankfurt am Main, mihale-wilson@wiwi.uni-frankfurt.de

(PdM). PdM concept aims at monitoring the physical conditions of specific objects like manufacturing plants or devices to predict failures in advance using statistical methodologies and Artificial Intelligence (AI). This form of maintenance is particularly cost-effective. However, PdM also requires a high degree of digitalization of the environment that needs high domain-specific expertise and knowledge on leveraging AI techniques in various contexts. After all, many PdM methods originate in industrial settings (e.g., manufacturing facilities and for smaller industrial parts like bearings, engines, electrical motors [So12, WY07], or aviation [A114]), which are highly digitalized and complex domains.

Interestingly, although repair and maintenance work alone accounts for one-third of the construction sector's output [AO20], and damages in appliances can substantially impact other infrastructure (as in the case of leakages), implementations of PdM for smart buildings and apartments remain challenging to achieve and are very rare. One potential reason, therefore, is the heterogeneity of infrastructure in smart buildings and apartments. Another major obstacle in transferring existing PdM approaches from industrial contexts to smart living areas is that PdM requires a high degree of domain knowledge and appropriate AI knowledge. Building on a systematic approach for PdM service design combined with various expert workshops, this work seeks to support practitioners in the smart housing industry (SHI) by providing a blueprint for realizing PdM for smart buildings and apartments. Our blueprint entails a proposition for a service solution and a decision support monitoring tool. It caters to the heterogeneity and complexity of smart living environments and proposes a **PdM as a Service solution**. After all, due to its stakeholders' group heterogeneity (i.e., building and apartment owners, managing companies, tenants, third party service companies) and their different needs within this context, PdM cannot be a one-fits-all solution. Instead, in the SHI context, PdM requires modularity and flexibility to match the various needs of the housing sector stakeholders.

Formally, this paper is structured as follows: Section two provides a brief overview of PdM in the service and SHI context. Section three presents the key insights from the systematic PdM service approach applied to the SHI – i.e., the challenges, objectives, and potential solutions. Finally, the last chapter summarizes the key insights of this work and gives an outlook for future work.

2 Theoretical Background and Research Setting

Motivated by the current shift in consumer consumption away from products towards solution-based services [Li06], this work posits that successful PdM in the SHI must follow a service approach. Services allow that individuals enjoy benefits through a temporary possession rather than ownership per se [LG04]. Moreover, individuals can book services to get work done for them, e.g., by outsourcing work to machines [Go99].

2.1 Product Service Systems

The concept of Product Service Systems (PSS) brings services and products together [Go99]. By definition, a PSS refers to a set of products and services that jointly fulfill users' needs. Furthermore, because PSS extends manufacturers' responsibility to the use phase until the end of product life [Zh12], it can achieve a range of benefits such as greater diversity of choices for customers, (mass) customization, additional value, and higher product quality [Mo02b]. Based on the benefits it provides, several scholars adapted and refined the PSS concept in various industries, e.g., aviation [Zh12], automotive industry [Wi06], or manufacturing [CBL06].

Notably, [CBL06] distinguish between product-orientated PSS, use-orientated PSS, and result-orientated PSS. A **product-orientated PSS** is characterized by the ownership of a material artifact being transferred to the buyer. An application of the PSS concept is product maintenance [CBL06, Wi06, Zh12]. As a product-orientated PSS, PdM can be an inherent service function of the product itself. Thus, PdM can support existing warranty and maintenance contracts and minimize costs [Le09]. An example would be house owners buying a boiler, including a PdM service that automatically notifies them if something is abnormal. In contrast, a **use-orientated PSS** is characterized by transferring only the rights of use to the customer, while the material artifact's ownership remains with the service provider [CBL06]. For instance, house owners rent a PdM solution for their existing boiler. This solution includes sensor hardware belonging to the services provided. The house owners only use the PdM services, which include the sensors for the time of the service contract. Besides, in the case of **result-orientated PSS**, a customer purchases the utility and not the use of a product. Following our example, the house owner would purchase a service for warm water without taking care of the respective infrastructure and maintenance solution.

To date, there is a trend towards smart PSS, as they combine smart digital technologies with physical, tangible products, intangible services, and business models of PSS [CHP18]. Digital technologies such as IoT, Cloud Computing, and Analytics enable the required input for PSS and their information management capabilities [Ar16]. In addition, modular smart PSS solutions can foster remote monitoring and forecast possible failures and maintenance needs because of PdM services' high degree of automation [CHP18].

To date, the literature on PdM applications in the industry is extensive [Al14, So12, We11, WY07]. However, scholarly work on PdM-related services is scarce. One notable scholarly effort on PdM-related services is the work of Lee and colleagues [Le09]. In their work, [Le09] present a methodology for PdM service design. Although Lee and colleagues' approach is a product-oriented PSS, we can also transfer it to other PSS types. Their approach for PdM service design [Le09] envisions eight steps that build our work's basic framework.

2.2 Systematic Approach for PdM Service Design

We follow [Le09] systematic approach for PdM service design and implement Predictive Maintenance as a Service (PdMaaS) in SHI. [Le09] structured PdM service design in eight steps: problem formulation; abstraction level and performance metrics identification; prognostic method selection; measurement selection and sensor strategy; monitoring strategy evaluation; experimental design planning; solution feasibility/selection; cost-benefit analysis (see Fig. 1).

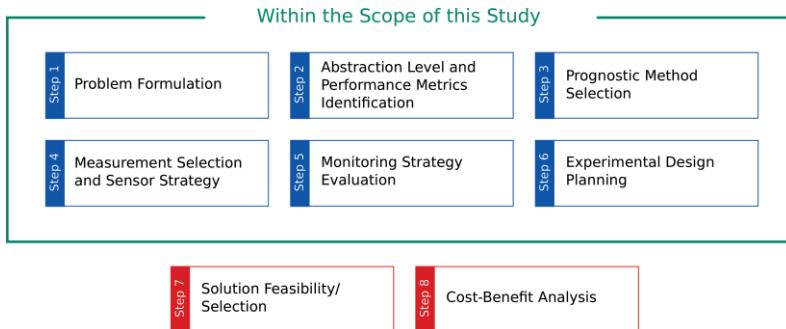


Fig. 1. Research methodology of our PdMaaS concept following [Le09].

Problem formulation represents one of the most crucial steps within this method. As the authors concede, successful PdM design requires that consumers' and companies' service needs are identified and translated into precise maintenance requirements. Accordingly, the problem formulation step intends to articulate companies' or consumers' service needs at its core. Such service needs can include service uptime needs; service to failure prevention; service for system streamlining; service for productivity improvement; service for information management; and closed-loop life cycle product management [Le09].

Notably, consumers' and companies' service needs are unique to companies' assets and products. Against the background that successful PdMaaS must closely mimic the industry's idiosyncrasies and companies' assets and products [Le09], we conduct our research in the context of an ongoing SHI research project (please see next section – Research Setting). At this point, we adopt Lee and colleagues' methodology because it is tailored specifically for the PdM setting. This way, we strive to address everything needed for PdM service design. We omit the last two steps of the systematic approach to leave it open for further discussions (e.g., within the MOC workshop) and future research and concentrate on a practice motivated PdM concept prototype. Within the next methodological step, we propose a proper experimental design and suitable prototypical artifact.

We carried out this work as part of a research project sponsored by the German government. Below we explain the goal, structure, and stakeholders in the research project ForeSight.

2.3 Research Setting

The research project ForeSight aims to develop and establish an Artificial Intelligence (AI)-based service platform for smart living [Ba19]. Within ForeSight, PdM is a central element for the smooth, optimal, and sustainable realization of Smart Living. The project's team entails experts from original equipment manufacturers (OEM) (for both smart and traditional products), smart home solution providers, research entities, and housing companies. We harness organizations' diversity within the project and their specific domain knowledge to conduct various workshops that help us articulate an SHI-specific PdM service design. Based on the input, we develop ideas on how PdM could look like and identify challenges and solutions for implementing the envisioned PdM for SHI. We invited 12 experts from various industries to participate in our workshops: Five participants work for housing companies that own and rent out a large number of "traditional" and smart apartments in the German housing market. Seven participants were domain experts working for smart housing solution providers or OEM vendors for smart home goods and devices. We designed and organized the workshops based on a modified version of the nominal group technique (NGT) to ensure timely and valuable results.

The NGT is a robust methodology for idea generation and prioritization and addresses the necessity of stakeholder involvement in a very early stage of the project [DVG75]. The topics of the four workshops were (1) identification of service needs for PdM in SHI, (2) objectives for a solution, (3) a smart service marketplace design for PdMaaS, and (4) the testing of our artifact. By combining ideas of the Improved Nominal Group Technique (INGT) proposed by [Fo89], we ensure proper preparation of all participants before the meetings. This preparation took place in the form of a short briefing on the workshops' topic several weeks before the workshops. Although we performed the workshops online, the NGT methodology ensures proper workshops and meaningful results [La07].

3 Predictive Maintenance as a Service in SHI

3.1 Problem Formulation

The first workshop's goal was to identify smart housing environments' service needs and formulate problems occurring when transferring a PdM solution to the SHI. We define smart houses as complex PSS systems of a plethora of interconnected sensors, actuators, hardware devices, and Artificial Intelligence (AI)-enabled software components. Our results reveal that **failure prevention, system streamlining, productivity improvement, and information management** stand out to be the most pressing SHI-related maintenance needs. While failure prevention adheres to the companies' needs to ensure that the complex systems of smart housing do not fail and thus bring the entire system down, system streamlining refers to companies' goal to prolong the functional life of components by continuously and intelligently monitoring and maintaining them. Similarly, service for productivity improvement relates to companies' goal of improving their productivity

through continuous operation readiness. In this context, companies strive to achieve reduced energy consumption and enhanced safety. Finally, companies' service for information management relates to maintaining operational readiness and requires optimal information flows for the stakeholders involved. Therefore, we will primarily focus on the information management in our prototype design to properly manage our AI-based service solution.

Out of these service needs, particular problems occur in the SHI context. SHI products and goods are often mass-produced and lack the option of customization. White goods such as washing machines or ovens rarely include PdM solutions. Other facilities such as elevators or HVAC systems that offer PdM do so in proprietary (silo) solutions that are product-related and lack interoperability. This way, current PdM solutions do not harness the plethora of information that (interconnected) sensors, actuators, hard- and software components can provide. The segregation of PdM solutions for single products or product solutions is surprising. After all, PdM in industries such as manufacturing and aviation reveals, there is much to be gained from combining data from various sources (e.g., sensors, hard- and software components) to improve the availability of machines, the cost of maintenance, productivity, and machine profitability [Mo02].

Regarding the predictive power of PdM, experts also agree that rich data from various sources is central for reliably detecting anomalies and thus predicting failures. Given that SHI is an emergent market, we focus on its peculiarities regarding the data measurement selection and sensor strategy in detail in a later section.

3.2 Abstraction Level and Performance Metrics Identification

Contrasting the challenges mentioned above in SHI (i.e., no PdM for certain goods, the existence of proprietary silo PdM solutions, and the heterogeneity of the SHI stakeholders) to the SHI service needs (i.e., failure prevention, system streamlining, productivity improvement, and information management), the workshop participants agree that efficient PdM in SHI should be overarching, intelligent and flexible.

In the industrial context, the basic idea of PdM is to determine the actual mean-time-to-failure of machines and other complex systems. Therefore, several aspects like the availability of machines, the cost of maintenance, the quality, productivity, and the machines' profitability can be optimized [Mo02]. For PdM in the industrial setting, modern methods use the information given by a system to determine the system's condition and predict when maintenance should be performed [AI14].

However, there are no clear, distinct aims for PdM in the SHI as in the producing industry, for example. Due to the sheer amount and heterogeneity of stakeholder groups involved, each stakeholder assesses the PdM performance differently and often displays competing interests. In this respect, PdM in SHI must serve the housing owners, the managing companies, the tenants, and third-party service companies. Performance metrics must

incorporate both efficiency (e.g., accuracy) and effectiveness, but also additional evaluation metrics like tenants' satisfaction with the solution, their perceived intrusiveness due to the digitalization, trust in the system, and the control of respective conditions and countermeasures. Thus, PdM in SHI requires modularity and flexibility to match the various needs of the housing sector's stakeholders. Ultimately, in the SHI context, getting the correct information at the right time to the right person is as crucial for PdM in SHI as identifying service needs at the optimal point in time.

3.3 Prognostic Method Selection

To be able to predict potential failures, PdM relies on detecting outliers in data sets. To this end, PdM commonly uses methods such as clustering or anomaly detection algorithms. In general, outliers can be errors or events. Errors are noisy data that may originate from a faulty sensor. Events refer to a change of a consistent state of the real world (Nesa et al., 2018). Outliers can be an extreme sensor reading, for instance, and such algorithms often work independently of the type of sensor. In this sense, the higher the amount of historical and real-time data and the higher accuracy of such outlier detection algorithms, the better the predictions of PdM routines. The presence of labeled data – i.e., on data that document failure events – improves the predictions' accuracy and interpretability. Due to the high availability of labeled historical data in the manufacturing setting, anomaly detection and failure prediction use mainly supervised machine learning techniques (SL) which learn from labels of failures (e.g., time series anomaly detection algorithms using neural networks like RNNs or LSTMs). In contrast to the manufacturing setting, the smart housing sector is a relatively young emergent market in which smart sensors and devices have relatively short life cycles (often not more than five years). Accordingly, historical data is almost inexistent.

Similarly, smart houses typically do not register anomalies. Thus, they do not produce labeled data, so the implementation of anomaly detection based on supervised learning is not feasible. A potential solution for the missing historical and labeled data is to use domain knowledge to verify whether the data at hand are non-anomalous or anomalous and use semi-supervised learning (SSL) instead of supervised learning. For SSL, not all data used to train algorithms needs to be labeled. Instead, SSL requires just a few labels that were made manually by domain experts. Potential algorithms include LSTM Autoencoders.

Another essential difference between the industrial and SHI setting is the environment and context in which PdM methods operate. For manufacturing sites, processes are usually standardized and follow a predefined schedule. In the residential context, there is no preschedule for activities. For instance, in SHI, different residents are using the laundry room and are washing their laundry at different times by using various washing cycles with different lengths. A PdM solution in the SHI must be aware of this status and must be able to recognize current situations and contexts as additional information besides the pure sensor measurements. Furthermore, PdM methods must operate on contextual data.

In this sense, anomaly detection algorithms need to monitor the data and the context itself. A contextual anomaly occurs, for instance, if users change their behavior. Imagine, for example, that residents of a specific apartment always turn on the ceiling lamp when they come home from work. If they suddenly stop turning on the ceiling lamp but turn on another lamp instead, PdM might erroneously label the ceiling lamp as a defect.

3.4 Measurement Selection and Sensor Strategy

The quality of anomaly detections and predictions about failure depends on the amount of data available. In general, the more data PdM systems can access, the more accurate its predictions are. Again, PdM in SHI hinges on several challenges unique to the smart housing context. One challenge is that the SHI often does not have the database necessary to conduct such predictions. SHI does not possess sensor densities as high as manufacturing and digitalization are scarce. Also, SHI does not always allow for retrofitting activities. While sensors can be retrofitted and attached closely meshed almost everywhere in a manufacturing plant in the industrial setting, new sensors in buildings and apartments can easily interfere with the tenants' aesthetics and living comfort. Besides esthetic reasons, the SHI does not allow retrofitting specific sensors and devices after completing a house's construction process. For instance, one can retrofit a leakage sensor to a production machine, but one cannot install it in a house where the walls and piping have already been built.

Another critical challenge is that in SHI, infrastructural objects or appliances are often black box solutions, and users rely on functionalities offered by the appliance manufacturers. Plants and machines are often customized to the factory's special needs or production line in the manufacturing context. Thus, factories can also require to receive the information needed to pursue PdM activities. However, in SHI, the hardware is typically not customizable. Washing machines, for instance, are white goods and cannot be modified easily. Therefore, information and data created by sensory in the washing machines cannot be accessed nor processed quickly for PdM.

Thus, one must use alternative data sources and AI-based PdM technologies to solve these challenges and to digitalize the SHI. For example, one can monitor refrigerators by analyzing the temperature or electricity flow [Ku18]. However, in SHI, adding those sensors to non-smart devices is cumbersome and expensive. A more straightforward solution would be to make use of so-called smart plugs that measure energy consumption. While in some use cases adding smart plugs can work well, this strategy is not appropriate for large energy consumers or appliances. Since smart plugs are often limited to standard socket sizes and a maximum power of around 4,000W, standard smart plugs do not work combined with three-phase current devices. In such cases, PdM could exploit aggregated energy consumption data provided by smart meters. After all, more and more residential buildings are installing smart meters that can capture energy and water consumption.

Smart meter disaggregation helps to break down the accumulated consumption data by using AI. Koutitas and Tassioulas [KT15] define smart meter disaggregation as a ‘methodology for recognizing individual appliance signal signatures from aggregated circuit readings’ (p. 1665). In other words, by using disaggregation methods, the energy consumption of specific devices within a building can be extracted from the total consumption. The same applies to the disaggregation of smart water meters. [CS13] use disaggregation to understand water consumption profiles for various activities (e.g., showers) and appliances (e.g., dishwashers or washing machines). If the disaggregation is approximately accurate, one can use the data to monitor specific devices and their consumption without needing to attach sensors in or next to the device. Thus, using disaggregation saves costs [KT15], is non-intrusive, and is generalizable to multiple households [Ko16].

Another helpful approach to generating a comprehensive data basis is substituting sensors. Combining several installed sensors can approximate another particular sensor's value. The sensors used to approximate one particular sensor's value can be either the same or a different type of sensor. For example, temperature sensors can approximate other temperature sensors in the same and adjacent room. Similarly, activity and motion detection, usually done by one individual infrared (IR) sensor, can be monitored by following activities in the context and actuators of other devices and furnishings. For example, without an IR sensor, an algorithm can deduce activity and motion from actuators in light switches, a refrigerator door, or a window panel.

3.5 Monitoring Strategy

AI-based PdM services are an excellent opportunity to optimize the sustainability of smart homes. However, without integrating decision support and control systems, this information is only helpful for theory and technicians. Decision-makers need to monitor and assess the failures and maintenance works’ impact and urgency. Hence, monitoring several different residential complexes and infrastructural objects associated with them requires dedicated monitoring solutions. Without a suitable visualization interface, it is almost impossible to keep track of AI-based computations on a wide array of different devices, apartments, and residential complexes. Additionally, in case of failures, one can take appropriate action by just notifying responsible persons for various events occurring within a complex and unstructured way.

Current software solutions like building management systems (BMS) or computer-aided facility management (CAFM) solutions often lack predictive features. To date, there is – according to our best knowledge – no suitable non-proprietary solution on the market. Solution providers of existing BMS and CAFM software could add a programmable interface to their software to integrate predictive features through external services. A critical drawback in doing so, however, is that integrating external information to existing software is cumbersome and adds more complexity to already very complex systems. Furthermore, it might not be possible to track and explain why an event occurs accurately.

Thus, instead of patching up legacy visualization products, it would be more suitable to implement a dedicated modular visualization dashboard that displays specific predictive and non-predictive events and physical conditions of the various entities and objects in a smart building. Such a decision tool would also display proper explanations of events – an increasingly important topic for AI. Also, a dedicated modular visualization dashboard can better address the needs of multiple stakeholders. After all, tenants, landlords, property managers, and external service providers such as artisans, water suppliers, or delivery services have different informational needs. Besides the pure visualization of events, a dedicated modular dashboard can also link to a marketplace, where stakeholders can also book and manage PdM services that best fit their needs in those use cases most relevant to them.

3.6 Prototypical Design for the Visualization of PdM

In this work, we limit ourselves to propose an artifact for the visualization of PdM. We do so since this artifact is of particular importance for decision-makers as an input for an AI-based management and process optimization. Additionally, the implementation of the monitoring tool is independent of the concrete anomaly detection and the preceding data generation process. A smart service marketplace offers different services that perform several PdM tasks in the smart housing context. A smart service can be, for instance, the monitoring of water pipes by using pre-installed hardware or PdM solutions for kitchen appliances. Such a marketplace is a multisided market, where different providers can offer their solution to a customer. Such a solution can be a PSS or product-independent service. Additionally, provided PSS solutions can be product-, use-, and result-based. Therefore, a possible provider can be an appliances manufacturer or a PdM software specialist, for example. Customers can be tenants, property and facility managers, or infrastructure providers. To ensure the required PdMaaS modularity, customers can choose smart services according to their preferences. Furthermore, the smart service marketplace is also a suitable platform for managing and monitoring such smart services. It should display and forecast the health status of an infrastructural object in the form of a central dashboard.

The dashboard's primary goal is to display the most critical or most common events regarding infrastructural objects' physical conditions while avoiding information overload. In addition, modularity and abstraction of various detail levels ensure a customizable solution to meet further information needs.

Based on a third workshop with the stakeholders of the ForeSight project, we prepared a design proposal for a monitoring tool (i.e., PdM dashboard) and a smart service marketplace. Finally, within a fourth expert workshop, the initial prototypical design was discussed and improved. Herein, we present the results from the fourth expert workshop. The experts in the workshop agreed that due to the multitude of stakeholders involved, the initial design of the core PdM dashboard should focus on three main stakeholder groups and their specific (informational) needs: landlords and property managers, tenants, and third-party technicians.

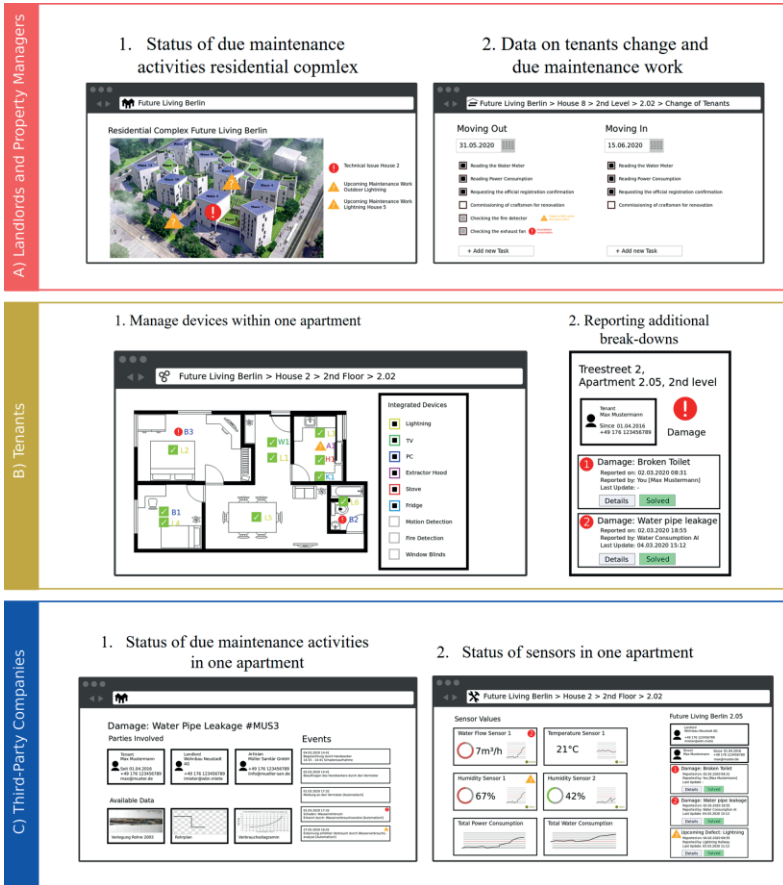


Fig. 2. PdM Visualization Prototype for Different Stakeholders

Landlords and property managers are in the first place interested in the physical conditions of an entire residential complex or building and in aggregated instead of single and detailed failures. The smart service marketplace’s dashboard can help these stakeholders keeping an overview of due maintenance work and help them optimize their maintenance processes. Knowing when tenants will change apartments, for instance, will allow landlords and managers to plan and execute upcoming maintenance works “in-between” tenants (see A in Fig. 2). While this procedure is typical for preventive maintenance work, it is often part of a standardized process to repair and renovate an apartment between different tenants. Therefore, we propose conducting PdM works also during the time intervals between tenant changes. Such PdM works would include any maintenance work that would be necessary in the upcoming few months. This procedure minimizes contact with the tenants and prevents any scheduling conflicts.

Tenants: Due to privacy regulations and different needs, tenants will have a different landing view on the smart service marketplace than landlords or property managers (see B in Fig. 2). Since their apartment is their central point of life, they are primarily interested in being informed on their own apartment's status, but sometimes also in the status of the common building (e.g., the status of the smart door at the entrance, status of the smart elevator). Additionally, they need to manage their smart service subscriptions and the individual sensors and devices within the apartment. Tenants require an intelligent interface to report anomalies or failures and get information about the repair status. Therefore, the tenant's view is more detailed. What is also crucial to tenants is that if failures are not detected automatically, they can still report issues manually.

Third-party companies: Like the landlords/managers and the tenants, third-party companies like external artisans and other service providers will be able to view information about the site and the mutual connections between devices and sensors. Especially the latter is helpful to detect the broken critical devices more efficiently. Therefore, service providers should see real-time and short-time historical data (i.e., data from the last hours before defect) of dedicated sensors. Additionally, service companies should have the opportunity to mark all problems as solved for all assigned problems. It can be beneficial for third-party companies to see further anonymized data of similar sites and sensors for further troubleshooting. With the help of digitalization, third-party companies can also remotely support tenants and monitor infrastructure.

4 Conclusion

The monitoring of a building's physical conditions and the prediction of upcoming maintenance work are essential for operating residential complexes – both estimations base on a distinct but related methodology. Although the idea of predicting failures originates from the manufacturing industry, this work transfers the PdM concept to the SHI by conceptualizing it as a Service (PdMaaS). Based on four interdisciplinary workshops with various domain experts from the SHI, this work presents a systematic evaluation of challenges and solutions of the application of PdMaaS in the SHI.

By closely considering the key objectives of a PdMaaS for SHI – i.e., modularity, flexibility, and applicability on different devices and appliances – we propose a prototype concept to visualize smart services in the context of PdM in the SHI. We limit our work to the visualization only as it is independent of the concrete implementation but enables an early demonstration to evaluate it among the different stakeholders independently of a specific underlying methodology. Nevertheless, this monitoring solution is crucial for decision-makers as an input for AI-based management and process optimization and needs a thorough evaluation.

Within the next step, we need to evaluate the feasibility of several algorithmic solutions for PdM in the SHI needs. Additionally, following [Le09], future research should also

consider the economic aspect of PdM in SHI by conducting comprehensive cost-benefit analyses for various PdM services in different residential contexts (e.g., large versus smaller housing developments, luxury versus social housing). Altogether, our work presents valuable insights that can serve as input in subsequent considerations on designing and implementing predictive maintenance as a service for the smart housing industry. Further, the insights presented in this paper can also directly enter into practitioners' ongoing efforts to adopt PdM for their industry- and business-specific needs. After all, by predicting damages and repairing or exchanging devices before damage occurs, any type of business can save substantial amounts of time and money.

Acknowledgment

This work is funded by the German Federal Ministry for Economic Affairs and Energy (BMWi) as part of the ForeSight project (Grant: 01MK20004J).

Bibliography

- [Al14] Alestra, S., Brand, C., Erofeev, P., Bordry, C., Burnaev, E., Papanov, A.: Rare Event Prediction Techniques in Application to Predictive Maintenance of Aircraft. In: Proceedings of ITaS conference. pp. 32–37, 2014.
- [AO20] Allcoat, J., Office for National Statistics: Output in the Construction Industry, <https://www.ons.gov.uk/file?uri=/businessindustryandtrade/constructionindustry/datasets/outputintheconstructionindustry/current/previous/v64/bulletindataset2v2.xlsx>, last accessed, 2020/06/15.
- [Ar16] Ardolino, M., Sacconi, N., Gaiardelli, P., Rapaccini, M.: Exploring the Key Enabling Role of Digital Technologies for PSS Offerings. *Procedia CIRP* 47, pp. 561–566, 2016.
- [Ba19] Bauer, J., Hoffmann, H., Feld, T., Runge, M., Hinz, O., Mayr, A., Förster, K., Teske, F., Schäfer, F., Konrad, C.: ForeSight - Platform Approach for Enabling AI-based Services for Smart Living. In: International Conference on Smart Homes and Health Telematics. pp.204–211. Springer, Cham, 2019.
- [CBL06] Cook, M.B., Bhamra, T., Lemon, M.: The transfer and application of Product Service Systems: from academia to UK manufacturing firms. *Journal of cleaner production* 14, pp. 1455–1465, 2006.
- [CHP18] Chowdhury, S., Haftor, D., Pashkevich, N.: Smart Product-Service Systems (Smart PSS) in Industrial Firms: A Literature Review. *Procedia CIRP* 73, pp. 26–31, 2018.
- [CS13] Cole, G., Stewart, R.A.: Smart meter enabled disaggregation of urban peak water demand: precursor to effective urban water planning. *Urban Water Journal* 10, pp. 174–194, 2013.
- [DVG75] Delbecq, A.L., Van de Ven, A.H., Gustafson, D.H.: Group techniques for program planning: A guide to nominal group and Delphi processes. Scott, Foresman, 1975.

- [Fo89] Fox, W.M.: The improved nominal group technique (INGT). *Journal of management development* 1989.
- [Go99] Goedkoop, M.J., Van Halen, C.J., Te Riele, H.R., Rommens, P.J., others: Product service systems, ecological and economic basics. Report for Dutch Ministries of environment (VROM) and economic affairs (EZ) 36, pp. 1–122, 1999.
- [Ko16] Kong, W., Dong, Z.Y., Ma, J., Hill, D.J., Zhao, J., Luo, F.: An extensible approach for non-intrusive load disaggregation with smart meter data. *IEEE Transactions on Smart Grid* 9, pp. 3362–3372, 2016.
- [KT15] Koutitas, G.C., Tassiulas, L.: Low cost disaggregation of smart meter sensor data. *IEEE Sensors Journal* 16, pp. 1665–1673, 2015.
- [Ku18] Kulkarni, K., Devi, U., Sirighee, A., Hazra, J., Rao, P.: Predictive maintenance for supermarket refrigeration systems using only case temperature data. In: 2018 Annual American Control Conference (ACC). pp.4640–4645. IEEE, 2018.
- [La07] Lago, P.P., Beruvides, M.G., Jian, J.-Y., Canto, A.M., Sandoval, A., Taraban, R.: Structuring group decision making in a web-based environment by using the nominal group technique. *Computers & Industrial Engineering* 52, pp. 277–295, 2007.
- [Le09] Lee, J., Chen, Y., Al-Atat, H., AbuAli, M., Lapira, E.: A systematic approach for predictive maintenance service design: methodology and applications. *International Journal of Internet Manufacturing and Services* 2, pp. 76–94, 2009.
- [LG04] Lovelock, C., Gummesson, E.: Whither services marketing? In search of a new paradigm and fresh perspectives. *Journal of service research* 7, pp. 20–41, 2004.
- [Li06] Lindahl, M., Sundin, E., Shimomura, Y., Sakao, T.: An outline of an interactive model for service engineering of functional sales offers. In: International Design Conference-DESIGN 2006. Bubrovnik, Croatia, 15-18 2006. pp. 897–904, 2006.
- [Mo02] Mobley, R.K.: An introduction to predictive maintenance. Elsevier, 2002.
- [Mo02b] Mont, O.K.: Clarifying the concept of product–service system. *Journal of cleaner production* 10, pp. 237–245, 2002.
- [So12] Soualhi, A., Clerc, G., Razik, H., Lebaroud, A.: Fault detection and diagnosis of induction motors based on hidden Markov model. In: 2012 XXth International Conference on Electrical Machines. pp. 1693–1699. IEEE, 2012.
- [We11] West, S.R., Guo, Y., Wang, X.R., Wall, J.: Automated fault detection and diagnosis of HVAC subsystems using statistical machine learning. In: 12th International Conference of the International Building Performance Simulation Association, 2011.
- [Wi06] Williams, A.: Product-service systems in the automotive industry: the case of micro-factory retailing. *Journal of cleaner production* 14, pp. 172–184, 2006.
- [WY07] Widodo, A., Yang, B.-S.: Support vector machine in machine condition monitoring and fault diagnosis. *Mechanical systems and signal processing* 21, pp. 2560–2574, 2007.
- [Zh12] Zhu, H., Gao, J., Li, D., Tang, D.: A Web-based Product Service System for aerospace maintenance, repair and overhaul services. *Computers in Industry* 63, pp. 338–348, 2012.

A Survey of Constraint Transformation Methods

Sven Löffler, Ilja Becker, Franz Kroll, Petra Hofstedt¹

Abstract: The solution performance of finite domain (FD) constraint problems can often be improved by either transforming particular constraints or sub-problems into other FD constraints like binary, table or regular membership constraints, or by transformation of the complete FD problem into an equivalent problem but of another domain, e.g. in a SAT problem. Specialized constraint solvers (like binary or SAT solvers) can outperform general constraint solvers for certain problems. However, this comes with high efforts for the transformation and/or other disadvantages such as a restricted set of constraints such specialized solvers can handle or limitations on the variables domains. In this paper we give an overview of CSP and constraint transformations and discuss applicability and advantages and disadvantages of these approaches.

Keywords: Constraint Programming; CSP; Refinement; Optimization; Regular Membership Constraint; Regular CSPs; Table Constraint; SAT; Binary Constraint

1 Introduction

Constraint programming (CP) is a powerful method to model and solve NP-complete problems in a declarative way. Typical applications of CP are among others rostering, graph coloring, optimization, resource management, planning, scheduling and satisfiability (SAT) problems [Ma98].

Because the search space of constraint satisfaction problems (CSPs) and constraint satisfaction optimization problems (CSOPs or COPs) is immensely big and the solution process often needs an extremely large amount of time we are always interested in improving the solution process. In practice there are often various ways to describe a CSP and consequently the problem can be modeled by different combinations of constraints, which results in the differences in resolution speed and behavior. For example, there is the possibility to represent a CSP with constraints, which are of the same kind. Thus we have the possibility to use a solver, solver settings or search strategies which are optimized for the used constraints.

The rest of this paper is structured as follows. In Section 2, we introduce the necessary definitions of constraint programming. In Sections 3, 4, and 5 we give a survey of existing transformations of a general CSPs into binary, boolean, and table and regular CSPs. Furthermore, in each of these three sections we show examples of transformations and discuss advantages and disadvantages. Finally, in Section 6 we summarise and explain future steps in our researches.

¹ Brandenburg University of Technology Cottbus-Senftenberg, Germany
Sven.Loeffler@b-tu.de, Ilja.Becker@b-tu.de, Franz.Kroll@b-tu.de, Hofstedt@b-tu.de

2 Preliminaries

In this section we introduce basic definitions and concepts of constraint programming (CP) and present the relevant constraints, which are used in the rest of the paper. We consider *CSPs*, which are defined in the following way.

A *constraint satisfaction problem (CSP)* is defined as a 3-tuple $P = (X, D, C)$ where $X = \{x_1, x_2, \dots, x_n\}$ is a set of variables, $D = \{D_1, D_2, \dots, D_n\}$ is a set of finite domains where D_i is the domain of x_i and $C = \{c_1, c_2, \dots, c_m\}$ is a set of constraints. At this, constraint $c_j = (X_j, R_j)$ is a relation R_j , which is defined over a set of variables $X_j \subseteq X$ [De03b; RBW06].

The *scope* of a constraint $c_j = (X_j, R_j)$ indicates the set of variables, which is covered by the constraint c_j : $scope(c_j) = X_j$ [De03b].

The relation R of a constraint $c = (X, R)$ represents a subset of the Cartesian product of the domain values $D_1 \times \dots \times D_n$ of the corresponding variables $X = \{x_1, \dots, x_n\}$. This can be expressed implicitly by a mathematical formula or (for FD constraints) explicitly by enumeration of the allowed tuples of domain values. In the following, we will use the explicit representation of constraints, i.e. by sets T of tuples of an ordered set of variables X . A CSP can only have a finite number of solutions because the number n of variables and the domain sizes are finite. Thus, it is possible to finitely enumerate all solutions of a CSP

Finally, we introduce two definitions of constraints relevant in this paper. Let a CSP $P = (X, D, C)$ and a subset X' of variables X of the CSP P be given.

For an ordered subset of variables $X' = \{x_1, \dots, x_n\} \subseteq X$, a positive *table* constraint $table(X', T)$ restricts any solution of the CSP to be compatible to the one of the given domain tuples of T . (For a negative constraint $table_n(X', T)$ a solution must be incompatible to all the given tuples in T .)

We use the notation of a *regular* constraint as a synonym for *regular membership* respectively *regular language membership* constraint. The *regular* constraint and its propagation [HPB04; Pe01; Pe04] is based on deterministic finite automaton (DFAs) [HU79]. Let $M = (Q, \Sigma, \delta, q_0, F)$ be a DFA, let $X' = \{x_1, \dots, x_n\} \subseteq X$ be an ordered set of variables with domains $D = \{D_1, D_2, \dots, D_n\}, \forall i \in \{1, \dots, n\} : D_i \subseteq \Sigma$. The regular constraint $regular(X', M)$ defines the allowed domain value tuples as:

$$\{(w_1, \dots, w_n) \mid \forall i \in \{1, \dots, n\}, w_i \in D_i, (w_1 w_2 \dots w_n) \in L(M)\} [\text{HK06}]$$

So, the concatenation of the values w_i of the variables $x_i, \forall i \in \{1, \dots, n\}$ must be accepted by the automaton M .

Based on the definition of the table and the regular constraint, we define a *table CSP* respectively a *regular CSP* as a CSP which contains only table or regular constraints.

Furthermore, there are *binary CSPs* which contain only constraints covering two or less variables and *boolean CSPs (SAT problems)* which contain only boolean variables and logical clauses of the form $(\bigvee_j (\neg)x_{ij})$ as constraints.

3 Transformations into Binary CSPs

Two common possibilities to transform FD-CSPs into binary CSPs are the dual [DP89] and the hidden transformations [Pe60; RPD90; SF94]. The goal of these transformations is to receive binary CSPs, i.e CSPs to which algorithms like path consistency (PC-1 und PC-2 in [De03a]) can be applied. Establishing path consistency is possible in polynomial time, but depending on the given CSP much transformation time is needed.

3.1 Dual Transformation

For the dual transformation of a CSP $P = (X, D, C)$ into a dual CSP $P^{dual} = (X^d, D^d, C^d)$ for each constraint $c_i = (X_i, T_i) \in C$ of the original CSP P a new dual variable $x_i^d \in X^d$ is created. The domain of each such variable contains all allowed tuples T_i of the corresponding constraint c_i . For each pair of original constraints c_i and $c_j \in C$, which cover at least one shared variable, a new binary dual constraint $c_{i,j}^d = (\{x_i^d, x_j^d\}, R_{i,j}^d) \in C^d$ is created. The relation $R_{i,j}^d$ defines all allowed tuples, which satisfy c_i and c_j .

It follows a formal definition of the dual transformation and an example, which illustrates the transformation of a CSP into a dual CSP.

Definition 1 (Dual transformation). *The dual transformation of a CSP $P = (X, D, C)$ is defined by $P^{dual} = (X^d, D^d, C^d)$, with:*

$X^d = \{x_1^d, \dots, x_m^d\}$ is a set of dual variables, where each variable x_i^d represents a constraint $c_i \in C$ of P .

$D^d = \{D_1^d, \dots, D_m^d\}$, with $D_i^d = T_i$ is the set of domains for the dual variables. For every dual variable $x_i^d \in X^d$ holds $D_i^d = T_i$, where T_i is the list of admissible tuples of $c_i = (X_i, R_i)$.

C^d is a set of dual constraints over the dual variables. For each pair of constraints $c_i, c_j \in C$ with $c_i \neq c_j$ and $scope(c_i) \cap scope(c_j) = X_{i,j} \neq \emptyset$ of P a dual constraint $c_{i,j}^d = (\{x_i^d, x_j^d\}, T_{i,j}^d) \in C^d$ is created. Each tuple $(a_i, b_j) \in T_{i,j}^{dual}$ contains a tuple $a_i \in T_i$ and a tuple $b_j \in T_j$, which have equal projections to their shared variables [Ba02].

Example 1 (Dual transformation). Given is the CSP P_1 $P_1 = (X, D, C)$ with

$$\begin{aligned} X &= \{x_1, x_2, x_3, x_4\} \\ D &= \{D_1 = \{0, 1\}, D_2 = \{0, 1\}, D_3 = \{0, 1, 2\}, D_4 = \{0, 1, 2\}\} \\ C &= \{c_1 = (x_1 \neq x_2), c_2 = (x_1 < x_3), c_3 = count(\{x_1, x_2, x_3, x_4\}, 2, 1)\} \end{aligned}$$

The *count*-constraint c_3 demands that the value 1 is assigned to 2 of the variables x_1, \dots, x_4 .

When transforming P_1 into a dual CSP, for each constraint $c_1, c_2, c_3 \in C$ a dual variable $X^d = \{x_1^d, x_2^d, x_3^d\}$ is created. The allowed tuples of c_1, c_2 and c_3 are enumerated for the domains of x_1^d, x_2^d , and x_3^d . For example $D_1^d = \{(0, 1), (1, 0)\}$.

Each pair $c_i, c_j \in C$ with at least one shared variable yields a new binary dual constraint $c_{i,j}^d$. For example we get $c_{1,2}^{dual} = (\{x_1^{dual}, x_2^{dual}\}, T_{1,2}^{dual})$, where tuple list $T_{1,2}^d$ represents the projection of the allowed tuples of c_1 and c_2 , such that the shared variable x_1 has the same value: $T_{1,2}^d = \{((0, 1), (0, 1)), ((0, 1), (0, 2)), ((1, 0), (1, 2))\}$. It follows an excerpt of the resulting dual CSP $P^{dual} = (X^d, D^d, C^d)$ with

$$\begin{aligned}
 X^d &= \{x_1^d, x_2^d, x_3^d\} \\
 D^d &= \{D_1^d = \{(0, 1), (1, 0)\}, \\
 &\quad D_2^d = \{(0, 1), (0, 2), (1, 2)\}, \\
 &\quad D_3^d = \{(0, 0, 1, 1), (0, 1, 0, 1), (0, 1, 2, 1), (0, 1, 1, 0), \dots, \}\} \\
 C^d &= \{c_{1,2}^d = (\{x_1^d, x_2^d\}, \{((0, 1), (0, 1)), ((0, 1), (0, 2)), ((1, 0), (1, 2))\}), \\
 &\quad c_{1,3}^d = (\{x_1^d, x_3^d\}, \{((0, 1), (0, 1, 0, 1)), ((0, 1), (0, 1, 2, 1)), \\
 &\quad\quad\quad ((0, 1), (0, 1, 1, 0)), \dots, ((1, 0), (1, 0, 1, 2)), \dots\}) \\
 &\quad c_{2,3}^d = (\{x_2^d, x_3^d\}, \{((0, 1), (0, 0, 1, 1)), ((0, 1), (0, 1, 1, 0)), \\
 &\quad\quad\quad ((0, 1), (0, 1, 1, 2)), ((0, 2), (0, 1, 2, 1)), \dots\})\}
 \end{aligned}$$

3.2 Hidden Transformation

Regarding the hidden transformation of a CSP $P = (X, D, C)$ into a hidden CSP $P^{hidden} = (X \cup X^h, D \cup D^h, C^h)$, the hidden variables X^h and their domains are created analogously to the dual variables in the above transformation. The hidden constraints C^h are binary constraints, each between an original variable $x_i \in X$ and a hidden variable $x_j^h \in X^h$ such that the constraint $c_j \in C$, now represented by the hidden variable x_j^h , contains the variable x_i , i.e. $x_i \in scope(c_j)$. In this way, a bipartite graph with the two sets of nodes X and X^h is created. The hidden constraints C^h guarantee that every variable is only assigned values, which fulfill the initial constraints, in which this variable involved.

Definition 2 (Hidden transformation). *Given a CSP $P = (X, D, C)$, the hidden transformation yields a CSP $P^{hidden} = (X^h \cup X, D^h \cup D, C^h)$, where X^h and D^h are created as X^d and D^d by the dual transformation. C^h is the set of binary hidden constraints, where each connects a hidden variable $x^h \in X^h$ with a variable $x \in X$ of the initial CSP P .*

For every hidden variable $x_j^h \in X^h$ and for every initial variable $x_i \in X$, which are connected via the initial constraint $c_j = (X_j, T_j)$ ($x_i \in scope(c_j)$), a hidden constraint $c_{i,j}^h \in C^h$ exists in C^h . Any such hidden constraint $c_{i,j}^h = (\{x_i, x_j^h\}, R_{i,j})$ specifies the admissible tuples $t \in T_j$ when a value a is assigned to the variable x_i ($R_{i,j} := t[x_i] = a$) [Ba02].

Example 2 (Hidden transformation). Given is again the CSP P_1 , which is supposed to be transformed into a hidden CSP. For each of the three constraints c_1, c_2 and c_3 a hidden variable $X^h = \{x_1^h, x_2^h, x_3^h\}$ and domain $D^h = \{D_1^h, D_2^h, D_3^h\}$ is created as by the dual transformation.

A new, binary hidden constraint $c_{i,j}^h$ must be generated for every pair of variables $x_i \in X$ and $x_j^h \in X^h$ with $x_i \in \text{scope}(c_j)$. We get the following constraints: $c_{1,1}^h = (\{x_1, x_1^h\}, T_{1,1}^h)$, $c_{2,1}^h = (\{x_2, x_1^h\}, T_{2,1}^h)$, $c_{1,2}^h = (\{x_1, x_2^h\}, T_{1,2}^h)$, $c_{3,2}^h = (\{x_3, x_2^h\}, T_{3,2}^h)$, $c_{1,3}^h = (\{x_1, x_3^h\}, T_{1,3}^h)$, $c_{2,3}^h = (\{x_2, x_3^h\}, T_{2,3}^h)$, $c_{3,3}^h = (\{x_3, x_3^h\}, T_{3,3}^h)$ and $c_{4,3}^h = (\{x_4, x_3^h\}, T_{4,3}^h)$.

The lists of tuples $T_{i,j}^h$ result as given above, e.g. $T_{1,1}^h = \{(0, (0, 1)), (1, (1, 0))\}$ associates with the first tuple the value 0 for x_1 with the valuation (0, 1) for constraint c_1 (and 1 for x_1 with (1, 0) for c_1 with the second tuple). This results in a new hidden CSP $P^{\text{hidden}} = (X \cup X^h, D \cup D^h, C^h)$ with:

$$\begin{aligned} X \cup X^h &= \{x_1, x_2, x_3, x_4, x_1^h, x_2^h, x_3^h\} \\ D \cup D^h &= \{D_1 = \{0, 1\}, D_2 = \{0, 1\}, D_3 = D_4 = \{0, 1, 2\}, \\ &\quad D_1^h = \{(0, 1), (1, 0)\}, D_2^h = \{(0, 1), (0, 2), (1, 2)\}, \\ &\quad D_3^d = \{(0, 0, 1, 1), (0, 1, 0, 1), (0, 1, 2, 1), (0, 1, 1, 0), \dots\} \\ C^h &= \{c_{1,1}^h = (\{x_1, x_1^h\}, T_{1,1}^h), c_{2,1}^h = (\{x_2, x_1^h\}, T_{2,1}^h), \\ &\quad c_{1,2}^h = (\{x_1, x_2^h\}, T_{1,2}^h), c_{3,2}^h = (\{x_3, x_2^h\}, T_{3,2}^h), \dots\} \end{aligned}$$

where $T_{1,1}^h = \{(0, (0, 1)), (1, (1, 0))\}$, $T_{2,1}^h = \{(0, (1, 0)), (1, (0, 1))\}$, $T_{1,2}^h = \{(0, (0, 1)), (0, (0, 2)), (1, (1, 2))\}$, $T_{3,2}^h = \{(1, (0, 1)), (2, (0, 2)), (2, (1, 2))\}$, ...

Advantages and Disadvantages of Binary Transformations. Both approaches have the *advantage*, that they allow the use of algorithms which require binary constraints, e.g. path consistency algorithms (PC-1 and PC-2 in [De03a]). On the other side both approaches have the *disadvantages*, that they are only reasonable when the whole CSP is transformed and they need to list all solutions of all constraints, which is mostly extremely time consuming.

The hidden transformation has the advantage over the dual transformation that the original variables remain. This allows a direct read off a solution, no variable transformation from hidden variables is necessary. The advantage of the dual transformation is that it needs less variables and constraints than the hidden transformation, in most cases.

4 Transformations into SAT Problems

SAT problems are capable of representing every other FD-CSP. This requires the transformation of all non-boolean variables into boolean ones, as well as a translation of all constraints based upon these new variables. The main advantage of this approach is the ability to leverage the power of modern SAT solvers. A variety of algorithms for the transformation of FD-CSPs into boolean CSPs is described in [Ga07; Pe15a; Wa00]. Following we describe some exemplary selected methods.

4.1 Direct Encoding

The direct encoding approach creates a new variable for each possible variable assignment. Depending on whether an original variable is assigned a certain value, the corresponding value variable is *True* (1) or not. For each illegal assignment of values to the variables X_i according to a constraint $c_i = (X_i, R_i)$ a clause over the previously created boolean values is generated, that rules out this assignment. Additional constraints are introduced that ensure, that only one value assignment representing variable per original variable can be true at a time.

Definition 3 (Direct encoding). *Given a CSP $P = (X, D, C)$ we define its direct encoding to be $P^{de} = (X^{de}, D^{de}, C^{de})$ where:*

$X^{de} = \{x_{i,j}^{de} \mid i \in \{1, \dots, |X|\}, j \in \{1, \dots, |D_i|\}\}$, is a set of boolean variables where each variable $x_{i,j}^{de}$ represents, whether the original variable x_i is assigned the j -th value of the domain D_i or not ($x_i = d_j \leftrightarrow x_{i,j}^{de} = 1$).

$D^{de} = \{D_{i,j}^{de} = \{0, 1\} \mid i \in \{1, \dots, |X|\}, j \in \{1, \dots, |D_i|\}\}$ is the set of the boolean domains of the newly created variables.

C^{de} is the set of constraints over the newly created boolean variables X^{de} . For each constraint $c = (X, T) \in C$ of the original CSP P , with $X = \{x_1, \dots, x_n\}$, a constraint $c^{de} \in C^{de}$ can be given in boolean form, that is equivalent to c . For each viable tuple $t_l = (v_{l,1}, \dots, v_{l,n}) \notin T$ of the constraint c a clause $\neg x_{1,v_{l,1}}^{de} \vee \neg x_{2,v_{l,2}}^{de} \vee \dots \vee \neg x_{n,v_{l,n}}^{de}$ is generated.

To ensure that the variables $X_i^{de} = \{x_{i,j}^{de} \mid \forall j \in \{1, \dots, |D_i|\}\}$ represent the possible variable assignments for x_i , one needs to ensure that exactly one variable in X_i^{de} becomes true (1). To that for each variable x_i so-called *addLeastOne* clauses $\bigvee_{v \in D_i} x_{i,v}^{de}$, as well as *atMostOne* clauses $\neg x_{i,v_1}^{de} \vee \neg x_{i,v_2}^{de}, \forall v_1, v_2 \in D_i$ with $i \neq j$ are added [Pr09].

Example 3 (Direct encoding). The CSP P_1 is transformed into a boolean CSP using direct encoding. For each variable x_1, x_2, x_3 and x_4 and for each of its domain values a new boolean variable must be created.

For each invalid tuple $t_i \notin T_i$ of each constraint $c_i, i \in \{1, 2, 3\}$ a clause is generated that excludes this tuple. From the invalid tuples $(0, 0)$ and $(1, 1)$ for c_1 therefore follow the clauses $c_{1,1} = \neg x_{1,0}^{de} \vee \neg x_{2,0}^{de}$ and $c_{1,2} = \neg x_{1,1}^{de} \vee \neg x_{2,1}^{de}$. The other constraints are transformed analogously.

Finally the *addLeastOne* and *atMostOne* constraints for the variables x_1, \dots, x_4 need to be created. E.g. for the original variable $x_1 \in X$ the resulting constraints are $addLeastOne(x_1) = (x_{1,0}^{de} \vee x_{1,1}^{de})$ and $atMostOne(x_1) = \{(\neg x_{1,0}^{de} \vee \neg x_{1,1}^{de})\}$.

The direct encoding of CSP P_1 leads to the boolean CSP $P^{de} = (X^{de}, D^{de}, C^{de})$ where:

$$\begin{aligned} X^{de} &= \{x_{1,0}^{de}, x_{1,1}^{de}, x_{2,0}^{de}, x_{2,1}^{de}, x_{3,0}^{de}, x_{3,1}^{de}, x_{3,2}^{de}, x_{4,0}^{de}, x_{4,1}^{de}, x_{4,2}^{de}\} \\ D^{de} &= \{D_{1,0}^{de} = D_{1,1}^{de} = D_{2,0}^{de} = \dots = D_{4,0}^{de} = D_{4,1}^{de} = D_{4,2}^{de} = \{0, 1\}\} \\ C^{de} &= \{c_{1,1} = (\neg x_{1,0}^{de} \vee \neg x_{2,0}^{de}), c_{1,2} = (\neg x_{1,1}^{de} \vee \neg x_{2,1}^{de}), \\ &\quad c_{2,1} = (\neg x_{1,0}^{de} \vee \neg x_{3,0}^{de}), c_{2,2} = (\neg x_{1,1}^{de} \vee \neg x_{3,0}^{de}), c_{2,3} = (\neg x_{1,1}^{de} \vee \neg x_{3,1}^{de}), \\ &\quad c_{3,1} = (\neg x_{1,0}^{de} \vee \neg x_{2,0}^{de} \vee \neg x_{3,0}^{de} \vee \neg x_{4,0}^{de}), \dots\}, \\ &\quad \text{addLeastOne:} \\ &\quad \cup \{(x_{1,0}^{de} \vee x_{1,1}^{de}), (x_{2,0}^{de} \vee x_{2,1}^{de}), (x_{3,0}^{de} \vee x_{3,1}^{de} \vee x_{3,2}^{de}), (x_{4,0}^{de} \vee x_{4,1}^{de} \vee x_{4,2}^{de})\} \\ &\quad \text{atMostOne:} \\ &\quad \cup \{(\neg x_{1,0}^{de} \vee \neg x_{1,1}^{de}), (\neg x_{2,0}^{de} \vee \neg x_{2,1}^{de}), (\neg x_{3,0}^{de} \vee \neg x_{3,1}^{de}), (\neg x_{3,0}^{de} \vee \neg x_{3,2}^{de}), \\ &\quad (\neg x_{3,0}^{de} \vee \neg x_{3,3}^{de}), (\neg x_{3,1}^{de} \vee \neg x_{3,2}^{de}), (\neg x_{3,1}^{de} \vee \neg x_{3,3}^{de}), (\neg x_{3,2}^{de} \vee \neg x_{3,3}^{de}), \dots\} \end{aligned}$$

4.2 Logarithmic Encoding

The idea behind logarithmic encoding is to encode the index of an assigned value for a variable in a binary representation with boolean variables. For each variable $x_i \in X$ of the original CSP P , exactly $r = \lceil \log_2 |D_i| \rceil$ variables $x_{i,r}^{le}, \dots, x_{i,1}^{le}$ are created, where the assignment for the variables $x_{i,r}^{le}, \dots, x_{i,1}^{le}$ represents the index of a value in the domain D_i of the original variable x_i in binary form.

Definition 4 (Logarithmic encoding). *Given a CSP $P = (X, D, C)$ we define its logarithmic encoding to be $P^{le} = (X^{le}, D^{le}, C^{le})$ where:*

$X^{le} = \{x_{i,j}^{le} \mid i \in \{1, \dots, |X|\}, j \in \{1, \dots, r\}\}$ with $r = \lceil \log_2 |D_i| \rceil$ is a set of boolean variables where each series of variables $x_{i,r}^{le}, \dots, x_{i,1}^{le}$ represents, that the original variable x_i is assigned the k -th value of the domain D_i . The assignments $b_{i,r} \dots b_{i,1}$ of the variables $x_{i,r}^{le}, \dots, x_{i,1}^{le}$ correspond to the binary representation of k .

$D^{le} = \{D_{i,j} = \{0, 1\} \mid i \in \{1, \dots, |X|\}, j \in \{1, \dots, \lceil \log_2 |D_i| \rceil\}\}$ is the set of the boolean domains of the newly created variables.

C^{le} is the set of constraints over the newly created boolean variables X^{le} . For each original constraint $c = (X, T) \in C$ with $X = \{x_1, \dots, x_n\}$, a constraint $c^{le} \in C^{le}$ can be given in

boolean form that is equivalent to c . For each tuple $t_1 = (v_{1,1}, \dots, v_{1,n}) \notin T$, that violates a constraint c , the following clause is created:

$$c_l^{le} = \bigvee_{i \in \{1, \dots, n\}, j \in \{\lceil \log_2 |D_i| \rceil, \dots, 1\}} \begin{cases} \neg x_{i,j}^{le} & b_{i,j} = 1 \\ x_{i,j}^{le} & , \text{ otherwise} \end{cases}$$

The values $b_{i,r} \dots b_{i,1}$ encode the binary representation of the index of the value $v_{1,i}$ in D_i .

Additionally one needs to ensure that for each domain whos size is not divisible by two, the binary representations of indices that do not correspond to values in the domain are invalid. For this each of these indices can be represented by a one-dimensional, negative tuple and be transformed into a clause accordingly [Ga07; IM94; Wa00].

Example 4 (Logarithmic encoding). Following the CSP P_1 will be transformed into a boolean CSP using logarithmic encoding. For the variables x_1 and x_2 , whose domain sizes are two, it suffices to create new variables $x_{1,0}^{le}$ and $x_{2,0}^{le}$ respectively. For the variables x_3 and x_4 two variables $x_{3,1}^{le}$ and $x_{3,0}^{le}$, and $x_{4,1}^{le}$ and $x_{4,0}^{le}$ respectively, are created accordingly to their domain size of three. The domains of all created variables equal $\{0, 1\}$. Because the original variables x_3 and x_4 cannot be assigned to value 3, it is necessary to add the two negative tuple clauses $c_{3,3}^n = \neg x_{3,0}^{le} \vee \neg x_{3,1}^{le}$, $c_{4,3}^n = \neg x_{4,0}^{le} \vee \neg x_{4,1}^{le}$.

For each of the three constraints c_1 , c_2 and c_3 a boolean constraint based on their respective negative tuple lists is created. For c_1 we get the negative tuple list $T'_1 = \{(0, 0), (1, 1)\}$, from which the constraints $c_{1,1}^{le} = x_{1,0}^{le} \vee x_{2,0}^{le}$ and $c_{1,2}^{le} = \neg x_{1,0}^{le} \vee \neg x_{2,0}^{le}$ are derived. The other constraints can be obtained analogously. Following is an excerpt of the transformed CSP $P^{le} = (X^{le}, D^{le}, C^{le})$ with:

$$\begin{aligned} X^{le} &= \{x_{1,0}^{le}, x_{2,0}^{le}, x_{3,1}^{le}, x_{3,0}^{le}, x_{4,1}^{le}, x_{4,0}^{le}\} \\ D^{le} &= \{D_{1,0}^{le} = D_{2,0}^{le} = D_{3,1}^{le} = D_{3,0}^{le} = D_{4,1}^{le} = D_{4,0}^{le} = \{0, 1\}\} \\ C^{le} &= \{(c_{3,3}^n = \neg x_{3,0}^{le} \vee \neg x_{3,1}^{le}), (c_{4,3}^n = \neg x_{4,0}^{le} \vee \neg x_{4,1}^{le}), \\ &\quad (c_{1,1}^{le} = x_{1,0}^{le} \vee x_{2,0}^{le}), (c_{1,2}^{le} = \neg x_{1,0}^{le} \vee \neg x_{2,0}^{le}) \\ &\quad (c_{2,1}^{le} = x_{1,0}^{le} \vee x_{3,1}^{le} \vee x_{3,0}^{le}), (c_{2,2}^{le} = \neg x_{1,0}^{le} \vee x_{3,1}^{le} \vee \neg x_{3,0}^{le}), \dots\}, \end{aligned}$$

4.3 Support Encoding

The support encoding approach utilizes the same mechanism as direct encoding for transforming variables and domains, but models the original constraints with clauses that represent the valid tuples. Further more, support encoding is only applicable on binary CSPs. After the definition follows an example that shows the application of the support encoding on the dual CSP P^{dual} of P_1 .

Definition 5 (Support encoding). Given a CSP $P = (X, D, C)$ we define its support encoding to be $P^{se} = (X^{se}, D^{se}, C^{se})$ where X^{se} and D^{se} are created like X^{de} and D^{de} in the direct encoding. C^{se} is the set of constraints over the newly created boolean variables X^{se} .

For each pair of variables $x_i, x_j \in X$ with $i \neq j$ of each original constraint $c = (X, T)$, $X = \{x_1, \dots, x_n\}$, $(x_i, x_j \in \text{scope}(c) = X)$ and each value $v \in D_i$ a new clause $\neg x_{i,v} \vee \bigvee_{w \in A} x_{j,w}$ is created. The set $A \subseteq D_j$ contains all values $v_k \in D_j$ which x_j can be instantiated to, such that at least one assignment $x_i = v$ and $x_j = v_k$ exists and satisfies c .

For each original variable $x_i \in X$ also the *addLeastOne* $\bigvee_{v \in D_i} x_{i,v}^{de}$ and *atMostOne* clauses $\neg x_{i,v}^{de} \vee \neg x_{i,w}^{de}, \forall v, w \in D_i, i \neq j$ must be added too [Ge02].

Example 5 (Support encoding). Given is the binary CSP P^{dual} . First the variables X^d are transformed into boolean variables X^{se} . The resulting variables are $X^{se} = \{x_{1,1}^{se}, x_{1,2}^{se}, x_{2,1}^{se}, x_{2,2}^{se}, x_{2,3}^{se}, x_{3,1}^{se}, x_{3,2}^{se}, \dots\}$, where each assignment of 1 to a variable $x_{i,j}^{se}$ represents, that the dual variables x_i^d gets assigned the j -th value of its domain. Thus, $x_{3,2}^{se}$ being assigned the value 1 represents the dual variable x_3^d being assigned the tuple $(0, 1, 0, 1)$.

For each variable $x^d \in X^d$ of the dual problem the *addLeastOne* and *atMostOne* clauses are created. For the variable x_2^d the resulting clauses are $x_{2,1}^{se} \vee x_{2,2}^{se} \vee x_{2,3}^{se}$, as well as $\neg x_{2,1}^{se} \vee \neg x_{2,2}^{se}$, $\neg x_{2,1}^{se} \vee \neg x_{2,3}^{se}$ and $\neg x_{2,2}^{se} \vee \neg x_{2,3}^{se}$. The clauses for x_1^d and x_3^d are created analogously.

Finally the constraints C^d need to be transformed. The original domains $D_i^d \in D^d$ are replaced by the domains $D'_i \in D'$, which only contain the indices of the domain values instead of the actual domain values ($D'_i = \{0, 1, \dots, |D_i^d| - 1\}, \forall i \in \{1, \dots, 3\}$). The constraint $c_{1,2}^d$ is transformed exemplary as follows. For each value $v \in D'_1$ of the dual value pair x_1^d, x_2^d the clause $\neg x_{1,v}^{se} \vee \bigvee_{w \in A} x_{1,w}^{se}$ is created. The resulting clauses are $c_{1,1} = \neg x_{1,0}^{se} \vee x_{2,0}^{se} \vee x_{2,1}^{se}$ and $c_{1,2} = \neg x_{1,1}^{se} \vee x_{2,2}^{se}$. Since both directions need to be taken into account, we also need to generate clauses for the variable pair x_2^d, x_1^d . This gives the clauses $c_{1,3} = \neg x_{2,0} \vee x_{1,0}$, $c_{1,4} = \neg x_{2,1} \vee x_{1,0}$ and $c_{1,5} = \neg x_{2,2} \vee x_{1,1}$.

4.4 More SAT Encodings

Additionally to the previously discussed encoding mechanisms there are more known transformations, that cannot be discussed in detail here. Following we provide brief summaries as well as references.

Minimal Support Encoding. The minimal support encoding was defined in [Ar08] and transforms a CSP into a SAT problem. Generally, the approach follows the methods of support encoding, the difference being that for each constraint c over the variables $\text{scope}(c) = \{x_1, x_2\}$ the support clauses are only given for one variable, x_1 or x_2 .

Regular Encoding. Another method to transform binary CSPs into boolean CSPs is regular encoding [AM04]. The main concept behind regular encoding is to replace each variable $x_i \in X$ of the original CSP with its domain being $D_i = \{0, \dots, m\}$ with a unary vector of boolean variables $\vec{U}_i = [x_{i,1}^{re}, \dots, x_{i,m}^{re}]$. Each vector \vec{U}_i is restricted so that a variable $x_{i,j}^{re}$ can only become 1 if all variables $x_{i,k}^{re}$ with $k < j$ also become 1. Each of these variables $x_{i,j}^{re}$ is called a regular variable.

Therefore each vector U_i takes the form $[1, 1, \dots, 1, *, *, *, \dots, *, 0, 0]$, where 1 represents variables that are assigned the value 1, * represents variables that weren't assigned a value yet, and 0 represents variables that were assigned 0. The number of 1s in \vec{U}_i needs to be equal to $d_i \in D_i$, which corresponds to the variable x_i .

A survey of the previously discussed encodings and further methods can be found in Chapter 4 of "Bridging Constraint Satisfaction and Boolean Satisfiability" [Pe15b].

Advantages and Disadvantages of SAT Transformations. All introduced SAT transformations have the *advantage*, that they allow the use of very well researched and fast SAT solvers. Also all transformations have the *disadvantage*, that they need much time for the transformation. This follows from the fact that in all approaches either all allowed tuples or all disallowed tuples must be processed. Even though a lot of information about the original CSP instance is usually lost during the translation stage and a large set of propositional clauses is produced, often needing much time in the process, SAT solvers sometimes outperform conventional CSP solvers [Pe15b].

It follows a discussion of the differences of the introduced transformations. Comparing support encoding with minimal support encoding, the first needs more constraints than the latter, but reaches arc-consistency if unit propagation is enforced on the transformed CSP. Logarithmic encoding needs less boolean variables than the other approaches, which is effectively noticeable if the domain sizes of the original CSP are very big. On the other hand, logarithmic encoding needs many more literals in the clauses, which slows down the solving speed of the SAT solver. The direct encoding approach and support encoding are usually inefficient in practice, since they often produce very large sets of clauses [Pe15b]. Thus, sometimes it is promising to use regular encoding which can represent equalities and inequalities with less clauses than the previously mentioned methods.

On the other hand, direct and logarithmic encoding both need to handle every tuple which does not satisfy a constraint, while support encoding needs to handle every tuple which does satisfy a constraint. Thus, direct and logarithmic encoding are more promising if the number of satisfying tuples of the constraints is small, while support encoding looks promising if the number of satisfying tuples of the constraints is small.

In contrast to support encoding, minimal support encoding does not reach arc-consistency if unit propagation is enforced. On the other hand, it can reduce the number of clauses. In the

end, the decision which transformation is the best one for a CSP depends on the structure, properties and connections of the used constraints.

5 Transformations into table and regular CSPs

For the transformation of a CSP $P = (X, D, C)$ into a table CSP or regular CSP every admissible tuple of each constraint must be listed and transformed. However, an important difference to the previous approaches is that now a constraint or subset of constraints is substituted by one single constraint, which reaches *generalized arc consistency (GAC)*. This can strengthen the propagation without changing the underlying solver.

The transformation of CSPs into table CSPs is called **tabulation**. Let be T_j the tuple list, which contains the tuples which satisfy the constraint $c_j = (X_j, T_j) \in C$. The constraint $table(X_j, T_j)$ (cf. Sect. 2) is the tabulation of the constraint c_j . So you can transform all constraints in C to receive a table CSP [Ge07].

The transformation of CSPs into regular CSPs is called **regularization**. Analogously to tabulation, let be T_j the tuple list, which contains the tuples which satisfy the constraint $c_j = (X_j, T_j) \in C$. A deterministic finite automaton (DFA) M_j can be created from the tuple list T_j . The constraint $regular(X_j, M_j)$ (again cf. Sect. 2) is the regularization of the constraint c_j . So you can transform all constraints in C to receive a regular CSP [LLH19].

Advantages and Disadvantages of table and regular Transformations. Both approaches have the *advantage*, that no prior transformations are needed. Because the table and regular constraint both reach generalized arc consistency (GAC), the regularization and the tabulation can increase the consistency level, if the constraint(s) to be substituted has/have a lower consistency level. Further advantages are, that no variables or domains must be transferred and it is not necessary to transform the whole CSP (partial transformations are possible). Thus, slow parts of a CSP can be transformed selectively. If the complete CSP is transformed a specialized table or regular solver can be used.

But both approaches in their basic version have the same disadvantage (as well as the aforementioned approaches), as all solutions of all constraints must be listed, which can be very time consuming. The regularization has one big advantage over every other here discussed transformation approach: For many global constraints there are direct transformations into regular constraints. For example the count constraint in our example CSP P_1 can be substituted directly by an DFA M_3 as represented in Figure 1, which can be created as explained in [De15; LLH18]. Using direct transformations allows us to avoid the enumeration of all admissible tuples of a constraint and, thus, leads to much faster transformation.

Furthermore, when comparing tabulation with regularization, tabulation has the advantage that a tabular constraint propagates faster if regularization can not find a small and compact

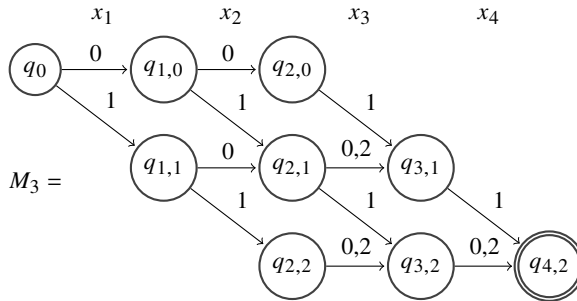


Fig. 1: The DFA M_3 which is equivalent to the count constraint c_3 of the original CSP P_1 .

DFA. On the other hand, if regularization finds a small and compact DFA, then regularization is applicable for more and bigger CSPs and propagates faster than a tabulated CSP.

6 Summary and Future Work

We presented and compared different transformations for the conversion of arbitrary CSPs into special CSPs, illustrated these by examples and discussed advantages and disadvantages.

The use of transformations usually leads to an acceleration of the search for a solution (due to the use of specific solvers or propagation algorithms). However the time needed for the transformations must be considered as well. Often the time needed for the transformation exceeds the time savings, that specialized algorithms can achieve. However, for certain CSPs the transformations can lead to considerable time savings. The following questions arise for future research: 1) How can the transformations be sped up? 2) How can one determine ahead of time, which transformations are reasonable for a certain CSP?

The direct transformation of global constraints into respective target constraints could be a solution with regard to the first question. For the transformation into regular CSPs exist direct transformations, which avoid the processing of all valid and invalid constraints. In many cases, this can reduce the time needed for the transformation dramatically.

Machine learning could be one approach for the recognition of a suitable transformation for a given CSP (question 2). In [LBH21], we show first promising steps towards an automated prediction, as to whether the original constraint or the transformation into a regular or tabular constraint is more promising.

References

- [AM04] Ansótegui, C.; Manyà, F.: Mapping Problems with Finite-Domain Variables into Problems with Boolean Variables. In: SAT 2004 - The 7th International

- Conference on Theory and Applications of Satisfiability Testing, Vancouver, BC, Canada, Online Proceedings. 2004.
- [Ar08] Argelich, J.; Cabiscol, A.; Lynce, I.; Manyà, F.: Encoding Max-CSP into Partial Max-SAT. In: 38th IEEE International Symposium on Multiple-Valued Logic (ISMVL), Dallas, Texas, USA. IEEE Computer Society, pp. 106–111, 2008.
- [Ba02] Bacchus, F.; Chen, X.; van Beek, P.; Walsh, T.: Binary vs. non-binary constraints. *Artif. Intell.* 140/1/2, pp. 1–37, 2002.
- [De03a] Dechter, R.: Consistency-Enforcing and Constraint Propagation. In: *Constraint processing*. Elsevier Morgan Kaufmann, chap. 3, pp. 51–84, 2003.
- [De03b] Dechter, R.: *Constraint processing*. Elsevier Morgan Kaufmann, 2003.
- [De15] Demassey, S.: Global Constraint Catalog, <http://sofdem.github.io/gccat/>, 07.04.2015, 2015.
- [DP89] Dechter, R.; Pearl, J.: Tree Clustering for Constraint Networks. *Artif. Intell.* 38/3, pp. 353–366, 1989.
- [Ga07] Gavarnelli, M.: The Log-Support Encoding of CSP into SAT. In: *Principles and Practice of Constraint Programming - CP, 13th International Conference*, Providence, RI, USA, September 23-27, Proceedings. Pp. 815–822, 2007.
- [Ge02] Gent, I. P.: Arc Consistency in SAT. In (van Harmelen, F., ed.): *Proceedings of the 15th European Conference on Artificial Intelligence, ECAI, Lyon, France*. IOS Press, pp. 121–125, 2002.
- [Ge07] Gent, I. P.; Jefferson, C.; Miguel, I.; Nightingale, P.: Data Structures for Generalised Arc Consistency for Extensional Constraints. In: *Twenty-Second AAAI Conference on Artificial Intelligence, Vancouver, British Columbia, Canada*. AAAI Press, pp. 191–197, 2007.
- [HK06] van Hoes, W.-J.; Katriel, I.: Global Constraints. In: *Handbook of Constraint Programming*. First, Chapter 6, Elsevier, Amsterdam, 2006, ISBN: 978-0-080-46380-3.
- [HPB04] Hellsten, L.; Pesant, G.; van Beek, P.: A Domain Consistency Algorithm for the Stretch Constraint. In (Wallace, M., ed.): *Principles and Practice of Constraint Programming - CP*. Vol. 3258. LNCS, Springer, pp. 290–304, 2004.
- [HU79] Hopcroft, J. E.; Ullman, J. D.: *Introduction to Automata Theory, Languages and Computation*. Addison-Wesley, 1979, ISBN: 0-201-02988-X.
- [IM94] Iwama, K.; Miyazaki, S.: SAT-Variable Complexity of Hard Combinatorial Problems. In: *Technology and Foundations - Information Processing, Volume 1, IFIP 13th World Computer Congress, Hamburg, Germany*, pp. 253–258, 1994.
- [LBH21] Löffler, S.; Becker, I.; Hofstedt, P.: ML-based Decision Support for CSP Modelling with Regular Membership and Table Constraints. In (Rocha, A. P.; Steels, L.; van den Herik, H. J., eds.): *13th International Conference on Agents and Artificial Intelligence, ICAART, Volume 2*. SCITEPRESS, 2021.

- [LLH18] Löffler, S.; Liu, K.; Hofstedt, P.: The Regularization of CSPs for Rostering, Planning and Resource Management Problems. In: Artificial Intelligence Applications and Innovations - 14th IFIP WG 12.5 International Conference, AIAI 2018, Rhodes, Greece, Proceedings. Pp. 209–218, 2018.
- [LLH19] Löffler, S.; Liu, K.; Hofstedt, P.: The Regularization of Small Sub-Constraint Satisfaction Problems. In (Hofstedt, P.; Abreu, S.; John, U.; Kuchen, H.; Seipel, D., eds.): Declarative Programming and Knowledge Management - Conference on Declarative Programming, DECLARE 2019, Unifying INAP, WLP, and WFLP, Cottbus, Germany, Revised Selected Papers. Vol. 12057. LNCS, Springer, pp. 106–115, 2019.
- [Ma98] Marriott, K.: Programming with Constraints - An Introduction. MIT Press, Cambridge, 1998, ISBN: 978-0-262-13341-8.
- [Pe01] Pesant, G.: A Filtering Algorithm for the Stretch Constraint. In (Walsh, T., ed.): Principles and Practice of Constraint Programming - CP. Vol. 2239. LNCS, Springer, pp. 183–195, 2001.
- [Pe04] Pesant, G.: A Regular Language Membership Constraint for Finite Sequences of Variables. In (Wallace, M., ed.): Principles and Practice of Constraint Programming - CP. Vol. 3258. LNCS, Springer, pp. 482–495, 2004.
- [Pe15a] Petke, J.: Bridging Constraint Satisfaction and Boolean Satisfiability. Springer, 2015, ISBN: 978-3-319-21809-0.
- [Pe15b] Petke, J.: SAT encodings. In: Bridging Constraint Satisfaction and Boolean Satisfiability. Artificial Intelligence: Foundations, Theory, and Algorithms, Chapter 4, Springer, 2015, ISBN: 978-3-319-21809-0.
- [Pe60] Peirce, C. S. (S.; Hartshorne, C.; Burks, A. W. (W.; Weiss, P.: Collected papers of Charles Sanders Peirce / edited by Charles Hartshorne and Paul Weiss. Belknap Press of Harvard University Press, Cambridge, 1960.
- [Pr09] Prestwich, S. D.: CNF Encodings. In (Biere, A.; Heule, M.; van Maaren, H.; Walsh, T., eds.): Handbook of Satisfiability. Vol. 185, Frontiers in Artificial Intelligence and Applications, IOS Press, pp. 75–97, 2009.
- [RBW06] Rossi, F.; Beek, P. v.; Walsh, T.: Handbook of Constraint Programming. Elsevier, Amsterdam, 2006, ISBN: 978-0-080-46380-3.
- [RPD90] Rossi, F.; Petrie, C. J.; Dhar, V.: On the Equivalence of Constraint Satisfaction Problems. In: 9th European Conference on Artificial Intelligence, ECAI 1990, Stockholm, Sweden, 1990. Pp. 550–556, 1990.
- [SF94] Sabin, D.; Freuder, E. C.: Contradicting Conventional Wisdom in Constraint Satisfaction. In: Eleventh European Conference on Artificial Intelligence, Amsterdam, The Netherlands. Pp. 125–129, 1994.
- [Wa00] Walsh, T.: SAT v CSP. In: Principles and Practice of Constraint Programming - CP, 6th International Conference, Singapore, Proceedings. Pp. 441–456, 2000.

Improving a Rule-based Fraud Detection System with Classification Based on Association Rule Mining

Michaela Baumann¹

Abstract: The detection of fraudulent insurance claims is a great challenge for insurance companies. Although the detection possibilities are getting better and better, fraudsters do not hesitate also using newer and more sophisticated methods. Apart from establishing new fraud detection systems, also the existing systems need to be updated and improved as best as possible. One common detection system is a rule-based expert system that checks predefined rules and gives alerts when certain conditions are met. Usually, the rules are treated separately and correlations within the rules are considered insufficiently. The work at hand describes how the classification based on association rule mining is used for improving such rule-based systems by bringing in relations between pairs of rules. The rule weights are determined through a genetic optimizer.

Keywords: Insurance fraud detection; Association rule mining; Expert system; Genetic optimizer; Classification

1 Introduction

The business model of insurance companies is the financial hedging of certain risks of their customers. The compensation in the collective, i.e., that many customers pay a relatively small amount of money and get a compensation in the event of a claim, makes this business worthwhile [Fa11, Ko13]. However, in all the working areas of an insurer, e.g., when designing products or calculating premiums, the notion of moral risk resp. moral hazard is present [He89]. The availability of an insurance may be disincentiving and delude an insured to take risks he or she would not have taken without the insurance [DB00]. Moral hazard may occur when a certain information asymmetry is present, i.e., when the insured knows more about its own risk-taking than the insurer does [Hö79, Sh92]. Insurance fraud constitutes a certain kind of moral hazard whereby its definition is not fixed per se. A very broad definition includes the excessive, involuntary, or unnecessary use of insurance benefits [Sc04]. A stricter definition of fraud is that it implies a criminal act with a serious misrepresentation, the intention to deceive, and the objective of obtaining an unjustified benefit [DZ02, De02, ED03, VD04]. Apart from the fact that a watertight proof of fraud is rarely possible [VD04], it is clear that insurance fraud leads to serious problems on the insurance market irrespective of the definition or type of fraud. The German insurance association (“Gesamtverband der Deutschen Versicherungswirtschaft”, GDV) estimates the

¹ NÜRNBERGER Versicherung, Anwendungsentwicklung BIACC, Ostendstraße 100, 90334 Nürnberg, Germany
michaela.baumann@nuernberger.de

damage caused by insurance fraud at 5 billion ($5 \cdot 10^9$) EUR in 2018 in Germany with about 10 % of all claims assessed as dubious [Fr18]. The damage is thereby not only a problem of the insurance company. Fraudulent claims increase the amount of claimed money which is passed on to all policyholders through an increase of the premiums or a reduction of the insurance coverage. However, dissatisfied or annoyed customers usually only blame the insurance company since the principle of insurance is often not clear to them [Kö15, VD04]. It is therefore necessary for insurance companies to detect and prevent insurance fraud as best as possible to compete on the market and indirectly maintain the (hopefully) good image of the company. However, this is a difficult task and resembles a tightrope walk since too much mistrust on the part of the insurers or even a general suspicion towards the insured has an extremely negative effect on the insurers' image. This is why fraud detection receives a lot of attention and especially during the last years, new methods from the fields of big data, machine learning, and AI have found their way into this field. The work at hand focuses on insurance fraud, more specifically on insurance fraud at claim time, and contributes to the efforts of fraud detection by showing a way of how a so-called expert system can be improved with a relatively little expense using association rule mining. Thereby, association rule mining is used to find correlations between existing rules of the expert system. Considering these correlations and adding them in form of additional rules to the original rule set might allow for a more accurate classification of insurance claims.

The paper is organized as follows: Section 2 gives a non-exhaustive overview of current methods for fraud detection. Section 3 shows the structure and the functioning of an expert system (Section 3.1), how classification with association rule mining works in general (Section 3.2), and how it is applied to the rule-based expert system (Section 3.3). In Section 4, the association rule mining is applied on a real-world expert system and it is shown how this technique can improve its expressiveness. Section 5 concludes the paper.

2 Related Work on Fraud Detection Methods

The ways of fraudulent behavior in the insurance sector as well as, e.g., credit card fraud, telecommunication fraud, or money laundering are manifold, as have to be the means to detect or prevent such behavior [BH02, Ko04, AMZ16]. When focusing on fraud at claim time, there are some general difficulties related to fraud detection. One difficulty is that fraud is not self-revealing [VD04], meaning that fraud has to be actively looked for and it has to be detected as early as possible. Furthermore, and this is especially important when trying to use supervised methods for fraud detection, it is difficult to actually prove fraud. Most target variables in supervised learning techniques flag suspicious claims, but not necessarily fraudulent claims. Methods for unsupervised fraud detection circumvent this problem and are mostly based on outlier detection and profiling [BH02].

One possibility of detecting claim fraud is to use rule-based expert systems. Such systems match every claim with a list of predefined indicators and when a claim shows a certain set of indicators, it is marked as suspicious [De02, VD04]. The construction of such systems

usually requires labeled data or reliable expert assessments and the rules themselves are of the form *If (condition) then (consequent)* [BH02]. Rule-based expert systems are, for example, used for detecting superimposed fraud cases in telecommunications networks [Hi09], for detecting computer (network) intrusions [GL91], for detecting anomalies in healthcare insurance claims [SA13], or for alerting fraud in the field of consumer credits [Le95]. An automatic re-engineering of rule-based expert systems may be conducted with cultural algorithms [SR97]. In such a setting, every rule is associated with a certain score, i.e., the rules are of the form *If (condition i) then add p_i to the fraud score* and the sum of the scores is a measure for the fraud likelihood. Apart from rule-based expert systems, there are also supervised machine learning methods for classifying claims into fraudulent and non-fraudulent ones such as logistic regression, k-nearest neighbors, decision trees, Bayesian neural networks, or support vector machines (SVMs) for classification [Vi02]. In particular, there exist neural networks for detecting management fraud [GC97, FCS95], neural and Bayesian networks for detecting fraud in communications networks [Ta98], decision trees and SVMs for detecting credit card fraud [SD11], or data mining models for detecting anomalies in healthcare insurance claims [SA13]. Fraud in the automobile insurance market is, e.g., predicted with (nested) multinomial logistic regression models [AAG99]. For training supervised models, not only traditional numeric features may be used but also, e.g., text features extracted from accident descriptions via text processing algorithms such as Latent Dirichlet Allocation [WX18]. When assessing the algorithms' performance, a particular focus lies on the Type I (false positive) and Type II (false negative) error rates [GC97]. As the imbalance of the data, i.e., a disproportionate ratio of observations in each class, is a big problem when applying supervised machine learning methods, there also exists a bundle of unsupervised methods for fraud detection such as an interactive, unsupervised machine learning approach for detecting fraud in the healthcare insurance [K GK15] or the so-called Peer Group Analysis, an unsupervised method for fraud detection where similar objects and their behavior around a target sequence is used for predicting the expected behavior of the target, for analyzing time series data with respect to outliers in financial data [FM06, BH01]. With a similar approach time series data of mobile phone users can be analyzed for fraudulent cases [BST01]. With unsupervised spectral ranking techniques outliers, for instance auto insurance fraudsters, can be ranked by their degree of deviation [Ni16]. Other anomaly detection methods imply a network analysis, e.g., social network analysis for detecting internet auction fraud [Ch11], fraudulent credit card transactions [Va15], automobile insurance fraud [ŠFB11], or money laundering [FCR17]. In addition to systems working with the internal data of an organization, fraud detection can also be efficiently supported with centralized insurance fraud bureaus [De02]. In Germany, e.g., there exists the HIS ("Hinweis- und Informationssystem"), an indication and information system of the German insurance industry, which supports the detection and prevention of insurance fraud [iHG]. Such centralized bureaus allow a certain degree of information pooling, whereby the insurers do not have direct access to the data to preserve honest competition and to fulfill all privacy issues [VD04]. Independent of the specific fraud detection methods, relevant fraud detection components must be chosen and it needs to be decided how they are combined into an organization's fraud detection system [FVB11].

During all efforts for detecting or preventing fraud, the most important issue is to keep the balance between effort and outcome [De02], in particular, the relationship between the insurer's investment for preventing/detecting fraud and the effort costs of the policyholder for different behavior [Ok13].

3 Adjusting the Rule-based Fraud Detection System

As stated in Section 1, we improve the detection of insurance fraud by extending an already existing rule-based expert system with very little effort. For doing this, we first characterize the rule-based system and show the functioning of a classification based on association rule mining (CARM) before applying the CARM on the rule-based system.

3.1 Characterization of Rule-based Expert Systems and Problem Description

The expert system, which is improved hereafter, with a CARM is a rule-based system similar to that one described by Sternberg and Reynolds [SR97]. During the car claims process, relevant information, e.g., the time and place of the accident, is recorded. As soon as the information is stored, the rule-based system goes into action. The claim data is matched independently with a set of rules. Each rule can either be triggered, i.e., its conditions are met, or not. For example, the accident may have happened during the weekend (or not). Furthermore, each rule has a certain weight. For each claim, the weights of the triggered rules are aggregated and when this aggregation meets a specified threshold, further actions are to follow. For example, a person responsible should have a closer look on the claim.

Definition 1 (rule-based system). A rule-based system is a vector (τ, \mathfrak{w}) , where $\tau = (r_1, r_2, \dots, r_k)$ is a vector of rules with $1 \leq k \in \mathbb{N}$. Each rule r_i has a weight $w_i > 0$ given in the weight vector $\mathfrak{w} = (w_1, \dots, w_k)$. A rule is a function which evaluates a claim $c \in \mathcal{C}$ to $r_i(c) = 1$ when the rule is fulfilled, i.e., the condition is true, and to $r_i(c) = 0$ when it is not fulfilled. The whole claim assessment is done via an aggregation function g which depends on the weight vector \mathfrak{w} : $g_{\mathfrak{w}}(c) = g(r_1(c) \cdot w_1, \dots, r_k(c) \cdot w_k) \geq 0$. An alert is signalled when $g_{\mathfrak{w}}(c) \geq l$ for some threshold $l > 0$.

An example for a rule-based system consisting of four rules could be: (a) If there are exactly two cars involved, assign 10 to the fraud score. (b) If the accident happened during the weekend, assign 5 to the fraud score. (c) If there is exactly one witness, assign 5 to the fraud score. (d) If the same insured reported another claim in the six months before, assign 30 to the fraud score. Written in the notation of Definition 1, we would get the rule vector $\tau_{ex} = (\text{There are exactly two cars involved, The accident happened during the weekend, There is exactly one witness, The same insured reported another claim in the six months before})$ with the corresponding weight vector $\mathfrak{w}_{ex} = (10, 5, 5, 30)$. Note that with the notation above, the rule-based system evaluates the claims in two steps. In

the first step, it is checked whether the condition of a rule is fulfilled or not leading to a rule conclusion of either 0 or 1, i.e., we get Boolean conditions. In a second step, each conclusion is multiplied with its respective weight. This leads to the same result as if the weight would have been directly part of the rule conclusion, but it allows to change the weighting in re-engineering phases [SR97] without having to change the rules themselves. One possibility for the aggregation function g is the summation [SR97], which we further assume if not otherwise stated. Continuing the example and assuming a certain claim c_{ex} , the rules could evaluate to $\tau_{ex}(c_{ex}) = (1, 0, 0, 1)$. Using the summation for g this would lead to a fraud score of $g_{w_{ex}}(c_{ex}) = 10 + 0 + 0 + 30 = 40$.

The weights for the rules as well as the rules themselves try to reflect the experts' valuable knowledge of fraudulent behavior. However, depending on the specific rule formulations and the number of rules involved, the above definition does not ensure that dependencies between the various rule conditions are considered; a claim adjuster would most likely take into account such dependencies during a manual checking [VD04]. In the rule-based system, every rule has a certain fixed weight which is added to the aggregation independently of the fulfillment of the other rules with g being the summation. When looking at real-world claim data, we may encounter three different types of dependencies between every two rule conditions: (i) no dependence, i.e., the two conditions are truly independent, (ii) a reinforcing dependence, i.e., the fulfillment of the two rules at the same time is more severe than their aggregated single weights, or (iii) a weakening dependence, i.e., the two rules each have a certain base weight but their simultaneous occurrence is not problematic (it may even be the normal case).

This observation also implies that weights may not only be positive, strengthening the suspiciousness of a claim, but also negative, i.e., mitigating its suspiciousness. In the four-rule-example, it could be the case that when there are exactly two cars involved and there is exactly one witness, i.e., the two rules are true at the same time, this is a more suspicious situation than the $10 + 5$ fraud score may suggest. On the other hand, exactly one witness for an accident that happened during the weekend may not be as suspicious as the $5 + 5$ fraud score proposes. The question arises of how to bring rule dependencies into the model while taking account of a certain cost-efficiency. This problem may be divided into two sub-questions. The first one is: (1) *How can we detect meaningful dependencies among the set of all rules?* The second one tackles the issue of economical efficiency. Any change to the existing system should be avoided; in particular, the surrounding systems should not be affected by the adaptation. (2) *How can we add the additional information to the system without changing its current operating mode?* The answer to Question (2) is more or less simple: As we only have rules and weights in the system, the additional information has to be added as either rules or weights. Notice that adjusting the function g towards a more involved function is not practical, since every time a rule is added or a weight is changed, the function g would otherwise have to be entirely redetermined, see Question (2). Further, the interpretability of the resulting function cannot be guaranteed when using sophisticated functions g , violating Question (1). Multi-level checks like "If the outcome of the first rule

is this, then check this in the next round, if the outcome is that, then check something else” would require a change of the system’s operating mode, which we want to avoid if possible. As weights are fixed for their respective rules, the information about dependencies between existing rules has to be added in form of new rules. Thereby, it is a straightforward way to establish rule combinations, i.e., two (or more) rules each are combined to new rules with an “and-link”(conjunction). The definition of a binary combination is:

Definition 2 (binary rule combinations). For a given rule vector \mathbf{r} and claims $c \in \mathcal{C}$ we define a binary rule combination as $\tilde{r}_{i,j} : \mathcal{C} \rightarrow \{0, 1\}$, $0 < i < j \leq k$, with $\tilde{r}_{i,j}(c) = r_i(c) \cdot r_j(c)$. The extended rule vector that contains all possible binary combinations is $\tilde{\mathbf{r}} = (r_1, \dots, r_k, \tilde{r}_{1,2}, \dots, \tilde{r}_{1,k}, \tilde{r}_{2,3}, \dots, \dots, \tilde{r}_{k-1,k})$.

With the definition above it is analogously possible to generate ternary and higher degree rule combinations. However, due to complexity and computational issues, we stick to binary combinations for the moment. When considering a rule vector \mathbf{r} of k rules, adding every possible binary combination would lead to additional $\binom{k}{2} = \frac{k(k-1)}{2}$ rules, of which many would probably not be meaningful. The identification of the meaningful ones is exactly the problem stated in Question (1). For finding them, we use the CARM explained next.

3.2 Classification Based on Association Rule Mining

Association rule mining is a very common technique in data mining, mainly originating from market basket analysis where predications like “When a customer buys products A and B, it is likely (with a probability of x %) that the customer also buys product C” are to be found [HGN00]. Such rules are derived from past purchases of all customers and then used to propose active customers further products based on their previous purchases or their current shopping cart. The single products are called *items* and the previous purchases are called *transactions*. Two major problems occur when applying association rule mining: first, the number of all possible rules grows exponentially with the number of items, and second, from the set of all mined rules the interesting ones have to be selected [HGN00]. Various algorithms exist to cope with these problems [ZZ02].

In more detail, the formal definition of association rule mining according to [AIS93] is the following: Let $\mathcal{I} = \{i_1, \dots, i_n\}$ be the set of all items. A subset of \mathcal{I} is called *itemset*. The set of all transactions is $\mathcal{T} = \{t_1, \dots, t_m\} = \{(tid_1, A_1), \dots, (tid_m, A_m)\}$ with tid_i being transaction identifiers and $A_i \subseteq \mathcal{I}$ being itemsets, $i = 1, \dots, m$. A *rule* is an implication $X \Rightarrow Y$ with $X, Y \subseteq \mathcal{I}$. Usually, only rules with $X \cap Y = \emptyset$ are of interest. To limit the number of rules to the interesting ones there exist two basic concepts. The *support* of a rule $X \Rightarrow Y$ is given through

$$\text{supp}(X \Rightarrow Y) := \text{supp}(X \cup Y) := \frac{|\{t \in \mathcal{T} | X \cup Y \subseteq t\}|}{|\mathcal{T}|} =: \frac{t_{X,Y}}{|\mathcal{T}|}$$

where $X \cup Y \subseteq t$ if $t = (tid, A)$ and $X \cup Y \subseteq A$. Interesting rules need to have a minimum support, i.e., their left and right hand sides need to appear together often enough. Rules that

do not have enough support are considered irrelevant and not generalizable. The second concept is that of *confidence* and measures the reliability of a rule, i.e., how often the rule holds given the left hand side is true:

$$\text{conf}(X \Rightarrow Y) := \frac{\text{supp}(X \cup Y)}{\text{supp}(X)} = \frac{|\{t \in \mathcal{T} | X \cup Y \subseteq t\}|}{|\{t \in \mathcal{T} | X \subseteq t\}|} =: \frac{t_{X,Y}}{t_X}$$

Interesting rules need to have a minimum confidence, too. There exist further measures for filtering out interesting rules, like interest or lift [Br97, HGH05]. The integration of classification and association rule mining is an approach to transfer the actually unsupervised data mining method of finding arbitrary association rules to a supervised method for finding rules for a classification [LHM98, CLZ05]. In the classification setting, we have labeled data where the class attribute together with its values is a subset of the set of items and in all transactions there is the class attribute together with exactly one of its values. During the rule generation process only a subset of all association rules is selected, namely those rules with only the class attribute on the right-hand side, i.e., the rules are of the form $X \Rightarrow \{\text{label}\}$ where “label” denotes the class attribute together with one of its values ($\text{label} \in \mathcal{I}$). This set of rules is referred to as class association rules (CARs) [LHM98]. The support and confidence are considered here as well. The generation of the classifier itself is not relevant for this paper as we are interested solely in the rules.

3.3 Extending the Rule-based Fraud Detection System

For extending the rule-based expert system (τ, \mathfrak{w}) with binary rule combinations according to Definition 2, we define $\mathcal{I} = \mathcal{I}_{\text{rules}} \cup \mathcal{I}_{\text{label}}$ with $\mathcal{I}_{\text{rules}} = \{r_1 = 0, r_1 = 1, \dots, r_k = 1\}$, $\mathcal{I}_{\text{label}} = \{\text{fraud} = 0, \text{fraud} = 1\}$, fix a support level and a confidence level, and take only those rules into account that fulfill both the confidence and the support level. Note that one claim is one transaction t . We further narrow the CARs to those with only “= 1” rule items and with at most two rule items, both on the left-hand side of the rules (the latter because of runtime issues). For instance, we have a set containing the following CARs:

- $\{\text{There are exactly two cars involved} = 1, \text{There is exactly one witness} = 1\} \Rightarrow \{\text{fraud} = 1\}$
- $\{\text{The same insured reported another claim in the six months before} = 1, \text{There is exactly one witness} = 1\} \Rightarrow \{\text{fraud} = 1\}$
- $\{\text{The same insured reported another claim in the six months before} = 1\} \Rightarrow \{\text{fraud} = 1\}$
- $\{\text{The accident happened during the weekend} = 1, \text{There is exactly one witness} = 1\} \Rightarrow \{\text{fraud} = 0\}$

Additionally, we delete those rules of the binary ones whose left-hand side is a superset of the left hand side of a CAR with the same consequent to further reduce the number of rules and to concentrate on the meaningful ones and prevent overfitting (superset criterion). In the above example, the second association rule would be deleted according to this criterion. The superset criterion is a heuristic approach proposed for pruning rule generation trees [CLZ05]. In our case, we can motivate it as follows: Assume two association rules $X_1 \Rightarrow Y$ and $X_2 \Rightarrow Y$ with $X_1 \subset X_2$, in our case $|X_1| = 1$ and $|X_2| = 2$ when regarding only unary and binary association rules, then it clearly holds that $\text{supp}(X_1 \Rightarrow Y) \geq \text{supp}(X_2 \Rightarrow Y)$, i.e., the more general rule $X_1 \Rightarrow Y$ has a higher support. For the confidence there is no such

inequality, however, when doing classification and the rule $X_1 \Rightarrow Y$ fulfills the minimum confidence requirement, i.e., from X_1 we can conclude the target, then the (single) rule on the left hand side is a strong indicator for the respective target value (fraud = 0 or fraud = 1). In the context of fraud detection this is usually only possible for very specific rules (provided that the minimum confidence is relatively high) with a very small support where a more specialized rule would not be meaningful. This is not a contradiction to the motivation of our approach of finding correlating fraud conditions that provide additional information to better classify the claims.

After having extended the rule vector τ to $\bar{\tau}$ (with all initial expert rules and only meaningful combinations), we have to extend and adjust the weight vector ω . The new weight vector $\bar{\omega}$ contains the weights for the old rules as well as the weights for the rule combinations. The alert threshold $l > 0$ may be maintained. For the set of all claims, the weights in $\bar{\omega}$ need to be optimal related to a fitness criterion so that only suspicious claims have an aggregated value above l . This is, of course, unlikely to be feasible, but the separation of suspicious and non-suspicious claims should be as good as possible while avoiding overfitting. One way of how to assign optimal weights is the usage of cultural algorithms [SR97], which are a specific kind of evolutionary algorithms. For weighting attributes for a k -nearest neighbors classification a similar, evolutionary optimization approach can be applied [KJD91]. Evolutionary algorithms are especially useful for large-scale optimization problems [EHG05]. The fitness criterion resp. objective function has to make sure that the number of correctly classified claims is maximal, e.g., it can be maximized using accuracy measures like the numbers of true positives and true negatives. Regarding the field of insurance fraud, it is desirable to have no false negatives, i.e., that all fraud attempts are discovered. At the same time, false positives cause huge costs for an insurer as the manual inspection of claims is time-consuming and therefore expensive. Depending on the rule base and its extent, it may be impossible to detect all attempts of fraud in the claim data. Hence, it is of special interest to reduce the false positives when extending an existing rule-based detection system while not worsening the false negatives. Such a focus can also be considered in the objective function.

4 Evaluating the Approach on a Real-World Data Set

For evaluating the improvement approach we consider a real-world data set of motor insurance claims from a German insurer. The data set consists of about 140,000 claims covering several years. An already existing rule-based expert system approximately working as stated in Definition 1 consists of about 110 unary, weighted rules. The data has been manually labeled during the claims' regular processing. However, this label only marks a certain suspiciousness, not a fraud. The data is highly unbalanced, which is very typical for fraud data as most of the claims an insurer gets are justified ones. Less than 5 % of the cases are marked as positive (suspicious). Thus, before conducting the algorithm for a CARM, we have to correct the imbalance in the data.

A first idea was to add further positive cases which are older than the original time horizon. Although this proved very promising, we omitted this possibility because of several problems. Some of the rules are not applicable to the old data, such as the rule “An electric car is involved” and would always evaluate to 0 in the old data leading to biased results. Furthermore, also the enriched data set would still have shown a great imbalance. Instead, we checked for several common sampling methods. When performing case generating algorithms like SMOTE [Ch02] or ROSE [MT14], these algorithms may actually deteriorate the results for extreme imbalances [Kr16]. When we performed such case generating algorithms we detected that technically impossible cases (like one contract is of two different types at the same time) were created. We also tried non-random undersampling methods like the OSS algorithm [KM97], which uses so-called Tomek links [To76] to remove redundant rows from the data. But the different non-random undersampling methods we tried only removed fewer than 1 % of the majority class cases and were, thus, not useful. This is why we decided to apply a combination of random over- and undersampling [LMT14] with a target proportion of 45:55 (minority class:majority class), which proved to be suitable after having tried different proportions in an extensive study, although the data set is likely to get biased [Kr16].

After having adjusted the class sizes, we perform the CARM using the R-package “arulesCBA” [JHG20]. We set the confidence to 90 % and the support to 0.2 % to get enough classification rules for the relatively high confidence value. We explicitly mined rules with one and two attributes on the left-hand side and deleted the binary ones fulfilling the superset criterion. This resulted in a total of 54 new rules, where 35 were positive rules, i.e. rules with a positive consequent (suspicious), and 19 negative rules (not suspicious). Considering the number of rows affected by the additional rules, we see that about 10.99 % of the claims fulfill at least one of the additional rules.

To compare the old and the extended rule base, we had to add weights for the new rules and analyze the performance of the two systems. The weights were determined using a genetic algorithm. To create equal conditions, we also adjusted the weights of the original rule-based system with the same genetic algorithm so that we can state that differences in the accuracy really stem from the new binary rules and not from a non-optimal weight assignment in the original system. For the optimization, we use the R-package “GA” [Sc19] with the original weights as one of the initial populations. For the weights, we allow values between b_{min} and b_{max} with $0 > b_{min} \in \mathbb{Z}_-$ and $0 < b_{max} \in \mathbb{Z}_+$. Since the optimization algorithm works on real values and not on integers, we take the floor of the generated weights [Sc]. As objective function to be maximized we chose $tpr^{w_{tpr}} * tnr^{w_{tnr}}$ where tpr is the rate of the true positives and tnr the rate of the true negatives. In order to place a higher importance on the true negatives resp. false positives, we set the weight for the true negatives to $w_{tnr} = 1 - w_{tpr} = 0.75$. Under this setting, we get the following results, that are summarized in Table 1: For the original rule set, the false positive rate is 12.85 % and the false negative rate is 57.78 %. The extended rule set gives a false positive rate of 10.68 % and a false negative rate of 58.27 %. This is a decrease of the false positive rate of about

16.89 % and an increase of the false negative rate of about 0.85 %. The false negative rate seems to be quite high, but considering how the rules are constructed and on what data they rely on, it was clear from the beginning that the rules are not able to catch all the fraudulent behavior. The main focus was on reducing the false positives (without increasing the false negatives, which we could not exactly achieve in our experiment) to save inspection costs and time and to prevent false suspicions. This is extremely important as the absolute number of false positives is times higher than the absolute number of false negatives. Considering this, the decrease in the false positive rate can be regarded as a success.

rule set	false positive rate	false negative rate
τ	12.85 %	57.78 %
$\bar{\tau}$	10.68 %	58.27 %

Tab. 1: Error rates for the original and the extended rule-based fraud detection system. The false positive error rate decreased by 16.89 % while the false negative rate increased by 0.85 %

As by construction, a genetic algorithm does not necessarily lead to the global optimum, we wanted to run the algorithm with different seeds and average the results, but this was prohibited by runtime issues in our experimental setting. However, it is particularly interesting that 82.86 % of the rules associated with a positive suspicion got on average a positive weight. The remaining 17.14 % got a weight of 0, so perhaps they have not been tested by the algorithm at all. 52.64 % of the rules associated with a negative suspicion got a negative weight, 36.84 % got a weight of 0, and 10.52 % a positive weight, but near 0. These results are summarized in Tab. 2. Thus, the rule weights and the right-hand side of the association rules are consistent for almost all cases, which confirms our approach.

target	positive weight	zero weight	negative weight
fraud = 1	82.86 %	17.14 %	0.00 %
fraud = 0	10.52 %	36.84 %	52.64 %

Tab. 2: Target value and sign of the weights. While none of the “fraud = 1” rule weights are not in line with the implication, there are 10.52 % of the “fraud = 0” rule weights not in line with the implication

5 Conclusions

This paper shows a method of how an existing rule-based expert system for fraud detection may be improved with relatively little construction effort. It uses the concept of association rule mining for classification to add meaningful new rules to the set of already existing rules. These new rules reflect dependencies between the old rules, which allows a more precise scoring of insurance claims. This approach is conducted on a real world example, for which the extended rule-based system yields satisfying results. Above all, the distinction between rules intended to strengthen the suspiciousness and those to weaken the suspiciousness could be observed after assigning new rule weights with help of a genetic optimizer. A faster method for the weight assignment would improve the feasibility of the approach. For its

evaluation, the rules resulting from the association rule mining algorithm have been taken over unmodified. Before, however, applying such an extension on a productive system, it is recommended that the fraud expert has a closer look at the automatically generated rules and at the assigned weights to estimate their meaningfulness, explainability, and quality. A particular focus should lie on the interplay of the original rules, the combinations and how the combinations increase or decrease the suspiciousness of the original rules.

Acknowledgement

The opinions expressed in this report are those of the author and not necessarily those of the author's employer. The author's employer does not guarantee the accuracy or reliability of the information provided herein. The author works as a Data Scientist at NÜRNBERGER Lebensversicherung AG, Nuremberg, Germany. The author thanks Michael H. Baumann (Universität Bayreuth) for valuable discussions and the referees for their helpful comments.

Bibliography

- [AAG99] Artís, Manuel; Ayuso, Mercedes; Guillén, Montserrat: Modelling different types of automobile insurance fraud behaviour in the Spanish market. *Insurance: Mathematics and Economics*, 24(1):67–81, 1999.
- [AIS93] Agrawal, Rakesh; Imieliundefinedski, Tomasz; Swami, Arun: Mining Association Rules between Sets of Items in Large Databases. *SIGMOD Rec.*, 22(2):207–216, June 1993.
- [AMZ16] Abdallah, Aisha; Maarof, Mohd Aizaini; Zainal, Anazida: Fraud detection system: A survey. *Journal of Network and Computer Applications*, 68:90–113, 2016.
- [BH01] Bolton, Richard J; Hand, David J: Unsupervised profiling methods for fraud detection. *Credit scoring and credit control VII*, pp. 235–255, 2001.
- [BH02] Bolton, Richard J.; Hand, David J.: Statistical Fraud Detection: A Review. *Statistical Science*, 17(3):235–249, 2002.
- [Br97] Brin, Sergey; Motwani, Rajeev; Ullman, Jeffrey D.; Tsur, Shalom: Dynamic Itemset Counting and Implication Rules for Market Basket Data. *SIGMOD Rec.*, 26(2):255–264, 1997.
- [BST01] Burge, Peter; Shawe-Taylor, John: An Unsupervised Neural Network Approach to Profiling the Behavior of Mobile Phone Users for Use in Fraud Detection. *Journal of Parallel and Distributed Computing*, 61(7):915–925, 2001.
- [Ch02] Chawla, Nitesh V; Bowyer, Kevin W; Hall, Lawrence O; Kegelmeyer, W Philip: SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16:321–357, 2002.
- [Ch11] Chiu, Chaochang; Ku, Yungchang; Lie, Ting; Chen, Yuchi: Internet Auction Fraud Detection Using Social Network Analysis and Classification Tree Approaches. *International Journal of Electronic Commerce*, 15(3):123–147, 2011.


- [CLZ05] Coenen, Frans; Leng, Paul; Zhang, Lu: Threshold Tuning for Improved Classification Association Rule Mining. In: *Advances in Knowledge Discovery and Data Mining*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 216–225, 2005.
- [DB00] Dembe, Allard E.; Boden, Leslie I.: Moral Hazard: A Question of Morality? *NEW SOLUTIONS: A Journal of Environmental and Occupational Health Policy*, 10(3):257–279, 2000.
- [De02] Derrig, Richard A.: Insurance Fraud. *The Journal of Risk and Insurance*, 69(3):271–287, 2002.
- [DZ02] Derrig, Richard A; Zicko, Valerie: Prosecuting insurance fraud – A case study of the massachusetts experience in the 1990s. *Risk Management and Insurance Review*, 5(2):77–104, 2002.
- [ED03] Ericson, Richard V.; Doyle, Aaron: The Moral Risks of Private Justice: The Case of Insurance Fraud. In: *Risk and Morality*, chapter 12, pp. 317–363. University of Toronto Press, 2003.
- [EHG05] Elbeltagi, Emad; Hegazy, Tarek; Grierson, Donald: Comparison among five evolutionary-based optimization algorithms. *Advanced engineering informatics*, 19(1):43–53, 2005.
- [Fa11] Farny, Dieter: *Versicherungsbetriebslehre*. VVW GmbH, 2011.
- [FCR17] Fronzetti Colladon, Andrea; Remondi, Elisa: Using social network analysis to prevent money laundering. *Expert Systems with Applications*, 67:49–58, 2017.
- [FCS95] Fanning, Kurt; Cogger, Kenneth O.; Srivastava, Rajendra: Detection of Management Fraud: A Neural Network Approach. *Intelligent Systems in Accounting, Finance and Management*, 4(2):113–126, 1995.
- [FM06] Ferdousi, Zakia; Maeda, Akira: Unsupervised Outlier Detection in Time Series data. In: *22nd International Conference on Data Engineering Workshops (ICDEW'06)*. 2006.
- [Fr18] Friedrich, Sara: , *Versicherungsbetrug – Du lügst!* <https://www.gdv.de/de/themen/positionen-magazin/du-luegst--39848>, 2018. Accessed: 2020-04-15.
- [FVB11] Furlan, Štefan; Vasilecas, Olegas; Bajec, Marko: Method for selection of motor insurance fraud management system components based on business performance. *Technological and Economic Development of Economy*, 17(3):535–561, 2011.
- [GC97] Green, Brian Patrick; Choi, Jae Hwa: Assessing the risk of management fraud through neural network technology. *Auditing*, 16:14–28, 1997.
- [GL91] Garvey, Thomas D; Lunt, Teresa F: Model-based Intrusion Detection. In: *Proceedings of the 14th national computer security conference*. volume 10, pp. 372–385, 1991.
- [He89] Heimer, Carol A: *Reactive Risk and Rational Action: Managing Moral Hazard in Insurance Contracts*. University of California Press, 1989.
- [HGH05] Hornik, Kurt; Grün, Bettina; Hahsler, Michael: arules – A computational environment for mining association rules and frequent item sets. *Journal of Statistical Software*, 14(15):1–25, 2005.

- [HGN00] Hipp, Jochen; Güntzer, Ulrich; Nakhaeizadeh, Gholamreza: Algorithms for association rule mining—a general survey and comparison. *ACM sigkdd explorations newsletter*, 2(1):58–64, 2000.
- [Hi09] Hilas, Constantinos S.: Designing an expert system for fraud detection in private telecommunications networks. *Expert Systems with Applications*, 36(9):11559–11569, 2009.
- [Hö79] Hölmstrom, Bengt: Moral Hazard and Observability. *The Bell Journal of Economics*, 10(1):74–91, 1979.
- [iHG] informa HIS GmbH: , Hinweis- und Informationssystem der Versicherungswirtschaft. <https://www.informa-his.de/his-online>. Accessed: 2020-05-19.
- [JHG20] Johnson, Ian; Hahsler, Michael; Giallanza, Tyler: . *arulesCBA: Classification Based on Association Rules*, 2020. R package version 1.1.6.
- [KGK15] Kose, Ilker; Gokturk, Mehmet; Kilic, Kemal: An interactive machine-learning-based electronic fraud and abuse detection system in healthcare insurance. *Applied Soft Computing*, 36:283–299, 2015.
- [KJD91] Kelly Jr, James D; Davis, Lawrence: A Hybrid Genetic Algorithm for Classification. In: *IJCAI*. volume 91, pp. 645–650, 1991.
- [KM97] Kubat, Miroslav; Matwin, Stan: Addressing the Curse of Imbalanced Training Sets: One-Sided Selection. In: *Icml*. volume 97. Nashville, USA, pp. 179–186, 1997.
- [Ko04] Kou, Yufeng; Lu, Chang-Tien; Sirwongwattana, S; Huang, Yo-Ping: Survey of fraud detection techniques. In: *IEEE International Conference on Networking, Sensing and Control*, 2004. volume 2, pp. 749–754, 2004.
- [Ko13] Koch, Peter: *Versicherungswirtschaft: Ein einführender Überblick*. VVW GmbH, 2013.
- [Kö15] Köneke, Vanessa; Müller-Peters, Horst; Fetchenhauer, Detlef et al.: *Versicherungsbetrug verstehen und verhindern*. Springer, 2015.
- [Kr16] Krawczyk, Bartosz: Learning from imbalanced data: open challenges and future directions. *Progress in Artificial Intelligence*, 5(4):221–232, 2016.
- [Le95] Leonard, Kevin J.: The development of a rule based expert system model for fraud alert in consumer credit. *European Journal of Operational Research*, 80(2):350–356, 1995.
- [LHM98] Liu, Bing; Hsu, Wynne; Ma, Yiming: Integrating classification and association rule mining. In: *KDD*. volume 98, pp. 80–86, 1998.
- [LMT14] Lunardon, Nicola; Menardi, Giovanna; Torelli, Nicola: . *ROSE: Random Over-Sampling Examples*, 2014. R package version 0.0-3.
- [MT14] Menardi, Giovanna; Torelli, Nicola: Training and assessing classification rules with imbalanced data. *Data Mining and Knowledge Discovery*, 28(1):92–122, 2014.
- [Ni16] Nian, Ke; Zhang, Haofan; Tayal, Aditya; Coleman, Thomas; Li, Yuying: Auto insurance fraud detection using unsupervised spectral ranking for anomaly. *The Journal of Finance and Data Science*, 2(1):58–75, 2016.
- [Ok13] Okura, Mahito: The relationship between moral hazard and insurance fraud. *Journal of Risk Finance*, 14:120–128, 2013.

- [SA13] Srinivasan, Uma; Arunasalam, Bavani: Leveraging Big Data Analytics to Reduce Healthcare Costs. *IT Professional*, 15(6):21–28, 2013.
- [Sc] Scrucca, Luca: , A quick tour of GA. <https://cran.r-project.org/web/packages/GA/vignettes/GA.html>. Accessed: 2020-06-14.
- [Sc04] Schiller, Jörg: Versicherungsbetrug als ökonomisches Problem: Eine vertragstheoretische Analyse. *Zeitschrift für die gesamte Versicherungswirtschaft*, 93:835–851, 2004.
- [Sc19] Scrucca, Luca: . GA: Genetic Algorithms, 2019. R package version 3.2.
- [ŠD11] Şahin, Yusuf G; Duman, Ekrem: Detecting credit card fraud by decision trees and support vector machines. In: *Proceedings of the International MultiConference of Engineers and Computer Scientists IMECS*. volume 1, p. 6, 2011.
- [ŠFB11] Šubelj, Lovro; Furlan, Štefan; Bajec, Marko: An expert system for detecting automobile insurance fraud using social network analysis. *Expert Systems with Applications*, 38(1):1039–1052, 2011.
- [Sh92] Shavell, Steven: On Moral Hazard and Insurance. In: *Foundations of Insurance Economics: Readings in Economics and Finance*. Springer Netherlands, Dordrecht, pp. 280–301, 1992.
- [SR97] Sternberg, M.; Reynolds, R. G.: Using cultural algorithms to support re-engineering of rule-based expert systems in dynamic performance environments: a case study in fraud detection. *IEEE Transactions on Evolutionary Computation*, 1(4):225–243, 1997.
- [Ta98] Taniguchi, M.; Haft, M.; Hollmen, J.; Tresp, V.: Fraud detection in communication networks using neural and probabilistic methods. In: *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP '98* (Cat. No.98CH36181). volume 2, pp. 1241–1244, 1998.
- [To76] Tomek, Ivan: Two Modifications of CNN. *IEEE Transactions on Systems, Man and Communications*, 6:769–772, 1976.
- [Va15] Van Vlasselaer, Véronique; Bravo, Cristián; Caelen, Olivier; Eliassi-Rad, Tina; Akoglu, Leman; Snoeck, Monique; Baesens, Bart: APATE: A novel approach for automated credit card transaction fraud detection using network-based extensions. *Decision Support Systems*, 75:38–48, 2015.
- [VD04] Viaene, S.; Dedene, G.: Insurance Fraud: Issues and Challenges. *The Geneva Papers on Risk and Insurance*, 29(2):313–333, 2004.
- [Vi02] Viaene, Stijn; Derrig, Richard A.; Baesens, Bart; Dedene, Guido: A Comparison of State-of-the-Art Classification Techniques for Expert Automobile Insurance Claim Fraud Detection. *Journal of Risk and Insurance*, 69(3):373–421, 2002.
- [WX18] Wang, Yibo; Xu, Wei: Leveraging deep learning with LDA-based text analytics to detect automobile insurance fraud. *Decision Support Systems*, 105:87–95, 2018.
- [ZZ02] Zhang, Chengqi; Zhang, Shichao: Association rule mining: models and algorithms. Springer-Verlag, 2002.

Organizing for temporality and supporting AI systems – a framework for applied AI and organization research

A data perspective on organizing temporary organizations combined with recent AI advancements and their potential combination.

Helge Nuhn ¹

Abstract: Both temporary forms of organizing (TO) and artificial intelligence have recently received increased practical and scholarly attention [HW19],[Zh19]. Their combination has not yet been subject of research or research application, nor is there a roadmap to the development of TO-specific AI applications. In relation to permanent organizations, TOs devote more time to organizing, planning, and adapting to change, but supporting organizing as a task is not yet in the scope of AI research. This article creates new links between the domains computer science and organizational theory. It reviews TOs as special vehicles for organizing endeavors, proposes relevant properties, reviews recent AI research advances, and synthesizes challenges for researching into AI-assisted organizing in TOs. A table of use cases along with proposed AI methods and required data proposes future research activities and is basis for a call for data sets, data challenges and metrics for assessing AI-assisted organizing, especially in temporary contexts.

Keywords: Temporary organizations, artificial intelligence, organizing, AI-assisted organizing, lifecycle

1 Introduction

Temporary organizations (TOs) are organizations that have a pre-determined endpoint in time, after which they dissolve [SS20]. Typically, the organization ceases to exist when it has delivered its proposed value. For TOs, organizing processes and structures, realizing the necessity of changes to it, and reacting correspondingly is vital. Compared to permanent organizations, the effort devoted to changing the organization instead of delivering value is higher [SW20]. Permanent organizations have in recent years, partly due to digitalization, experienced increased market pressure to adapt to changes quickly. As temporary organizations are typically housed within or across permanent ones, they are affected by their environment similarly [HW19].

Methods of artificial intelligence have also gained popularity in recent years. Advancements in computing power as well as algorithmics lead to approaches that generate insights previously impossible to obtain from available data. Many of these are

¹ Wilhelm Büchner Hochschule, FB Wirtschaftsingenieurwesen und Technologiemanagement, Hilpertstr. 31, 64295 Darmstadt, helge.nuhn@wb-fernstudium.de,  <https://orcid.org/0000-0002-8835-676X>

based on the idea of artificial neural networks (NN, [TV16]). These have proven to be able to recognize vague or hidden patterns in text, video, structured data, or their combinations, and to generate data that is of comparable quality as the input data [Br20], [Zh21], [Ga20], [Ra21].

Change in organizations has become a novel regular state, as opposed to the exception to the rule. With line-organizations being too inflexible to yield to change; projects are instantiated to induce change, cause agility and make structure more adaptive [Ma08]. Methods of artificial intelligence in general, and those relying on a neural network paradigm in specific, have influenced permanent organizations in several aspects already [TV16]. Meta-analyses reviewed the use of the applied technical methodologies as well as the areas of application. However, most applications revolve around business-related topics like customer or supplier classification, production optimization, financial or risk assessment, process control. To our knowledge, no study to date has taken into consideration the specific application in contexts of managing and organizing for temporality.

From a theoretical perspective, structuration theory aims at explaining the relation between actors and the structures they are confronted with, or part of [Ma08]. Actions are always to be seen before the context of structures, and structures as well as actions feedback onto each other. Actors may be aware of changes they induce, related to as “intentional” actions, of unintended side-effects may be part of the systematic answer of the structures, which is induced by “unintentional” actions of an actor. In a theoretical context of temporary organizations [LS95], it becomes clear that unintentional changes become more relevant, as knowledge about structures is inherently less stable, less documented and less assessable by members of temporary organizations.

From a practical perspective, other selected problems that are specific to temporary organizations are (cf. [SS20] for these and many others) that the objectives and goals of TOs are often unclearly defined and may only exist schematically in the minds of a few people. Also, TOs change comparatively fast in terms of structure, processes and people involved. Compared to line-organizations that aim at maximum efficiency, temporary organizations’ primary concern is effectiveness, which roots in cross-functional collaboration and causes communication efforts to expand, competing for bandwidth.

The problem domain in temporary organizing is therefore dissimilar from the business-related problem domains that have so far been tackled via artificial intelligence approaches. Both in theory and practice, a gap exists, that this article proposes to be filled with the help of a research framework.

2 Life cycles of TOs

‘A temporary organization can be defined as an aggregate of individuals temporarily collaborating for a shared cause’ [NHW17] (pg. 255). Temporary organizations have been attributed more specific properties, like novelty of the task they are designed to tackle, uniqueness of the endeavor [Ba10]. The most specific one is arguably, that the ending of the organization is defined, more or less clearly, and graspable for every person that is part of the TO [HW19].

TOs hence form and dissolve. The in-between has been described in different approaches. [LS95] Lundin and Söderholm relate to the Project Management Institute (PMI) to partition a TO’s lifetime into the phases as described in table 1.

These phases have been widely adopted in TO research and have been used on various occasions to discuss in more detail about phenomena that TOs produce. But the phase labels alone remain too abstract to deduce problems whose solving process may benefit from AI methods. Therefore, a more descriptive phase-labelling is proposed, along its back-reference to the Lundin-Söderholm phase naming:

Phase name	Lundin-Söderholm	Description
Instantiation	Concept	Definitional begin-point, giving a cause to a TO.
Vision and goal formulation	Concept	Translating the TO cause into different formats, helping to convey an idea of endpoint of the TO and outcome caused by its actions.
Organizing	Development	Structuring the problem and solution space in relation to endpoint and outcome, deduction of organizational structure (teams, sub-teams, roles, etc.)
Planning	Development	Logically assigning work to teams and roles, forecasting resource usage and deducing time-consumption.
Value creation, coordination, and communication	Implementation	Resources working on assigned work with the aim to long-term reach what was understood as goal and objective. Passing information for reasons of coordinating work and provision educts for subsequent combination into products or services the TO is supposed to produce.
Intervention and adaptation		Adaptation of the TO system (structure, processes) in reaction to external or internal impulses and realization thereof.
Termination	Termination	Termination phase

Tab. 1: Proposed TO phase framework

Like Lundin and Söderholm described for their phases, the actions taken and decisions made in each phase are not without overlap. For example, instantiation and vision and goal formulation are intertwined as different levels of detail of visions are being formulated. The presented approach foresees an individual phase because TOs typically have to spend

substantial amounts of time on identifying requirements and expectations as to the overall objective and outcome, which is in part re-negotiated during that phase [LS95].

3 TO-specific challenges and TO latent space dimensionality

At the very core of any TO lies the challenge, that it is new [SW20]. However, TOs are typically not completely “new” in terms of the people that interact with each other. They make implicit use of shared mental models that their members previously established, putting more emphasis on a relational perspective of TOs [Ro20]. These models help to address some challenges easier or quicker. However, in comparison with permanent organizations, the relevance and impact on effectiveness of shared mental models, mindsets etc. is greater than that of structures and processes [Ma20].

Mindsets, shared beliefs or values as beliefs [Ma20] are latent constructs and not easily defined; however, they seem to be what make e.g. agile teams perform often better from the start [GT19] and are therefore highly relevant. As such, they are unfortunately not directly observable. But they leave behind patterns in the actions and the artefacts or TOs. These patterns should be reoccurring throughout all types of TOs, but they will probably not be found repetitively within the lifetime of one TO. This makes data gathering on them very difficult, as one has to compare (many) TOs in similar life-stages and in similar situations and between similar actions. Measuring the patterns in TO organizing is not per se a high-dimensional problem. Miler and Gaida only identified 50 potential items with which to assess agile mindsets in organizing work in temporary contexts [MG19].

Proposition 1: Latent constructs of low dimensionality exist that cause patterns in organizing TOs along their life cycle and are comparable between TOs, e.g. mindsets.

In practice, structured approaches like project management methodologies, or even simple heuristics like copying known structures as a basis for organizing temporary teams [Mi16], help modeling TOs on a level that is abstract enough for integral consideration yet at the same time detailed enough to allow for effective organizing. Often, they come with templates for specific artefacts that serve as (non-IT) tools in organizing, like organizational charts.

Even special types of assisting software like project management suites offer support in digitalizing TO organization. However, diligently setting these systems up is itself a time-consuming task. In addition, IT systems may be helpful organizing tools for some stakeholders, but at the same time are not adequate channels of communicating TO properties for others.

Proposition 2: Although solutions exist for storing organizational structures of TOs digitally, their level of detail is typically limited. For an encompassing view on TO organization, additional unstructured data, like that of human communication, must be considered as well.

As soon as TOs start their existence and organizing takes place, feedback of the environment and from internal feedback loops is perceived and processed. The TO is then in a collective state of permanent learning: every TO member learns about facts, opinions, beliefs, visions, objectives, plans of others that they interact with as well as micro-political dynamics, internal and external relationship dynamics, requirements etc. [AS01]. TOs constantly adapt in the light of those continuously gained new insights. Thus, the TO's organizational structure and its plans could be updated continuously. Due to the many different dimensions of organizing and planning, many artefacts must be adapted, quality-assured, communicated, received, acknowledged, understood, and accepted among all affected members of the TO and potentially outside stakeholders.

Proposition 3: The effectiveness and efficiency of inducing changes to the organization and plans of a TO is dependent on effective and efficient communication of the necessary changes to each of them. This in turn relies on proper, consistent documentation.

4 TO-specific artefacts

Project management frameworks define in varying levels of detail artefacts that are used to facilitate the statement and communication of structures, processes, plans and ideas. E.g. the Project Management Body of Knowledge, the PM2 framework, the V-Model XT, or Prince2 all define deliverables that are used to give meaning and direction to TOs. On the other hand, many TOs likely start off without selecting a project management standard. In these cases, TO management relies on self-made documentation that serves similar purposes: define structures and plans.

Drawing from prior work experience, numerous experts in the field agree on the point that none of the given professional project management frameworks is unique concerning the hard organizing artefacts. All of them handle goals, requirements, structures, processes, responsibilities and so on in compatible ways that would be translatable from one framework to the other with little complications.

Proposition 4: Regardless of the standards employed, TO organization and plan artefacts are semantically similar, while their formats may differ.

Some examples of artefacts used for organizing and planning as observed in practice are listed in the following. The listing is naturally incomplete but was enhanced by descriptions regarding typical forms of information contained within them, with the help of TO and project experts.

Artefacts	Description	Typical formats
Project charter	Mostly textual description of the projects initial situation, problem setting, goal, expected outcomes, approach.	Written text document, potentially enriched by 2D graphical imagery as alternate means of communicating relevant relations

Artefacts	Description	Typical formats
Organizational charts	lines and boxes, names of persons, organizations, potentially responsibilities	2D representation, boxes confining individual actors or groups of actors, lines representing relations
Approach descriptions, process descriptions	high-level process descriptions with phases, steps, responsibilities, differing formats possible (BPMN, etc.)	2D representation to include time as (abstract) dimension, shapes may have specific meaning
WBS	Named work packages, responsibilities, forecast for work effort / duration, due dates, buffers, dependencies	2D representation similar to organizational charts, boxes/shapes confine work packages and related information; spreadsheets and tables with similar data
GANTT charts	Similar to WBS, with added time dimension	2D representation to WBS, dependencies may be made explicit through lines between shapes
Work plans	Tables with work packages, dates, amount of work, available time, start and end dates, deadlines, dependencies, responsibilities	Spreadsheets/tables from spreadsheet or planning software, potentially exported in raw formats, potentially also screenshotted and pasted in emails (2D graphical representation from underlying data)
Work packages descriptions	Mostly textual descriptions of work packages, outlining approaches and necessary steps to produce outputs	Textual descriptions, potentially with header information in tabular form, enumerations
Status reports	Reports with header information about reporting cycles, traffic lights symbology for status, prose text or bullet-point formatted information about previous and upcoming activities, risks and decision needs, information on work packages and milestones	Mixed formats, prose, enumerations, imagery, symbology, 2D graphic representation

Tab. 2: TO artefacts, descriptions and analyses of structural formatting aspects

These artefacts may be embedded within other communication structures like emails, meetings, presentations and thus have communicational contexts that they relate to.

All these artefacts are logically connected through the management frameworks that define them. So, there is an underlying logical structure that defines the relation between all these documents. It is dynamic and prone to changes over time. As an example: the TO management is bound to report, therefore status reports exist that relate to work packages, working times, milestone dates, etc. This information can be described as data models of graph-like structures that build up as the TO starts to organize and adapts as TOs reacts to external impulses.

Proposition 5: All organizing artefacts are bound together by a logical structure. Functions exist that reflect upon the rules according to which the TO is organized and continues to adapt its organization.

Such structures can be represented as knowledge graphs [Al20]. Such graphs help computerized methods access and process information stored within them – and deep NN approaches have been shown to successfully operate on such graphs. Researchers have recently shown that creating knowledge graphs helps with strategic planning in text-based gaming problem settings. Therefore, potential is seen that such approaches may also help in real-life planning [Ad20]. The context elaborated here, however, calls for specifically adapted, complex graph structures with extensive attributes and properties that reflect upon organization of TOs – TO organizing graphs (TOOG) – in the sense that it holds encoded information about its structures and plans.

Proposition 6: TO organizing graphs should be conceptualized. They are the basis of intelligent systems which make use of various AI methods to support organizing.

Proposition 7: Intelligent systems are proposed that support TO organizing in such ways, that:

- They generate TOOG from any TO organizing artefact.
- They generate useful and overall-consistent organizational artefacts.
- They identify conflicting relations between two TOOG efficiently and act as consistency checkers.
- They merge non-conflicting TOOGs efficiently.
- They propose decisions to bring about consistency when conflicting TOOGs are compared.

In the following section, the article reviews selected recent advancements in AI research that show potential for further development into functions on TOOG. Afterwards the article continues to map such approaches to TO lifecycle stages and useful implementation scenarios, before requirements towards corresponding data sets and TOOG are discussed in more detail.

5 Review of recent advances in AI methods

Stunning research advances have taken place in the discipline of processing both image data and written, natural language. The immense progress in accuracy, precision, as well as reliability largely rely on deep learning methodologies or deep neural networks. Noteworthy sub-categories of image processing through artificial neural networks are image labelling, image captioning, image (semantic) segmentation, object detection and bounding box generation within images e.g. with YOLO or COCO, and others [Zh21]. AI methods that are applied are artificial neural networks (NN), convolutional NN, recurrent NN, LSTM cells, deep NN, but also transformer approaches have been applied successfully [Do20]. Typical inputs can range from “low” parameters spaces, with e.g. 27x27 pixels of 7-bit greyscale imagery in the typical MNIST data set for handwritten

numbers of which 50.000 labelled examples exist, up to high resolutions of 640x480 in more than 200k RGB-colored pictures e.g. for the COCO data set [Li15].

Text analysis and knowledge extraction is the text-based correspondence to this type of image processing described. Through text analysis, knowledge from within the texts can be extracted, e.g. by means of named entity recognition, disambiguation, alignment, or linking, instance extraction, fact extraction, relationship extraction and more. Rule-based approaches have exploited just as the more recent transformer-based deep NN approaches [AI20].

Equally intriguing is the build-up of AI capabilities of processing text, i.e. natural written language. With hundreds of billions of parameters, the latest autoregressive models can produce written texts, while only consuming few words, or even questions, as inputs. OpenAI's GPT-3 is one example of this [Br20], just as open-sourced alternatives like GPT-Neo or GPT-NeoX. After using convolutional or recurrent neural networks, also LSTM techniques had been employed in natural language processing tasks. Often, auto-encoder approaches have been used. A major leap forward was attained by introducing the transformer models, that allows algorithms to learn and choose for themselves what context to consider while selecting words and sequences that produce the final text. It is a specific type of auto-encoder or at least relies on encoder-decoder mechanisms inherently [Va17]. In a similar fashion, data-to-text approaches, e.g. based on the ROTOWIRE data set [Wi17] attempt to create review texts for NBA basketball games from given raw datapoints, that appeal to the reader as if humans had written them. Other typical task types of natural language processing include classification, (sequence) labelling, summarization, but also question-answering, part-of-speech tagging, fact-verification, relational extraction, topic modeling, named entity recognition, translation, and many more. Publicly available data sets and code exist in high numbers (<https://paperswithcode.com/>).

The combination of both techniques discussed before has recently shown astonishing results: text input and image generation were combined to generate images from text. As an example, given the input “a tapir made of accordion” generates an image of a fictitious animal with a typical tapir face, four legs and a torso that looks like the typical accordion bellows [Ra21]. Other examples of translation of diverse inputs into images have been explored before, for example drawing realistic images from sketches [Gh19] (<https://arnabgho.github.io/iSketchNFill/>), [Ga20].

In TO organizing, helpful use cases would be to get relevant results for a request like “a work breakdown structure, containing only work packages that external companies account for”. One might critically remark that this use case might as well be solved by a mere filter on a table of work packages in a project management IT system. A counter argument, however, would be that GPT-like networks contain such factual table-data within their billions of parameters just as well.

Another example for a combinatorial challenge of working on image/graphical as well as textual data is the research by [Co21]. They attempted to identify structures in tables (given as 2D visual structures) to extract the relevant data from the individually identified cells. They too rely on CNNs and YOLO, amended by OCR methods to obtain very good results on recall, precision, and F1-Scores.

Looking at it from a user-perspective, works like that of [Gh19] prove that AI research is far from being theoretical only. Interactive user interfaces allow for direct interaction with the AI algorithm, allowing for the creation of meaningful output in a (wo)man/machine co-creative manner. However sketchy the input of the user may be, the AI model takes care that the output reconciles both user requirements given as an input, as well as semantic soundness as good as possible. See for example StyleGAN by Härkönen et al., StyleFlow by Abdal et al. [Hä20] [Ab20]. The TO Organizing Graph proposed in the previous section would attempt to build such underlying structures for AI algorithms that attempt similar things to those discussed in the previous paragraphs. A corresponding use case would e.g. be to obtain organizational charts with different views, like different levels of detail, or structural choices like the order of sub-projects displayed.

6 Use cases of AI in TO contexts

The presented research shows how usable AI methods have become for practical use cases, and that there are data sets, data challenges as well as metrics for researchers to work with and measures against. TO research needs an AI research roadmap, including corresponding data sets and challenges for researchers to respond to.

In the following, this article lists on one page a conceptual model that combines TO lifecycle considerations with one to two challenges that, AI models should have the potential to tackle successfully. Implications for data sets are amended, indicating challenges to data set compilation for future TO/AI research as proposed above.

The challenges described below can be grouped into areas of analysis and generation. Analysis serves the aspect of understanding input and creating underlying TOOGs [Hu20]. Generation focuses on text generation, image generation or a combination of both in the sense of combined artefact generation. Also, translation/transformation of inputs in diverse formats are proposed. Multiple challenges describe varying inputs, like text or graphical documents that are to be translated into other document types. For some challenges, this process is described as reliant on underlying data structures like the TOOG described above.

Lifecycle phase	Challenges	Proposed methods	AI	Data requirements
Vision and goal formulation	Create summaries of various lengths of project charters, combining	knowledge extraction, training on	graph graph	Project charters and additional information

Lifecycle phase	Challenges	Proposed methods	AI	Data requirements
	textual descriptions of project endeavors.	neural network in conjunction with learned embeddings of textual descriptions, GPT-like translation (e.g. into a modeling format)		(proposals, contracts), masked/tagged according to relevance for different levels of details for project descriptions
Organizing	Create work breakdown structures, organizational charts, graphical depictions of planned approaches (one-page summaries) from textual descriptions, sketches, tables in conjunction with available text of vision and goal formulations.	knowledge graph extraction, training on graph neural network in conjunction with auto-regressive text-generating NN models, GPT-like translation (e.g. into a modeling format)		All of the above, plus work breakdown structures, work package descriptions and org charts labelled according to accuracy of depiction
Planning	Create time-plans (tables with dates, GANTT charts, milestone lists, etc.) that rely upon conventional knowledge of typical lengths of phases or work package (i.e. workshop series typically last weeks, not hours, workshops last hours not, months), but also on several textual inputs (project charters, work package descriptions) allow for real-time regulation.	Graph neural networks with underlying knowledge graphs from previous phases, neural rendering or auto-regressive text-generating NN models, GPT-like translation (e.g. into a modeling format)		All of the above
Value creation, coordination, and communication	a) Screening texts (emails) for risks regarding timelines; observe	Multi-label/multi-class classification of text, especially conversations		Conversation in chats/emails, documentation like protocols,

Lifecycle phase	Challenges	Proposed methods	AI	Data requirements
	email/chat/document content for sentiment analysis, risks, compatibility of technical detail levels; b) Summarizing weekly/monthly activities for status reports in different levels of detail.	with multiple participants; relation extraction regarding causalities wrt/ to time planning; auto-regressive text-generating NN models, GPT-like translation		text documents like concepts, presentations
Intervention and adaptation	Process proposed changes to TO organizations (plans and structures), in textual or graphical depictions, work breakdowns, timelines, milestones and dates and create consistent adaptations of existing artefacts.	Knowledge graph regression, image segmentation, semantic extraction, extractive text summarization, question answering, fact verification		Segmentation masked inputs of intervening graphical content, like counter-proposed project plans or organizational structures; labelled texts that propose different organizing as compared to current organizational structure.
Termination	Summarization of the core results of the TO in various lengths and depths of detail, e.g. for a final project report. Qualitative review of vision, goals, organizational structures, plans as they changed over time.	Text summarization, text regression, data to text / table / image translation		All deliverables mentioned above

Tab. 3: TO Lifecycles, AI challenges and methods and data requirements

7 Data requirements, collection and further research

To tackle the challenges outlined in this article, useful data sets need to be compiled, and data challenges need to be thought up that help develop useful algorithms and define novel methodological approaches that generate value for contexts of organizing for temporality.

The data requirements described in the table above summarize as follows. Artefacts like organization charts, project charters, work package descriptions, visual representations of approach concepts, documented communication, external plans, etc. must be collected in an artefact collection. The files should be categorized and attributed with meta-data that can be referenced. The meta-data should encompass artefact type, timestamps, authorship (pseudonymous), keywords, short descriptions, comments, communication, plans whether they effectively influenced the TO, whether they did so partly, or not at all, and more.

The references would occur from within a TO organizing graph. This type of data structure serves as a database that holds information about the organizational structure of the TO in a more formalized form. It should contain information about beginning and end of the TO, so it can synchronize with timestamps within the artefact collection. The TOOG should be versioned, so that ideally for every timestamp of each artefact within the artefact collection, there would be a version. More versions should be allowed, because some artefacts would cause the factual TO organization to change with delay.

Viewing it from the perspective of some of the described challenges, parts of the data set would need to be enhanced even further, e.g. through masking of visual plans or concepts so that e.g. algorithms could be guided to specific parts that developed influence onto the TO organization. The data set and data set creation could also be made subject to AI algorithms, in that it would profit from automated, consistent pseudonymization.

8 Conclusion

This article laid out the specific management challenge of organizing temporal endeavors and the dynamic nature of temporary organizations. Typical artefacts that are created while organizing in TOs have been listed and described. Through a series of propositions, it was argued that AI methods can support TO organizing tasks. Recently developed AI methods have been reviewed in the view of their potential to provide valuable output for TO organizing tasks. Corresponding use cases in terms of challenges for future researchers have been proposed. The necessity for a unique data set for researching the use and potential of AI in TO organizing was argued, basic requirements towards such a data set were proposed. This article calls for further development of such a data set and consequently corresponding AI research.

9 References

- [Ab20] Abdal, Rameen, Peihao Zhu, Niloy Mitra, and Peter Wonka. „StyleFlow: Attribute-conditioned Exploration of StyleGAN-Generated Images using Conditional Continuous Normalizing Flows.“. 2020.
- [Ad20] Adhikari, Ashutosh, et al. „Learning Dynamic Knowledge Graphs to Generalize on Text-Based Games.“ ArXiv abs/2002.09127 (2020).
- [Al20] Al-Moslmi, Tareq, Marc Gallofré Ocaña, Andreas L. Opdahl, and C. Veres. „Named Entity Extraction for Knowledge Graphs: A Literature Overview.“ IEEE Access 8 (2020): 32862-32881.
- [AS01] Sense, A. J. & Antoni, M. (2001). Learning within and across projects: A comparison of frames. In Quality Management and Organizational Development Conference (QMOD2001)
- [Ba10] Bakker, René M. „Taking Stock of Temporary Organizational Forms: A Systematic Review and Research Agenda.“ International Journal of Management Reviews (Wiley) 12 (11 2010): 466–486.
- [Br20] Brown, Tom B., et al. „Language Models are Few-Shot Learners.“ Language Models are Few-Shot Learners. 2020. arXiv:2005.14165
- [Co21] Colter, Zachary, M. Fayazi, Zineb Benameur-El, Serafina Kamp, Shuyan Yu, and Ronald Dreslinski. „Tablext: A Combined Neural Network And Heuristic Based Table Extractor.“ ArXiv abs/2104.11287 (2021).
- [Do20] Dosovitskiy, A., et al. „An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale.“ ArXiv abs/2010.11929 (2020).
- [Ga20] Gao, Chengying, Qi Liu, Qi Xu, Limin Wang, Jianzhuang Liu, and Changqing Zou. „SketchyCOCO: Image Generation From Freehand Scene Sketches.“ 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2020.
- [Gh19] Ghosh, Arnab, et al. „Interactive Sketch & Fill: Multiclass Sketch-to-Image Translation.“ Interactive Sketch & Fill: Multiclass Sketch-to-Image Translation. 2019.
- [GT19] Gregory, Peggy, and Katie Taylor. „Defining Agile Culture: A Collaborative and Practitioner-Led Approach.“ 2019 IEEE/ACM 12th International Workshop on Cooperative and Human Aspects of Software Engineering (CHASE). IEEE, 2019.
- [Hä20] Härkönen, Erik, Aaron Hertzmann, Jaakko Lehtinen, and Sylvain Paris. „GANSpace: Discovering Interpretable GAN Controls.“ arXiv, 2020.
- [HW19] Henning, Christian H. C. A., and Andreas Wald. „Toward a wiser projectification: Macroeconomic effects of firm-level project work.“ International Journal of Project Management 37 (2019): 807-819.
- [Hu20] Huang, Zhicheng, Zhaoyang Zeng, Bei Liu, Dongmei Fu, and J. Fu. „Pixel-BERT: Aligning Image Pixels with Text by Deep Multi-Modal Transformers.“ ArXiv abs/2004.00849 (2020).

- [Li15] Lin, T.-Y.; Maire, M.; Belongie, S.; Bourdev, L.; Girshick, R.; Hays, J.; Perona, P.; Ramanan, D.; Zitnick, C. L. & Dollár, P. (2015), 'Microsoft COCO: Common Objects in Context'.
- [LS95] Lundin, Rolf A., and Anders Söderholm. „A theory of the temporary organization.“ 11 (1995): 437-455.
- [Ma08] Manning, S. „Embedding Projects in Multiple Contexts: A Structuration Perspective.“ 2008.
- [Ma20] Martinsuo, Miia. „The Management of Values in Project Business: Adjusting Beliefs to Transform Project Practices and Outcomes.“ *Project Management Journal* (SAGE Publications) 51 (6 2020): 389–399.
- [MG19] Miler, J. & Gaida, P. (2019), On the Agile Mindset of an Effective Team An Industrial Opinion Survey, in 'Proceedings of the 2019 Federated Conference on Computer Science and Information Systems', IEEE
- [Mi16] Miterev, M.; Engwall, M. & Jerbrant, A. (2016), 'Mechanisms of Isomorphism in Project-Based Organizations', *Project Management Journal* 48, 24 - 9.
- [NHW17] Nuhn, Helge F. R., Sven Heidenreich, and Andreas Wald. „Performance Outcomes of Turnover Intentions in Temporary Organizations“. *European Management Review* (Wiley) 16 (11 2017): 255–271.
- [Ra21] Ramesh, Aditya, et al. „Zero-Shot Text-to-Image Generation.“ *Zero-Shot Text-to-Image Generation*. 2021.
- [Ro20] Rodriguez-Lluesma, Carlos, Pablo García-Ruiz, and Javier Pinto-Garay. „The digital transformation of work: A relational view.“ *Business Ethics: A European Review* (Wiley) 30 (10 2020): 157–167.
- [SS20] Samimi, E., and Sydow, J.. „Human resource management in project-based organizations: revisiting the permanency assumption.“ *The International Journal of Human Resource Management* 32 (7 2020).
- [SW20] Sydow, J. & Windeler, A. (2020), 'Temporary organizing and permanent contexts', *Current Sociology* 68(4), 480--498.
- [TV16] Tkáč, Michal, and Robert Verner. „Artificial neural networks in business: Two decades of research.“ *Applied Soft Computing* (Elsevier BV) 38 (1 2016): 788–804.
- [Va17] Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, L. & Polosukhin, I. (2017), 'Attention Is All You Need'. ArXiv 1706.03762
- [Wi17] Wiseman, Sam, Stuart M. Shieber, and Alexander M. Rush. „Challenges in Data-to-Document Generation.“ *Challenges in Data-to-Document Generation*. 2017.
- [Zh21] Zhou, Jie, et al. „Graph Neural Networks: A Review of Methods and Applications.“ *Graph Neural Networks: A Review of Methods and Applications*. 2021.

Kontinuierliche Evaluierung eines KI-basierten interaktiven Systems

Niclas Mummert¹, Olga Lange², Paul Cioflica³, Tobias Reutemann⁴


Abstract: Eine kontinuierliche Evaluierung eines KI-basierten interaktiven Systems wird in diesem Beitrag am Beispiel eines Fahrerassistenzsystems zur Ermittlung von Nutzerpräferenzen über eine Auswahl vom Fahrmodus (Comfort, Normal Sport und Sport+) aus dem Automobilbau vorgestellt. Das zu evaluierende System enthält Machine Learning Modelle, welche anhand Straßen (Asphalt, Pflastersteine, Feldweg) und Nutzerinformationen einen Fahrmodus vorschlägt. Herausforderungen der Evaluierung entstehen vor allem durch die Eingaben der Nutzenden an der Benutzungsschnittstelle (User Interface). Dieser Beitrag stellt mögliche Evaluierungsmethoden zur kontinuierlichen Erfassung der Qualität dieses KI-basierten interaktiven Systems dar und geht der Frage nach Zusammenhängen zwischen dem Handeln der Nutzenden und der Veränderung der ML-Modellen nach.

Keywords: Evaluierung, KI-basiertes System, kontinuierliche Evaluierung, Machine Learning (ML), User Interface (UI), Benutzungsschnittstelle.

1 Einleitung

Die Problemstellung dieser Ausarbeitung bedient sich an der kontinuierlichen Evaluierung der Gebrauchstauglichkeit von KI-basierten interaktiven Systemen. KI-basierte interaktive Systeme enthalten viele dynamische Systemparameter, die auch die Ausgaben an der Benutzungsschnittstelle beeinflussen. Die zentrale Frage dieser Ausarbeitung ist: Wie kann eine kontinuierliche Evaluierung von einem KI-basierten interaktiven System gesichert werden? Ferner werden Zusammenhänge zwischen dem Handeln der Nutzenden und des ML-Modells betrachten.

¹ Duale Hochschule Baden-Württemberg, Fakultät für Technik, 70178 Stuttgart, niclas.mummert@bertrandt.com,  <https://orcid.org/0000-0002-7839-3938>

² Duale Hochschule Baden-Württemberg, Fakultät für Technik, 70178 Stuttgart, olga.lange@lehre.dhbw-stuttgart.de,  <https://orcid.org/0000-0002-3584-9288> ³ Bertrand Technologie GmbH, 71297 Mönshiem, paul.cioflica@bertrandt.com

⁴ Bertrand Technologie GmbH, 71297 Mönshiem, tobias.reutemann@bertrandt.com

2 Implementierung, praktische Umsetzung und Bewertung der Evaluierungsmethoden

Bei der Evaluierung der Machine Learning Modelle wird der Fokus vor allem auf die Funktionalität gelegt. Das bedeutet, es werden Modultests zur Absicherung der Funktionalität der Machine Learning Modelle erstellt. Diese Tests sind in einer Pipeline automatisch ausführbar und liefern ein zu erwartendes Ergebnis. Bei der Evaluierung der Benutzungsschnittstelle werden Probandentest, welcher eine Einführung in das System mit einer Aufgabe, eine Beobachtung mit Eye-Tracking und eine Befragung mit System Usability Scale (SUS) beinhaltet und mit 12 Probanden durchgeführt wurde. Zudem ist beobachtet worden, dass die Nutzerfreundlichkeit der Benutzungsschnittstelle in Abhängigkeit zu der Treffergenauigkeit des Machine Learning Modells steht. Je höher die Genauigkeit des Machine Learning Modells, desto präziser sind die Vorschläge eines Fahrmodus für den Nutzenden. Hier wurde eine höhere System Usability (SU) von 70% beobachtet als bei unpräziseren Vorschlägen. Es wird vermutet, dass die Genauigkeit des Machine Learning Modells bei häufigeren Eingaben der Nutzenden noch höheren wird, was in einen höheren SU-Wert hinausläuft. Bei dem Eye-Tracking-Verfahren wurde beobachtet, dass die dynamische Veränderung der Fahrmodus durch Machine Learning Modelle einen großen Einfluss bei der Entscheidung als auch bei der Entscheidungsfindung der Probanden hat. Diese Kette von Abhängigkeiten wird in Abbildung 1 dargestellt.

Nr.	ML-Modell	Genauigkeit, %
1	Convolutional Neural Network (CNN) zur Erkennung von Fahrbahndeckentypen	70%
2	K-Nearest Neighbour Klassifikator zur Ermittlung eines Fahrmodus ohne Nutzerdaten	80%
3	K-Nearest Neighbour Klassifikator zur Ermittlung eines Fahrmodus mit Nutzerdaten	85%

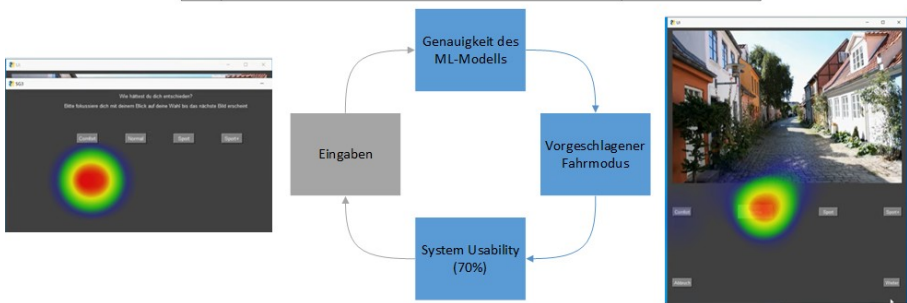


Abb. 1: Zusammenhänge zwischen dem Handeln der Nutzenden und der Genauigkeit des Machine Learning Modells

Anschließend wurden die Probandentests kontinuierlich über einen Monat mit jeweils 3 Probanden angewendet. Dabei wurden keine neuen Erkenntnisse über die Qualität des Systems gewonnen. Wird ein System häufig verändert, dann kann das Anwenden von Probandentests die Qualität sicherstellen. Die Zusammenhänge zwischen dem Handeln der Nutzenden und des ML-Modells sind durch die Kette der Abhängigkeiten gegeben.

Workshop: Konzeptionelle
Herausforderungen für die KI:
Nachhaltigkeit

Künstliche Intelligenz und Nachhaltigkeit

Klaus Mainzer¹

Abstract: Die Informatik hat sich in den letzten Jahrzehnten zu einem der größten Treiber des gesellschaftlichen Wandels entwickelt. Bedingt durch den rasanten Fortschritt in der Informationstechnik finden Informatiksysteme eine rasante Verbreitung. Die Infrastrukturen des Internets und World Wide Webs schaffen völlig neue Formen der Interaktion und Kommunikation zwischen Menschen, Maschinen und Infrastrukturen (Internet der Dinge). Produkte und Dienstleistungen auf Basis von Softwaresystemen greifen um sich. Informatik verändert in vielen Anwendungsgebieten die Sicht auf die Welt. Damit steht die Informatik als Treiber des Wandels auf den ersten Blick in einem krassen Gegensatz zu einer statisch verstandenen Nachhaltigkeit.

Keywords: Nachhaltigkeit, Künstliche Intelligenz.

1 Zur Einführung: Informatik und Nachhaltigkeit

Nachhaltigkeit ist ein positiv besetzter Begriff – versteht man darunter doch historisch die Bewahrung von Werten im Sinne eines Wirtschaftens im Heute nicht zum Nachteil der Generationen von morgen. Vor dem Hintergrund des starken Wandels durch die Informatik scheinen Nachhaltigkeit und Informatik in einem strikten Gegensatz zu stehen.

Deutlich wird das bereits bei dem klassischen Anliegen der Nachhaltigkeit der Bewahrung der Umwelt. Informationstechnik ist einer der Energiefresser. Eine Initiative in Richtung eines schonenderen Umgangs mit den Ressourcen ist das Konzept der Green IT. Spannender aber als das durchaus wichtige Thema der Green IT ist die Frage, in welcher Weise Informatik und Informatiksysteme gestaltet werden können, dass sie eben nicht im Gegensatz zu einer umfassend verstandenen Nachhaltigkeit stehen. Das ist das Ziel nachhaltiger Innovationen, die dynamisch die soziale Welt verändern und dabei ökologische und ökonomische Ziele gleichermaßen beachten. Wichtig ist aber auch die kultur- und sozialwissenschaftliche Frage, was Menschen überhaupt wollen und welche Zukunft sie wünschen. Hier stellt sich die Frage nach der Macht von Software- und Hardwarekonzernen bei der Gestaltung einer zukünftigen Welt. Wie abhängig machen wir uns von deren Vorgaben?

Informatik ist die neue Ingenieurwissenschaft der Kultur- und Sozialwissenschaften. Sie verändert unsere Sicht auf die Welt, unser Denken. Steht das im Gegensatz zu Nachhaltigkeit im Sinne eines wertorientierten Handelns zur Bewahrung der Umwelt und Gestaltung der Zukunft? Hier findet sich eine Fülle interdisziplinärer Fragen, die zu einer Kooperation von Informatik mit Kultur- und Sozialwissenschaften führen muss, damit

nachhaltige Innovationen möglich werden.

2 Was kann Künstliche Intelligenz?

Künstliche Intelligenz (KI) beherrscht längst unser Leben, ohne dass es vielen bewusst ist. Smartphones, die mit uns sprechen, Armbanduhren, die unsere Gesundheitsdaten aufzeichnen, Arbeitsabläufe, die sich automatisch organisieren, Autos, Flugzeuge und Drohnen, die sich selbst steuern, Verkehrs- und Energiesysteme mit autonomer Logistik oder Roboter, die ferne Planeten erkunden, sind technische Beispiele einer komplex vernetzten Welt intelligenter Systeme. Sie zeigen uns, wie unser Alltag von KI-Funktionen bestimmt ist. Um den Beitrag von KI zur Nachhaltigkeit zu untersuchen, muss zunächst geklärt werden, was KI ist und was sie kann (Mainzer 2019).

2.1 Symbolische KI: Logik und Deduktion

In einer ersten Phase orientierte sich KI an formalen (symbolischen) Kalkülen der Logik, mit denen Problemlösungen regelbasiert abgeleitet werden können. Man spricht deshalb auch von symbolischer KI. Ein typisches Beispiel ist das automatische Beweisen mit logischen Deduktionen, die sich mit Computerprogrammen realisieren lassen. Automatisierung bedeutet bis zu einem bestimmten Grad auch Autonomie, da Computerprogramme die Beweistätigkeit eines Mathematikers übernehmen. Wissensbasierte Expertensysteme sind Computerprogramme, die Wissen über ein spezielles Gebiet speichern und ansammeln, aus dem Wissen automatisch Schlussfolgerungen ziehen, um zu konkreten Problemen des Gebietes Lösungen anzubieten. Im Unterschied zum menschlichen Experten ist das Wissen eines Expertensystems aber auf eine spezialisierte Informationsbasis beschränkt ohne allgemeines und strukturelles Wissen über die Welt (Puppe 1988; Mainzer 1990).

Um ein Expertensystem zu bauen, muss das Wissen des Experten in Regeln gefasst werden, in eine Programmiersprache übersetzt und mit einer Problemlösungsstrategie bearbeitet werden. Die Architektur eines Expertensystems besteht daher aus den folgenden Komponenten: Wissensbasis, Problemlösungskomponente (Ableitungssystem), Erklärungskomponente, Wissenserwerb, Dialogkomponente. In dieser Architektur werden zugleich die Grenzen symbolischer KI deutlich: Fähigkeiten, die nicht oder nur schwer symbolisch erfasst und regelbasiert simuliert werden können, bleiben der symbolischen KI verschlossen. Mit Blick auf Nachhaltigkeit können Expertensysteme eingesetzt werden, um notwendiges Spezialwissen aus unterschiedlichen Domänen

zusammenzustellen und auszuwerten. Beweis- und Kontrollverfahren tragen zur Sicherheit von Informatiksystemen bei.

2.2 Subsymbolische KI: Statistik und Induktion

Sensorische und motorische Fähigkeiten werden nicht aus Lehrbuchwissen logisch abgeleitet, sondern aus Beispielen erlernt, trainiert und eingeübt. So lernen wir, uns motorisch zu bewegen und in einer Vielzahl sensorischer Daten Muster und Zusammenhänge zu erkennen, an denen wir unser Handeln und Entscheiden orientieren können. Da diese Fähigkeiten nicht von ihrer symbolischen Repräsentation abhängen, spricht man auch von subsymbolischer KI. An die Stelle der formalen Schlüsse der Logik tritt nun die Statistik der Daten. Beim statistischen Lernen sollen allgemeine Abhängigkeiten und Zusammenhänge aus endlich vielen Beobachtungsdaten durch Algorithmen abgeleitet werden (Vapnik 1998). An die Stelle der Deduktion in der symbolischen KI tritt also in der subsymbolischen KI die Induktion. Dazu können wir uns ein naturwissenschaftliches Experiment vorstellen, bei dem in einer Serie von veränderten Bedingungen (Inputs) entsprechende Ergebnisse (Outputs) folgen. In der Medizin könnte es sich um einen Patienten handeln, der auf Medikamente in bestimmter Weise reagiert.

Dabei nehmen wir an, dass die entsprechenden Paare von Input- und Outputdaten unabhängig durch dasselbe Zufallsexperiment erzeugt werden. Statistisch sagt man deshalb, dass die endliche Folge von Beobachtungsdaten $(x_1, y_1), \dots, (x_n, y_n)$ mit Inputs x_i und Outputs y_i ($i = 1, \dots, n$) durch Zufallsvariablen $(X_1, Y_1), \dots, (X_n, Y_n)$ realisiert wird, denen eine Wahrscheinlichkeitsverteilung $P_{X,Y}$ zugrunde liegt. Algorithmen sollen nun Eigenschaften der Wahrscheinlichkeitsverteilung $P_{X,Y}$ ableiten. Ein Beispiel wäre die Erwartungswahrscheinlichkeit, mit der für einen gegebenen Input ein entsprechender Output auftritt. Es kann sich aber auch um eine Klassifikationsaufgabe handeln: Eine Datenmenge soll auf zwei Klassen aufgeteilt werden. Mit welcher Wahrscheinlichkeit gehört ein Element der Datenmenge (Input) eher zu der einen oder anderen Klasse (Output)? Wir sprechen in diesem Fall auch von binärer Mustererkennung.

Die derzeitigen Erfolge des Machine Learning scheinen die These zu bestätigen, dass es auf möglichst große Datenmengen ankommt, die mit immer stärkerer Computerpower bearbeitet werden. Die erkannten Regularitäten hängen dann aber nur von der Wahrscheinlichkeitsverteilung der statistischen Daten ab:

Statistisches Lernen versucht, ein probabilistisches Modell aus endlich vielen Daten von Ergebnissen (z.B. Zufallsexperimente) und Beobachtungen abzuleiten.

Statistisches Schließen versucht umgekehrt, Eigenschaften von beobachteten Daten aus

einem angenommenen statistischen Modell abzuleiten.

In der Automatisierung statistischen Lernens nehmen neuronale Netze mit Lernalgorithmen eine Schlüsselrolle ein. Neuronale Netze sind vereinfachte Rechenmodelle nach dem Vorbild des menschlichen Gehirns, in denen Neuronen mit Synapsen verbunden sind. Die Intensität der neurochemischen Signale, die zwischen den Neuronen ausgesendet werden, sind im Modell durch Zahlengewichte repräsentiert. Probabilistische Netzwerke haben experimentell eine große Ähnlichkeit mit biologischen neuronalen Netzen. Werden Zellen entfernt oder einzelne Synapsengewichte um kleine Beträge verändert, erweisen sie sich als fehlertolerant gegenüber kleineren Störungen wie das menschliche Gehirn z.B. bei kleineren Unfallschäden. Das menschliche Gehirn arbeitet mit Schichten paralleler Signalverarbeitung. So sind z.B. zwischen einer sensorischen Inputschicht und einer motorischen Outputschicht interne Zwischenschritte neuronaler Signalverarbeitung geschaltet, die nicht mit der Außenwelt in Verbindung stehen.

Tatsächlich lässt sich auch in technischen neuronalen Netzen die Repräsentations- und Problemlösungskapazität steigern, indem verschiedene lernfähige Schichten mit möglichst vielen Neuronen zwischengeschaltet werden. Die erste Schicht erhält das Eingabemuster. Jedes Neuron dieser Schicht hat Verbindungen zu jedem Neuron der nächsten Schicht. Die Hintereinanderschaltung setzt sich fort, bis die letzte Schicht erreicht ist und ein Aktivitätsmuster abgibt (Hornik/ Stinchcombe/ White 1989).

Wir sprechen von überwachten Lernverfahren, wenn der zu lernende Prototyp (z.B. die Wiedererkennung eines Musters) bekannt ist und die jeweiligen Fehlerabweichungen daran gemessen werden können. Ein Lernalgorithmus muss die synaptischen Gewichte so lange verändern, bis ein Aktivitätsmuster in der Outputschicht herauskommt, das möglichst wenig vom Prototyp abweicht. Beim nicht-überwachten Lernen entdeckt ein Algorithmus selbstständig Ähnlichkeiten und Korrelationen von Daten, um die Datenmasse (Big Data) nach diesen Kriterien zu strukturieren. Deep Learning bezieht sich auf die Tiefe des Netzes, d.h. die Anzahl der Schichten in neuronalen Netzen, um so die Fähigkeiten zur Mustererkennung in großen Datenmassen zu steigern.

Ein effektives Verfahren besteht darin, für jedes Neuron der Outputschicht die Fehlerabweichung von tatsächlichem und gewünschtem Output zu berechnen und dann über die Schichten des Netzwerks zurückzuverfolgen. Wir sprechen dann von einem Backpropagation-Algorithmus. Die Absicht ist, durch genügend viele Lernschritte für ein Vorgabemuster den Fehler auf Null bzw. vernachlässigbar kleine Werte zu vermindern. Es bedarf keiner weiteren Erläuterung, dass die Digitalisierung der Umwelt mit einem gewaltigen Datenaufkommen verbunden ist, das durch statistische Lernverfahren zu bearbeiten ist, um Muster, Strukturen und Trends zu erkennen. Menschliche Gehirne

wären bei diesem Datenaufkommen überfordert, um z.B. Überwachungs- und Frühwarnsysteme einzurichten.

2.3 Vom statistischen zum kausalen Lernen

Statistisches Lernen und Schließen aus Daten reichen aber nicht aus. Wir müssen vielmehr die kausalen Zusammenhänge von Ursachen und Wirkungen hinter den Messdaten erkennen (Pearl 2009). Diese kausalen Zusammenhänge hängen von den Gesetzen der jeweiligen Anwendungsdomäne unserer Forschungsmethoden ab, also den Gesetzen der Physik, den Gesetzen der Biochemie, der Klimaforschung, der Life Sciences etc. Wäre es anders, könnten wir mit den Methoden des statistischen Lernens und Schließens bereits die Probleme dieser Welt lösen: Statistisches Lernen und Schließen ohne kausales Domänenwissen ist blind – bei noch so großer Datenmenge (Big Data) und Rechenpower!

Neben der Statistik der Daten bedarf es zusätzlicher Gesetzes- und Strukturannahmen der Anwendungsdomänen, die durch Experimente und Interventionen überprüft werden. Kausale Erklärungsmodelle (z.B. Klimamodelle oder Populationsmodelle) erfüllen die Gesetzes- und Strukturannahmen einer Theorie (z.B. Thermodynamik oder komplexe dynamische Systeme):

Beim kausalen Schließen werden Eigenschaften von Daten und Beobachtungen aus Kausalmodellen, d.h. Gesetzesannahmen von Ursachen und Wirkungen, abgeleitet. Kausales Schließen ermöglicht damit, die Wirkungen von Interventionen oder Datenveränderungen (z.B. durch Experimente und Interventionen) zu bestimmen.

Kausales Lernen versucht umgekehrt, ein Kausalmodell aus Beobachtungen, Messdaten und Interventionen (z.B. Experimente) abzuleiten, die zusätzliche Gesetzes- und Strukturannahmen voraussetzen.

Bei komplexen Systemen wie neuronalen Netzen mit z.B. Tausenden von Elementen und synaptischen Verbindungen erlauben zwar die Gesetze der statistischen Physik, globale Aussagen über Trend- und Konvergenzverhalten des gesamten Systems zu machen. Die Zahl der empirischen Parameter der einzelnen Elemente ist jedoch unter Umständen so groß, dass keine lokalen Ursachen ausgemacht werden können. Das neuronale Netz bleibt für uns eine „Black Box“. Vom ingenieurwissenschaftlichen Standpunkt aus sprechen Autoren daher von einem „dunklen Geheimnis“ im Zentrum der KI des Machine Learning: *“. . . even the engineers who designed [the machine learning-based system] may struggle to isolate the reason for any single action”* (Knight 2017).

Zwei verschiedene Ansätze im Software Engineering sind denkbar:

1. Testen zeigt nur (zufällig) gefundene Fehler, aber nicht alle anderen möglichen.
2. Zur grundsätzlichen Vermeidung müsste eine formale Verifikation des neuronalen Netzes und seiner zugrundeliegenden kausalen Abläufe durchgeführt werden.

Zusammengefasst folgt: Machine Learning mit neuronalen Netzen funktioniert, aber wir können die Abläufe in den neuronalen Netzen nicht im Einzelnen verstehen und kontrollieren. Heutige Techniken des Machine Learning beruhen meistens nur auf statistischem Lernen, aber das reicht nicht für sicherheitskritische Systeme, die nachhaltig sein sollen. Daher sollte Machine Learning mit Beweisassistenten und kausalem Lernen verbunden werden. Korrektes Verhalten wird dabei durch Metatheoreme in einem logischen Formalismus garantiert (Mainzer 2021; Mainzer/Schuster/Schwichtenberg 2021).

2.4 Von der symbolischen und subsymbolischen KI zur hybriden KI

Die Modelle subsymbolischer KI erinnern an die Organisation des Lernens im menschlichen Organismus: Verhalten und Reaktionen laufen dort ebenfalls weitgehend unbewusst ab. „Unbewusst“ heißt, dass wir uns der kausalen Abläufe des durch sensorielle und neuronale Signale gesteuerten Bewegungsapparats nicht bewusst sind. Das lässt sich mit Algorithmen des statistischen Lernens automatisieren. In kritischen Situationen reicht das aber nicht aus: Um mehr Sicherheit durch bessere Kontrolle im menschlichen Organismus zu erreichen, muss der Verstand mit kausaler Analyse und logischem Schließen eingreifen. Dieser Vorgang sollte im Machine Learning durch Algorithmen des kausalen Lernens und logischer Beweisassistenten automatisiert werden. Ziel ist daher eine hybride KI, in der analog zum menschlichen Organismus symbolische und subsymbolische KI verbunden werden.

3 Digitale Kreislaufwirtschaft als Beispiel für Nachhaltigkeit

In den folgenden Empfehlungen werden die Potentiale der Digitalisierung und künstlichen Intelligenz für Kreislaufwirtschaft herausgestellt. In einer traditionellen linearen Wirtschaft werden Rohstoffe in Produkte für Konsum und Verbrauch umgewandelt, die Restmüll erzeugen. Linearität bedeutet, dass mehr Konsum mehr Rohstoffverbrauch erfordert und mehr Restmüll erzeugt. Es kommt darauf an, Verbrauch und Konsum in einen Kreislauf zurück zu koppeln (Recycling), um Rohstoffverbrauch und Umweltbelastung mit z.B. Restmüll zu minimieren. Verbrauchte Produkte und Restmüll sollten so weit als möglich wieder dem Kreislauf zugeführt werden. Die Natur selbst hat solche Kreislaufprozesse im Lauf der Evolution erfolgreich erprobt.

Kreislaufwirtschaft stellt daher einen wesentlichen Schritt für eine nachhaltige Wirtschaft dar. Bisher kann das Potential nur für bestimmte digitale Technologien in eingeschränkten Fallanalysen demonstriert werden. Um das Potential digitaler Technologien für geschlossene Ressourcenzyklen umfassend zu verstehen, müssen ihre Möglichkeiten und Wechselwirkungen mit soziotechnischen Anwendungen systematisch untersucht werden. Dabei darf allerdings nicht vergessen werden: Jeder soziotechnische Fortschritt muss mit geeigneten Kommunikationsstrategien begleitet werden, da diese Veränderungen und Eingriffe (i) durch die Gesellschaft akzeptiert und getragen und (ii) in Geschäftsmodelle umgesetzt werden müssen, um nachhaltig zu sein. Zur digitalen und KI-gestützten Kreislaufwirtschaft werden folgende Empfehlungen gegeben (TUM Senior Excellence Faculty 2021):

4 Handlungsempfehlungen:

- 1) *Als Grundlage und Voraussetzung für eine nachhaltige Anwendung von KI-Tools benötigt Deutschland zunächst eine mächtige digitale Infrastruktur und Digitalisierung, die stärker ausgebaut und gefördert werden muss, besonders im öffentlichen Sektor und in der mittelständischen Industrie.*

Kreislaufwirtschaft als zentrales Beispiel der Nachhaltigkeit besteht aus sehr komplexen Prozessen mit hohem Datenaufkommen, die ein Ressourcenmanagement mit geschlossenen Rückkopplungsschleifen erfordern. Daher kann Kreislaufwirtschaft nur erfolgreich sein, wenn sie durch effektive digitale Infrastrukturen unterstützt wird. Diese Forderung ist eine notwendige Bedingung für die digitale Umsetzung einer Kreislaufwirtschaft.

- 2) *Neben konkreten Technologien und Infrastrukturen muss die Bedeutung von Standards und Normen berücksichtigt werden.*

Standards und Normen unterstützen die Effektivität von Netzwerken, da technische Lösungen durch Normung aufeinander abgestimmt werden und ihre Diversität minimiert wird. Gesetzliche und ethische Normen fördern Vertrauen und Akzeptanz einer vertrauenswürdigen IT und KI.

- 3) *Vor neuen kostspieligen Investitionen sollten die bereits existierenden digitalen Tools darauf getestet werden, wie weit sie bereits Ressourcenmanagement in der Kreislaufwirtschaft unterstützen und verbessern können.*

Digitalisierung bedeutet zunächst, digitale Tools von IT-Ausrüstung, IT-unterstützte Kommunikation und Kontrolle zu nutzen, die zur Zeit noch nicht hinreichend im öffentlichen Sektor und im Mittelstand zur Verfügung stehen.

- 4) *Kosten und Nutzen in Unternehmen und öffentlicher Verwaltung sollten klar und deutlich genannt werden.*

Die Digitalisierung der Kreislaufwirtschaft muss sich in Unternehmen und im öffentlichen Sektor rechnen, um als ökonomisch sinnvoll akzeptiert zu werden.

- 5) *Multilateraler Datenaustausch entlang den Versorgungs- und Lieferungsketten und den Lebenszyklen von Produkten ist notwendig.*

Kreislaufwirtschaft ermöglicht nachhaltiges Ressourcenmanagement, Abfallreduktion und Recycling von Ressourcen. Größere Transparenz von Produktentstehung trägt dazu bei, den Verbrauch zu reduzieren und Unternehmen zu bewegen, den Ressourcenerwerb zu verändern.

- 6) *Neben etablierten IT-Tools sollten neue vielversprechende Technologien wie Blockchain angewendet, getestet und entwickelt werden.*

Blockchain besteht aus seiner Kette von verschlüsselten Datenblöcken, die durch kryptographische Codes verbunden und erweitert werden können. Durch diese Ketten von Datenblöcken können z.B. Lieferungs-, Versorgungs- und Produktionsabläufe detailliert digital abgebildet werden, um zu jedem Zeitpunkt ihre Transparenz, Kontrolle und ihren Zugriff zu garantieren. Blockchain ist eine neue IT-Technologie, die alle Arten von Ressourcen mit Tokens kennzeichnen kann. Dadurch erhalten sie eindeutige digitale Identifikationen (ähnlich einer digitalen Münze), die alle Beteiligten verfolgen und nutzen können. Das macht den Wert der Ressourcen mehr ersichtlich, erleichtert ein neues System für die Bepreisung und den Handel von Ressourcen und setzt Anreize, um ein zirkuläres Verhalten zu übernehmen. Da viele Anwendungen von Blockchain noch in einer Testphase sind und breite Anwendung benötigen, sollte mehr Forschung und Entwicklung von Blockchain für nachhaltige Kreislaufwirtschaft betrieben werden.

- 7) *Neben etablierten IT-Tools sollten aber auch neue vielversprechende Technologien wie KI angewendet, getestet und entwickelt werden.*

KI öffnet neue Möglichkeiten für eine breite Palette der Nachhaltigkeit: In der Landwirtschaft können Drohnen oder sensorbasiertes Monitoring angewendet werden, um die Lebensgrundlagen von Pflanzen sowohl in ökonomischer als auch ökologischer Sicht besser einschätzen zu können. In der Produktion lässt sich der Energieverbrauch durch intelligente Netzwerke und Robotik reduzieren. Der Lebenszyklus von Produkten kann durch vorausschauende Wartung ausgeweitet werden. Im Recycling und Abfallmanagement kann KI die Identifikation und das Sortieren von Abfall verbessern. Für die Effizienz des Bauens und das Energiemanagement bietet KI verbesserte Systemkontrolle, Regulation des Heizens, Kühlens und der Ventilation. Kurz: Maschine

learning sollte zur Beschleunigung und Optimierung von Versorgungsketten beitragen und die Zirkulation von Produkten, Komponenten und Materialien unterstützen.

- 8) *Eine Schlüsselforderung ist die Entwicklung nachhaltiger Geschäftsmodelle, die auf digitaler zirkulärer Ökonomie beruhen.*

Es gibt bereits vielversprechende Fallstudien zur Kreislaufwirtschaft durch Digitalisierung: Beispiele sind nachhaltige Hauswirtschaft (urban housing), nachhaltige Smart Cities, Metallrecycling und Plastikrecycling.

- 9) *Die ökologische, ökonomische und gesellschaftliche Balance von digitalen Tools und Infrastrukturen, d.h. ihre ökologische, ökonomische und gesellschaftliche Gesamtbilanz in der Kreislaufwirtschaft ist entscheidend.*

Beispielsweise muss die Effizienz digitaler Anwendungen und ihre ökonomischen und ökologischen Kosten des Energieverbrauchs ausgeglichen sein, um die Vorteile einer Kreislaufwirtschaft nutzen zu können.

5 Bibliography

- [HSW89] Hornik, Kurt / Stinchcombe, Maxwell / White, Halbert: Multilayer feedforward networks are universal approximators neural networks, in: Neural Networks 2 1989, 359-366.
- [Kw17] Knight, Will: The dark secret at the heart of AI, in: MIT Technology Review. April 11 2017, 1-22. 2000.
- [Mk90] Mainzer, Klaus: Knowledge-based systems. Remarks on the philosophy of technology and artificial intelligence, in: Journal for General Philosophy of Science 21 1990, 47-74.
- [Mk19] Mainzer, Klaus: Künstliche Intelligenz. Wann übernehmen die Maschinen? Berlin 2. Aufl. 2019.
- [Mk21] Mainzer, Klaus: Statistisches und kausales Lernen im Machine Learning, in: Philosophisches Handbuch der künstlichen Intelligenz (Hrsg. Klaus Mainzer). Berlin 2021
- [MSS21] Mainzer, Klaus; Schuster, Peter; Schwichtenberg, Helmut (Hrsg.): Proof and Computation. From Proof Theory and Univalent Mathematics to Program Extraction and Verification, World Scientific Singapore 2021.
- [Pj9] Pearl, Judea: Causality: Models, Reasoning, and Inference. Cambridge (Mass.) 2009.
- [Pf88] Puppe, Frank.: Einführung in Expertensysteme. Berlin 1988.
- [TUF21] TUM Senior Excellence Faculty (Hrsg.), TUM Forum Sustainability: Circular Economy, TUM (Technische Universität München) University Press: München 2021.

[Vv88] Vapnik, Vladimir N.: Statistical Learning Theory. New York 1998.

**Workshop: Künstliche Intelligenz
für kleine und mittlere Unternehmen
(KI-KMU 2021)**

Blueprint for a Production-Ready Information Retrieval System based on Multi-Modal Embeddings

André Ebert¹, Anika Apel², Piotr Chodyko², Kyle Hiroyasu², Festina Ismaili², Hyein Koo²,
Julia Kronburger¹, Robert Pesch¹

Abstract: Deep Learning models for mapping documents from different domains, e.g., text, images, and audio, into a common vector space, enable a seamless information retrieval between the different domains and, thus, significantly improve the user experience of many expert tools. Despite various models for multi-modal mappings presented in scientific literature, the implementation and integration remain a challenge within the industry, especially for small or medium-sized companies. Reasons are, that developing such retrieval systems for production use-cases is a non-trivial task, requiring scalable, reliable, and cost-efficient infrastructure, services as well as adequate Deep Learning models. We present a generic and flexible blueprint architecture, targeting the development of a production-ready image-text retrieval search system using *Kubernetes*, *MLflow*, *Elasticsearch*, and integrating state-of-the-art Deep Learning models.

Keywords: Information Retrieval; Image-to-Text; Multi-Modal Embeddings; Deep Learning; Artificial Intelligence; Data-Science to Production

1 Introduction

With the increasing adoption of smart devices and the explosion of available data, methods to search and retrieve information have grown in importance. In the most familiar settings, text-based retrieval systems such as modern-day search engines, allow searching large corpora of documents to identify relevant information. Moreover, research has focused on multi-modal retrieval, which is the task of retrieving information from different forms of media to enable seamless information retrieval from various domains, e.g., texts, images, or audio records. Thus, for example, users of an image-text-retrieval system can search for the most relevant text documents using a photo as query input or vice-versa. Machine Learning and especially Deep Learning models have proven to be a powerful tool for such retrieval systems [?, ?, ?, ?, ?]. Numerous blog posts demonstrate the simplicity of implementing highly sophisticated models from the literature with millions of parameters using modern Machine Learning frameworks like *PyTorch* and *TensorFlow* in simple

¹ inovex GmbH, Ludwig-Erhard-Allee 6, 76131 Karlsruhe, Germany
aebert@inovex.de, jkronburger@inovex.de, rpesch@inovex.de

² Technische Universität München, Fakultät für Informatik, Boltzmannstr. 3, 85748 Garching, Germany
anika.apel@tum.de, piotr.chodyko@gmail.com, kylehiroyasu.tum@gmail.com, festina.ismaili@gmail.com, hyeinkoo@gmail.com

notebooks. Unfortunately, the transformation of these prototypes into a data product in the industry proves to be a non-trivial task. Apart from rigorous model evaluation, advanced data engineering, software engineering, and IT-operations know-how are required to address the hidden technical debts in Machine Learning systems [?], e.g., for building and maintaining a serving infrastructure, addressing scalability issues, introducing *Continuous Integration and Continuous Development (CI/CD)* pipelines, including model versioning, dealing with prediction uncertainties, and model as well as system monitoring.

This work describes the implementation of an image-text-retrieval system based on deep multi-modal embeddings and its deployment within a scalable, lean, and robust architecture. In this scenario, an image is used as a query to retrieve text documents with a high relevance. Besides objectives like rapid response times and the quality of results, the system's production readiness, as well as technical details concerning its architecture, are of importance. In this context, the presented system delivers the blueprint for deploying and operating an information retrieval system based on deep multi-modal embeddings, using in-cloud and on-premise based container environments.

2 Deep Learning Approaches for Multi-Modal Information Retrieval

Deep Learning based image-text-retrieval approaches can be categorized based on the way the final cross-modal embeddings are obtained [?]: (i) *Pairwise Learning*, (ii) *Interaction Learning*, and (iii) *Attributes Learning*.

Pairwise Learning methods aim at embedding texts and images in a joint vector space. The goal is to obtain semantically similar vector embeddings for images and texts. These approaches consist commonly of two branches, one for encoding images and another one for encoding texts. The outcomes of both branches are vectors, whose similarity can be calculated. The measurement of their similarity provides a supervisory signal, that is used to improve the encoding during training. Prominent examples of such methods are *VSE++* and *DCMP* [?, ?]. In contrast, *Interaction Learning* methods like *Oscar* or *CAMP* [?, ?] tend to achieve high accuracy rates compared to other approaches. Images and texts are not encoded separately and information flows between both branches before obtaining the embeddings in the joint vector space. This enables learning of correspondences between particular parts of images and texts. However, such interactions make these methods prohibitively expensive in terms of computation. Methods of *Attributes Learning*, such as *VSRN* and *ACMM* [?, ?], obtain high-level feature texts and images to compute correlations between those. Compared to interaction methods there is no flow of information between text and image branches when creating vector representations.

In case of the production-ready system targeted by this work, the integrability in a target architecture and the fast retrieval of decent results is decisive. Thus, the sheer accuracy scores of a model is not the only selection criterion, as described in the following section.

3 Requirements

In order to cover the demands of a highly efficient, generic, and capable platform for multi-modal information retrieval, its fundamental requirements are specified in prior. In order to address specific needs and preconditions as they can be found within small and medium sized businesses, the requirements were derived in collaboration with the consortial partners of the Service-Meister project as follows:

Generalizability: The architecture design allows easy customization to enable the implementation of a diverse range of different applications.

Affordability: Despite the in general demanding features of a state-of-the-art platform relying on Deep Learning, the necessary expenses must scale with the actually used resources and also be affordable for small- and medium-sized businesses.

Scalability: In order to handle a sparse as well as a massive amount of incoming requests, the architecture must be implemented in a scalable way. New resources, such as storage, virtual environments, and processing power are acquired on-demand and without any manual intervention.

Modularity and Extensibility: These aspects cover two crucial challenges: 1) Machine Learning models must be switchable without additional effort in order to deploy such with a better performance in a seamless way, 2) these models must be provided and integrated in a standardized and auditable way. A generic and easy to customize design, which allows the applicability of the architecture to different use cases must be ensured.

Fast Response Time and Usability: Both of these issues are of fundamental importance when it comes to user acceptance of the provided application. Ideally, response times for requests and searches are nearly real-time and a comfortable user interface is provided.

Accuracy and Result Quality: Models generating results must be highly performant and adjustable, in case of new information becomes available. Therefore, capable resources for retraining and evaluation of models must be in reach, also during active service usage.

Robustness and Restorability: The whole architecture must be robust in terms of workload as well as in terms of restorability. Trained models are versioned, all code is deployed in an automatized way and the execution environment, as well as resources, are persisted to ensure the possibility of recovery from scratch.

In the following, the selected Deep Learning model for computing multi-modal embeddings as well as a blueprint architecture for implementing a retrieval system based on such embeddings is presented.

4 VSE++ for Deep Multi-Modal Embeddings

VSE++[?] is a state-of-the-art and well documented (in the scientific literature, on *GitHub*, and various blog posts) multi-model retrieval model based on *Pairwise Learning* for transforming images and text documents into a joint vector space. Such embeddings can be

efficiently pre-computed, stored, and indexed for later retrieval. Thus, *VSE++* is used as Deep Learning model for the presented architecture.

In detail, *VSE++* transforms an image i with an image encoder $\phi(i, \Theta_\phi)$, whereas a caption c is transformed by a text encoder $\Psi(c, \Theta_\psi)$ to obtain the corresponding vector embeddings. Θ_ϕ and Θ_ψ express the model parameters of the image and the text encoder, respectively. Subsequently, these embeddings are mapped into a joint vector space:

$$f(i, W_f, \Theta_\phi) = W_f^T \phi(i, \theta_\phi), \quad (1)$$

$$g(c, W_f, \Theta_\psi) = W_f^T \psi(i, \theta_\psi). \quad (2)$$

The similarity within the joint vector space $s(i, c)$ can be measured by the *Cosine Similarity*. The model is trained using a loss function focused on hard negatives. Hence, for an image-text pair (i, c) one wants to maximize its dissimilarity with the most similar incorrect image and caption in the batch. The dimensionality of the joint embedding space is set to $n = 1024$. Therefore, a *Resnet152* is used to encode images and a *GRU* without pre-training is used to encode text [?]. Before feeding images into the encoder, they become centered, randomly cropped, and resized to a shape of 224×224 . In order to preprocess related captions, a vocabulary was determined by using available training data. The text is tokenized using *NLTK* [?], one-hot encoded with respect to our vocabulary, and then passed through an embedding layer to obtain word embeddings of a dimension of 300, which serves as the input of the text encoder.

5 System Architecture of a Deep Learning based Retrieval System

Within this section, an overview of the proposed blueprint architecture for deep, multi-modal information retrieval is given. Figure 1 illustrates its individual components as well as the flow of information between them. In the current use-case, the architecture uses a state-of-the-art Deep Learning model to create multi-modal embeddings. These can be utilized to retrieve text objects corresponding semantically to the given image objects. In general, the architecture consists of two core components: 1) a *Google Cloud* based *Kubernetes* cluster for the orchestration of docker images and the scalable provision of docker containers and 2) an on-premise *GPU Cluster* for the training of Deep Learning models.

5.1 Cloud-based Services in the Kubernetes Cluster

A fundamental corner-stone for orchestrating and deploying a data-product into production is a scalable and flexible container system, which allows rapid up and downscaling for the optimal utilization of resources in an automatized fashion. Currently, *Kubernetes*³

³ <https://kubernetes.io>

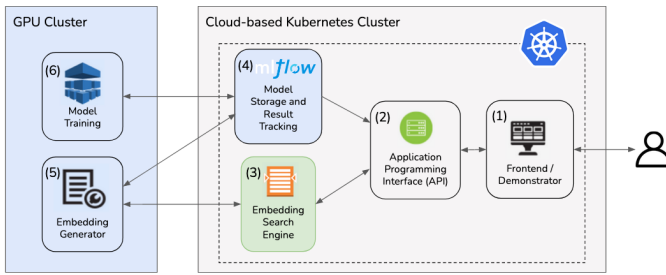


Fig. 1: Overview of the proposed blueprint architecture.

is the de-facto standard for deploying and orchestrating *Docker*⁴ containers. A managed *Kubernetes* cluster as Infrastructure-as-a-Service (IaaS) can be easily deployed in all major hyperscaler environments (e.g., *Microsoft Azure*⁵, *Amazon Web Services (AWS)*⁶, and *Google Cloud Platform (GCP)*⁷). In this work, a managed *Kubernetes* cluster is deployed in the GCP, relying on four small nodes having two virtual CPUs and 8GB memory, each. The cluster serves the 1) *Frontend/Demonstrator*, the 2) *Application Programming Interface (API)*, the 3) *Embedding Search-Engine*, and the 4) *Model Store and Result Tracking* component. Load balancing is done with *Kubernetes Ingress* by routing external *HTTP(S)* traffic to internal services. Thereby, load balancing enforces a distributed network traffic across multiple workers and it is assured, that a single service never carries an excessive demand. In addition, it offers the flexibility to add or remove nodes and resources based on the current demand.

1) Frontend and Demonstrator with Vue.js In order to allow rapid testing and the acquisition of user feedback concerning the retrieval system’s result quality, an adequate frontend is developed and provided to service users.

2) Application Programming Interface (API) To expose the functionalities of the data-product, a backend is implemented. It accepts *REST* requests, encapsulates the computation of embeddings, and queries the *Embedding Search Engine*. It is implemented by using *FastAPI* and operated by using *Uvicorn* — an *Asynchronous Server Gateway Interface (ASGI)* server, which is an asynchronous, performance optimized, and Python based web server. The entire backend is bundled as a *Docker* image and deployed into the *Kubernetes* cluster. There are two exposed interfaces: (i) to compute embeddings for a given image and (ii) to search for similar documents by using the embeddings.

3) Embedding Search-Engine By using *Pairwise Learning* models with joint embedding spaces for images and text, it is possible to pre-compute and store all vector representations of the underlying retrieval database at once. The retrieval step can be therefore formulated

⁴ <https://www.docker.com/>

⁵ <https://portal.azure.com>

⁶ <https://aws.amazon.com>

⁷ <https://cloud.google.com>

as: (i) computing an embedding of an image and (ii) finding the most similar text embedding among all pre-computed embeddings in the collection of relevant documents. To search for most-similar documents using a given embedding, there are several popular solutions: *Faiss* [?], *ScaNN* [?], (experimental) extensions on ordinary SQL databases such as PostgreSQL⁸, and even approaches based on *SQLLite* such as *Magnitude* [?]. For the architecture presented within this work, *Elasticsearch*⁹ was used. Since *Elasticsearch* version 7.3, embeddings can be indexed as vectors and advanced queries on vectors based on the *Cosine Similarity*, the dot product, the *l1*, and the *l2* norm have been introduced. Thereby, pre-computed embeddings from *VSE++* can be stored directly in the search engine. Besides its popularity and adaption in many enterprise environments, *Elasticsearch* scales easily and showed promising performance results within the experiments in Section 6 for indexing and retrieving in a collection of ten thousands of embeddings.

4) Model storage and tracking of results with MLflow Experiments and trained models for computing multi-modals embeddings are stored in the model management service *MLflow*¹⁰. In particular, the following three features of *MLflow* are of specific importance: *MLflow Tracking*, *MLflow Models*, and the *Model Registry*. Trained models are serialized in the environment-independent *torchscript* format, a unified framework to deploy models for inference in production without requiring the source code or an installed version of *PyTorch*. These models can be easily tagged and used for generating embeddings. *MLflow* comes with a *Python Application Programming Interface*, a user interface, as well as extensions for *PyTorch*, which simplifies the integration into the proposed system.

IT-Engineering and Operations All source code of the above mentioned workflows is versioned in *Gitlab*. Automated CI/CD processes build and test docker images as well as deploy them into *Kubernetes* or the *GPU Cluster*.

5.2 Deep Learning Infrastructure

The model training is done on an on-premise *GPU Cluster*: standard Linux machines (Intel i7-5930K) with 64GB RAM containing two *nVidia Titan X* graphics cards with 12GB video memory, each. Parallel GPU computation is realized by using *Cuda*¹¹. Depending on data and information policies as well as the capabilities of hosting a *GPU Cluster*, the model training can also be migrated to a cloud service as well. Within the *GPU Cluster*, there are two relevant components: the 5) *Embedding Generator*, and the 6) *Model Training* (see Figure 1).

⁸ For example *pgvector*, see <https://github.com/ankane/pgvector>

⁹ <https://www.elastic.co>

¹⁰ <https://www.mlflow.org/>

¹¹ <https://developer.nvidia.com/cuda-zone>.

5) Embedding Generator Here, the latest, stable model from *MLflow* is retrieved and applied to generate embeddings from a given collection of documents. Subsequently, these embeddings are stored within the search engine.

6) Training of Machine Learning Models with PyTorch Workflows for training and evaluating *VSE++* model are based on *PyTorch Lightning5*. For model training itself, workflows adapted by the *VSE++* authors are used. Newly trained models are evaluated based on the retrieval metrics *Precision@k* and *Recall@k* (*P@K*, and *R@k*).

6 Results and Discussion

Datasets	Size (images)	Train	Validation	Test
Flickr30k	31.783	29.783	1.000	1.000
COCO	123.287	113.287	5.000	5.000
Total	155.070	143.070	6.000	6.000

Tab. 1: Dimensions and splitting of the combined dataset for model training and evaluation, consisting of *MS COCO* and *Flickr30k*.

In the following, the features of the proposed blueprint architecture introduced in Section 5 are reviewed, discussed, and matched in terms of fulfilling the requirements stated in Section 3.

To evaluate the performance we trained *VSE++* and applied the proposed system onto the *Microsoft Common Objects in Context (MS COCO)* and the *Flickr30k* datasets, which can be regarded as the state-of-the-art datasets for performance measurement within this domain. They are consisting of 155.070 matched images and text documents (see Table 1).

Generalizability The application is decomposed into a set of manageable services which are faster to develop, easier to understand, and can be updated independently. Thereby, the underlying micro-service architecture allows an easy transformation and adaption to a broad variety of different applications and services. Moreover, the modularity of containers orchestrated by *Kubernetes* within the *GCP* enables a fast exchange and up scaling of all functionalities.

Affordability Table 2 gives a brief overview of the average monthly costs for the proposed architecture. These costs provide a broad overview, while exact pricing may change due to selected hyperscaler, used computing engines, storage accounts, etc. They were accrued from running all the services of the entire application for one month without auto-scaling. The major costs are associated with components delivered by the *GCP Compute engine*. Additionally, expenses for the on-premise *GPU Cluster* must be considered, which would be added to the expenses of the cloud platform if hosted there. In a production environment, the utilization of features like *Kubernetes* auto-scaling allows the available resources to both, grow and shrink, which leads to correlating costs into both directions. Thereby, the

Service	Resource Type	Subtotal
Compute Engine	E2 Instance Core	153.36€
Compute Engine	E2 Instance Ram	82.20€
Compute Engine	Storage PD Capacity	26.47€
Compute Engine	Network Load Balancing	17.82€
Compute Engine	Misc. Networking Services	14.58€
Cloud Storage	Download Worldwide Destinations	18.88€
Cloud Storage	Standard Storage	0.71€
Cloud Storage	Regional Standard Class B Operations	0.09€
	Total	314.11€

Tab. 2: A high-level overview of the average operational costs for one month.

expenses scale with the demand to the platform, which means if a business is small or a new business idea is formed, it can be brought into production in an affordable way. If the demand gets higher due to heavy usage of the service, it may become more expensive, but also the revenue created by the provided service may be larger.

Scalability By leveraging the *Kubernetes* capabilities, the blueprint benefits from technology agnostic frameworks. Thereby, the flexibility to deploy any application resource or web framework needed for the provision of a service is given. Due to features like service discovery, load balancing, horizontal scaling, and self-healing, the utilization of *Kubernetes* is another important factor for achieved scalability, high-availability, and fault tolerance.

Modularity and Extensibility Developing Deep Learning based applications and bringing them into production raises several new challenges. The Machine Learning lifecycle, a complex and iterative process including data collection, model training, verification, deployment, and monitoring must be considered [?]. To ensure a fluent model deployment, reproducibility, and an easy expansion of the application, standardized model versioning and model management is important. To overcome these challenges, *MLflow Models* and *Model Registry*, both part of *MLflow*, are used. Thereby, trained models can be packaged, versioned, and utilized in an easy fashion. To deploy them in a unified way, they become serialized with *torchscript*. Thereby, the serialized model is self-contained including all requirements and the process itself is generic. Due to the versioning within the Model Registry, the deployment of most recent models, as well as a rollback to previously deployed versions, are feasible. Moreover, tagging and seamless deployment of models into the system without any additional effort is possible.

Fast Response Time and Usability In order to make a user accept a service, it must be able to offer high quality results and to deliver results with a rapid response time. In the context of this work, the fast retrieval of the textual results is based on the similarity of joint vector embeddings that offer the advantage of pre-computation. Due to only processing incoming request documents, the response time is lowered significantly. In order to store and search for pre-computed embeddings, *Elasticsearch* which is capable of storing and retrieving

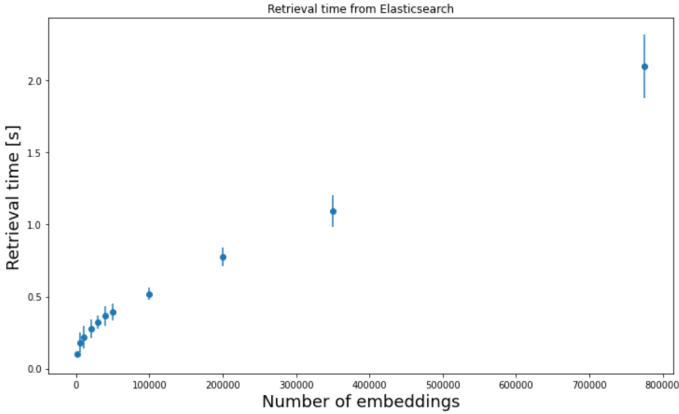


Fig. 2: Resulting retrieval time of *Elasticsearch* as an embedding search engine containing a varying number of pre-computed multi-modal embeddings.

dense vectors in a distributed way is used. Figure 2 shows the result retrieval times for a different number of pre-computed and stored vector embeddings. Within a range 0 – 10.000 embeddings, the retrieval time takes significantly less than 1s. For a number of more than 70.000 embeddings, the retrieval time ranges around 20 seconds, which is still comparably low concerning the huge amount of stored information. These results prove the adequate capabilities of the embedding search engine.

Additionally, the integration of *Elasticsearch* into *Kubernetes* is comparably easy, it is well-established, matured, and offers a decent documentation as well as support. Next to a fast response time and a decent result quality, the provision of an intuitive, appealing, as well as functional user interface is mandatory for a high usability and the acceptance of a service by potential users.

Model	R@1	R@3	R@5	P@1	P@3
VSE++	0.343	0.542	0.634	0.343	0.388

Tab. 3: Results for training *VSE++* on a combined dataset consisting of *COCO* and *Flickr30k*, see Figure 3.

Accuracy and Result Quality In order to deliver high quality results, models can continuously be trained on new data, even during times of active service operation. This is ensured by providing a separate computation instance in terms of the *GPU Cluster*. With this feature in combination with the options of versioning and delivering new models via *MLflow*, an optimal result quality can be guaranteed.

Table 3 shows the results of the *VSE++* model for the retrieval metrics *Precision@k* for $k=1$ and $k=3$ as well as *Recall@k* with $k=1$, $k=3$, and $k=5$. The achieved results of the re-implemented models used in the blueprint architecture resemble with the ones achieved by authors of *VSE++* [?].

Robustness and Restorability In terms of robustness, the architecture is backed by the *Kubernetes* auto scaling and load balancing features. Due to the appliance of state-of-the-art *DevOps* tools like *Gitlab*, and *Gitlab CI* the complete restorability of the needed infrastructure is ensured. In order to also preserve the underlying resource architecture, tools like *Terraform* could be applied. Moreover, the usage of unit and integration tests during the deployment process assures the correct functionality of the published code.

7 Conclusion and Outlook

We presented a scalable, modular, and highly responsive multi-modal retrieval system — a retrieval system for seamless information retrieval from different domains such as text, image, and audio — based on deep learned embeddings and loosely coupled services in a *Kubernetes* cluster. Despite the *Kubernetes* cluster was deployed in a *IaaS* manner into the *GCP*, its design is generic and can be deployed in any other cloud or on-premise environment. GPU intensive Deep Learning training and indexing were conducted on dedicated hardware; this can also be done in the cloud, either directly on the *Kubernetes* cluster with additional GPU resources, or via services such as Google Colab¹². The multi-modal retrieval step is solved by using the VSE++ Deep Learning model, which maps documents from different domains into a common vector space. Thereby, semantically similar features are mapped to similar locations. The *Pairwise Learning* approach VSE++ provides reasonable good retrieval results and requires comparable few computational power for computing embeddings. To evaluate our implementations and benchmark the system the *COCO* and the *Flickr30k* dataset (consisting of matched images and text documents) were used. Finally, with *Elasticsearch*, *MLflow*, and an implemented REST-based backend the draft for a general, scalable, and cost-efficient architecture for a multi-modal retrieval system was realized. The presented architecture can be utilized and extended for any multi-modal information retrieval system. Thus, this blueprint can be used as a generic guideline by small as well as medium-sized companies to bring their Deep Learning applications into production.

Acknowledgement The findings and concepts presented in this work are partially funded by the *Service-Meister* project - funded by the *Innovationswettbewerb KI* and the *Bundesministerium für Wirtschaft und Energie* of the Federal Republic of Germany.



References

- [ACP19] Ashmore, Rob; Calinescu, Radu; Paterson, Colin: Assuring the machine learning lifecycle: Desiderata, methods, and challenges. arXiv preprint arXiv:1905.04223, 2019.
- [BKL09] Bird, Steven; Klein, Ewan; Loper, Edward: Natural language processing with Python: analyzing text with the natural language toolkit. Ó'Reilly Media, Inc.", 2009.

¹² <https://colab.research.google.com/>

- [Ch20] Chen, Jianan; Zhang, Lu; Bai, Cong; Kpalma, Kidiyo: Review of Recent Deep Learning Based Methods for Image-Text Retrieval. In: 2020 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR). IEEE, pp. 167–172, 2020.
- [Fa18] Faghri, Fartash; Fleet, David J; Kiros, Jamie Ryan; Fidler, Sanja: VSE++: Improving Visual-Semantic Embeddings with Hard Negatives. In: Proceedings of the British Machine Vision Conference (BMVC). 2018.
- [Gu20] Guo, Ruiqi; Sun, Philip; Lindgren, Erik; Geng, Quan; Simcha, David; Chern, Felix; Kumar, Sanjiv: Accelerating Large-Scale Inference with Anisotropic Vector Quantization. In: International Conference on Machine Learning. 2020.
- [He16] He, Kaiming; Zhang, Xiangyu; Ren, Shaoqing; Sun, Jian: Identity mappings in deep residual networks. In: European conference on computer vision. Springer, pp. 630–645, 2016.
- [HW19] Huang, Yan; Wang, Liang: Acmm: Aligned cross-modal memory for few-shot image and sentence matching. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 5774–5783, 2019.
- [JDJ19] Johnson, Jeff; Douze, Matthijs; Jégou, Hervé: Billion-scale similarity search with GPUs. IEEE Transactions on Big Data, pp. 1–1, 2019.
- [Li19] Li, Kunpeng; Zhang, Yulun; Li, Kai; Li, Yuanyuan; Fu, Yun: Visual semantic reasoning for image-text matching. In: ICCV. 2019.
- [Li20] Li, Xiujun; Yin, Xi; Li, Chunyuan; Zhang, Pengchuan; Hu, Xiaowei; Zhang, Lei; Wang, Lijuan; Hu, Houdong; Dong, Li; Wei, Furu et al.: Oscar: Object-Semantics Aligned Pre-training for vision-language Tasks. In: European Conference on Computer Vision. Springer, pp. 121–137, 2020.
- [Pa18] Patel, Ajay; Sands, Alexander; Callison-Burch, Chris; Apidianaki, Marianna: Magnitude: A Fast, Efficient Universal Vector Embedding Utility Package. In: Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations. Association for Computational Linguistics, Brussels, Belgium, pp. 120–126, November 2018.
- [Sc15] Sculley, David; Holt, Gary; Golovin, Daniel; Davydov, Eugene; Phillips, Todd; Ebner, Dietmar; Chaudhary, Vinay; Young, Michael; Crespo, Jean-Francois; Dennison, Dan: Hidden technical debt in machine learning systems. Advances in neural information processing systems, 28:2503–2511, 2015.
- [Wa19] Wang, Zihao; Liu, Xihui; Li, Hongsheng; Sheng, Lu; Yan, Junjie; Wang, Xiaogang; Shao, Jing: Camp: Cross-modal adaptive message passing for text-image retrieval. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 5764–5773, 2019.
- [ZL18] Zhang, Ying; Lu, Huchuan: Deep cross-modal projection learning for image-text matching. In: Proceedings of the European Conference on Computer Vision (ECCV). pp. 686–701, 2018.

Brezel-Cast: Verkaufsprognose von Backwaren

Nico Döring¹, Jonathan Kreiss¹, Thomas Schuster ¹ und Raphael Volz ¹


Abstract: In diesem Papier diskutieren wir die Anwendbarkeit von Verfahren der künstlichen Intelligenz zur Prognose von Absatzzahlen für eine Bäckerei mittlerer Größe. Dabei wird beschrieben, wie bei der Entwicklung zusätzliche Daten (Kontextinformationen) zur Prognose genutzt werden. Daraufhin werden zwei Verfahren des maschinellen Lernens trainiert und im Ergebnis miteinander verglichen. Neben einer abschließenden Bewertung und Ausblick auf zukünftige Verbesserungen, wird zudem eine Einschätzung zum Einsatz im Produktivbetrieb abgegeben.

Keywords: Künstliche Intelligenz, Maschinelles Lernen (ML), Absatzprognose, Nachhaltige Produktion

1 Einleitung

Künstliche Intelligenz (KI) durchdringt immer mehr Bereiche unseres Lebens und Unternehmen setzen zunehmend auf KI-Verfahren wie das maschinelle Lernen (ML). Das Potential von ML-Verfahren ist neben der Verbesserung von Produkten und Dienstleistungen auch eine Steigerung der Prozesseffizienz. Während viele Anwendungsfälle derzeit von Großunternehmen beschrieben werden, setzen inzwischen auch zunehmend mittelständische Betriebe auf ML. Ein typischer Anwendungsfall für den Einsatz von ML ist die Prognose künftigen Kundenverhaltens wie zum Beispiel von Verkaufszahlen bestimmter Produkte. Besonders bei schnelllebigem Konsumgütern wie Nahrungsmitteln, bringt dies viele Vorteile mit Blick auf eine effiziente, nachhaltige Produktion und Logistik. Dies gilt nicht nur für große Supermarktketten, sondern auch für kleinere Betriebe (Bäckereien). In diesem Papier zeigen wir, wie Verfahren des maschinellen Lernens (ML) in einer Bäckerei eingesetzt werden kann, um Absatzzahlen prognostizieren zu können. Wir zeigen, wie die Verkaufsmenge einzelner Produkte pro Filiale prognostiziert und so die Menge von Retouren und Abverkäufe reduziert werden kann, um die Produktion nachhaltiger zu gestalten und Ausschuss zu reduzieren.

¹ Hochschule Pforzheim, Tiefenbronner Straße 65, 75175 Pforzheim,
{doeringi, kreissjo, thomas.schuster, raphael.volz}@hs-pforzheim.de,

 <https://orcid.org/0000-0002-9539-1627>, <https://orcid.org/0000-0003-2816-1618>

2 Verwandte Arbeiten

Die Vorhersage von Lebensmitteln und auch von Backwarenverkäufen ist eine Aufgabe der Zeitreihenprognose. Im Vergleich zu traditionellen Methoden zur Verkaufsprognose, können Ansätze des maschinellen Lernens eine höhere Prognosegenauigkeit, Flexibilität und Skalierbarkeit über große Datensätze bieten [Ta19]. [Ne10] vergleichen empirisch ML-Ansätze für Zeitreihenprognosen. Im Vergleich performen Neuronale Netze und Gaussian Process Regression am besten. K-Nearest Neighbour erzielt hingegen eine mittelmäßige Qualität der Prognose, zeigt im Vergleich aber einen geringeren Rechenaufwand. Aufgrund der Aufwände und der Qualität bei der Verkaufsprognose von salzigen Snacks eines Supermarktes in [Ba15] kann sich Random Forest als Verfahren im Vergleich durchsetzen. [BS11] zeigen erfolgreich die Anwendbarkeit von Entscheidungsbaumalgorithmen zur Prognose der Nachfrage in Food Courts. In [LI17] werden basierend auf Deep Learning und meteorologischen Daten Verkäufe von Lebensmitteln einer japanischen Supermarktkette prognostiziert. [SV19] nutzen ebenfalls Kontextinformationen (z.B. Wetterdaten, Ferientage) um die Nachfrage von Parkplätzen in Parkhäusern zu prognostizieren.

3 Ausgangssituation und Prognoseansatz

Gemeinsam mit einer mittelständischen Bäckerei, die derzeit mehrere Filialen betreibt und durch den zentralen Produktionsstandort täglich frisch beliefert, sollte untersucht werden wie genau Absatzzahlen für einzelne Produkte pro Filiale prognostiziert werden können und ob diese Vorhersage besser ist als die der eigenen Mitarbeiter. Somit kann die Produktion der gesamten Kette effizient auf die Nachfrage ausgerichtet werden – also sowohl Über- als auch Unterproduktion reduziert werden. Im Projekt wurden exemplarisch die Verkäufe von Brezeln analysiert. Die zentrale Forschungsfrage war daher: Kann ein ML-Modell die Verkaufsmenge von Brezeln für den nächsten Tag, pro Filiale genauer prognostizieren als der Bäckermeister?

Um diese Frage zu beantworten, wurden Algorithmen aus dem überwachten maschinellen Lernen evaluiert. Die Datengrundlage waren Verkaufszahlen der letzten vier Jahre (2016-2019), welche aus dem Kassensystem einer Filiale der Bäckerei extrahiert wurden. Im Hinblick auf die Algorithmen und durch Literaturanalyse ähnlicher Fragestellungen, wurden Regressionsansätze, Entscheidungsbaumalgorithmen (besonders Random Forest Verfahren) und Neuronale Netze als vielversprechend erachtet. Für die Prognose wurde untersucht wie geeignet die beiden Algorithmen K-Nearest Neighbour (kNN) und Random Forest sind, um die Verkaufsmenge der Brezeln vorherzusagen.

4 Konzeption und Entwicklung der Absatzprognose

In diesem Abschnitt beschreiben wir die Entwicklung der Absatzprognose über die

komplette Verarbeitungstrecke der Daten. Dazu gehen wir zuerst auf die Vorverarbeitung der Daten ein und vergleichen anschließend die beiden Ansätze miteinander. Der Abschnitt endet mit einer abschließenden Bewertung der Modelle.

4.1 Datenaufbereitung und Modellierung

Zunächst wurden die Daten des Kassensystems (CSV-Dateien) mit Hilfe von pandas eingelesen. Der Datensatz umfasste Verkaufszahlen einer Filiale im Zeitraum von 2016 bis 2019. Die Parameter des Datensatzes wurden gefiltert und auf, für die Prognose bedeutenden, reduziert. Dies waren: Verkaufsdatum und zugehörige Verkaufsmenge jeder Transaktion. Anschließend wurden Ausreißer und Stornierungen (negative Verkaufswerte) bereinigt. Zusätzlich wurden Transaktionen eines Tages aggregiert und zu $n = 1461$ ($365 \cdot 4 + 1$) Beobachtungen aufsummiert.

Um die Verkaufszahlen besser prognostizieren zu können, wurde der Datensatz um Kontextinformationen angereichert, also durch weitere, erklärende Variablen erweitert. Dazu gehörte beispielsweise Informationen über den jeweiligen Wochentag, ob es sich an einem Tag um einen Ferientag handelt und Wetterdaten. Die Wetterdaten (durchschnittliche Tagestemperatur und Niederschlagsmenge pro Tag) wurden über das Informationsangebot des Deutschen Wetterdienstes (DWD) [DWD21] bezogen, welches historische und auch regionale Wettermessungen kostenlos bereitstellt.

4.2 Datenanalyse

Zur Bewertung der unabhängigen Variablen wurden zunächst einige Visualisierungen vorgenommen. Zunächst wurde die Verkaufsmenge pro Tag anhand einiger Diagramme wie z.B. Histogramm, Liniendiagramm und Boxplots visualisiert. Das Ziel war es die Verkaufsdaten zu untersuchen und Trends, wie zum Beispiel Saisonalitäten zu entdecken. Anhand eines Histogramms wurde beispielsweise festgestellt, dass an den meisten Tagen, zwischen 60 und 175 Brezeln verkauft werden.

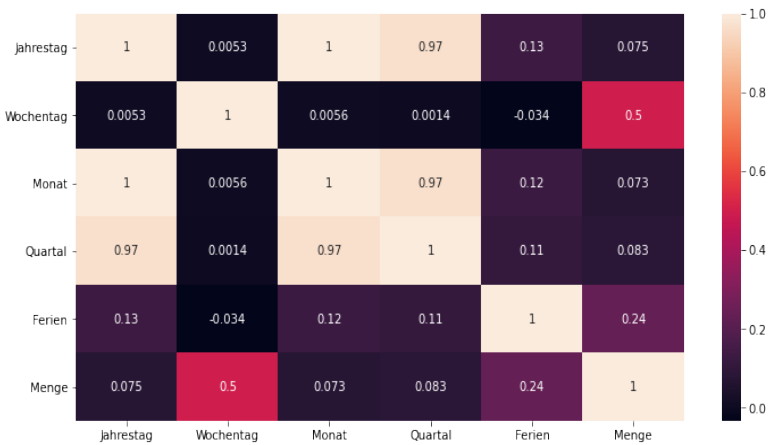


Abb. 1: Spearman Korrelationskoeffizient Verkaufsmenge zu Parametern

Anschließend wurde mit Hilfe des Spearman Koeffizienten die Korrelation zwischen der Verkaufsmenge und den einzelnen Parametern ermittelt (siehe Abb. 1). Erwartungsgemäß zeigte sich die höchste Korrelation (0,5) zwischen dem Wochentag und der Verkaufsmenge. Diese entspricht einer positiven mittelstarken Korrelation. Der Boxplot in Abbildung 2 verdeutlicht ebenfalls diesen Zusammenhang. Wie erwartet ist die Verkaufsmenge an Freitagen (4), Samstagen (5) und Sonntagen (6) höher als an anderen Wochentagen. Ein ähnliches Bild zeigte sich an Feiertagen.

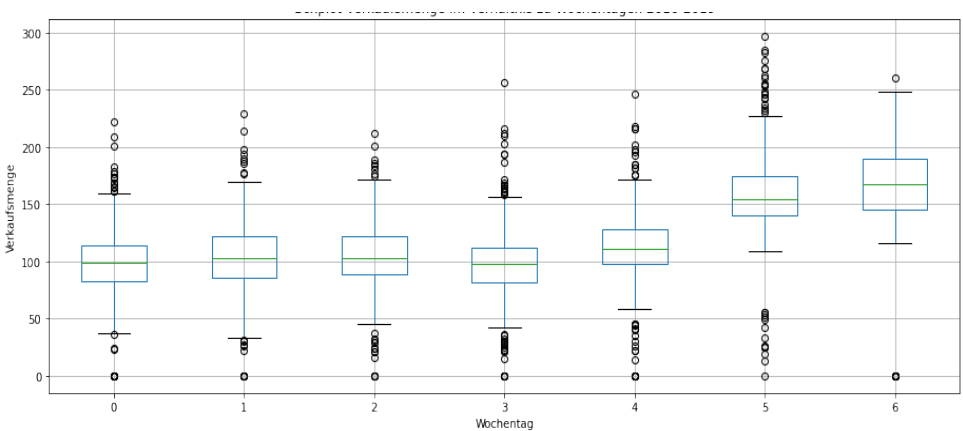


Abb. 2: Brezel Verkaufsmenge pro Wochentag von Montag (0) - Sonntag (6)

4.3 Modell Training

Das Trainingsverfahren wurde, durch ein Sliding Window erweitert [HHS17]. Dabei

wird ein Zeitfenster rollierend über dem Datensatz verschoben und durch die aus dem Fenster entnommenen Datenpunkte (Lag Variablen) erweitert. Dieser Ausschnitt an Vergangenheitswerten liefert somit zusätzliche Information für eine Prognose.

K-Nearest-Neighbour: Die Daten wurden in Trainings- und Testdaten aufgeteilt und durch eine Datenskalierung transformiert. Da der kNN Algorithmus nur den Hyperparameter k zum Tuning bereitstellt, wurde mit Hilfe des GridSearchCV-Algorithmus der optimale Wert $k=4$ ermittelt. Nach Training und Prognose der Werte aus dem Testdatensatz ergaben sich die Evaluierungsparameter aus Tabelle 1.

Evaluierungsparameter	Wert
Mean Absolute Error (MAE)	21.07
Prognosegenauigkeit in % aus MAE	78.64 %
Mean Squared Error (MSE)	768.89
Root Mean Squared Error (RMSE)	27.73

Tab. 1: Gemittelte Out-of-Sample Fehlerrate des kNN Modells

Random Forest: Die Daten wurden auch hier aufgeteilt. Um die Hyperparameterauswahl zu optimieren und eine Kreuzvalidierung zu ermöglichen, wurde eine randomisierte Suche (RandomSearchCV) und in einem zweiten Schritt der GridSearchCV-Algorithmus auf die eingegrenzten Parametermöglichkeiten angewandt. Als Baseline diente eine Random Forest Prognose mit den Standard-Hyperparametern (Default Modell) aus der scikit-learn Bibliothek. Die Prognose mit dem Default Modell erreichte eine Prognosegenauigkeit von 83.85%. Mit den Parametern aus der randomisierten Suche verbesserte sich das Ergebnis um 1.27% auf 85.12%. Eine Auswahl durch den GridSearchCV-Algorithmus, führte zu keiner weiteren Steigerung. Daher wurde die Annahme gestellt, dass das Nutzenoptimum erreicht wurde. Die wichtigsten Hyperparameter des Modells am Nutzenoptimum (Finales Modell) sind 1400 klassifizierende Bäume und eine maximale Baumtiefe von 10. Die Ergebnisse sind in Tabelle 2 zusammengefasst.

Evaluierungsparameter	Default Modell	Finales Modell	Verbesserung
Mean Absolute Error (MAE)	18.71	16.98	1.73
Prognosegenauigkeit in % aus MAE	83.85%	85.12%	1.27%
Mean Squared Error (MSE)	603.34	495.17	108.17
Root Mean Squared Error (RMSE)	24.56	22.25	2.31

Tab. 2: Gemittelte Out-of-Sample Fehlerrate des Random Forest Modells

4.4 Evaluation

Aus unserem Vergleich der Machine Learning Algorithmen hat das Finale Modell des Random Forest mit einer Prognosegenauigkeit von 85.12% am besten abgeschnitten. Der kNN-Algorithmus konnte nur eine Prognosegenauigkeit von 78.64% liefern. Im Vergleich ist somit die Prognose des Random Forrest im Schnitt um 6.48% besser. Auch konnte das MSE beim Random Forest deutlich gesenkt werden, was für die Robustheit

des Modells spricht. Die Ergebnisse können nun genutzt werden, um weitere Anpassungen vorzunehmen und die Prognosegenauigkeit des Random Forests zu verbessern. Zudem kann es als Baseline für Vergleiche mit weiteren Verfahren dienen.

Evaluierungsparameter	kNN	Random Forest	Verbesserung
Mean Absolute Error (MAE)	21.07	16.98	4.09
Prognosegenauigkeit in % aus MAE	78.64%	85.12%	6.48%
Mean Squared Error (MSE)	768.89	495.17	273.72
Root Mean Squared Error (RMSE)	27.73	22.25	5.48

Tab. 3: Vergleich kNN zu Random Forest

5 Fazit und Ausblick

Zur Prognose des Absatzes von Backwaren wurden in diesem Projekt zwei Algorithmen, der Random-Forest-Algorithmus und der K-Nearest-Neighbour, untersucht, getestet und miteinander verglichen. Das Ziel war es die Verkaufsmenge der Brezeln, einer Bäckereifiliale möglichst genau für einen Tag vorherzusagen. In ersten Analysen kristallisierte sich heraus, dass der Wochentag einen starken Einfluss auf die Verkaufsmenge hat. Nachdem die Daten entsprechend aufbereitet und mit Kontextinformationen (Wochen- und Feiertage sowie Wetterdaten) angereichert wurden, wurden die Modelle trainiert. Nach Hyperparameteroptimierung und Evaluation der beiden getesteten Algorithmen, erzielte der Random Forest-Algorithmus das beste Ergebnis, mit einer Prognosegenauigkeit von über 85% und einer Abweichung von nur 17 Brezeln gegenüber den tatsächlichen Absatzzahlen. Von der Bäckerei wissen wir, dass man sich bei einer manuellen Prognose etwa um die doppelte Menge verschätze. Somit könnte künftig die manuelle Einschätzung der Mitarbeiter ersetzt werden. In einem nächsten Schritt können weitere Verfahren trainiert und mit den bereits trainierten Modellen verglichen werden. Darüber hinaus lässt sich die Verkaufsprognose nun auch auf andere Artikel und Filialen sowie zusätzliche Verkaufsstrategien erweitern. Die interessanteste Erweiterung wäre die direkte Anbindung an das Kassensystem, sodass die Prognose mit dem Sliding Window Ansatz jeweils tagesaktuell mit den entsprechenden Verkaufsdaten durchgeführt werden kann. Für Prognosen in der Praxis ist die Komplexität des eingesetzten Verfahrens, Modells und Benutzerschnittstelle abzuwägen. So kann ein leicht erklärbares Verfahren die Akzeptanz und das Vertrauen in das System fördern. Eine Reduzierung von Features vereinfacht zudem die Verwaltung von Datenabhängigkeiten. Die Benutzerschnittstelle soll einfach gehalten werden, z.B. jeweils eine Spalte mit dem historischen Durchschnitt der Verkaufszahlen des Produkts pro Wochentag und der angepassten KI-Prognose.

Literaturverzeichnis

- [Ba15] Bajari, P.; Nekipelov, D.; Ryan, S. P.; Yang, M.: Machine learning methods for demand estimation. *American Economic Review*, 105 (5), 481-85, 2015.

-
- [BS11] Bozkir, A. S.; Sezer, E. A.: Predicting food demand in food courts by decision tree approaches. *Procedia Computer Science* Vol. 3, pp. 759-763, 2011.
- [DWD21] DWD, Deutscher Wetterdienst, <https://opendata.dwd.de/>, zugegriffen: 17/05/2021.
- [HHS17] Hota, H. S.; Handa, R.; Shrivastava, A. K.: Time Series Data Prediction Using Sliding Window Based RBF Neural Network. In: *International Journal of Computational Intelligence Research* Vol. 13 No. 5, pp. 1145-1156, 2017.
- [LI17] Liu, X. und Ichise, R.: Food Sales Prediction with Meteorological Data – A Case Study of a Japanese Chain Supermarket. In: *Data Mining and Big Data: Second International Conference. Lecture Notes in Computer Science* Vol. 10387, Springer International Publishing AG, Fukuoka Japan, pp. 93-104, 2017.
- [Ne10] Nesreen, K. A. et al.: An Empirical Comparison of Machine Learning Models for Time Series Forecasting, *Econometric Reviews* 29:5-6, pp. 594-621, 2010.
- [SV19] Schuster, T. und Volz, R.: Predicting Parking Demand with Open Data. In: *Digital Transformation for a Sustainable Society in the 21st Century. I3E 2019. Lecture Notes in Computer Science* Vol. 11701, Springer, Cham, pp. 218-229, 2019.
- [Ta19] Tarallo, E. et al.: Machine Learning in Predicting Demand for Fast-Moving Consumer Goods: An Exploratory Research. In: *IFAC-PapersOnLine* Vol. 52 Issue 13, pp. 737-742, 2019.

Chameleon: A Semi-AutoML framework targeting quick and scalable development and deployment of production-ready ML systems for SMEs

Johannes Otterbach,¹ Thomas Wollmann²

Abstract: Developing, scaling, and deploying modern Machine Learning solutions remains challenging for small- and middle-sized enterprises (SMEs). This is due to a high entry barrier of building and maintaining a dedicated IT team as well as the difficulties of real-world data (RWD) compared to standard benchmark data. To address this challenge, we discuss the implementation and concepts of Chameleon, a semi-AutoML framework. The goal of Chameleon is fast and scalable development and deployment of production-ready machine learning systems into the workflow of SMEs. We first discuss the RWD challenges faced by SMEs. After, we outline the central part of the framework which is a model and loss-function zoo with RWD-relevant defaults. Subsequently, we present how one can use a templatable framework in order to automate the experiment iteration cycle, as well as close the gap between development and deployment. Finally, we touch on our testing framework component allowing us to investigate common model failure modes and support best practices of model deployment governance.

Keywords: Semi-AutoML, production systems, IT infrastructure, ML development and deployment, computer vision

1 Introduction

Modern Machine Learning and Artificial Intelligence systems hold the promise to disrupt many workflows of small- and medium-sized enterprises (SMEs) and propel them forward in the digitization efforts. This will result in more efficient use of resources as well as more targeted fulfillment of customer demands. Despite the promise of the technology, the development and deployment of state-of-the-art Machine Learning systems remains challenging for SMEs: If Machine Learning projects are being tackled at all, they are often terminally stuck in the proof-of-concept (PoC) stage as scaling and productionizing them requires a team of dedicated software- and data-engineers as well as data scientists. As a consequence many SMEs are cut-off from the technological progress, reserving their use to big companies and thus inhibiting their competitiveness [**oecd-report**]. In addition to this, SMEs adopting external AI systems incur risks due to technical debt, new attack surfaces and reliability of the technology regarding model generalization and robustness, adversarial

¹ Merantix Labs GmbH, Max-Urich-Strasse 3, 13355 Berlin, Germany johannes.otterbach@merantix.com

² Merantix Labs GmbH, Max-Urich-Strasse 3, 13355 Berlin, Germany thomas.wollmann@merantix.com

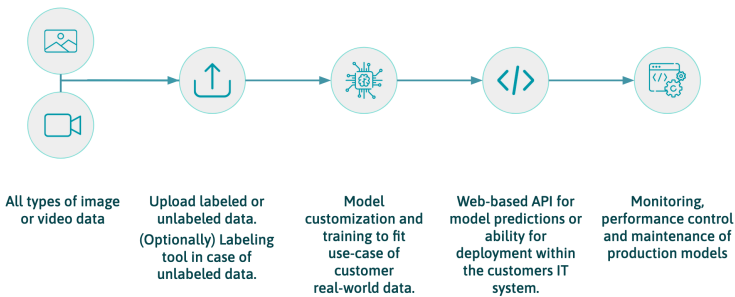


Fig. 1: Functional diagram of the Chameleon semi-AutoML framework. The framework is generally applicable to any data domain and provides optional labeling tools. It then ingests and prepares the data for model training with minimal effort. The resulting model is ready to be deployed and monitored without further tweaks to the environment.

attacks, dataset bias, and other complications in connection with Machine Learning systems [[ai-safety-concrete](#); [technicalDebt-ml](#); [reliable-ml](#)].

In this paper we present Chameleon, a semi-AutoML [**AutoML**] software framework developed at Merantix Labs, to address these challenges. The framework allows to tackle the diversity in workflow and use-cases of SMEs in a highly scalable manner by providing tools to efficiently orchestrate a big training infrastructure, abstracting common use-cases, deploying monitoring tools, and finally scale and virtualize the production deployment for inference. In addition to this the framework provides an optional labeling tool and incorporates standard testing procedures to assess the vulnerability of the Machine Learning system to common problems outlined above (see Fig. 1).

As a result of this, the cost of developing and deploying a robust Machine Learning model into production is significantly reduced and the burden of hiring and maintaining a dedicated Machine Learning-engineering team can be outsourced from the SMEs to specialized companies.

While the Chameleon framework is very flexible and can be applied to any data-modality, we focused on the development of computer vision use-cases such as image classification, object detection, image and instance segmentation.

The report is structured as follows: First we will discuss the different challenges of real-world data (RWD) that SMEs face and show how they differ from standard benchmarks available to academic research. We then dive into the overall architecture of Chameleon that allows us to quickly iterate on customer projects and finally show how the testing framework is integrated into the platform.

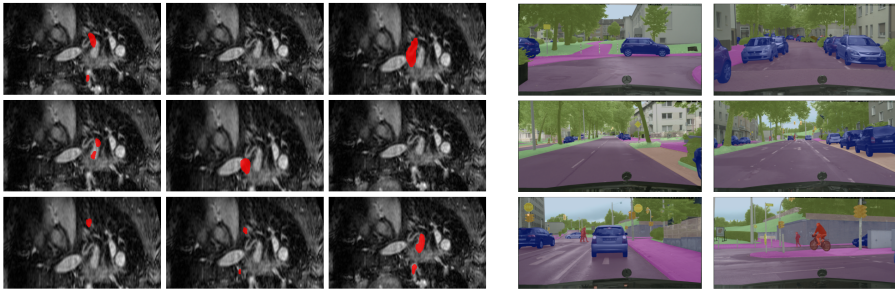


Fig. 2: (left) Prostate segmentation examples taken from [medical-decathlon], (right) Cityscapes segmentation samples taken from [cityscapes]. While both tasks are segmentation tasks from a technical perspective, the label density is significantly higher and more balanced in the Cityscapes dataset, while the label density is low in the prostate segmentation task, but the data is much more regular due to similar poses. Evaluating on the standard benchmark of Cityscapes hence does not necessarily make predictions of algorithmic performance on different data such as medical segmentation.

2 Data Challenges for SMEs in Computer Vision Use-Cases

Computer vision is a central and maybe the oldest field of study in the modern Deep Learning era [Imagenet2015]. It fielded the recent breakthrough of Artificial Intelligence techniques with the Imagenet challenge in 2012 [Imagenet2015; AlexNet2012] and heralded a decade of innovations ranging from natural language processing, computer vision, and video processing to speech recognition and optimization using techniques from (un-)supervised and generative modeling to reinforcement learning.

However, the development of these techniques heavily relies on the availability of open-source datasets that serve as benchmark metrics against which to test new algorithms. Within the computer vision domain, such datasets are, Imagenet [Imagenet2015], PascalVOC [Everingham10], Flickr 100M [yfcc100m], CoCo [msft-coco], Cityscapes [cityscapes], etc. As a general rule of thumb public datasets are assembled from common data modalities on the internet. As such they do not represent distributions from highly specialized domains such as medical image analysis [grand-challenge], remote sensing or visual quality control. Moreover, pre-training on these public datasets does not necessarily transfer to these specialized domains [liu2017detecting] diminishing the utility of the public datasets. Consequently, it is important to understand the differences between these benchmark datasets and the data found at SMEs.

For example X-ray, Magnetic Resonance Imaging (MRI) or visual quality control (VQC) data is typically more structured as they are acquired in a controlled environment and tend to have the same position of body or tool parts in the image as opposed to natural images. In addition the data is not always obtained using natural image techniques but are false-color images generated through post-processing of frequency and intensity data of specialized equipment such as Light Detection and Ranging (LIDAR), MRI or fluorescence microscopy

[[wollmann201770](#)]. These dataset biases often allow for more efficient modeling approaches [[Maier_2019](#)] and require different data-preprocessing techniques compared to the ones developed in the general Machine Learning community.

A second example within the VQC domain is the detection of small defects. Typically this use-case is formulated as an image segmentation task, where we do not only require the prediction if an image contains a defect, but in addition, also where the defect is located within the image. Since by nature these defects are small we face a strong class imbalance problem where most of the segmentation is the no-defect ground truth, while we only have a few defect pixels, which often are also broken up into several classes. Similar types of problems also exist in the medical image domain. This stands in contrast to public image segmentation benchmark datasets such as Cityscapes or CoCo, which are developed for natural scene understanding and display significantly reduced class imbalance (see Fig. 2). In addition the strong imbalance and small defect regions lead to pronounced instabilities in standard segmentation losses [[reinke2021common](#)].

As a result of these differences, architectures, loss functions, training curricula and other techniques reaching state-of-the-art (SOTA) performance on standard benchmarks do not always transfer to RWD use-cases. To quickly assess the transferability of new developments to datasets of SMEs, Chameleon incorporates a model zoo of different computer vision architectures and loss functions. As outlined below, those components are templatable, allowing for fast iteration and experimentation to reduce time to deployment of performant models. Moreover, these components come with relevant defaults that are reflective of the dataset difference between the benchmark and RWD domains.

3 Chameleon framework

The IT industry has seen an increasing number of frameworks for developing Machine Learning models, such as PyTorch [[pytorch](#)], Tensorflow [[tensorflow](#)], Keras [[keras](#)], JAX [[jax](#)] for Python, Owl [[owl](#)] in OCaml, Flux [[flux](#)] for Julia, flashlight [[flashlight](#)] for C++. However, the field is still lacking tools for the end-to-end development of Deep Learning pipelines, similar in spirit to packages such as scikit-learn [[scikit-learn](#)] or Spark [[spark](#)], and many SOTA applications are only available in form of Jupyter Notebooks [[jupyter](#)] in more-or-less maintained open-source projects. One of the reasons for the lack of end-to-end solutions is the significantly larger infrastructure overhead of training Deep Learning models, as they often require accelerators, e.g. GPUs, cloud-storage for large datasets in the range of several GB up to PB (at the point of writing this report), monitoring tools as well as inference deployment and integration. The result is a large gap in the ability to deploy a working proof-of-concept model into production [[scully2015](#)].

We address this shortcoming by building a semi-AutoML pipeline that can be orchestrated and deployed with minimal efforts. The core of the pipeline is formed by Kubernetes to deploy containers into cloud-based solution such as the Google cloud, AWS or Azure. The

containers are templated through a setup and are built during the Continuous Integration and Deployment (CI/CD) step. Once this setup is in place the deployment of Machine Learning systems becomes much easier as we can break down the steps of a development cycle into individual container deployments and the CI step ensures stability and reproducibility of the development [graetz].

We chose to separate the Extract-Transform-Load (ETL) stages from the training and inference stage of the pipeline. The ETL stage comprises the data-preprocessing steps and allows quick and repeated experimentation with different data-augmentation and data-enhancement approaches, necessary to address the RWD challenges outlined in the previous section. In addition it allows to build reusable components for ETL that can in many cases be reused end-to-end for new customer use cases. To facilitate fast iteration of model developments, we equipped the core Machine Learning module with several additional abstractions: Loss functions and model types as well as training loops and inference code-paths. In this way the training and inference can be orchestrated through a Terraform template. In the ideal scenario, this results in minimal code that needs to be developed to integrate a new data source into the pipeline and moves the development of high-performance Machine Learning systems closer to low-code environments. In addition to these benefits, the templating lets us auto-generate good pipelines based on well-chosen and field-tested heuristics. In order to address more dynamic data environments, where data can become stale over time, we can integrate the models into a CI/CD cycle to regularly roll out updates to the models without performance regressions.

4 Testing Framework

Responsible deployments of Machine Learning pipelines into production require following a reliable and responsible model governance process [model-cards; ai-safety]. A central part of this process is the investigation of generalization, biases and robustness to adversarial attacks. This is often a manual process since it requires experimentation with the trained model and subjecting it to a variety of different scenarios. Using Chameleon, we automate many pieces of this process in order to reduce the manual labour and speed up the investigation process. For example we implement data-augmentation to simulate different image conditions, such as fog or blur to simulate common perturbations and corruptions [hendrycks2019robustness]. We also support standard white-box adversarial attacks, such as Fast-Gradient-Sign-Method (FGSM) [goodfellow2014explaining], iterative-FGSM [kurakin2017adversarial] and others, on segmentation data [xie2017adversarial; fischer2017adversarial]. Finally we evaluate the model performance under these perturbations and visualize the results (see Fig. 3). This process speeds up the analyst in assessing the model's capabilities and improves modeling as it surfaces model failure modes early, allowing to develop various mitigation strategies throughout the full pipeline such as model ensembling for segmentation mask predictions. In addition, the inference path in the Chameleon framework along with the ETL pipeline lets us easily investigate the

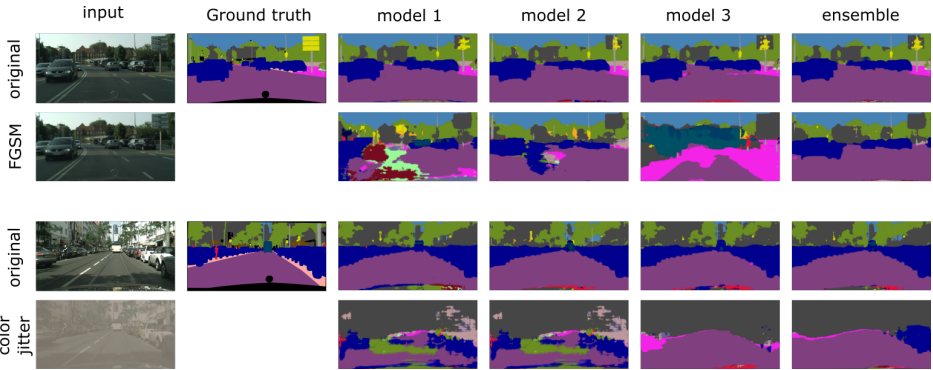


Fig. 3: Gradient-based and color jitter adversarial attack against image segmentation models. Each individual model is strongly fooled by the adversary as can be seen by the predicted label map of the outputs. However, using an ensemble of the models results in a significantly more robust prediction making it harder for the attacker to fool the system. The testing framework of Chameleon facilitates such investigations at a much faster rate.

generalization of the model to other datasets within a relevant domain or even outside the domain. Doing this highlights complementary failure modes and informs the SMEs about limitations of the model so they can be monitored in production.

5 Summary

SMEs face a large gap between developing PoCs of modern Machine Learning systems and their deployment into production cycles. Not only is developing SOTA models expensive in terms of compute as well as human resource, but SMEs also face a challenge in modeling due to large differences between their own data and standard benchmarks of the Machine Learning community. To close this gap, we developed the Chameleon framework which allows us to quickly integrate and iterate on modeling real-world data using a zoo of models and training procedures. Due to our templatable and containerized semi-AutoML framework, we can quickly deploy models into a production workflow, reducing the integration burden of customers and saving them from hiring and maintaining a dedicated Machine Learning team. We incorporate an automated testing framework to assess common Machine Learning model failure modes, such as adversarial attacks and image defects, thus ensuring best practices for model deployment governance and increasing trust in the capabilities of automated decision engines.

6 Acknowledgements

The research leading to these results is funded by EU ECSEL Project *SECREDAS Cyber Security for Cross Domain Reliable Dependable Automated Systems* (Grant Number: 783119).

Energieeffizientes Kaltstartverhalten spanender Werkzeugmaschinen

Kompensation thermischer Einflüsse mit Methoden des Maschinellen Lernens

Deborah Walz¹, Andreas Wächter², Stefan Tomov¹, Konrad Heimbach³ und Matthias Weigold²

Abstract: Die Kompensation thermischer Einflüsse und daraus resultierender geometrischer Verlagerungen spielt eine bedeutende Rolle bei der Gewährleistung einer hohen Bearbeitungsqualität von Werkstücken in Zerspanungsprozessen. Übliche Vorgehensweisen zur Reduktion thermischer Verlagerungen während der Produktion gehen mit einem erheblichen Energiebedarf einher oder modellieren die komplexen Zusammenhänge thermischer Einflüsse nur ungenügend. Methoden des Maschinellen Lernens stellen einen vielversprechenden Ansatz zur Modellierung dar. Es wird eine Lösung angestrebt, die aufwandsarm auf Produktionsmaschinen ähnlicher Bauart übertragen werden kann. Derzeit ist ungeklärt, ob eine explizite oder implizite Modellierung der zeitlich multivarianten Daten eine zufriedenstellende Lösung bietet. Als besonders herausfordernd stellt sich die Verfügbarkeit von ausreichend vielen Datenbeispielen zur Modellierung der relevanten Größen dar.

Keywords: Machine Learning, Neural Network, Energy Efficiency, Compensation, Machining

1 Einleitung

Im Rahmen des laufenden Forschungsprojekts ETA im Bestand werden anwenderfreundliche Technologien und Methoden zur Steigerung der Energieeffizienz im Fabrikbestand entwickelt. Im Betrachtungsrahmen liegt die metallverarbeitende Industrie, die eine prägende Rolle in der deutschen Industrielandschaft einnimmt [Em13]. Für unterschiedliche Wirtschaftszweige der metallverarbeitenden Industrie strebt das Projekt Werkzeuge zur Realisierung von Einsparpotenzialen von 25–40 % durch die Verknüpfung von Einzelprozessen mit der Produktionsumgebung an. Auf Prozessebene spielen energieeffizienz-steigernde Lösungen von Maschinenherstellern, beispielsweise von Werkzeugmaschinen wie Dreh- und Fräsmaschinen, eine relevante Rolle. In diesem Zusammenhang werden Energieeffizienzpotenziale im Pausen- und Kaltstartverhalten von Werkzeugmaschinen in der Zerspanung adressiert.

¹ anacision GmbH, Albert-Nestler-Straße 19, 76131 Karlsruhe, deborah.walz@anacision.de

² Technische Universität Darmstadt, Institut für Produktionsmanagement, Technologie und Werkzeugmaschinen, Otto-Berndt-Straße 2, 64287 Darmstadt, a.waechter@ptw.tu-darmstadt.de

³ EMAG GmbH & Co. KG, Austraße 24, 73084 Salach, kheimbach@emag.com

Die sukzessive Abschaltung von Nebenaggregaten wie Kühlschmierstoffpumpen in Pausenzeiten stellt eine wirkungsvolle Maßnahme zur Energieeinsparung dar [AKS10, Bre12]. Allerdings wird die Temperaturverteilung in der Maschine unter anderem durch den Kühlschmierstoff beeinflusst [WB06]. Der Stillstand der Kühlschmierstoffzirkulation kann lokale Temperaturänderungen und in der Folge thermisch bedingte Verformungen in der Maschine hervorrufen. Resultierend können geforderte Bearbeitungsqualitäten gegebenenfalls nicht mehr erreicht werden. Um eine hohe Bearbeitungsgenauigkeit zu erzielen, werden beim Einschalten und in Pausenzeiten üblicherweise Warmfahr- und Warmhalteprogramme an Werkzeugmaschinen gefahren, in denen die Maschinenachsen bewegt werden. Obwohl keine Bearbeitung stattfindet, gehen diese Programme mit erheblichem Energiebedarf der Maschine einher. Beispielhaft können durch den Verzicht von täglich 180 Minuten Warmhaltedauer bei einer Leistungsaufnahme von 4 kW pro Tag und Maschine 12 kWh eingespart werden.

Das Ziel ist Methoden zu entwickeln, die den Verzicht auf Warmfahr- und Warmhalteprogramme ermöglichen oder diese verkürzen, um so die aus diesen Programmen resultierenden Verbräuche zu reduzieren. Hierbei wird der Ansatz verfolgt, auftretende Verlagerungen maschinenseitig während der Bearbeitung zu kompensieren. Durch Bildung eines im Betrieb ausführbaren Modells sollen die Zusammenhänge der verschiedenen Parameter zu der Maschinenverformung beschrieben und zu beliebigen Zeitpunkten für die Kompensation nutzbar gemacht werden. Hierzu werden neuartige Methoden des Maschinellen Lernens hinsichtlich ihrer Eignung analysiert. Im weiteren Vorhaben wird untersucht, ob die Zusammenhänge thermischer Einflüsse an spanenden Werkzeugmaschinen mit hinreichender Güte zur Kompensation abgebildet werden können und mit welchen Datenmengen und welchem Aufwand sich diese Modelle auf ähnlich konstruierte Maschinen mit ähnlichen Zielgrößen übertragen lassen.

2 Konstruktive und steuerungsseitige Maßnahmen zur thermischen Kompensation

Bei der Herstellung präziser Geometrien mittels Zerspanung spielt das thermische Verhalten der genutzten Werkzeugmaschine eine wesentliche Rolle. Darum findet die Stabilisierung des thermischen Verhaltens von Werkzeugmaschinen bereits in der Konstruktion besondere Beachtung.

Um die thermisch bedingte Verlagerung an den für die Zerspanung relevanten Punkten gering zu halten, werden die Komponenten nach ihrem Wärmedehnungsverhalten symmetrisch angeordnet [DHS07]. Weiter werden in der Konstruktionsphase die Wärmeübergangs- und Wärmeausdehnungskoeffizienten der eingesetzten Materialien beachtet, um den Wärmegang und die einhergehende Verlagerung abzuschwächen, zu verzögern oder gezielt zu nutzen [DHS07]. In dieser Phase der Entwicklung werden auch modellbildende Verfahren wie die Finite-Elemente-Methode eingesetzt, um Bauteile frühzeitig konstruktiv in Bezug auf das thermische Verhalten optimieren zu

können [Sc11]. Weiterhin werden Maschinenkomponenten wie Motoren gekühlt [DHS07]. Eine verbreitete Regelstrategie ist es, die Temperatur des Kühlmediums der Umgebungstemperatur nachzuführen und so ausgeprägte Kältenester und eine inhomogene Temperaturverteilung zu vermeiden. Zuletzt wird der verbleibenden thermischen Verlagerung steuerungsseitig entgegengewirkt. Hierzu wird meist auf Basis eines Temperatursensors eine Anpassung des Maschinenkoordinatensystems einer Achse vorgenommen. Der Zusammenhang zwischen Temperatur und Verlagerung wird experimentell ermittelt [WB06] und mathematisch linear hinterlegt. Dieses Vorgehen führt zu guten Ergebnissen, wenn die gemessene Stelle einen dominanten Einfluss auf die thermische Verlagerung hat und die Bearbeitungsparameter konstant sind. Bei nicht gleichbleibender Maschinenbelastung oder unterschiedlichen Werkstück- und Bearbeitungsparametern liefert diese Art der Kompensation kein hinreichendes Ergebnis.

3 Kompensation thermischer Einflüsse mittels Maschinellen Lernen

Das übergeordnete Ziel ist die Entwicklung eines Modells, das mithilfe von gemessenen Sensordaten die Auswirkungen des thermischen Verhaltens der Maschine abbildet.

3.1 Voruntersuchungen des komplexen thermischen Verhaltens

Es wird der Ansatz verfolgt, die auftretende Verlagerung maschinenseitig während der Bearbeitung zu kompensieren. Das Ausmaß und die Richtung der Verlagerung und somit der notwendigen Korrekturwerten hängen von vielfältigen Parametern und deren zeitlichen Verlauf ab [WB06]. Solche Parameter sind beispielhaft die Drehzahl der Maschinenspindel, Temperaturunterschiede zwischen Maschinenkomponenten, deren zeitlicher Verlauf, sowie der eigentliche Bearbeitungsprozess für ein spezifisches Werkstück. Zur Erfassung dieser Größen ist verschiedenartige Sensorik nötig. Die experimentelle Untersuchung der Zusammenhänge an vier betrachteten Maschinen zeigt, dass sich die Einflussgrößen heterogen und zeitlich variant auf das Verformungsverhalten auswirken.

Abb. 1 zeigt dies in einer Messreihe über einen Zeitraum von 66 h ausgehend von einer kalten Maschine mit nachfolgend konstanter Belastung. Beim thermischen Verhalten sind tageszeitabhängige Schwankungen der Umgebungstemperatur sowie eine Aufheizphase der Motorspindel und der Werkzeugaufnahme erkennbar, wobei zudem alle Temperaturwerte mit der Umgebungstemperatur schwanken. Die infolge des Temperaturgangs auftretenden Verlagerungen unterscheiden sich hinsichtlich der Betrachtungspunkte. An der Werkstückspindel ist beim Einschalten eine schnelle Verlagerung in positiver x-Richtung erkennbar wobei die Position im weiteren Verlauf annähernd konstant bleibt. An der Werkzeugträgerkonsole findet beim Einschalten eine deutlich langsamere Verlagerung in negativer x-Richtung statt, die im weiteren Verlauf

mit der Umgebungstemperatur schwankt, wobei zudem ein schleichender Abwärtstrend erkennbar ist. Ausschlaggebend für die Qualität des Werkstücks ist die Überlagerung dieser beiden Einzelverlagerungen. Die Verlagerung zwischen Werkzeug und Werkstück wird anfangs durch die Spindelverlagerung und im weiteren Verlauf durch die Verlagerung des Werkzeugträgers dominiert. Ein direkter Zusammenhang einer thermischen Messgröße zur Verlagerung ist nicht erkennbar.

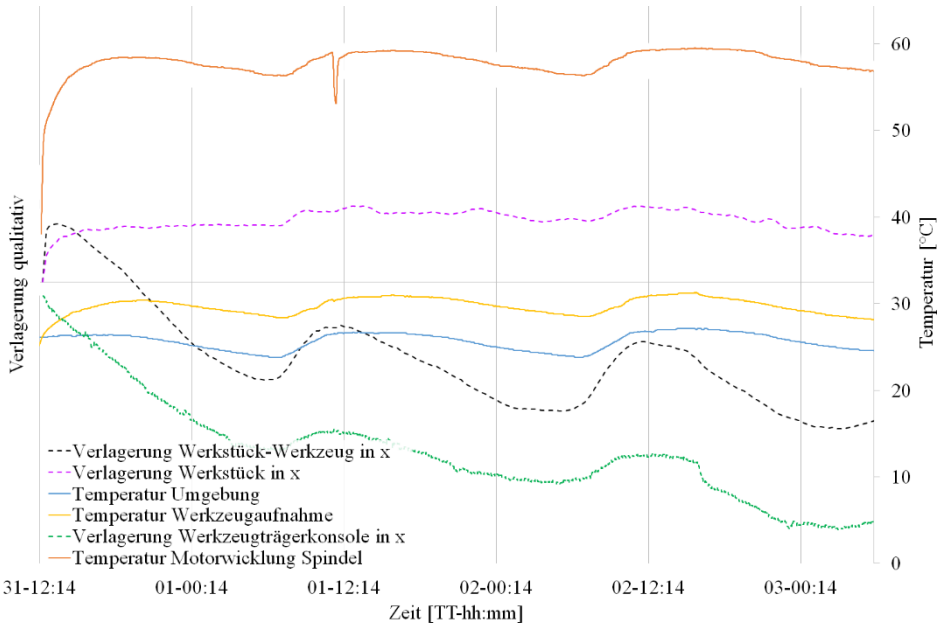


Abb. 1 Verlagerung mit Einschaltverhalten und konstanter Belastung über 66 h

3.2 Diskussion vorliegender Rahmenbedingungen und Daten

Als Grundlage für die Modellierung werden N Sensordaten, welche das thermische Verhalten der Maschine charakterisieren, zu verschiedenen Zeitpunkten $t \in \{1, \dots, T\}$ gemessen. Hierbei entspricht $t = 1$ dem ausgeschalteten Zustand und $t = T$ dem warmgefahrenen Zustand der Maschine. Die restlichen Zeitpunkte $t \in \{2, \dots, T - 1\}$ bilden den Übergang zwischen diesen beiden Zuständen der Maschine ab. Da das thermische Verhalten über die verschiedenen Zeitpunkte hinweg durch mehrere Sensordaten abgebildet wird, bietet es sich an, die N Sensordaten zu jedem Zeitpunkt als N -dimensionale Vektoren zu beschreiben. Hierdurch lassen sich die vorliegenden Sensordaten zu jedem Zeitpunkt als multivariate Zeitreihen auffassen:

$$X_t = (X_{1,t}, \dots, X_{N,t}) \in \mathbb{R}^N, \quad t \in \{1, \dots, T\}. \quad (1)$$

Weiterhin können Sensordaten kurze Messlücken sowie unterschiedliche Abtastraten beinhalten. Durch eine Imputation werden diese Messlücken durch plausible Werte ersetzt. Eine Angleichung der Abtastrate und Zeitstempel der einzelnen Werte wird durch Methoden der Harmonisierung wie Up-/Downsampling erreicht.

3.3 Modellierung mit künstlichen neuronalen Netzen

Klassische Modelle sind nicht ausreichend, um die Korrekturwerte bei der Bearbeitung spezifischer Werkstücke unter Einbeziehung der hohen Parametervielfalt und Systemkomplexität mit geringer Latenz herzuleiten. Insbesondere die zeitliche Änderung multivariater Einflussparameter sowie der kontinuierliche Verlauf der Kompensation in bestimmter Richtung müssen berücksichtigt werden. Um einen Ansatz zu entwickeln, welcher unter diesen Anforderungen eine hohe Güte erreicht, werden modell-basierte Verfahren betrachtet, deren Zielgröße die Verlagerung der Maschine am Werkstück in x -, y - und z -Richtung bildet. Somit lassen sich aus diesem Modell direkt Korrekturparameter ableiten und an die Steuerung übergeben.

Die verfügbare Datenmenge, der Grad des a-priori nötigen Fachwissens, die Komplexität des Modells (einstellbar durch Modellparameter) sowie das Modellierungsverfahren an sich sind Stellschrauben, welche Einfluss auf die Modellgüte haben und deren Zusammenspiel im Laufe des Projekts untersucht wird. Beispiele für mögliche Modellierungsverfahren multivariater Zeitreihen sind rekurrente neuronale Netze oder autoregressive Ansätze. Da eine modellbasierte Übertragbarkeit auf ähnliche Maschinen mit geringem Anpassungsaufwand angestrebt wird, wird zudem durch Ansätze des Transfer Learnings evaluiert, inwieweit ein trainiertes Modell für ähnlich konstruierte Maschinen – womöglich mit einer deutlich reduzierten Datenmenge - nachgelernt werden kann [Be11, SQ10]. Ist die verfügbare Datenmenge nicht groß genug, um eine hohe Güte zu erreichen, werden weniger datenintensive Modellierungsansätze oder Verfahren erprobt, sowie weiteres Fachwissen über thermische Zusammenhänge integriert.

Neuronale Netze bieten eine besonders hohe Flexibilität in der Konfiguration der Modellkomplexität, der Modellierung mehrere Zielvariablen sowie einer Übertragbarkeit auf neue Maschinen. Die Übertragbarkeit wird durch Festhalten der Modellparameter in frühen Schichten eines neuronalen Netzes bei gleichzeitigem Nachlernen der höheren Schichten erreicht. Aus diesen Gründen werden in einem ersten Schritt neuronale Netze untersucht, wobei anzumerken ist, dass diese für die Abbildung stark nicht-linearen Verhaltens viele Datenpunkte benötigen. Andererseits können zeitliche Zusammenhänge durch manuelles Modellieren mit Fachexperten in die Modelle integriert werden. So kann beispielsweise bekanntes Verhalten bei Temperaturgängen an spezifischen Messpunkten dem Modell in Form von Aggregationen zur Verfügung gestellt werden. Ein Nachteil hierbei ist der hohe manuelle Aufwand, welcher sowohl für die initiale Modellbildung als auch für die Übertragbarkeit auf ähnliche Maschinen zu beachten ist.

4 Zusammenfassung und Ausblick

Um das thermische Verhalten einer Werkzeugmaschine stabil zu halten, werden Warmfahr- und Warmhalteprogramme eingesetzt, welche mit einem erheblichen Energiebedarf der Maschine einhergehen. Zudem wird mit klassischen Methoden aus dem Maschinenbau versucht, die thermisch bedingte Verlagerung durch Korrekturwerte während der Bearbeitung zu kompensieren. Diese Korrekturwerte hängen jedoch von einer Vielzahl an Parametern ab, welche zudem über die Zeit variieren, sodass als Konsequenz die Verlagerungen ohne energetisch aufwändige Warmfahr- und Warmhalteprogramme nicht hinreichend gut kompensiert werden können.

Unter Verwendung von künstlichen neuronalen Netzen soll ein Modell entwickelt werden, das die Auswirkungen des thermischen Verhaltens der Maschine abbildet. Dieses Modell kann für eine Kompensation der Verlagerung der Maschine zu einem beliebigen Zeitpunkt genutzt werden, wodurch sich der Energiebedarf der Maschine deutlich verringert, indem auf Warmlauf- und Warmhalteprogramme verzichtet werden kann oder diese verkürzt werden können. Die Herausforderungen bestehen darin, mit den verfügbaren Daten hinreichend genaue Modelle zu erstellen, die zudem über eine hohe Generalisierbarkeit verfügen, sodass die Modelle auf ähnlich konstruierte Maschinen mit geringem Aufwand übertragbar sind.

Die modellbasierte Anwendung fokussiert eine einzelne Maschine. Um das Modell während des Betriebs einer Maschine zu nutzen, werden die gemessenen Sensordaten einem Industrie-PC zur Verfügung gestellt. Auf diesem sollen mit dem hinterlegten Modell die Korrekturwerte zur Kompensation der thermischen Verlagerung berechnet und anschließend direkt in die Steuerung der Maschine integriert werden. Eine Integration weiterer Systeme, wie ERP oder andere Maschinen, ist nicht nötig. Hieraus kann sich ein günstiges Nutzen-Kosten-Verhältnis auch für Unternehmen mit niedrigen bis mittleren Reifegraden der Digitalisierung und Vernetzung ergeben.

5 Danksagung

Die Autoren bedanken sich für die finanzielle Unterstützung des Projekts ETA im Bestand (Förderkennzeichen: 03EN2048A-I) durch das Bundesministerium für Wirtschaft und Energie (BMWi) und die Projektbetreuung durch den Projektträger Jülich (PtJ).

6 Literaturverzeichnis

[AKS10] Abele, E.; Kuhrke, B.; Rothenbücher S.: Maximierung der Energieeffizienz von Werkzeugmaschinen. Zerspanungstechnik 09/10, S. 26-29, 2010.

- [Be11] Bengio, Y. et al.: Deep Learners Benefit More from Out-of-Distribution Examples. In (JMLR W&CP): Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, Fort Lauderdale, S. 164-172, 2011.
- [Bre12] Brecher, C. et al.: Effizienzsteigerung von Werkzeugmaschinen durch Optimierung der Technologien zum Komponentenbetrieb – EWOTeK, Apprimus, Aachen, 2012.
- [DHS07] Donmez, M.A.; Hahn, M.H.; Soons, S.A.: A Novel Cooling System to Reduce Thermally-Induced Errors of Machine Tools. In: CIRP Annals 56/1, S. 521-524, 2007.
- [Em13] Emec, S. et al.: Analyse von Potenzialen der Material- und Energieeffizienz in ausgewählten Branchen der Metall verarbeitenden Industrie, 2013.
- [GFS07] Graves A.; Fernández S.; Schmidhuber J.: Multi-dimensional Recurrent Neural Networks. In (de Sá J.M., Alexandre L.A., Duch W., Mandic D.): Artificial Neural Networks – ICANN 2007. Lecture Notes in Computer Science, Vol 4668. Springer, Berlin u.a., 2007
- [HS97] Hochreiter S.; Schmidhuber J.: Long Short-Term Memory. Neural Computation 9/97, S. 1735-1780, 1997.
- [Sc11] Scharschmidt, K.-H.: Modellbasiertes Verfahren zur Kompensation thermoelastischer Verlagerungen von Werkzeugmaschinen, PZH Produktionstechnisches Zentrum, Hannover, 2011.
- [SQ10] Pan, S. J.; Yang, Q.: A Survey on Transfer Learning. IEEE Transactions on Knowledge and Data Engineering, 22/10, S. 1345-1359, 2010.
- [WB06] Weck, M.; Brecher, C.: Werkzeugmaschinen 5. Messtechnische Untersuchung und Beurteilung, dynamische Stabilität, Springer, Berlin u.a., 2006.

From Physical to Virtual: Leveraging Drone Imagery to Automate Photovoltaic System Maintenance

Maximilian Lowin¹, Domenic Kellner¹, Tobias Kohl¹, Cristina Mihale-Wilson¹

Abstract: Optimizing the maintenance of large-scale infrastructure can be a significant cost driver for small and medium-sized enterprises (SMEs). This paper presents a feasible approach to combine data from real-world physical structures collected through an automated maintenance process with cloud-based AI services to generate a meaningful virtual representation of such structures. We use photovoltaic systems as an exemplary physical structure and thermal imaging, collected through scheduled drone monitoring. With help of these unstructured data sources, we demonstrate our approach's applicability. Our solution artifact provides a lightweight AI application that is adoptable for other problem spaces, enabling an easier knowledge transfer from research to SMEs. By combining Cloud Computing with Machine Learning, the artifact identifies present and emerging damages of physical objects. It provides a virtual representation of the object's state and empowers a meaningful visualization.

Keywords: Digital Twin; Machine Learning; Predictive Maintenance; Photovoltaic System; Internet of Things, Process Automation, Neural Networks, Visualization

1 Introduction

Ineffective maintenance is a major cost driver [Mo02]. Failures of expensive infrastructure especially influences small and medium-sized enterprises (SMEs). Hence, optimizing the timing of maintenance cycles by predicting upcoming failures and thus minimize cost as well as resources is particularly important for SMEs. One emerging Predictive Maintenance method is based on Digital Twins (DT). The DT method has therefore seen adoption in a wide variety of different industry sectors [Sh18]. DT has a data-centric view, which enables the system itself to review its own contextual information and make operational decisions based on those data-driven analytics. With the advent of Internet of Things (IoT) devices and interconnected structures such as in smart cities, Predictive Maintenance based on DT will become ubiquitous. Although DT has many advantages, one of the main reported challenges by businesses concerning the adoption of DT's in Predictive Maintenance processes is data availability [EBA20]. This is especially true for SMEs, since they cannot rely on large physical structures and surplus resources. Often businesses are legally required to provide regular cost intensive on-site inspections for their physical assets, so that Predictive Maintenance based on DT

¹ Goethe University Frankfurt, Theodor-W.-Adorno-Platz 4, D-60629 Frankfurt am Main, Germany, {lowin, kellner, tokohl, mihale-wilson}@wiwi.uni-frankfurt.de

would be more than useful. Against the background that large-scale physical structures are not trivial to realize as virtual entities, this paper aims to utilize object detection and the DT concept to create virtual representational models for real-world physical structures. We show how to achieve this based on the example of photovoltaic (PV) systems. Specifically, we present a prototype artifact that automatically utilizes thermal drone images and Artificial Intelligence (AI) to show a complete virtualized representation of the physical structures condition.

To evaluate our artifact directly in the application context of SMEs, we tested a pilot project on-site and evaluated its validity in collaboration with solar park operators and a medium-sized technical support company. We contribute to the automation of creating digital representations of physical systems to build more sustainable business models for SMEs. While our solution artifact provides a lightweight AI application focused on the digitalization of solar parks, SMEs can adapt this approach to other problem spaces and industries i.e. allowing them to implement the functionality needed in their business context. Thus, our approach improves and optimizes current processes by saving time and costs.

We structured the remainder of the paper as follows. The next section provides an overview of the existing state of knowledge on Predictive Maintenance and Digital Twins. Afterwards, we explain our produced artifact in more detail, covering data collection, object detection, and the DT visualization. Finally, after evaluating our artifact, we conclude with implications of our work and possible future research paths.

2 Related Work and Research Context

2.1 Design Science Research Methodology

We draw on the Design Science Research (DSR) Contribution Framework presented by [GH13]. The DSR Contribution Framework allows us to frame our work and prioritize specific steps required to achieve our research goal. Second, for implementation, we draw on the Design Science Research Methodology (DSRM) by [Pe07]. The DSRM provides helpful principles, practices, and procedures consistent with prior Information Systems literature, suitable to evaluate automation processes of SMEs. It starts with identifying the problem and defining the artifact's objectives. We incorporate the existing state of knowledge for the design and development as suggested by [GH13]. Altogether, our artifact encompasses three layers: a data collection layer, a processing layer, and a visualization layer. After presenting each of the three layers, we follow [Pe07] to evaluate and demonstrate the proposed artifact's value.

2.2 Digital Twin

The DT represents the formation of a two-way relationship between a physical structure and a virtual model. The DT concept has mostly seen wide adoption in manufacturing industries as it establishes a relationship between the physical object and sometimes multiple virtual models. However, to bring the concept of DTs into various other industries a wide adoption by SMEs will be crucial.

The DT supports the creation of virtual representations allowing recording of activities throughout the whole life-cycle [Sc17]. Originally, the DT started as a Computer-Aided Design (CAD) description-based model of a product or so-called "Thing" with descriptive information added to it. However, the concept evolved to a more encompassing data-centric view by adding actionable components [GV17]. These components enable one of the DTs' core capabilities to predict an unexpected event's physical response before it occurs [Sc17]. Whereas CAD description-based models were mostly static representations, the DT allows for dynamic interactions and simulations, resulting in actionable knowledge becoming available before real physical events occur [GV17]. DTs are not just singular models but rather a set of linked data artifacts and simulation models. By being an evolving "living" model for a physical asset, the DT offers an effective way of real-time interaction and integration in a so-called cyber-physical system (CPS) [Sh18]. This actionable knowledge allows us to monitor so-called emergent behavior. Especially for automated maintenance based on drone imagery, recognizing emergent behavior is a focal point why a DT is particularly relevant [GV17].

Therefore, in our artifact, we use the DT to include all types of information about the physical structure, each of the structures' assembled units, and their operational conditions. Thereby the DT provides access to the critical information for diagnostics and helps to initiate the maintenance processes. [GV17]

2.3 Automated Maintenance Processes

Maintenance is a significant cost driver in almost all industries with tangible assets [WY07]. As advanced analytics methods like Predictive Maintenance (PdM) became popular in the late 1980s due to ineffective maintenance methods [Mo02a] they are now omnipresent in many industries. PdM uses a system's information to determine the system's condition and predicts the optimal maintenance schedule [Al14]. As such, data has become of ubiquitous significance for PdM success. While this information can originate from various sources such as IoT devices and sensors by sensing internal information like temperature, vibration, voltage, or current [Ha10] it is almost nonexistent for larger physical structures as tangible assets. Maintenance of these assets also often legally requires businesses to provide regular on-site inspections. These inspections bind many resources and are a significant cost driver.

To implement automated PdM, rich sensor data is a prerequisite that enables anomaly and potential failure detection through a combination of Machine Learning (ML)

techniques and Cloud Computing. ML uses historical data to identify patterns in data and learn from them to formulate predictions and insights relevant to various stakeholders [Be20]. In practice, ML methods help to detect unusual behavior like anomalies by identifying data records that deviate from the non-anomalous distribution of the sensor data [CBK09]. To date, mostly industrial settings such as manufacturing or aviation dispose of the data necessary to optimize their PdM activities [Al14]. However, due to PdM's huge potential, transferring the concept to further applications can prove beneficial, too.

Besides pure forecasts for optimal maintenance windows, businesses are also interested in automating the maintenance process as much as possible. With current efforts focused on the automation of scheduling the maintenance process rather than automating the maintenance itself [Yü20], this work seeks to advance the PdM process automation by presenting an automated monitoring solution for PdM.

3 Artifact Description

We conduct this research in the context of PV systems. Implementing Predicted Maintenance and automated monitoring solutions on PV panels is beneficial since analysts can often predict specific failures at a certain time in advance [Be17]. Traditional maintenance approaches analyze each panel's output data, e.g., by incorporating the produced power or the present voltage at each panel [De18]. An essential downside of such traditional maintenance approaches is that it is very work-intensive, it requires modern measuring technology, and accurate information for each panel separately. Besides, it also requires expert knowledge on the technical set-up and the conversion of solar energy into electricity.

Newer maintenance approaches are more independent of the solar panel's technology and use infrared (IR) thermography to inspect solar panels. Based on their thermal characteristics [BS15], such approaches can identify faulty modules and panels. A single PV module consists of multiple cells that convert light into electricity. Faulty cells or panels heat up due to the inability to convert solar energy to power [Sp12]. Under infrared thermography, the heated spots appear like "hotspots". These so-called hotspots may occur due to cell failure, interconnection failure, partial shading, or mismatched cells [De12]. Undetected hotspots in a module may lead to a degradation of cell properties and damage the entire module. IR cameras can reveal modules and cells with anomalies [Sp12]. Because such anomaly detection does not require detailed information on the modules or advanced expert knowledge, thermography is a non-intrusive maintenance technique with high potential cost benefits.

Drones equipped with IR technology enhance the maintenance process due to their versatility. Regular on-site inspections are a legal necessity for solar park operators. Combining these inspections with drone monitoring would enable a faster and more detailed maintenance process. Drones can quickly inspect field plants and access roofs

without further safeguarding inspectors [BS15]. We use both technologies and enhance them with Machine Learning to automate the process and improve its applicability and efficiency.

3.1 Concept Architecture

To present our artifact in a meaningful way, we split the description into separate layers (see Fig. 1), including the data collection, processing, and visualization layer: The **data collection layer** encompasses the initial thermal images as an unstructured data source. The **processing layer** incorporates cloud computing and ML to detect anomalies like reduced operational capacities in unstructured data. Finally, the **visualization layer** transforms the unstructured data into the virtual entity of the DT. Thus, the layer enables visualization of the asset structure as well as the condition and status of each assembled unit to provide helpful decision support.

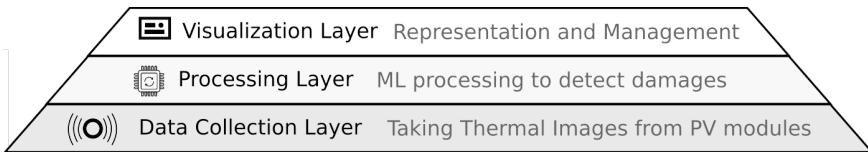


Fig. 1: Concept Architecture (Layer Structure)

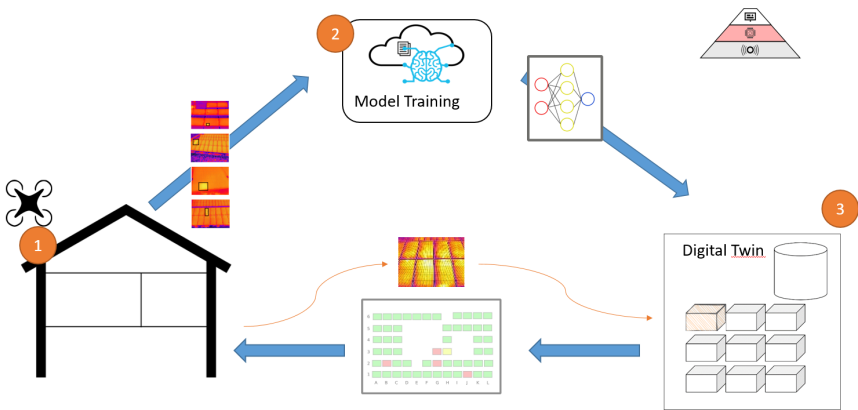


Fig. 2: Orchestration Architecture: AI-Modeling and Digital Twin Provisioning

For physical structures such as PV systems, the idea is to create a complete virtual model or DT of the respective structures. As with many areas of application, the most pressing issue is the availability of data. We gather the necessary data via an automated

maintenance process based on thermal drone imagery to address this issue. Using such an external unstructured data source, we are not bound to utilize only certain PV module types or technologies. However, processing the unstructured data into a virtual entity poses various challenges: Firstly, we must incorporate the overall system's physical constructional composition and map each units' diagnostic condition on the observed model. Secondly, we have to observe and determine damages or patterns on each assembled PV unit. In the end, the visualization layer needs to highlight the assets condition based on the prognosticated health of all assembled units. Based on the thermal images, the processing layer creates a detailed prediction for each unit's operational condition and allows for actionable insights (e.g., corrective maintenance processes), see Fig. 2.

3.2 Data Collection & Processing Layer

In the **data collection layer**, we stored several thermal drone images from various solar panels. Mainly, we collected thermal images of panels produced by 'Schott Solar' and 'CanadianSolar' as well as inverters from 'SMA', 'Fronius', and 'Emerson Control Techniques'. We collected 237 images showing up to 8 damages per picture. While some pictures show damaged cells, others display no failures or anomalies. The solar panels were mainly located in solar parks and are between 6 and 21 years old. Given the high density of PV panels, solar parks are the optimal location to build a reasonable data layer.

In the **processing layer**, we detect where damages in a thermal drone image exactly appear by employing state-of-the-art Computer Vision (CV), i.e., object detection based on pre-trained deep learning models. Pre-trained models enable a lightweight knowledge transfer from research to SMEs through out-of-the-box models. To be more precise, we use YOLO_v5 (You Only Look Once - Version 5), a real-time object detection algorithm. We fine-tune the algorithm by optimizing the weights in the last layers. As a short explanation, the YOLO v5 algorithm is one of the most efficient object detection algorithms in CV. It builds on a Convolutional Neural Network (CNN), capable of real-time predictions with high accuracy in most cases. First, it applies the CNN to the complete input image before it divides the given image into different regions where YOLO_v5 predicts bounding boxes and probabilities. Afterward, these probabilities are weighted to identify the region where the object, i.e., damage, appears [Re16]. We compared the YOLO v5 algorithm to other models (e.g., EfficientDet and MobileNet) but YOLO v5 yields the best accuracy scores and processing performance - which is critical when conducting automated maintenance. We use 111 thermal drone images (including several damages) with a pre-processed size of 416x416. The training set entails 78 thermal drone images. The validation set comprises 21 images. Finally, the test set entails 12 images. We trained the model for 3,000 epochs using a batch size of 32. In Fig. 3, we present our models' output results, i.e., annotated thermal images, and we can observe that it can capture certain solar panel damages. We are currently only using one class/object called 'broken', which indicates whether a panel is damaged or

not. Future work can extend the number of classes to distinguish between different kinds of damages solar panels can exhibit. For example, a typical damage is a defect bypass diode, where a whole substring of the panel sees a loss in efficiency (see left panel Fig. 3). Besides, only single cells in the panel can be damaged, too (see right panel Fig. 3). Additionally, other relevant damages include cell cracks and defective junction boxes.

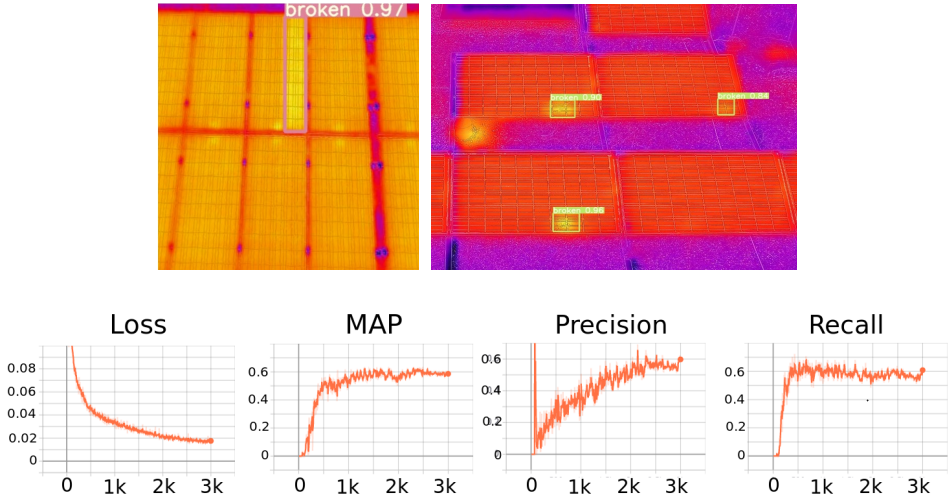


Fig. 3: Exemplary Output and Performance Metrics of YOLOv5 during Training

3.3 Visualization Layer

A powerful visualization solution supports technicians and decision-makers to comprehensively monitor several solar plants and PV modules. However, it is necessary to assign the faulty solar panels to their DT representation to visualize the ML phase results. To this end, we leverage the fact that state-of-the-art drones tag and store the pictures they take together with additional meta-information like the time and location of each image. In our example, the location consists of the drone's exact location and the camera's exact location as three-dimensional GPS coordinates, including the height of the drone relative to the departure point. Additionally, it stores the camera angle from the GPS coordinate as pitch, yaw, and roll. We can compute each solar panel's approximate position on a picture by applying linear algebra based on this information (see Fig. 4).

Finally, we can visualize the solar panels' DT in a central dashboard. Such a dashboard (see Fig. 4) provides a high-level overview of all modules' status and their need for action. Possible information includes basic data like the manufacturer of the modules and year of manufacture, performance measures like the energy produced, manually recorded notes like damages, and damages detected automatically. The dashboard can enrich this information with additional details and explanations, e.g., the remaining

lifetime of a module, information regarding the type of damage (like cell crack or defect bypass diode), why the ML model classified a module as damaged, and recommendations for actions. Thus, it is a helpful extension for various stakeholder groups. On the one hand, various related industries such as the housing industry can adopt this dashboard to their existing Computer-Aided Facility Management systems to extend their infrastructure documentation. On the other hand, solar park operators and technical support companies can focus on the predictive components to improve their maintenance process.



Fig. 4: From a Digital Twin Representation to a Comprehensive Dashboard

3.4 Artifact Evaluation

To establish convincing evidence that our artifact provides a practical approach for modeling a virtual entity of the physical structure, we evaluated the artifact's performance in terms of validity and utility in a real-world setting. We observed several solar park sites for validity and used the artifact to create fitting virtual entities for the physical structures. Overall, we observed three different solar parks and 12 residential roofs equipped with PV modules. We collected 237 thermal images and identified 166 faulty cells. Overall, the tests on PV solar parks corroborate the efficiency and value of the proposed artifact. When looking at our ML model's performance metrics, we focus on the Mean Absolute Precision (MAP), Precision, and Recall, as these are the most reliable metrics in CV tasks. The MAP is 63%, Precision 58%, and Recall 63% on the test set. Although our models' results promise a good performance, training the ML model based on a different algorithm or more input data might be beneficial. However, different CV algorithms require significantly higher computational power and longer

training times. Thus, they might be unsuitable for fast and efficient deployment and the hardware stack of most SMEs. Moreover, when comparing our model's results to EfficientDet and MobileNet, YOLOv5 still outperformed both lightweight models in terms of MAP (EfficientDet: 28.26%; MobileNet: 18.33%). Notably, the presented approach's identified damages might be incomplete or include so-called false positives (i.e., working modules wrongly classified as damaged). Against this background, our artifact is a lightweight support solution to significantly reduce expert knowledge needs, even though it does not replace expert knowledge altogether.

Overall, our artifact presents both advantages and disadvantages. One shortcoming of the described artifact lies in the frequency of data collection. The data is not real-time, i.e., permanent measurable, but instead based on automated maintenance cycles that might only occur several times a year. However, this is sufficient for most SMEs, as most of them currently have less control over such an infrastructure. On the beneficial side, our artifact can partly automate the full process by using autonomous flight capabilities of drones based on a predefined route combined with automated scheduling. Since flight scheduling depends on multiple factors, e.g., on the maintenance need and external conditions like wind, sunlight, and clouds, ML can predict an optimal timing. Due to legal restrictions, fully automated drone-based maintenance is currently not possible but by combining the drone flight with the required on-site inspections, we can decrease the necessary timeframe as well as staff requirements. Our proposed semi-automated process can decrease maintenance efforts, reduce costs, and increase structured information available for further analysis and decision-making.

To assess whether the artifact also fulfills necessary utility criteria in the problem space, we contacted the responsible technical operations managers and solar park operators of several sites. By showcasing the performance and discussing the results, we established a confidence level in our solution. The transformation of the whole process into separate services enables the transformation of PV systems into virtual entities. Thus, stakeholders can select the relevant services they want to use, increasing the presented system's usefulness. Since most PV system operators currently do not rely on a digital representation of their modules, they demand such a solution. Furthermore, our artifact provides a lightweight AI application that could be further adopted for other problem spaces. SMEs could adopt this approach utilizing pre-trained image recognition models to build specific solutions in their area of expertise. This solution lowers the barrier of entry for digitalization and implementing smart maintenance processes enabling sustainable business models also for SMEs.

4 Conclusion

Particular context areas and industries still lack the availability of structured qualitative data to analyze the condition of physical assets. Modern technologies like cloud computing, Digital Twins, and Machine Learning can help SMEs to tackle this task by

integrating a virtual entity for physical structures. We demonstrated the applicability of such a transformation process on the example of PV modules. PV modules are rarely digitalized, and their maintenance is very work-intensive. With a drone equipped with a thermal camera, we reveal an improvement of current maintenance processes by using Computer Vision. Thus, we can transform solar panels and their unstructured data consisting of thermal images into their DTs. This information is useful for further processing, like predicting maintenance needs and automatic scheduling.

We contributed to the theory of processing drone imagery to identify the respective DTs based on location data. Additionally, we provide valuable insights into the process of optimizing PV maintenance. Last, our results serve as input for SMEs decision-makers on how to collect and evaluate PV modules' respective physical conditions and improve current processes.

While our demonstration focuses on PV maintenance, a more abstract representation would also enable an operationalization of the artifact into other context areas. This approach would thus allow for significant knowledge transfer between current research and SMEs and enables sustainable business models through automation. Potential context areas include all those with larger physical structures and a lack of digitalization. Possible applications are virtual identities for buildings, public infrastructure, or traffic surveillance. Such a knowledge transfer would also significantly increase the value and validity of the researched approach.

Further research can either transfer our approach to a different context area or improve certain steps of our approach. Due to the modularity of possible services, future work could segment issues and research questions and address them individually without needing to address the whole implementation.

Acknowledgements

This work was funded by the BMWi through the ForeSight project. Additionally, we like to thank Oberhessische Versorgungsbetriebe AG (OVAG), RMS Rhein Main Solarpark GmbH, and GreenVesting GmbH & Co. KG for their time and valuable support. The authors state no conflict of interest.

Bibliography

- [Al14] Alestra, S., Brand, C., Erofeev, P., Bordry, C., Burnaev, E., Papanov, A.: Rare Event Prediction Techniques in Application to Predictive Maintenance of Aircraft. In: Proceedings of ITaS conference, pp. 32–37, 2014.
- [Be17] Betti, A., Trovato, M.L.L., Leonardi, F.S., Leotta, G., Ruffini, F., Lanzetta, C.: Predictive Maintenance in Photovoltaic Plants with a Big Data Approach. 33rd European Photovoltaic Solar Energy Conference and Exhibition; pp. 1895-1900, 2017.
- [Be20] Bell, J.: Machine Learning: Hands-On for Developers and Technical Professionals. John Wiley & Sons, 2020.

- [BS15] Buerhop-Lutz, C., Scheuerpflug, H.: Inspecting PV-plants using aerial, drone-mounted infrared thermography system. 2015.
- [CBK09] Chandola, V., Banerjee, A., Kumar, V.: Anomaly detection: A survey. *ACM Comput. Surv.* 41, article 15, pp. 1-58, 2009.
- [De12] Denio, H.: Aerial solar Thermography and condition monitoring of photovoltaic systems. In: 2012 38th IEEE Photovoltaic Specialists Conference, pp. 613–618, 2012.
- [De18] De Benedetti, M., Leonardi, F., Messina, F., Santoro, C., Vasilakos, A.: Anomaly detection and predictive maintenance for photovoltaic systems. *Neurocomputing* 310, pp. 59–68, 2018.
- [EBA20] Errandonea, I., Beltrán, S., Arrizabalaga, S.: Digital Twin for maintenance: A literature review. *Computers in Industry* 123, article 103316, 2020.
- [GH13] Gregor, S., Hevner, A.: Positioning and Presenting Design Science Research for Maximum Impact. *MIS Quarterly* 37, pp. 337–356, 2013.
- [GV17] Grieves, M., Vickers, J.: Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems. In: Kahlen, F.-J., Flumerfelt, S., and Alves, A. (eds.) *Transdisciplinary Perspectives on Complex Systems: New Findings and Approaches*, pp.85–113. Springer International Publishing, Cham, 2017.
- [Ha10] Hashemian, H.M.: State-of-the-art predictive maintenance techniques. *IEEE Transactions on Instrumentation and measurement* 60, pp. 226–236, 2010.
- [Mo02a] Mobley, R.K.: *An introduction to predictive maintenance*. Elsevier, 2002.
- [Pe07] Peffers, K., Tuunanen, T., Rothenberger, M.A., Chatterjee, S.: A design science research methodology for information systems research. *Journal of management information systems* 24, pp. 45–77, 2007.
- [Re16] Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: *You Only Look Once: Unified, Real-Time Object Detection*. Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016.
- [Sc17] Schleich, B., Anwer, N., Mathieu, L., Wartzack, S.: Shaping the digital twin for design and production engineering. *CIRP Annals* 66, pp. 141–144, 2017.
- [Sh18] Shubenkova, K., Valiev, A., Shepelev, V., Tsiulin, S., Reinau, K.H.: Possibility of Digital Twins Technology for Improving Efficiency of the Branded Service System. In: 2018 Global Smart Industry Conference (GloSIC), pp. 1–7, 2018.
- [Sp12] Spagnolo, G.S., Vecchio, P.D., Makary, G., Papalillo, D., Martocchia, A.: A review of IR thermography applied to PV systems. In: 2012 11th International Conference on Environment and Electrical Engineering, pp. 879–884, 2012.
- [WY07] Widodo, A., Yang, B.-S.: Support vector machine in machine condition monitoring and fault diagnosis. *Mechanical systems and signal processing* 21, pp. 2560–2574, 2007.
- [Yü20] Yürüşen, N.Y., Rowley, P.N., Watson, S.J., Melero, J.J.: Automated wind turbine maintenance scheduling. *Reliability Engineering & System Safety* 200, article 106965, 2020.

German Question Answering in CRM Systems

Sophie Schäfer,¹ Thorsten Zylowski²

Abstract: We present a Question Answering (QA) approach within the context of Customer Relationship Management (CRM) systems in German which we applied to several use cases. However, we aim to generalize the suggested approach, so it can be applied to other domains. In this course, the subjects of tabular QA, extractive QA and Frequently Asked Questions (FAQ) are examined. Moreover, we reveal our findings regarding fine-tuning transformer models for QA and propose an automatic labelling mechanism that can be integrated into QA systems in order to simplify the creation of training data. We evaluated on various CRM-related data sources. The evaluation of the fine-tuned extractive QA pipeline resulted in an F1-score of 76.93 %. A qualitative analysis by domain experts showed that the tabular QA pipeline with translation and column name mapping leads to an accuracy of 40 %.

Keywords: Natural Language Processing; Question Answering; Information Retrieval; Customer Relationship Management

1 Introduction

Operating computers with textual or spoken language via virtual assistants, such as Siri or Alexa, becomes increasingly popular as it facilitates work or daily life tasks. Not only the usage of language to control computers but also the amount of available data increases rapidly. Searching through large document collections and databases to retrieve the desired information is time-consuming and often involves using query languages. Although document ranking functions facilitate the retrieval of relevant information, they do not provide specific answers to the users. Therefore, it would be of great advantage to be able to pose a question in natural language, which is a communication form users are familiar with, and receive the exact answer. This is the task of Question Answering.

In this paper, we examine how QA can be of benefit in the CRM use case in the German language. Due to their manifold sources and communication channels, CRM systems contain heterogeneous data, structured and unstructured, which complicates finding useful information. We assume that QA can alleviate this issue by providing exact answers to questions posed in human language. However, the diversity of data sources in CRM calls for a diverse and adaptive QA system that is able to handle (semi-)structured as well as unstructured data.

¹ CAS Software AG, Future Labs, CAS-Weg 1 - 5, 76131 Karlsruhe, Deutschland sophieschae1@gmail.com

² CAS Software AG, Future Labs, CAS-Weg 1 - 5, 76131 Karlsruhe, Deutschland thorsten.zylowski@cas.de

2 Related Work

The emerging topic of transformer-based QA is increasingly moving into the centre of attention. Neural language models, like Google's well-known Bidirectional Encoder Representations from Transformers (BERT) [De19], can be fine-tuned for the specific task of QA. They are then often referred to as Machine Reading Comprehension (MRC) models or readers and are used to extract spans from paragraphs to answer a given question. Some approaches combine document ranking functions with MRC models in an Information Retrieval based QA approach [Ch17][KF19]. Regarding QA over tables or knowledge-based QA, many research activities involve semantic parsing, where questions in natural language are transformed into logical forms [Cl10][LJK13]. Google researchers recently introduced a groundbreaking approach, the Table Parser (TAPAS)[He20]. To further tailor a QA system to a company's needs, it can be beneficial to extend the notion of QA by integrating Frequently Asked Questions (FAQ)-Retrieval. FAQ-Retrieval aims at facilitating the acquisition of information "by allowing users to pose natural language queries over an FAQ collection" [KŠ18]. To accomplish this, the approach by Sakata et al [Sa19] includes using a transformer to compare the word embeddings from the user's query to the questions in the FAQ dataset. The FAQ retriever then outputs a ranked list of question and answer pairs, starting with the FAQ pair, where the question best matches the user query. Although QA is widely researched, QA in German within CRM at medium-sized enterprises is not particularly researched and therefore this work offers new insights.

3 Question Answering Architecture

One of the major goals of using a CRM system is retrieving information that helps the user to fulfil his or her current task. In the proposed architecture (Fig. 1.) the user has a single access point for retrieving relevant information by typing a question in natural language. This question goes through a query classification step that tries to infer the corresponding use case. Possible use cases are, for example, finding general information on a company website, retrieving structured sales information or looking up the answer in a curated FAQ database.

The classified query is sent to the corresponding QA pipeline. Due to the fact that the retrievable data can be stored in an unstructured, semi-structured and structured form, several QA pipelines are needed. Each of these data types requires different QA techniques, which are described in detail in the following section. The answer is generated by one of the pipelines and is returned to the user. It contains facts that answer the question as well as additional information.

The QA architecture includes an annotation mechanism for continuous generation of training data which is described in section 3.4 in detail.

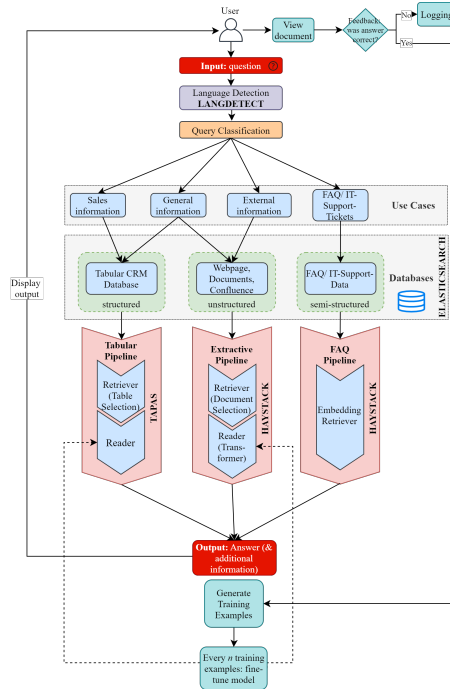


Fig. 1: QA System Architecture

3.1 QA Pipelines

Extractive or free text QA describes the task of retrieving an exact answer to a question from a given context. In the course of this work, the cutting-edge open-source framework Haystack [Pi21b] developed by the German company deepset has been deployed. Haystack offers scalable and highly customizable QA systems utilizing SOTA neural language models. As detailed in their documentation an extractive QA pipeline consists of the three main components document store, retriever and reader. The document store holds the document collection to be queried, in the CRM use case we used Elasticsearch to index the documents. The retriever applies a document ranking function (in our case BM25) to filter the documents given a question. Then the resulting candidate documents are passed to the reader, the MRC model we fine-tuned on domain-specific QA data, that "reads" the candidate documents to extract the answer to the question.

Especially when selling more complex products like software, companies often have frequently asked questions (FAQ) sections, comprising the questions that other users have asked before or are likely to ask in the future, and the corresponding answers. In this case, we examine a website that provides questions or problem descriptions concerning CRM products and solutions addressing them. FAQs are often causal type questions (how, why)

and therefore require a relatively long answer or explanation. Thus, it is sufficient to retrieve the whole answer, instead of extracting shorter spans from the answer. Again, for FAQ-style QA we employ the framework Haystack. A so-called sentence transformer [RG19] computes the similarity of the embeddings of the user questions with those of the questions present in the FAQ dataset. As a result, the pipeline returns the question-answer pair that semantically best matches the user query. In this work, a multilingual model³ trained on German and English data is used.

One of the currently most promising approaches to query tables is the above-mentioned TAPAS [He20]. TAPAS enriches the BERT architecture with additional embeddings containing information about the tabular structure. Moreover, TAPAS comes with two classification layers to select table contents and predict aggregations on multiple cells. As TAPAS is laid out for the English language, we implemented a translation mechanism based on transformer models by Helsinki-NLP [TT20] to translate the questions as well as the table contents to English before passing them to TAPAS. To further improve the model performance, we created a mapping function. There are multiple ways to phrase a question since a lot of words have synonyms. However, experience has shown that TAPAS performs better when certain keywords from the table, such as column names are included in the question. To mitigate this issue, for each table, a schema is automatically created containing synonyms and their lemmas from WordNet[Pr10] for every column name. For example, when a user asks for the “postal code” of a city, while the corresponding column name is “zip code”, the function would automatically replace the keyword “postal code” in the question with “zip code” to ensure better matching.

3.2 Model Fine-tuning

In order to fit transformer models in NLP for a downstream task like extractive QA, transfer learning can be applied in multiple stages. In this work, we leverage deepset’s recently published pioneering German language model GELECTRA-large-GermanQuAD [MRP21], which achieved a new SOTA on German QA benchmarks. The model has been fine-tuned on a novel German SQuAD-style dataset, GermanQuAD. We further fine-tuned the model using the framework FARM [Pi21a] on a small manually labelled German dataset. The dataset comprises 1107 training examples, including CRM-related information, and was annotated with a tool specifically designed for this purpose⁴.

3.3 Prototype

For the practical realisation of our concept, we implemented our architecture as a web app based on the Flask⁵ framework (Fig. 2.) which can be easily integrated into CRM

³ <https://huggingface.co/T-Systems-onsite/cross-en-de-roberta-sentence-transformer>

⁴ <https://annotate.deepset.ai/>

⁵ <https://palletsprojects.com/p/flask/>

systems. The user has the possibility to enter a question and choose the desired number of returned answers. The extracted answers are displayed with additional context which helps the user to assess the correctness of the answer. Two buttons below each answer allow further inspections and serve as a feedback and labelling mechanism for continuous model improvement.

CAS
SOFTWARE
Question Answering

Ask a question:

Number of answers
1

Enter your question here...

Select product for FAQ

Search

You asked: *was ist genesisWorld?*
These are the search results:

detected language: de
answer time: 2.05 seconds

Top 1 Answer: die Nr. 1 CRM-Software für Ihr Unternehmen

Context:
cas Aktuelles Presse Support Kontakt Jobs Sprache Customer Centricity CRM & xRM Branchen
Lösungen Referenzen Unternehmen CAS genesisWorld | Überblick Highlights Funktionen
Kundenstimmen Kontakt Support CAS genesisWorld – DIE NR. 1 CRM-SOFTWARE FÜR IHR
UNTERNEHMEN Revolutionieren Sie Ihr Kundenmanagement – mit der leistungsstarken CRM-
Standardlösung CAS genesisWorld. Durch flexible Module passt sich die CRM-Software exakt an Ihre
Bedürfnisse an. Intelligente Schnittstellen sorgen für eine nahtlo

Relevance: 78.56 %

Correct View document

Fig. 2: QA System Prototype

3.4 Labelling Mechanism

Annotating training data for extractive QA manually is rather time-consuming. In order to mitigate the problem of lack of non-English QA data, an automated labelling mechanism is introduced. The user is encouraged to provide feedback, whether one of the delivered answers is accurate. In case the user marks an answer as correct, a SQuAD-style training example is created. The MRC model can be fine-tuned on that data following a continual learning approach, tailoring it directly to the user's needs. In the CRM use case, it can be beneficial to create specific user roles. For example, employees in the sales department are assumed to ask different kinds of questions than software developers. Hence, the feedback of employees with a similar user profile can be collected separately from user feedback of other user profiles allowing each user group to obtain a model, which is specifically fine-tuned for their daily work.

4 Evaluation

For the evaluation of the extractive QA pipeline, the performance of the combination of retriever and reader is measured. For testing purposes, a German manually annotated domain dataset with 272 questions is used. Since it mainly contains data from company websites including company-specific terms like product names, it can be considered CRM-related data. The parameters chosen in the evaluation represent a trade-off between answer accuracy and time consumption. For the retriever, the top-k value is set to 5, which means that for each query 5 relevant documents are retrieved that are to be read by the reader. The reader looks for 3 relevant answers per document. The evaluation shows, that the combination of retriever and reader achieves a top-1 F1-Score of 76.93 %. The retriever on its own reaches a recall of 93.75 %, i.e. for 255 of 272 queries the correct document was retrieved. The reader on its own achieves a top-1 F1-Score of 82.69 %.

For the proof of concept of the modified tabular pipeline, domain experts were asked to pose questions on five tables, extracted and shortened to around 70 cells from the CRM database. The questions and tables were then run through the tabular pipeline. In the last step, the candidates evaluated if the returned answers were correct. The total evaluation of the tables leads to an overall accuracy of 40.00 %.

As there currently are no German FAQ datasets publicly available, especially in the CRM context, the evaluation is deployed on a subset of the above-mentioned FAQ website comprising 50 question-answer pairs on six products. The evaluation is accomplished by posing issues or questions and simultaneously selecting the concerned product, which is the expected flow from the implemented architecture. Hence, the query is always performed on an even smaller subset of the provided dataset, which narrows down the answer possibilities and thereby improves the results. In the course of the assessment, 20 questions or problem descriptions were posed by employees. Afterwards, they assessed, whether the top-1 retrieved question-answer pair was the most relevant to their question. They stated that 16 out of these 20 were correctly retrieved as the top 1 answer, leading to an overall top-1-accuracy of 80.00 %.

5 Conclusion

This work demonstrates that extractive QA combined with FAQ and tabular QA is possible within the CRM context in the German language. It allows the user to find the desired information faster and more precisely. By enabling the user to ask a question in human language, a natural and intuitive workflow is encouraged. It is shown, that especially extractive QA with the domain-customized version of GELECTRA-large-GermanQuAD and FAQ QA already achieve proper results for a high percentage of questions. For German tabular QA, in this work, a significant improvement was accomplished, yet the resulting German pipeline is still rather immature. The architecture suggested within this paper can be easily modified for application to other domains, where different data types occur. In

conclusion, this work proves the concept of multi-source QA in CRM at a medium-sized company and the developed prototypical QA system succeeds at revealing the merits as well as the need for German QA in CRM.

References

- [Ch17] Chen, D.; Fisch, A.; Weston, J.; Bordes, A.: Reading Wikipedia to Answer Open-Domain Questions, 2017, arXiv: 1704.00051 [cs.CL].
- [Cl10] Clarke, J.; Goldwasser, D.; Chang, M.-W.; Roth, D.: Driving Semantic Parsing from the World's Response. In: Proceedings of the Fourteenth Conference on Computational Natural Language Learning. Association for Computational Linguistics, pp. 18–27, 2010, URL: <https://www.aclweb.org/anthology/W10-2903>.
- [De19] Devlin, J.; Chang, M.-W.; Lee, K.; Toutanova, K.: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers). Association for Computational Linguistics, Minneapolis, Minnesota, pp. 4171–4186, June 2019, URL: <https://www.aclweb.org/anthology/N19-1423>.
- [He20] Herzig, J.; Nowak, P. K.; Müller, T.; Piccinno, F.; Eisenschlos, J. M.: TAPAS: Weakly Supervised Table Parsing via Pre-training, 2020, URL: <http://arxiv.org/pdf/2004.02349v2>.
- [KF19] Kratzwald, B.; Feuerriegel, S.: Putting Question-Answering Systems into Practice: Transfer Learning for Efficient Domain Customization. *ACM Transactions on Management Information Systems* 9/4, pp. 1–20, 2019, ISSN: 2158-656X, URL: <http://arxiv.org/pdf/1804.07097v2>.
- [KŠ18] Karan, M.; Šnajder, J.: Paraphrase-focused learning to rank for domain-specific frequently asked questions retrieval. *Expert Systems with Applications* 91/, pp. 418–433, 2018, ISSN: 09574174.
- [LJK13] Liang, P.; Jordan, M. I.; Klein, D.: Learning Dependency-Based Compositional Semantics. *Computational Linguistics* 39/2, pp. 389–446, 2013, ISSN: 0891-2017.
- [MRP21] Möller, T.; Risch, J.; Pietsch, M.: GermanQuAD and GermanDPR: Improving Non-English Question Answering and Passage Retrieval, 2021, arXiv: 2104.12741 [cs.CL].
- [Pi21a] Pietsch, M.; Soni, T.; Chan, B.; Möller, T.; Kostić, B.: FARM (Version 0.6.2), 2021, URL: <https://github.com/deepset-ai/FARM>.

- [Pi21b] Pietsch, M.; Soni, T.; Chan, B.; Möller, T.; Kostić, B.: Haystack (Version 0.7.0), 2021, URL: <https://github.com/deepset-ai/haystack/>.
- [Pr10] Princeton University: About WordNet. Ed. by Princeton University, 2010, URL: <https://wordnet.princeton.edu/>.
- [RG19] Reimers, N.; Gurevych, I.: Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In: Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, Nov. 2019, URL: <https://arxiv.org/abs/1908.10084>.
- [Sa19] Sakata, W.; Shibata, T.; Tanaka, R.; Kurohashi, S.: FAQ Retrieval using Query-Question Similarity and BERT-Based Query-Answer Relevance, 2019, URL: <http://arxiv.org/pdf/1905.02851v2>.
- [TT20] Tiedemann, J.; Thottingal, S.: OPUS-MT — Building open translation services for the World. In: Proceedings of the 22nd Annual Conference of the European Association for Machine Translation (EAMT). Lisbon, Portugal, 2020.

A systematic literature review of machine learning canvases


Lukas-Walter Thiée  ¹

Abstract: The use of machine learning technology is still significantly lower in small and medium sized enterprises than in large enterprises. It seems that there are specific challenges in the implementation of data-driven methods, that hinder SMEs in their adoption. One approach to support the initialization and execution of such methods is the use of boundary objects, e.g., canvases, serving as a visual communication document. As it is not clear which approaches are being pursued in detail and how they are interrelated, in this paper, a systematic literature review is being presented, that identifies and analyzes 18 canvas artifacts. These canvases represent the status quo and they can be grouped into four distinct categories of different foci. The aggregation of the fields and questions provides an essence of canvas contents, to point out gaps and ultimately to expand the canvas approach as well as ML adoption.

Keywords: Machine Learning, Canvas, Literature Review, SME

1 Introduction and Approach

The extent to which companies apply machine learning (ML) has increased in recent years and with over a quarter of German companies using multiple ML technologies today [ID20], it is far from a niche technology. While larger companies have been taking advantage of data-driven analytics for some time, mainly with a focus on process optimization [WG19], [VD18] and product development [WG18], for small and medium sized companies (SMEs) in Germany it is still difficult to start and use ML applications in their businesses [Mi20a]. As of 2020 only 10% of the small companies utilize multiple ML technologies and over 22% admit that ML applications are still not a topic in the company at all [ID20]. Furthermore, only 29% of SMEs assess ML as a driver for innovation and product development [ID20], which could indicate an imbalance in the assessment of opportunities and challenges of data science (DS), artificial intelligence (AI) in particular. Nevertheless, two out of five SMEs are beginning to plan digitalization projects and another 29% are considering to do so [Kf20]. Even though studies have shown benefits of data usage for SMEs [Mo18], SMEs have been focusing descriptive approaches and conventional business intelligence [BF18], rather than leveraging predictive or prescriptive data analytics to its full potential [Co16]. It seems that to the same extent as SMEs were challenged with the introduction of information technology in its beginnings [Gh11], they are now challenged to implement advanced data analytics and data strategies

¹ Leuphana Universität Lüneburg, Institut für Wirtschaftsinformatik, Universitätsallee 1, 21335 Lüneburg,
lukas-walter.thiее@leuphana.de,  <https://orcid.org/0000-0002-1998-376X>

[VLF15]. However, in practice there seem to be specific challenges for SMEs regarding the initialization and successful execution of ML projects [BvK20]. A regional survey in Germany identifies high investment cost, data collection and processing, shortage of specialists and AI maturity as the main barriers for companies [Mi20b]. One of the main reasons for the slow adoption of ML technology in SMEs is probably the lack of resources [Kf17]. Compared to larger companies, SMEs have less financial power and fewer or no skilled personnel for specific tasks in ML projects. Consequently, this gives rise to a variety of challenges. Lack of management focus and experience [Mo17], lack of internal and external experts [Ru15], or complexity of (sequential) data [FRL17] are exemplary challenges that can impede both project initiation and success. One of these challenges is the detailed description and expedient guidance in the use case definition and execution phase of an ML project. Naturally, these projects include multiple stakeholders at varying levels of data-literacy [Ke21], which unfortunately promotes the development of communication silos [ID20]. This communication barrier, internally and externally, must be overcome. Therefore, in addition to general methods that foster ML adoption, such as design thinking workshops, specific tools have been proposed, such as the DUCAR process model for smart picking of ML use cases [Sc20], or procedures to identify and prioritize use cases [Ku17], [Sc16]. Also the use of Guided Analytics [Bo19] has been proposed, so that non-experts can access ML tools [Qi18]. Nevertheless, this approach currently still mainly focuses model tuning. Another promising approach to facilitate ML communication and execution is the use of canvases. “A canvas is just a visual chart to describe a complex object, in a better way than a simple text document. [...] blocks are arranged on the chart in a way that makes visual sense” [Do19]. Canvases are a kind of boundary object, i.e., a document that serves as a communication platform between multiple stakeholders, motivates cross-disciplinary collaboration [KPD18], and provides common identity [SG89]. Originating from the Business Model Canvas by Osterwalder and Pigneur [OP10], this approach has been further developed and analyzed in many ways, for example in data-driven business models [Ha14]. However, it is not clear which of these canvas approaches is most promising for initial ML projects.

Research Approach and Contribution. For this reason, it is essential to review relevant resources systematically. Comparing different data-related canvases not only streamlines research in this field, but also provides useful guidance in business practice. A compilation of existing tools and a possible extension that can easily be applied in practice, can help bridging between the aforementioned corporate silos, i.e., decision-makers, developers, domain experts, and external partners, as well as promote the execution of ML tasks. Furthermore, the result of this review will be the foundation of conducting empirical research in longitudinal case studies to investigate the applicability and usefulness of such boundary objects. The detailed research questions for this paper are: (1) Which canvas models, that address ML or AI implementation, are available, and which contents do they cover?; and (2) Where are gaps and what are potential extensions of these canvases in order to address specific challenges and needs in initial ML projects? Providing answers to these questions might support initialization and successful execution of ML projects in small organizations. Therefore, the goal for this review is to find relevant canvas tools, compare these tools systematically, and potentially enhance them.

2 Methodology: Literature search process documentation

In order to find existing evidence, identify gaps, and build an appropriate background on the topic a systematic literature review is being conducted [Ki04]. The review shall identify relevant sources within a defined scope, synthesize the findings [WW02], and most of all provide guidance on research and practice [Sc15]. As rigor and reproducibility are key to qualitative IS reviews [Br09], the search process is documented in the following. The selection of databases (see Tab. 1) comprises ten international renowned databases, such as the Web of Science Core Collection, the ACM digital library and IEEE Xplore. The latter belong to the most impactful databases in computer science [KJZ16]. The selection of databases was inspired by the journal rating for IS literature from VHB JOURQUAL3 [HS15], which also comprises IS conference proceedings. The AIS Electronic Library features important IS outlets, such as MISQ or BISE. Furthermore, two online archives, Google Scholar and ArXiv.org; as well as three German online libraries were included, because they basically supplement the corpus of resources. To further integrate practitioner views, the Harvard Data Science Review, which is a non-peer-reviewed open access journal of the MIT Press, was also considered.

Database	Results	Dupl.	TAK	
AISeL	4	4	4	
EBSCOhost BSC	2	2	2	Backward:
ACM Digital Library	2	1	0	+15
Emerald Insight	0	0	0	
ScienceDirect	3	3	1	
SCOPUS	8	4	1	Forward:
IEEE Xplore	1	1	1	+1
Web of Science	15	14	2	
Google Scholar	3	0	0	
ArXiv	1	0	0	Language:
Taylor&Francis	3	1	0	-2
Springer Link	3	3	1	
HMD	0	0	0	
Duncker & Humblot	0	0	0	Artifact Identity:
RonPub	0	0	0	-1
Harvard DS Review	0	0	0	
Total	45	33	12	25

Tab. 1: Databases and literature search process results

The formulation of the search terms is very important in this matter, because there are often no clear dividing lines between the terms, e.g., the phrases “data mining”, “data-driven”, “digitalization”, or “big data analytics” all have a large intersection, which is mainly due to the fact that in the field of IS there is a plethora of terms and abbreviations that describe the topic area or sub-areas. Not even the distinction between AI, ML and deep learning is unambiguous. For example, an initial search in the EBSCOhost Business Source library, yields over 30,000 results. Therefore, we limit the search to the search string: *{ "Machine Learning Canvas" OR "Artificial Intelligence Canvas" OR "ML Canvas" OR "AI Canvas" }*. This initially excluded phrases such as digital canvas or data

science canvas. The query was individually adapted to the requirements of the respective database. No keyword was integrated that was specifically suitable for SMEs, since it was assumed that so far rather general approaches are available. The query was executed on all fields, i.e., title, abstract, keywords and full text, and the time span for results was limited to the years 2000 to 2021. The search was conducted end of March 2021. As the search term is rather strict, there were 45 initial hits within the included databases (see Tab. 1). Eliminating duplicates left 33 results and screening titles, abstracts, and keywords (TAK) reduced the results to 12 articles. The screening excluded hits that didn't relate to ML or AI canvases or projects, such as the "Business Ethics Canvas" [VHR20]. A backward search through the listed references as proposed by Webster and Watson [WW02] complements the search and adds 15 articles to the results. The backward search was performed in such a way that a TAK screening was performed for references that showed a promising title related to canvases, such as "Data-Driven Business Models" [BT19]. Two articles were excluded from the results, because they were not written in English and another article was excluded, because it was a foundational research article to a different article in the results, describing the same artifact, "ML-Process Canvas" [Zh19]. The results all originated from the time window of the years 2016-2021. This period is very recent, therefore, conducting a forward search in the mentioned databases didn't reveal any further relevant results. However, a sample forward search on Google yielded citations on the professional online network LinkedIn, which produced another result [Sc18]. In total there were 25 relevant articles left, which contained 18 canvases for in-depth analysis, which is the answer to part one of research question one (Tab. 2).

Year	Source	Canvas Artifact	#Fields	Structure
2016	[MK16]	Data Canvas: Data-Need Fit	10	F, D, E
2017	[SN17]	Data Value Map	14	F, S, D
2017	[He17]	Digitalization Canvas	9	F, S, Q
2018	[AGG18]	AI Canvas	7	F, Q
2018	[DR18]	AI Canvas	8	F, Q
2018	[Sc18]	The ML Canvas (Big Data MBA Version)	+2	F, D, Q
2018	[Sc18]	Hypothesis Development Canvas v1.1	10	F, Q
2019	[BT19]	Data Insight Generator	6	M, D, Q
2019	[Do19]	Machine Learning Canvas v0.4	10	F, D, Q
2019	[KM19]	Data Innovation Board	14	F, S, Q
2019	[KMK19]	Data Collection Map	12	F, E
2019	[Za19]	AI Project Canvas	9	F, S, Q
2020	[EL20]	AI performance canvas (prototype)	10	F, S, Q
2020	[FBP20]	Data Product Canvas	7	F, Q, E
2020	[HST20]	Key Activity Canvas	10	M, Q, E
2020	[KS20]	Canvas for the use of AI (author)	7	F, Q
2020	[Zh20]	ML Lifecycle Canvas	6	M, D, Q
2021	[Ke21]	Enterprise AI Canvas	12	F, S, Q

Notes: F = Fields, M = Matrix, S = Sections, E = Examples, D = Descriptions, Q = Questions

Tab. 2: Literature search results – 18 canvas artifacts

3 Analysis and Results

The analysis of the canvases is performed in two parts, structural analysis and content categorization. The structural analysis considers the layout and the process of filling. For this purpose, both the artifacts themselves and a full-text analysis are performed. The process in the full-text analysis followed in this qualitative content analysis is based on inductive formulation of categories [Ma10], that were given by the captions in the canvases. Also, it is being determined whether and which roles or persons are being addressed to fill the respective canvas and if the filling is self-sufficient and detailed guidelines published. Further, if applicable, the scientific derivation and evaluation of the artifacts are investigated.

Structural analysis. Almost all artifacts are structured as a canvas with fields that have a descriptive title accompanied by either questions or examples (see Tab. 2, Structure column). A few use a 2-d matrix arrangement with rows and columns. Some artifacts provide sections or arrows in the layout to guide the user in filling in the fields, e.g., the “Key Activity Canvas” provides dotted arrows to visualize interactions between the customer, the company, and the partners [HST20], or Zawadzki draws arrows that indicate how to go from section to section in the “AI Project Canvas” [Za19]. 4 articles provide written guidelines how to proceed in filling in the canvas fields, such as “Explore, Ideate, Evaluate” [KM19], “Think, Validate, Know” [BT19], design loops [Zh20], or agile development [He17]. Whereas 3 others point to the iterative nature of the tool [Do19], [Sc18], [HST20], the remaining 11 do not specify the process (once or iterative), so that a single pass of the filling has to be assumed. In terms of size, or number of fields respectively, the 18 artifacts range from 7 fields in the “AI Canvas” [AGG18] to 14 fields in the “Data Innovation Board” [KM19]. The highest number of cells, results from one of the matrix approaches, namely 21 cells in the “Key Activity Canvas” [HST20].

As an overarching finding, it can be noted that all articles see data-driven projects, and the respective canvas in particular, as an interdisciplinary task. Nevertheless, 5 of the 18 articles don’t specify the person or department which should fill in the canvas. 5 others only mention general terms, such as “business stakeholders” [Sc18], “heterogeneous stakeholder groups” [BT19], “pioneers” [DR18], or “different departments and diverse expertise” [MK16]. Specifically mentioned are Data and AI project managers, IT departments, domain experts, service design teams, data science teams, “senior executives, middle management, frontline staff, business stakeholders, technology stakeholders and customers” [SN17], or in general “managers, who provide the glue between everyone” [Do19]. Regarding the scientific derivation, 6 articles explicitly name Design Science Research, Action Design Research, Research through Design, or Design Thinking as methodological procedures in their articles, e.g., applying questionnaires, triangulation, or design principles as methods. 5 of these also describe the evaluation procedure in their research, which is predominantly a focus group workshop. Also, the ontology, which the artifact is based upon, is mentioned in 2 articles, namely for the “Data Collection Map” [KMK19], and for the “Data Innovation Board” [KM19]. 2 other articles describe the interviews and workshops conducted as part of a case study [He17], [MK16]. [DR18] and

[Do19] ground their artifacts on the Business Model Canvas by [OP10]. However, nearly half of the results (8/18) don't specify the scientific method, which the artifact is based upon, and these contributions also don't mention the evaluation technique applied.

Categorization. In the second part of the analysis, the individual canvases and their core contents and objectives are being investigated. 6 of the 18 canvases explicitly label the canvas with the term AI, e.g., the “AI Canvas” by Agrawal et al. (2018), and 3 assign ML to their artifact, e.g., the “ML Lifecycle Canvas” by Zhou et al. (2020). The naming already indicates that different foci are being set. As the artifacts all belong to the same realm of data science, categorizing them is ambitious due to the proximity of their contents. Nevertheless, four categories can be proposed, as summarized in Fig. 1. This methodological procedure is in line with Webster and Watson's (2002) call for a concept-centric approach in IS literature reviews [WW02].

ML / DS Focus	(AI) Project Focus	Data Value Focus	Data Source Focus
<ul style="list-style-type: none"> • [Do19], [Sc18], [Sc18], [Zh20], [HST20] 	<ul style="list-style-type: none"> • [AGG18], [DR18], [Za19], [KS20], [Ke21], [EL20], [He17] 	<ul style="list-style-type: none"> • [FBP20], [SN17], [MK16], [KM19] 	<ul style="list-style-type: none"> • [BT19], [KMK19]

Fig. 1: Categories of canvases with different thematic foci

Machine Learning / Data Science Focus. The first category is formed by canvases with a technical focus on machine learning and data science. It includes the “Machine Learning Canvas v0.4” [Do19] and its extension the “Machine Learning Canvas (Big Data MBA Version)” [Sc18], as well as the “Hypothesis Development Canvas” [Sc18], the “ML Lifecycle Canvas” [Zh20], and the “Key Activity Canvas” [HST20]. These artifacts belong to this category, because they describe concrete machine learning or data science process steps, e.g., the definition of inputs and outputs or engineering of corresponding features. “The Machine Learning Canvas [is] the first step towards making sure you connect what ML can do to your organization’s objectives, and towards assessing feasibility. It should be filled in before starting any implementation work, and even before Exploratory Data Analysis” [Do19]. This artifact is the most technical, as it contains fields like “ML task” or “Features”. Dorard also integrates metrics – offline and online – and value proposition into his canvas. However, it is supposed to be an initial document, and “the canvas results [have to be] translated to a technical specification document” [Ma19] later on. Schmarzo builds upon Dorard’s canvas and proposes two additional fields, namely Prescription and Automation in the “Big Data MBA Version”, to adopt the canvas to “data science requirements” [Sc18]. This extension shall form the path from a small ML use case to an integrated, scaled application. Nevertheless, it’s not obvious how to really do that. Schmarzo also proposes the “Hypothesis Development Canvas v1.1” in order to facilitate „collaboration between the business stakeholder and the data science team to identify the hypothesis requirements that underpin data science engagement success” [Sc18]. Another rather technical approach is the “ML Lifecycle Canvas”. It’s a “conceptual design tool featuring the holistic visualization of cooperation among ML, users, and scenarios during the ML lifecycle” [Zh20]. This canvas is unique in the regard,

that it provides a detailed question list and “persona cards” to fill in the canvas. Questions like “Is there any ML model feasible for completing the required tasks?” illustrate the level of detail regarding the final ML design process. The authors align the questionnaire with questions from “existing guidebooks on human-AI interaction” [Zh20], like Google PAIR [Go19] or [Am19]. The “Key Activity Canvas” [HST20] is a matrix arrangement and integrates three views, Customer, Company and Partners, for the “methodological assistance (key activities) guiding the actual conceptualization of necessary activities in analytics-based services”.

(AI) Project Focus. The second category includes canvases with a holistic view on AI projects, “AI Canvas” by [AGG18] and the eponymous by [DR18], “AI Project Canvas” [Za19], “Canvas for the use of AI” [KS20], “Enterprise AI Canvas” [Ke21], “AI performance canvas” [EL20], and “Digitalization Canvas” [He17]. In principle, these could synonymously be applied to ML projects. However, the distinction from the first category arises from the fact that a focus is placed on the overall project rather than on the technical details, e.g., the cost and revenue structure of an AI project. The “AI Canvas” [AGG18] is supposed to support corporate decision-making through prediction models. “How can you decide whether employing a prediction machine will improve matters? The AI Canvas is a simple tool that helps you organize what you need to know into seven categories [Prediction, Judgment, Action, Outcome, Input, Training, Feedback] in order to systematically make that assessment” [AGG18]. It’s rather designed “for a non-technical audience” [EL20]. Technical details and business integration are not being covered [Ke21]. This is also true for the “AI Canvas” by Dewalt and Rands. They try to connect a business opportunity via a strategy with a solution using AI models. It’s the pathway to “Become an AI Company in 90 Days” [DR18]. Similarly, based on the original Business Model Canvas the “AI Project Canvas” by Zawadzki is helpful for “project managers” [Ke21] and is intended to “structure and convey the holistic idea of your AI project to others” [Za19]. Another way to “determine the relevance of Artificial Intelligence for your company” [KS20] is the “Canvas for the use of AI”, which is embedded in a corporate transformation process towards AI. The approach by Kerzel “Enterprise AI Canvas” [Ke21] features two parts, one with a technical focus and one with a business focus. It’s supposed to “bring business and data science experts together and systematically evaluate potentially new business opportunities” [Ke21]. Although part 2 “model and data view” shows similarities to Dorard's approach, technical details on modeling are omitted. The AI Project Focus category also includes the “AI performance canvas”, whose main objective is the “collaborative construction of performance goals for data & AI products in organizations” [EL20]. In this approach, there is a strong focus on feasibility of the product and legal/compliance issues in data governance. Nevertheless, only a prototype canvas is being presented and the “trigger questions” of the fields are not yet published. The “Digitalization Canvas” [He17] is the only canvas specifically designed for SMEs and corresponds to a holistic digitalization strategy. The approach promotes easy-to-implement and strategic data projects. While the canvas itself isn’t self-sufficient, there are detailed questionnaires within the case study. “The Digitalization Canvas together with a project portfolio defines a concrete roadmap to digitalization and summarizes the arguments to apply for the necessary budget” [He17].

Data Value Focus. The third category are canvases with a data value or data product focus, namely the “Data Product Canvas” [FBP20], the “Data Value Map” [SN17], the “Data Canvas – Data-Need Fit” [MK16], and the “Data Innovation Board” [KM19]. These canvases seek to identify and implement the value of data through tangible customer or user benefits and information gain through data. The “Data Product Canvas” [FBP20] can be used when an organization in an initiation phase aims to develop ideas for a data product. Similarly, the “Data Canvas and Data-Need Fit are intended to spark a discussion on available data in organizations among diverse stakeholders. The Data Canvas provides trigger questions and a visual representation that help to develop a common understanding of available data” [MK16]. “A Data-Need Fit is found when data sources contribute gain creators and pain relievers that users find valuable” [MK16]. “To facilitate a shared understanding for data initiatives” [SN17] is also the main objective of the “Data Value Map”. This canvas focuses the process from data creator to data user and emphasizes data governance topics, such as data principles and access, and business related topics such as cost reduction or revenue generation. Finally, the “Data Innovation Board” [KM19] features three design thinking steps, namely Exploration, Ideation, and Evaluation, and promotes the description of performance goals in “a visual collaboration tool that anyone can work with”. The artifact is clearly supposed to facilitate initial progress in a data-driven project, and the authors note that “it is a beginner’s tool [, which needs] to be accompanied by other visual tools with more specific views on technology and algorithms” [KM19].

Data Source Focus. The fourth category includes canvases that focus on explaining the data source and the data processing, in order to gain a better understanding of data as an asset: the “Data Insight Generator” [BT19] and the “Data Collection Map” [KMK19]. The “Data Insight Generator” [BT19] is a workshop canvas that connects key data resources with a value proposition for data-driven business models. It contains columns for Pipes, Analytics and Insight. The process of filling in is guided by the rows Think, Validate, and Know, which are to be processed one after the other. Finally, the “Data Collection Map was designed as an entry point in the ideation process of data-driven use cases. Hence, the purpose of the tool is to get people to think about data (e.g. clicks and engagement metrics) instead of IT systems (e.g. Google Analytics) and to raise the necessary data awareness about the available data resources within the organization” [KMK19]. It’s basically an add-on to the “Data Innovation Board” by Kronsbein and Mueller (2019).

Categorization of the fields and questions. As it is not sufficient to only categorize the artifacts on a title level, the fields are being analyzed. In order to conceptualize the core content of the canvases and thus answer the second part of research question one, all fields (or headers in matrix patterns) are being captured and groups as well as top categories are being proposed. This should clarify where the canvases overlap and where the focus has been placed so far. In total there are 163 fields in the results, e.g., the “Machine Learning Canvas v0.4” [Do19] contains 10 fields that contain one or more questions to guide the filling (see Tab. 2). Logically, this count includes multiple entries. Therefore, the fields that cover a similar area in terms of content are grouped together. In the next step, multiple entries are eliminated and fields are combined that either describe exactly or almost the

same thing. For example, fields like “Data Sources”, “AI data base”, “Metadata”, and all the various data types from the “Data Collection Map” [KMK19], belong to the same group of “Data Sources” (see Tab. 3). As most fields contain more than one guiding question or example and not every field headline describes the same content, in order to refine the assignment, all guiding questions or examples are also individually examined. Ultimately, 39 groups and 11 top categories can be created. The content intersection of the final clusters compiles the name of their top category, e.g., the groups “Data Quality”, “Data policies”, and “Data lifecycle” build the top category “Data Governance”. The result of this assignment is presented in Tab. 3.

Cat.	Group	#	Example Question
Business & Value	Strategy	13	“What trends, market facts are relevant for the topic [...]?” [KM19]
	Risks	10	“What risks are associated with the use of AI for our industry?” [KS20]
	Operat. value	7	“How does the use-case generate value?” [Ke21]
	Revenue	6	“How will the project generate revenue?” [Za19]
	Cost	9	“Will the project reduce internal costs [...]?” [Za19]
Product & Customer	Product/Service	14	“Which potentials in the production area can be leveraged?” [KS20]
	Delivery	7	“In which form do we provide the data service to our users [...]?” [FBP20]
	Customer	16	“Who is our customer?” [FBP20]
	Gains/Value	9	“What is the customer value of the hypothesis?” [Sc18]
	Pains/Needs	9	“What customer pain is the AI project solving?” [Za19]
Organi- zation	Implementation	8	“How might we implement the idea?” [KM19]
	Internal skills	7	“Is the required know-how for the implementation available inhouse?” [He17]
	Stakeholders	3	“Sponsor: Which senior manager is responsible?” [Ke21]
	Domain	2	“Which domain expertise is needed?” [Ke21]
	Partners	8	“Which external services and products are required?” [He17]
Tech nol.	Systems	7	“Which systems are required and already available to handle data?” [Ke21]
	Infrastructure	7	“How are the models served? Edge, on-premise or Cloud?” [Ke21]
	Integration	4	“Which networks along the value chain are necessary?” [KS20]
Data charact.	Data types	25	“What kind of data do we need for training?” [Zh20]
	Data sources	9	“Which raw data sources can we use (internal and external)?” [Do19]
	Data availability	14	“What data is currently collected in the organization?” [KM19]
	Data collection	8	“How might we collect the needed data?” [KM19]
	Data pipelines	3	“Which interfaces can I use to combine this data?” [BT19]
Data gov.	Data quality	6	“How is the validity of the data, [...] consistency, and completeness?” [He17]
	Data policies	9	“Are there any compliance requirements [...]?” [He17]
	Data lifecycle	5	“Determining the definition, [...] and retirement of data.” [SN17]
Pre- proc.	Data preparation	4	“What do we have to do to prepare the data [...]?” [HST20]
	Features	2	“Which features are likely important?” [Ke21]
	Inputs	4	“What are the model inputs?” [DR18]
Modeling	Learning	6	“Is there any ML model suitable for the available dataset?” [Zh20]
	Analytics	5	“With which data analytics methods do we generate insights [...]?” [FBP20]
	Interpretation	4	“How can we interpret the mined patterns?” [HST20]
	Prediction	5	“What should be predicted?” [Ke21]
	Decision	9	“Prescription: Once we have a prediction, what do we do?” [Sc18]
Evaluation	KPI Model	3	“Which key metric are you optimizing for?” [Za19]
	KPI Business	6	“Outcome: What are your metrics for task success?” [AGG18]
	Improvements	3	“How can you use the outcomes to improve the algorithm?” [AGG18]
	Automation	6	“When do we create/update models with new training data?” [Do19]
	Live / Ex-post	6	“Methods and metrics to evaluate the system after deployment [...].” [Do19]

Tab. 3 Categories and groups of all canvas fields and questions with examples

4 Discussion and Conclusion

The primary contribution of this review is the categorization of existing canvases. The review generates overview for initial ML projects. It complements project management approaches in software development, such as the v-model or scrum. Using one of the categorized canvases is a solid starting point. Practitioners can use the four categories as guidance and pick one of the mentioned canvases, e.g., if they want to explore their data, the canvases from the category "Data sources" will help. Although "an over emphasis on technology" [SN17] has been mentioned as a potential barrier, we feel that diving deeper into the technical details, i.e., data processing and modeling, is key to foster ML adoption. Therefore, with regard to research question two, the canvases with a ML/DS Focus are most suitable. SMEs can use the canvases and/or the catalog of questions to promote cooperation, e.g., with research or consultancy. The canvases are all standalone artifacts for valid use cases. Nevertheless, there is still room for improvement for three reasons. The first reason is, that there are still gaps regarding content. E.g., no detailed questions for hyper parameter tuning, visualization of results, or concept drift [We19] could be identified. Also scalability and feasibility checks were not mentioned. The second reason regards the applicability of the canvases. Clear guidelines on who, when, and how to use the canvas are needed, describing the explicit benefit. Otherwise, their usefulness is mitigated and scientific artifacts will not be favored against gray literature. A balance between detailed description, e.g., as in Google PAIR [Go19], and self-sufficiency, will provide the greatest benefit. Eventually, the evaluation benchmark of the canvases must be: "What's the artifact from the artifact?" Meaning that using the canvas facilitated building an ML application that provides value. The third reason is concerning the fact that "canvases help us ask the right questions, but they don't provide the answers" [DR18]. In order to lift the canvas approach from ideation to application, providing answers to the questions is necessary to foster initial ML adoption. E.g., guiding the software tool chain or the model selection as in the scikit learn cheat sheet² might be fruitful. Other potential extensions of the canvas approach could be the integration of cloud service platforms, e.g., AWS, or other MLaaS providers [RGC15], or questions regarding "ground truth" [SB14]. Future research may include three key points. First, the question list has to be compared to a) the challenges of ML adoption in SMEs and b) to existing process descriptions, such as CRISP-DM [WH00]. Second, a comprehensive canvas for initial ML projects can be conceptualized from the findings of this review. However, the level of detail of this canvas shouldn't be sacrificed for generalizability, which could be prevented by a layered approach. And third, this concept can then be used and evaluated in empirical research, especially in workshops and case studies. The conceptualization of a new artifact and its evaluation would also address the inherent limitations of this review paper, as personal bias and experience could not totally be omitted, especially in the categorization parts. Qualitative research through interviews regarding ideation, communication, and problem-solving could build the basis to assess the added and perceived value of the canvases.

² https://scikit-learn.org/stable/tutorial/machine_learning_map/index.html

Bibliography

- [AGG18] Agrawal, A.; Goldfarb, A.; Gans, J.: A Simple Tool to Start Making Decisions with the Help of AI. <https://hbr.org/2018/04/a-simple-tool-to-start-making-decisions-with-the-help-of-ai>, accessed 16 Apr 2021.
- [Am19] Amershi, S. et al.: Guidelines for Human-AI Interaction. In (Brewster, S. et al. Eds.): CHI 2019. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, May 4-9, 2019, Glasgow, Scotland UK. ACM, New York, NY, pp. 1–13, 2019.
- [BF18] Becker, W.; Feichtinger, C.: Data Analytics in jungen Unternehmen. Ergebnisse einer Online-Befragung. *Der Betriebswirt* 2/59, pp. 26–31, 2018.
- [Bo19] Bourcetet, A. et al.: Guided Machine Learning for Business Users. BLED 2019 Proceedings, 2019.
- [Br09] Brocke, J. vom et al.: Reconstructing the giant: on the importance of rigour in documenting the literature search process: Proceedings of the 17th European Conference on Information Systems, 2009.
- [BT19] Babett Kühne; Tilo Böhmman: Data-Driven Business Models - Building the Bridge between Data and Value. Twenty-Seventh European Conference on Information Systems (ECIS2019), Stockholm-Uppsala, Sweden., 2019.
- [BvK20] Bauer, M.; van Dinther, C.; Kiefer, D.: Machine Learning in SME: An Empirical Study on Enablers and Success Factors. AMCIS 2020 Proceedings, 2020.
- [Co16] Coleman, S. et al.: How Can SMEs Benefit from Big Data? Challenges and a Path Forward. *Quality and Reliability Engineering International* 6/32, pp. 2151–2164, 2016.
- [Do19] Dorard, L.: The Machine Learning Canvas. A handbook for innovators and visionary managers striving to design tomorrow's Machine Learning systems. DRAFT version 0.1 released on 12 February 2019, 2019.
- [DR18] Dewalt, K.; Rands, R.: Become an AI Company in 90 Days: The No-Bullshit Guide for Understanding AI, Identifying Opportunities, and Launching Your First Product. Prolego, 2018.
- [EL20] Engel, M.; Lang, F.: A Pilot Study on Designing a Data & AI Performance Canvas. AMCIS 2020 Proceedings, 2020.
- [FBP20] Fruhwirth, M.; Breidfuss, G.; Pammer-Schindler, V.: The Data Product Canvas - A Visual Collaborative Tool for Designing Data-Driven Business Models. BLED 2020 Proceedings, 2020.
- [FRL17] Funk, B.; Rettenmeier, M.; Lang, T.: Deep Learning auf sequenziellen Daten als Grundlage unternehmerischer Entscheidungen. *Wirtschaftsinformatik & Management* 5/9, pp. 16–25, 2017.
- [Gh11] Ghobakhloo, M. et al.: Information Technology Adoption in Small and Medium-sized Enterprises: An Appraisal of Two Decades Literature. *Interdisciplinary Journal of Research Business* 7/1, pp. 53–80, 2011.
- [Go19] Google PAIR: People + AI Guidebook. <https://pair.withgoogle.com/guidebook/>,

accessed 16 Apr 2021.

- [Ha14] Hartmann, P. et al.: Big Data for Big Business? A Taxonomy of Data-driven Business Models used by Start-up Firms. Cambridge Service Alliance, 2014.
- [He17] Heberle, A. et al.: Digitalization Canvas - Towards Identifying Digitalization Use Cases and Projects. *Journal of Universal Computer Science* 11/23, pp. 1070–1097, 2017.
- [HS15] Hennig-Thurau, P. D. T.; Sattler, P. D. H.: VHB-JOURQUAL3. <https://vhbonline.org/vhb4you/vhb-jourqual/vhb-jourqual-3/tabellen-zum-download>.
- [HST20] Hunke, F.; Seebacher, S.; Thomsen, H.: Please Tell Me What to Do – Towards a Guided Orchestration of Key Activities in Data-Rich Service Systems. In (Hofmann, S.; Müller, O.; Rossi, M. Eds.): *Designing for Digital Transformation. Co-Creating Services with Citizens and Industry. 15th International Conference on Design Science Research in Information Systems and Technology, DESRIST 2020, Kristiansand, Norway, December 2–4, 2020, Proceedings*. Springer International Publishing; Imprint: Springer, Cham, pp. 426–437, 2020.
- [ID20] IDG Business Media GmbH: Studie Machine Learning 2020. <https://www.lufthansa-industry-solutions.com/de-de/newsroom-downloads/news/maschinelles-lernen-wird-in-deutschen-unternehmen-standard/>, accessed 4 Feb 2021.
- [Ke21] Kerzel, U.: Enterprise AI Canvas Integrating Artificial Intelligence into Business. *Applied Artificial Intelligence* 1/35, pp. 1–12, 2021.
- [Kf17] KfW Bankengruppe: Digitalisierung der Wirtschaft: breite Basis, vielfältige Hemmnisse. Unternehmensbefragung. <https://www.kfw.de/PDF/Download-Center/Konzernthemen/Research/PDF-Dokumente-Unternehmensbefragung/Unternehmensbefragung-2017-%E2%80%93-Digitalisierung.pdf>, accessed 7 May 2021.
- [Kf20] KfW Bankengruppe: Unternehmensbefragung 2020. Anteil der Digitalisierungsplaner stagniert auf hohem Niveau. <https://www.kfw.de/PDF/Download-Center/Konzernthemen/Research/PDF-Dokumente-Unternehmensbefragung/Unternehmensbefragung-2020-%E2%80%93-Digitalisierung.pdf>, accessed 7 May 2021.
- [Ki04] Kitchenham, B.: Procedures for Performing Systematic Reviews. Keele University Technical Report TR/SE-0401, 2004.
- [KJZ16] Kehrer, S.; Jugel, D.; Zimmermann, A.: A systematic literature review of big data literature for EA evolution. *Digital Enterprise Computing 2016, Lecture Notes in Informatics (LNI)*, 2016.
- [KM19] Kronsbein, T.; Mueller, R.: Data Thinking: A Canvas for Data-Driven Ideation Workshops. *Hawaii International Conference on System Sciences 2019 (HICSS-52)*, 2019.
- [KMK19] Kayser, L.; Mueller, R.; Kronsbein, T.: Data Collection Map: A Canvas for Shared Data Awareness in Data-Driven Innovation Projects. *Proceedings of the 2019 Pre-*

ICIS SIGDSA Symposium, 2019.




- [KPD18] Kollwitz, C.; Perez Mengual, M.; Dinter, B.: Cross-Disciplinary Collaboration for Designing Data-Driven Products and Services. Pre-ICIS SIGDSA Symposium on Decision Analytics Connecting People, Data & Things, San Francisco 2018, 2018.
- [KS20] Kreutzer, R. T.; Sirrenberg, M.: AI Challenge - How Artificial Intelligence Can Be Anchored in a Company. In (Kreutzer, R. T.; Sirrenberg, M. Eds.): Understanding Artificial Intelligence. Fundamentals, Use Cases and Methods for a Corporate AI Journey. Springer International Publishing; Imprint: Springer, Cham, pp. 235–273, 2020.
- [Ku17] Kuschicke, F. et al.: A Data-based Method for Industrial Big Data Project Prioritization. Proceedings of the International Conference on Big Data and Internet of Thing, pp. 6–10, 2017.
- [Ma10] Mayring, P.: Qualitative Inhaltsanalyse. In (Mey, G.; Mruck, K. Eds.): Handbuch qualitative Forschung in der Psychologie. VS Verlag, Wiesbaden, pp. 601–613, 2010.
- [Ma19] Marin, I.: Data Science and Development Team Remote Communication: the use of the Machine Learning Canvas: 2019 ACM/IEEE 14th International Conference on Global Software Engineering (ICGSE 2019). Montreal, Quebec, Canada, 25-26 May 2019. IEEE, Piscataway, NJ, pp. 18–21, 2019.
- [Mi20a] Mittelstand 4.0-Kompetenzzentrum Dortmund: Digital in NRW - Kompetenz für den Mittelstand. Künstliche Intelligenz im Mittelstand - Potenziale und Anwendungsbeispiele, 2020.
- [Mi20b] Ministerium für Wirtschaft, Arbeit und Wohnungsbau Baden-Württemberg: MONITORING-REPORT. Wirtschaft DIGITAL Baden-Württemberg 2020. https://www.wirtschaft-digital-bw.de/fileadmin/media/Dokumente/Studien/Monitoring_BaWue_2020_Langfassung.pdf.
- [MK16] Mathis, K.; Köbler, F.: Data-Need Fit – Towards Data-Driven Business Model Innovation. ServDes. 2016, Fifth Service Design and Innovation conference, 2016.
- [Mo17] Moeuf, A., Lamouri, S., Pellerin, R., Eburdy, R., Tamayo, S.: Industry 4.0 and the SME: a technology-focused review of the empirical literature. Proceedings of 7th International Conference on Industrial Engineering and Systems Management IESM, 2017.
- [Mo18] Mohd Selamat, S. A. et al.: Big data analytics-A review of data-mining models for small and medium enterprises in the transportation sector. WIREs (Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery) 3, 2018.
- [OP10] Osterwalder, A.; Pigneur, Y.: Business model generation. A handbook for visionaries, game changers, and challengers. Wiley&Sons, New York, 2010.
- [Qi18] Qian Yang et al.: Grounding Interactive Machine Learning Tool Design in How Non-Experts Actually Build Models. undefined, 2018.
- [RGC15] Ribeiro, M.; Grolinger, K.; Capretz, M. A.: MLaaS: Machine Learning as a Service: 2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA). IEEE, pp. 896–902, 2015 - 2015.

- [Ru15] Russegger, S. et al.: Big Data und Data-driven Business für KMU. Digital networked Data, 2015.
- [SB14] Shalev-Shwartz, S.; Ben-David, S.: Understanding machine learning. From theory to algorithms. Cambridge University Press, Cambridge, New York, NY, Port Melbourne, Delhi, Singapore, 2014.
- [Sc15] Schryen, G.: Writing Qualitative IS Literature Reviews—Guidelines for Synthesis, Interpretation, and Guidance of Research. Communications of the Association for Information Systems 1/37, 2015.
- [Sc16] Schmarzo, B.: Big data MBA. Driving business strategies with data science. Wiley, Indianapolis, Indiana, 2016.
- [Sc18] Schmarzo, B.: Data Science “Paint by the Numbers” with the Hypothesis Development Canvas. <https://www.linkedin.com/pulse/data-science-paint-numbers-hypothesis-development-canvas-schmarzo/>, accessed 27 Apr 2021.
- [Sc20] Schäfer, F. et al.: Smart Use Case Picking with DUCAR: A Hands-On Approach for a Successful Integration of Machine Learning in Production Processes. Procedia Manufacturing 51, pp. 1311–1318, 2020.
- [SG89] Star, S. L.; Griesemer, J.: Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. Social Studies of Science 3/19, pp. 387–420, 1989.
- [SN17] Sammon, D.; Nagle, T.: The Data Value Map: A framework for developing shared understanding on data initiatives. ECIS 2017: 25th European Conference on Information Systems, pp. 1439–1452, 2017.
- [VD18] VDMA Software und Digitalisierung Ed.: Quick Guide Machine Learning im Maschinen- und Anlagenbau, Frankfurt, 2018.
- [VHR20] Vidgen, R.; Hindle, G.; Randolph, I.: Exploring the ethical implications of business analytics with a business ethics canvas. European Journal of Operational Research 3/281, pp. 491–501, 2020.
- [VLF15] Vossen, G.; Lechtenböcker, J.; Fekete, D.: Big Data in kleinen und mittleren Unternehmen: Eine empirische Bestandsaufnahme (Nr. 135). Arbeitsberichte des Instituts für Wirtschaftsinformatik, Westfälische Wilhelms-Universität Münster, 2015.
- [We19] Weber, C. et al.: A New Process Model for the Comprehensive Management of Machine Learning Models. Proceedings of the 21st International Conference on Enterprise Information Systems ICEIS 2019, pp. 415–422, 2019.
- [WG18] Wolff, D.; Göbel, R.: Digitalisierung. Wie Die Digitalisierung Unsere Lebens- und Arbeitswelt Verändert. Segen oder Fluch. Springer, Berlin, Heidelberg, 2018.
- [WG19] WGP Wissenschaftliche Gesellschaft für Produktionstechnik e.V.: WGP-Standpunkt KI in der Produktion. Künstliche Intelligenz erschliessen für Unternehmen, accessed 7 May 2021.
- [WH00] Wirth, R.; Hipp, J.: CRISP-DM: Towards a standard process model for data mining. Proceedings of the Fourth International Conference on the Practical

Application of Knowledge Discovery and Data Mining, 2000.

- [WW02] Webster, J.; Watson, R. T.: Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly* 2/26, pp. 13–23, 2002.
- [Za19] Zawadzki, J.: Introducing the AI Project Canvas - Towards Data Science. <https://towardsdatascience.com/introducing-the-ai-project-canvas-e88e29eb7024>, accessed 15 Apr 2021.
- [Zh19] Zhou, Z. et al.: ML-Process Canvas: A Design Tool to Support the UX Design of Machine Learning-Empowered Products. CHI'19 Extended Abstracts, 2019, Glasgow, Scotland UK, 2019.
- [Zh20] Zhou, Z. et al.: ML Lifecycle Canvas: Designing Machine Learning-Empowered UX with Material Lifecycle Thinking. *Human-Computer Interaction* 5-6/35, pp. 362–386, 2020.

Reifegradmodelle zur Bewertung Künstlicher Intelligenz in kleinen und mittleren Unternehmen

Thomas Schuster ¹, Lukas Waidelich ¹ und Raphael Volz ¹

Abstract: Die digitale Transformation stellt Unternehmen vor große Herausforderungen und begünstigt disruptive Veränderungsprozesse. Als Schlüsseltechnologie nehmen Anwendungen der Künstlichen Intelligenz (KI) eine übergeordnete Rolle ein. Führungskräfte müssen daher in der Lage sein, Notwendigkeit und Fortschritt des Einsatzes von KI im Unternehmen regelmäßig bewerten zu können. Hierfür eignen sich sogenannte Reifegradmodelle (RM), die eine Einordnung der Anwendungen und Produkte im Kontext der KI ermöglichen. Dieser Beitrag verfolgt das Ziel den Status Quo der aktuellen Forschung bestehender KI-RM aufzuzeigen und einzuordnen. Gleichzeitig sollen notwendige zusätzliche Perspektiven eines mittelstandsgerechten KI-RM identifiziert werden und als Grundstein für weitere Forschungsaktivitäten dienen.


Keywords: Künstliche Intelligenz, KI, Reifegrad, Reifegradmodell, KMU

1 Einführung

Verschiedene Faktoren haben die rasante Weiterentwicklung der künstlichen Intelligenz (KI) begünstigt und deren kommerzielle Anwendung beflügelt. Leistungsfähige KI Lösungen kommen in immer mehr Anwendungsbereichen zum Einsatz und erschließen damit große Potenziale. KI Anwendungen beinhalten die inhärente Fähigkeit eine entscheidende Rolle bei der Wettbewerbsdifferenzierung einzunehmen [KW19]. Studien zeigen, dass KI als wichtigstes Handlungsfeld für unternehmerischen Erfolg angesehen wird. Dem gegenüber steht ein inhomogenes Vorgehen bei der konstruktiven Einführung von KI in Unternehmen. Speziell kleine und mittlere Unternehmen (KMU) stehen vor großen Herausforderungen [Bk21]. Unsere Arbeit zielt darauf ab, die derzeit konfuse Lage bestehender KI-Reifegradmodelle (KI-RM) zu entzerren und eine Lösung zu präsentieren, sodass Unternehmen eine verlässliche Einordnung ihres aktuellen KI-Entwicklungsstands vornehmen können. Unsere Forschungsfragen für dieses Übersichtspapier sind daher:

- Welche KI-RM existieren und welche Schwerpunkte setzen diese?
- Wie helfen KI-RM den aktuellen Stand im eigenen Umfeld zu identifizieren?
- Wie helfen KI-RM nächste Schritte zur KI-Entwicklung zu bestimmen?

¹Hochschule Pforzheim, Tiefenbronner Straße 65, 75175 Pforzheim,
{thomas.schuster, lukas.waidelich, raphael.volz}@hs-pforzheim.de,

 <https://orcid.org/0000-0002-9539-1627>, <https://orcid.org/0000-0003-2602-0942>,
<https://orcid.org/0000-0003-2816-1618>

Zunächst werden Grundlagen von KI als auch von RM erläutert, um den Rahmen für die Analyse und den Entwicklungsansatz zu setzen. Danach eine Methodik als mehrstufiges Vorgehensmodells skizziert und anschließend die Umsetzung der Methodik, im Detail beschrieben. Dies beinhaltet die Problemdefinition, den Vergleich bestehender KI-RM und die Festlegung einer Entwicklungsstrategie. Schließlich werden ein Fazit und Rück-schlüsse auf weitere, notwendige Forschungsaktivitäten beschrieben.

2 Grundlagen

2.1 Künstliche Intelligenz

Das Konzept der Künstliche Intelligenz (KI) existiert seit langem. Zahlreiche KI-Initiativen bildeten bereits 1950er- und 1960er-Jahren eine erste Hochphase. Fehlender Fortschritt leiteten in den 1970er-Jahren den KI-Winter ein, der mit Enttäuschungen und Desillusionierungen verbunden war [FSX18]. In den vergangenen Jahren wird die Diskussion um KI besonders durch die fortgeschrittene Leistungsfähigkeit der Hardware und die Kombination unterschiedlicher Verfahren wieder verstärkt diskutiert und führt zur aktuellen Blüte der Disziplin. Damit einhergehend sind zahlreiche neue Forschungsgebiete entstanden und KI wird inzwischen in fast allen Branchen eingesetzt. KI-Anwendungen kommen bspw. bei der Krebsfrüherkennung, dem autonomen Fahren oder bei Chatbots zum Einsatz. Andererseits finden auch Debatten über Risiken statt; dazu zählen Arbeitsplatzverlust, fehleranfällige Algorithmen oder Verlust von Privatsphäre [Br05, Ca20].

Die KI-Forschung hat ebenfalls eine Evolution durchlaufen und umfasst heute zahlreiche Abstufungen von schwacher bis starker KI und zahlreiche Disziplinen wie Maschinelles Lernen oder Deep Learning [Ri19]. Bei der schwachen KI übernimmt die KI eine unterstützende Tätigkeit, während bei einer starken KI selbstständiges kognitives Verhalten ohne menschliches Zutun vorausgesetzt wird. Die höchste Form der KI ist die künstliche Superintelligenz, die aktuell nur in der Theorie möglich erscheint [Ri19].

2.2 Reifegradmodelle

Reifegradmodelle (RM) zielten zunächst darauf ab Organisationen bei der Verbesserung ihrer Software-Prozesse zu unterstützen [Pa09]. Inzwischen haben sich RM domänenübergreifend etabliert und finden in vielen Bereichen Anwendung, Beispiele sind Informationssysteme (IS), Entwicklung, Nachhaltigkeit, Gesundheit oder Geschäftsprozessmanagement, digitale Transformation und Industrie 4.0 [Da20]. Allgemein werden RM in innovativen, neuartigen und komplexen Fachgebieten eingesetzt, die einen systematischen und strukturierten Ansatz erfordern [Ya20]. Spezifische KI-RM werden vornehmlich von Personen aus der Praxis und nicht aus dem wissenschaftlichen Umfeld definiert.

RM dienen als Messinstrument zur Bewertung der Fähigkeiten einer Organisation [De05], schaffen initiales Bewusstsein für eine Thematik, geben eine valide Standortbestimmung durch Bewertung der eigenen Fähigkeiten wieder und können Innovationsprozesse anstoßen [We12]. RM können deskriptive, präskriptive oder vergleichende Ausprägung enthalten. Ein deskriptives RM wird zur Ist-Analyse genutzt und untersucht bestehende Ausprägungen nach vorgegebenen Kriterien. Bei der Zielerreichung eines bestehenden Reifegrads oder Vorgabe entsprechender Verbesserungsmaßnahmen handelt es sich um ein präskriptives RM. Zuletzt ermöglicht ein vergleichendes RM ein Benchmark zwischen internen oder externen Entitäten. RM können die genannten Ausprägungen auch vereinen.

3 Forschungsmethodik

Als Forschungsmethode kommt der Design Science Research (DSR) Ansatz nach Hevner et al. [He04] zum Einsatz. Das methodische Vorgehen bildet die Basis für die Konzeption eines Design-Artefakts (DA) und der Entwicklung des fundierten KI-RM. Unsere Forschungsarbeit orientiert sich an den sieben Schritten, die Becker, Knackstedt und Pöppelbuß [BPP09] (siehe Abbildung 1) zur Entwicklung von Reifegradmodellen vorschlagen und zielt darauf ab ein neues DA in Form eines KI-RM zu entwickeln.

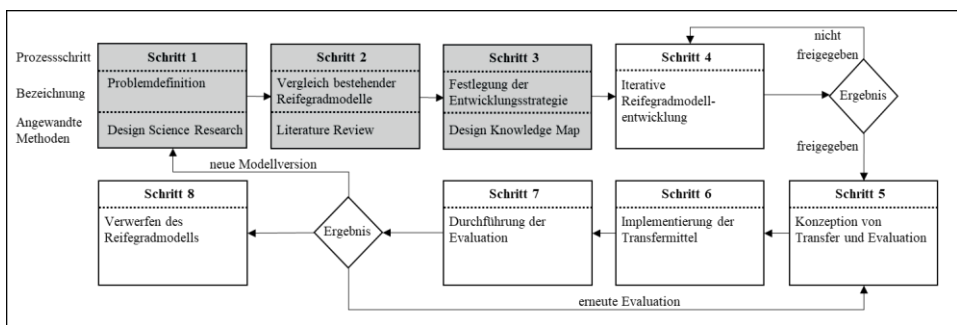


Abb. 1: Vorgehensmodell zur Entwicklung eines Reifegradmodells angelehnt an [BPP09]

Dieses Papier behandelt drei der sieben vorgegebenen Schritte des RM-Vorgehensmodells: (1) *Problemdefinition*, (2) *Vergleich bestehender RM*, (3) *Festlegung der Entwicklungsstrategie*. Die weiteren Schritte (4) *iterative RM-Entwicklung*, (5) *Konzeption von Transfer- und Evaluation*, (6) *Implementierung der Transfermittel* und (7) *Durchführung der Evaluation* befinden sich aktuell in Durchführung und sollen in zukünftige Forschungsarbeiten artikuliert werden. Abbildung 1 zeigt die notwendigen Prozessschritte für das Design-Artefakt und benennt die angewendeten Methoden je Prozessschritt.

3.1 Problemdefinition

Zu Beginn der RM-Entwicklung steht die initiale Definition der Problemstellung [BPP09]. Ein essenzielles DSR-Ziel ist Wissen und Verständnis zu erlangen, die die Entwicklung

und Implementierung von technologiebasierten Lösungen für bisher ungelöste Geschäftsprobleme ermöglichen [He04]. Ein Problem soll als Differenz zwischen dem Ist-Zustand und dem Soll-Zustand z.B. im Kontext eines IS verstanden werden.

3.2 Vergleich bestehender Reifegradmodelle

Bevor ein Vergleich bestehender KI-RM durchgeführt werden konnte, müssen diese zuerst im Rahmen des zweiten Prozessschritts [BPP09] (siehe Abbildung 1) identifiziert werden. Dafür wurde eine systematische Literaturrecherche durchgeführt die sich im Wesentlichen an die Methodik von Brocke et al. [Br09] orientiert und fünf Phasen umfasst: 1) die Definition der grundlegenden Forschungsparameter [Co88], (2) die Definition der Suchbegriffe und Suchbegriffskombinationen für die Datenbankrecherche, (3) die Auswahl der durchsuchten Datenbanken, (4) die Analyse der Suchergebnisse und (5) die Rückwärtssuche nach weiteren Informationsquellen. Die grundlegenden Forschungsparameter sind in tabellarischer Form zusammengefasst:

Fokus	Die Schwerpunkte liegen auf Forschungsergebnissen und praxisorientierten Anwendungsfällen von KI-RM
Ziel	Identifikation von bestehenden, wissenschaftliche fundierten und anwendungsorientierten KI-RM Modelle für Unternehmen
Perspektive	Die Perspektive der Recherche ist neutral.
Umfang	Selektive Literaturrecherche mit fokussiertem Untersuchungsgebiet
Organisation	Konzeptionell ausgerichtet, Mustererkennung in Bezug auf KI-RM
Zielgruppe	Die Recherche adressiert (IS-)Forscher und Anwender für KI-RM

Tab. 1: Definition der Rechercheparameter mittels Taxonomie nach Cooper [Co88]

Die Definition der Suchbegriffe für die Recherche umfassen die Begrifflichkeiten *Artificial Intelligence* und *Maturity Model* oder *Readiness Model* oder *Capability Model* sowie deren Kombination. Die Auswahl der Datenbanken für Konferenz- und Journalpapiere erfolgte durch Anfragen an Google Scholar, SpringerLink und ACM, wobei besonders IS-Konferenzen und Journale mit Peer-Review Verfahren betrachtet wurden, die nachfolgend um technische KI-Berichte erweitert wurden. Weiterhin wurden die ersten 50 Ergebnisse je Datenbank bzw. Suchmaschine der letzten zehn Jahre untersucht, nach Titel und Abstract analysiert, Duplikate und nicht relevante Ergebnisse exkludiert. Ferner wurde eine Vorwärts- und Rückwärtssuche relevanter Werke vorgenommen und der Stand um die neuen Werke erweitert. Insgesamt konnten 15 KI-RM Ansätze gefunden werden, die Artefakte von RM aufweisen. Die Ergebnisse werden in Abschnitt vier präsentiert.

3.3 Festlegung der Entwicklungsstrategie

Das gewählte Vorgehensmodell [BPP09] sieht nach der Analyse bestehender Modelle eine strategische Grundsatzentscheidung vor. Die Optionen umfassen den Entwurf eines neuen Modells, Verbesserung eines bestehenden Modells, Kombination mehrerer bestehender

Modelle zu einem neuen Modell, Übertragung von Strukturen oder Anwendung von Inhalten aus bestehenden Modellen auf neue Domänen [BPP09]. Bei einer Neuentwicklung eines KI-RM sollen Stärken wie Inhalt und/oder Struktur bestehender Modelle berücksichtigt und Schwachstellen gezielt angegangen werden [De05].

Als Entscheidungsgrundlage sollen die identifizierten KI-RM in die sogenannte Design Knowledge Map (DSM), welche der DSR-Domäne zuzuordnen ist, eingruppiert werden. Jedes DSR Projekt hat ein Startpunkt, der auf bestehendem Design Knowledge (DK) aufbaut. Die DSM ermöglicht die Einordnung des bestehenden Wissens anhand eines Neun-Felder-Portfolio der analysierten KI-RM nach den Kriterien Projektierbarkeit (in den Problemraum), Tauglichkeit (als Lösungsraum) und Zuversicht (in die Evaluation) gemäß den DSR Grundsätzen [Br20]. Die Kriterien können wiederum verschiedene Ausprägungen (niedrig, mittel, hoch) aufweisen.

4 Ergebnisse identifizierter KI-Reifegradmodelle im Vergleich

4.1 Problemdefinition

Der Einfluss von KI als wirtschaftlicher Erfolgsfaktor ist anerkannt [KW19, Wo21]. Laut einer aktuellen Studie sehen mehr als zwei Drittel der deutschen Unternehmen KI als die wichtigste Zukunftstechnologie, acht von 100 Unternehmen setzen bereits auf KI-Lösungen und etwa ein Drittel plant einen zeitnahen Einsatz [Bk21]. Dessen ungeachtet befindet sich die Kommerzialisierung von KI in einer frühen Phase [BGV21]. Besonders KMU sind beim Einsatz von KI zurückhaltend [Bk21]. Die Einführung von KI-Anwendungen wird in Organisationen als Herausforderung angesehen. Dieser Umstand kann darauf zurückgeführt werden, dass KI vielfach noch nicht umfassend in die Geschäftsstrategie eingebunden ist. Zugleich entsteht unternehmensintern oftmals Lücken zwischen Organisationseinheiten, die KI bereits verstehen und anwenden und solchen, die dies noch nicht tun. Derzeit nehmen KI-Lösungen eine Schlüsselrolle bei der kontinuierlichen Verbesserung von Produkten, Geschäftsprozessen und -modellen ein und leisten damit einen wertvollen Beitrag zur Unternehmensentwicklung [De05, Ya20].

Ein KI-RM dient Unternehmen zur Standortbestimmung. Führungskräfte können sich einen Überblick des KI-Fortschritts verschaffen, somit die Leistungsfähigkeit der Organisation kritisch betrachten und schlussendlich optimierende Maßnahmen ergreifen. Hierzu muss der Reifegrad bestimmt werden [ACM19]. Mit unserem KI-RM zielen wir darauf ab Unternehmen dazu zu befähigen die KI-Kommerzialisierung voranzutreiben.

4.2 Vergleich bestehender Reifegradmodelle

Im Rahmen der Recherche konnten 15 verschiedene KI-RM Ansätze identifiziert werden.

Der Schwerpunkt liegt auf den drei vollständigen und auf wissenschaftlicher Basis entwickelter KI-RM. Diese werden in Tabelle 2 gegenübergestellt. Zudem existieren wissenschaftliche Artefakte, die Elemente eines KI-RM beinhalten, die ebenfalls untersucht werden sollen. Ferner brachte die Suche zahlreiche KI-RM hervor, die auf keiner empirischen Basis stehen, eine mangelnde Dokumentation aufweisen und eher als Beratungsangebote von KI-Lösungen von Unternehmen verstanden werden können.

Autor/en	Alsheibani et. al [ACM19]	Yams et. al [Ya20]	Holmstrom [Ho21]
RM-Name	Artificial Intelligence Maturity Model	AI Innovation Maturity Index	AI Readiness Framework
Fokus	KI allgemein	KI allgemein Innovationsmanagement	KI allgemein
Dimensionen	KI-Funktionalitäten Datenstruktur Personal Organisation	Strategie Ökosystem Mindset Organisation Technologie Daten	Technologie Aktivitäten Grenzen Ziele
Reifegradphasen	initial prüfend bestimmt verwaltet optimiert	grundlegend experimentierend operativ forschend integriert	kein gering moderat hoch exzellent
Zweck	Identifikation und Basis für Verbesserung KI-Status Quo im Unternehmen	Syst. Unterstützung von KI-Integration in Innovationssystemen	Unterstützung bei der Entwicklung von KI-Geschäftsideen

Tab. 2: Vergleich der identifizierten KI-RM aus der Literatur

Die drei ganzheitlich und wissenschaftlich fundierten KI-RM sind innerhalb der vergangenen drei Jahre publiziert worden und nehmen einen eher allgemeingültigen KI-Fokus ein. Die KI-RM umfassen sowohl unterschiedlich ausgeprägte Dimensionen als auch Reifegradphasen, welche per Definition vorhanden sein müssen, um als RM klassifiziert zu werden.

Weiterhin konnten auch Werke identifiziert werden, die KI-RM Artefakte enthalten. Diese dienen entweder als inhaltliche Vorarbeit oder vereinen essenzielle Elemente eines KI-RM. Die Arbeit von Vakkuri et al. sieht konkreten Forschungsbedarf und thematisiert den Bedarf eines etablierten KI-RM und gibt Impulse ethische Dimensionen zu berücksichtigen [Va21]. Daneben beschreiben drei gefundene Publikationen mannigfaltige Dimensionen für potenzielle KI-RM, die teilweise von empirischen Studien hergeleitet wurden (siehe Tabelle 3) [JWW21, MG21]. Weitere Veröffentlichungen behandeln zumeist fünfstufige KI-Reifegradphasen, ohne dabei auf ein konkretes RM zu verweisen oder konkrete KI-Dimensionen zu nennen [EI19, NSP19, Ha20].

	Gentsch [Ge19]	Mikalef und Gupta [MG21]	Jöhnk et al. [JWW21]
Dimensionen	Strategie Organisation Entscheidungen Daten Analyse	Daten Technologie Basis-Ressourcen Technische Fähigkeiten Geschäftliche Fähigkeiten	Strategie Ressourcen Wissen Kultur Daten

Tab. 3: Vergleich der KI-Dimensionen aus Vorarbeiten (KI-RM Artefakte)

Zuletzt konnte durch die Suche weitere KI-RM in Form von Berichten gefunden werden, die zwar formal die Anforderungen an ein KI-RM erfüllen jedoch nicht im Detail betrachtet werden. Die Ansätze von Gartner, Intel und Ovum stehen entweder nicht auf einer wissenschaftlichen Basis oder weisen wissenschaftliche Mängel auf (mangelhafte Dokumentation und nicht reproduzierbares Vorgehen). Eine weitere KI-Studie bezieht sich auf den Reifegrad von Verwaltungen auf internationaler Ebene und soll daher nicht analysiert werden [SSP20]. Dieser Vergleich dient als Basis für weitere KI-RM Entwicklung und ist gleichzeitig Voraussetzung für die Festlegung der Entwicklungsstrategie.

Viele RM versuchen den komplexen Sachverhalt anhand einiger weniger KI-Dimensionen abzubilden (siehe Tabelle 2 und 3), greifen die Thematik damit eher generalistisch auf und stiften durch die oberflächliche Betrachtung nur einen geringen Mehrwert. Die vollständigen KI-RM nehmen kaum eine Differenzierung nach Kriterien wie Branchenzugehörigkeit oder Unternehmensgröße vor. Sogenannte Key Performance Indicators (KPI) zur Einordnung in die Reifegradphasen der jeweiligen KI-Dimensionen konnten ebenfalls nicht identifiziert werden. Ein in der Praxis bewährtes KI-RM, das bereits eine Evaluationsphase erfolgreich durchlaufen hat, konnten wir ebenfalls nicht identifizieren. Ferner vermischen wir für die jeweiligen Reifegradphasen Handlungsmaßnahmen, also konkrete Schritte wie sich die bewertete Organisation kontinuierlich verbessern kann, um das nächste Reifegradlevel erreichen zu können. Schließlich sollte auch eine Aufwand-Nutzen Abwägung für jede KI-Maßnahme im Unternehmen getroffen werden. Nicht jede Erreichung des nächsthöheren KI-Reifegrades steht in einem betriebswirtschaftlichen Verhältnis, sollte also sorgfältig überprüft und im KI-RM berücksichtigt werden.

4.3 Festlegung der Entwicklungsstrategie

Die drei identifizierten KI-RM aus Tabelle 2 werden anhand der drei Kriterien Projektierbarkeit des Problemraums, Tauglichkeit des Lösungsraums und Zuversicht in die Evaluation in die DKM eingeordnet. Das Ergebnis ist in Abbildung 2 dargestellt. Wie bereits angemerkt ist der DSR Bereich KI-RM noch neu, kann jedoch schon auf grundlegende DK aufbauen. Die beiden Publikationen der Autoren Alsheibani, Cheung, Messom [ACM19] und Holmstrom [Ho21] weisen das angesprochene Grundlagenwissen im Bereich Problem- und Lösungsraum auf. Die Zuversicht in die Evaluation muss durch die Praxisvalidierung erst noch erbracht werden, jedoch finden sich vergleichbare Ansätze im Bereich der KI-RM Artefakte. Die Publikation von Yams et. al [Ya20] weist unseres Erachtens eine höhere Projektierbarkeit des Problemraums auf, da diese einen spezifischeren

Fokus setzt und aussagekräftigere Dimensionen anbietet. Damit einher geht aber eine geringere Zuversicht in die Evaluation.

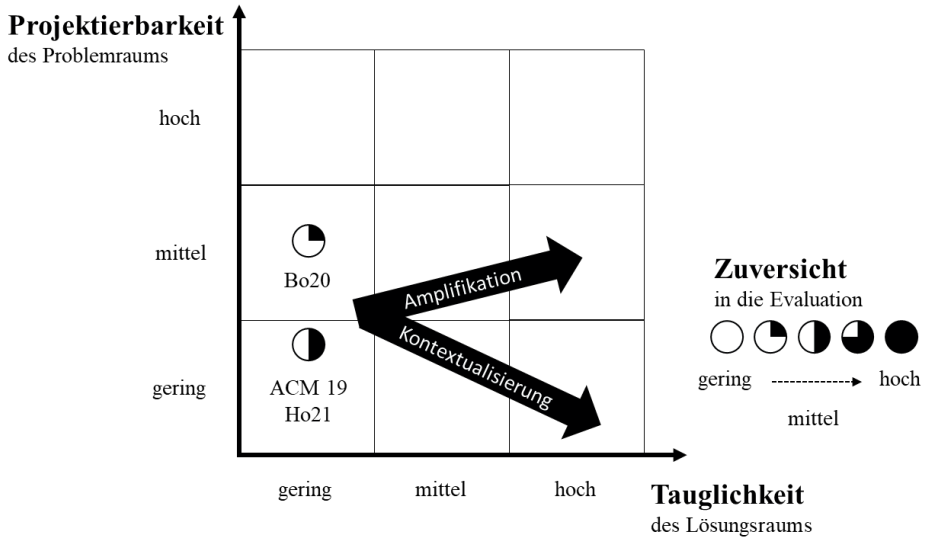


Abb. 2: Einordnung der identifizierten Vorarbeiten in DKM und strategische Ausrichtung [Br20]

Für unsere weiterführende Arbeit verfolgen wir den Ansatz der Verbesserung eines bestehenden Modells durch die Kombination mehrerer bestehender Modelle zu einem neuen Modell. Die Dimensionen des KI-RM sollten grundsätzlich überdacht werden. Aus unserer Sicht sollten die Dimensionen Datenschutz und Ethik in einem KI-RM Berücksichtigung finden. Neue Technologien bringen verstärkt neue Herausforderungen in Bezug auf personenbezogene Daten mit sich und stellen uns vor ethischen Fragestellungen. Aus diesem Grund sollten diese Aspekte betrachtet werden und sich im KI-RM wiederfinden. Besonders KMU stehen beim Thema KI vor großen Herausforderungen und drohen den Anschluss zu verlieren [Bk21]. Hier sehen wir großes Handlungspotenzial und die Möglichkeit eine Forschungslücke zu schließen, indem der aktuelle Stand der KI-RM mit KMU-Fokus weiterentwickelt wird. Hierbei kommen zwei zentrale Strategien mit Stoßrichtung Erhöhung der Tauglichkeit des Lösungsraums in Frage (siehe Abbildung 2). Bei der Kontextualisierung wird die Projektierbarkeit des Problemraums reduziert, da lediglich eine besondere Problemstellung betrachtet wird. Ein Beispiel könnte ein spezifisch ausgerichtetes KI-RM für KMU oder eine Branchenfokus sein. Die Amplifikation zielt darauf ab, das Niveau der Projektierbarkeit des Problemraums zu halten und die Tauglichkeit des Lösungsraums zu steigern. Dies ist mit einem höheren Aufwand im Vergleich zur Kontextualisierung verbunden. Als Beispiel kann ein branchenübergreifendes KI-RM genannt werden. Die Strategie der Kontextualisierung auf KMU ergibt in diesem Fall Sinn. Darüber hinaus könnte ein mittelstandsgerechtes KI-RM mit entsprechenden Handlungsempfehlungen und schlüssigen Erklärungen konzipiert werden.

5 Zusammenfassung und Ausblick

Diese Arbeit unterstreicht sowohl die Notwendigkeit als auch das Bewusstsein für KI-RM aus einer wissenschaftlichen und praxisorientierten Perspektive. Wir konnten insgesamt 15 teils vielversprechende Ansätze mittels der Literaturrecherche ausfindig machen. Nach Untersuchung der recherchierten KI-RM können wir konkludieren, dass die Mehrheit der Ansätze methodische und wissenschaftliche Mängel beinhalten oder wichtige KI-RM Elemente vermissen lassen. Lediglich drei Modelle können überzeugen. Bei näherer Betrachtung und Einordnung in die Methodik DKM kann festgestellt werden, dass diese einen relativ geringen Lösungsraum abdecken und Defizite bei der Evaluation aufweisen. Damit leistet unsere Arbeit einen Beitrag zur Standortbestimmung vorhandener KI-RM und weist gleichzeitig wertvolle Weiterentwicklungspotenziale aus.




Zusammenfassend erscheint es sinnvoll die bestehenden und als vielversprechend identifizierten KI-RM weiterzuentwickeln. Dabei sollten die aufgeführten Kritikpunkte adressiert und Datenschutzaspekte sowie ethische Fragestellungen einfließen. Wir forcieren die Fortsetzung des KI-RM Entwicklungsprozesses nach Becker, Knackstedt und Pöppelbuß [BPP09] mit den Phasen vier bis sieben unter Einhaltung der Kontextualisierungsstrategie nach Brocke et al. [Br20]. Besonders im Bereich KMU-gerechtes KI-RM sehen wir Chancen, einen Mehrwert für die Durchdringung von KI-Anwendungen in mittelständischen Unternehmen leisten zu können. Eine ausführliche Evaluation mit der Zielgruppe ist in diesem Kontext unabdingbar und schließt sich wie konkrete Handlungsempfehlungen zur korrekten Anwendung des KI-RM an.

Literaturverzeichnis

- [ACM19] Alsheibani, S.; Cheung, Y.; Messom, C.: Towards an Artificial Intelligence Maturity Model: From Science Fiction to Business Facts. In: 23rd Pacific Asia Conf. on Information Systems, PACIS, 2019.
- [Bk21] Bitkom, <https://www.bitkom.org/Presse/Presseinformation/Kuenstliche-Intelligenz-kommt-in-Unternehmen-allmaehlich-voran>, Stand: 27.04.2021.
- [BKP09] Becker, J.; Knackstedt, R.; Pöppelbuß, J.: Developing Maturity Models for IT Management. *Bus. Inf. Syst. Eng.* 1, S. 213–222, 2009.
- [Br05] Brey, P.: Freedom and Privacy in Ambient Intelligence. In *Ethics and Information Technology*, 7; pp. 157–166, 2005.
- [Br09] Brocke, J.; Simons, A.; Niehaves, B.; Reimer, K.; Plattfaut, R.; Cleven, A.: Reconstructing the giant: on the importance of rigour in documenting the literature search process (2009). In: *Proc. European Conf. on Information Systems*, 161, 2009.
- [Br20] vom Brocke, J.; Winter, R.; Hevner, A.; Maedche, A.: Accumulation and Evolution of Design Knowledge in Design Science Research: A Journey Through Time and Space. In *Journal of the Association for Information Systems*, 21; S. 520–544, 2020.
- [Ca20] Carr, S.: AI gone mental: engagement and ethics in data-driven technology for mental health. In *Journal of mental health (Abingdon, England)*, 29; S. 125–130, 2020.

- [Da20] Dahlin, G.: What can we learn from process maturity models – A literature review of models addressing process maturity. In *International Journal of Process Management and Benchmarking*, 10; S. 495, 2020.
- [De05] De Bruin, T.; Rosemann, M.; Freeze R.; Kulkarni, U.: Understanding the main phases of developing a maturity assessment model. In: *16th Australasian Conf. on Information Systems, ACIS*, S. 8-19, 2005.
- [El19] Ellefsen, A. et al.: Striving for excellence in AI implementation: AI Maturity Model Framework and Preliminary Research Results In: *Logforum* 15 (3), S. 363–376, 2019.
- [FSX18] Fang, J.; Su, H.; Xiao, Y.: Will Artificial Intelligence Surpass Human Intelligence? In *SSRN Electronic Journal*, 2018.
- [Ge19] Gentsch, P.: *AI in Marketing, Sales and Service*. Springer International Publishing, Cham, 2019.
- [Ha20] Haefner, N. et al.: Artificial intelligence and innovation management: A review, framework, and research agenda. In: *Technological Forecasting and Social Change* 162, S. 120392, 2020.
- [He04] Hevner, A.; March, S.; Park, J.; Ram, S.: Design Science Research in Information Systems Research. *MIS Quarterly* 28(1), 2004.
- [Ho21] Holmström, J.: From AI to digital transformation. The AI readiness framework. In *Business Horizons*, 2021.
- [JWW21] Jöhnk, J.; Weißert, M.; Wyrki, K.: Ready or Not, AI Comes— An Interview Study of Organizational AI Readiness Factors. In: *Bus Inf Syst Eng* 63 (1), S. 5–20, 21.
- [KW19] Knapp, P.; Wagner, C.: Künstliche Intelligenz schafft neue Geschäftsmodelle im Mittelstand. In: *Künstliche Intelligenz*. Springer Gabler, Berlin, S. 161-172, 2019.
- [MG21] Mikalef, P.; Gupta, M.: Artificial intelligence capability: Conceptualization, measurement calibration, and empirical study on its impact on organizational creativity and firm performance. In: *Information & Management* 58 (3), S. 103434, 2021
- [NSP19] Niewiadomski, P.; Stachowiak, A.; Pawlak, N.: Knowledge on IT Tools Based on AI Maturity. In: *Procedia Manufacturing* 39 (15), S. 574–582, 2019.
- [Pa09] Paulk, M.: A History of the Capability Model for Software. In *ASQ Software Quality Professional*, 12(1),2009.
- [Ri19] Richter, A.; Gačić, T.; Kölmel, B.; Waidelich, L.: Künstliche Intelligenz und potenzielle Anwendungsfelder im Marketing. In: *Dialogmarketing Perspektiven 2018/2019*. Springer Fachmedien, Wiesbaden, S. 31-52, 2019.
- [SSP20] Shearer, E.; Stirling, R.; Pasquarelli, W.: *Government AI Readiness Index 2020*. In *Oxford Insights*, 2020.
- [Va21] Vakkuri, V. et al.: Time for AI (Ethics) Maturity Model Is Now. In: *Proceedings of the Workshop on Artificial Intelligence Safety*, virtuell, Februar, 2021.
- [We12] Wendler, R.: The maturity of maturity model research: A systematic mapping study. In *Information and Software Technology*, 54; S. 1317–1339, 2012.
- [Ya20] Yams, N.; Richardson, V.; Shubina, G.; Albrecht, S.; Gillblad, D.: Integrated AI and Innovation Management: The Beginning of a Beautiful Friendship. *Technology Innovation Management Review* 10, S. 5-18, 2020.

Neuromorphic Vision mit Spiking Neural Networks zur Sturzerkennung im betreuten Wohnen

Sven Nitzsche¹ , Brian Pachideh¹ , Victor Pazmino¹ , Norbert Link², Christoph Schauer², Lukas Theurer², Valentin Haas³, Philipp Marquardt³, Sergey Biniaminov³, Jürgen Becker⁴

Abstract: Zur schnellen Erkennung von gefährlichen Stürzen in betreuten Wohnsituationen können klassische Kameras kombiniert mit künstlichen neuronalen Netzen (Artificial Neural Networks) verwendet werden. Solche Lösungen haben allerdings eine hohe elektrische Leistungsaufnahme und erfordern daher eine dauerhafte Stromversorgung. Dies macht die Integration in bestehende Räume aufwendig. Im Rahmen des Projekts EmbeddedNeuroVision wird daher eine extrem energieeffiziente Lösung basierend auf neuromorphen Kameras und Spiking Neural Networks erforscht, die die elektrische Leistungsaufnahme um mehrere Größenordnungen senken kann. Die energieeffiziente Verarbeitung schafft neue Möglichkeiten für batteriebetriebene Visionssysteme, die nicht nur im betreuten Wohnen, sondern auch in industriellen und Smart-City-Anwendungen flexibler eingesetzt werden können.

Keywords: Spiking Neural Networks; Neuromorphic Vision; Event-based Vision; Neuromorphic Computing; Human Action Recognition; Machine Learning; Neuronale Netze; Low Power; Edge AI

1 Einleitung

Klassische Kameras kombiniert mit künstlicher Intelligenz ermöglichen diverse Anwendungen wie z.B. automatisierte Gesichtserkennung zur Zugangskontrolle, Fahrassistenzsysteme und unterstütztes Wohnen. Beim unterstützten Wohnen können etwa Räume mit Kameras und künstlicher Intelligenz (KI) ausgestattet werden, um verschiedenste menschliche Aktivitäten zu erkennen. Diese reichen von alltäglichen Aktivitäten wie Gehen, Sitzen, Stehen oder Liegen bis hin zu Gefahrensituationen wie Stürzen. Insbesondere bei Senioren können Stürze eine große Gefahr darstellen, wenn sich keine weiteren Personen im Haushalt befinden. Intelligente Kamerasysteme können in solchen Situationen als Alternative zu Notfallknöpfen dienen, lassen sich im Gegensatz zu diesen aber unauffälliger in das alltägliche Leben integrieren und funktionieren auch, wenn die gestürzte Person selbst nicht mehr handlungsfähig ist. Zudem muss das System nicht aktiv von einer Person angelegt werden

¹ FZI Forschungszentrum Informatik, Haid-und-Neu-Str. 10-14, 76131 Karlsruhe, Deutschland, {nitzsche, pachideh, pazmino}@fzi.de

² Inferics GmbH, Benzstr. 3, 76185 Karlsruhe, Deutschland

³ HS Analysis GmbH, Haid-und-Neu-Str. 7, 76131 Karlsruhe, Deutschland

⁴ Institut für Technik der Informationsverarbeitung (ITIV), Karlsruher Institut für Technologie, Deutschland

und kann entsprechend nicht vergessen werden, wie dies etwa bei am Körper getragener Sensorik der Fall ist, etwa Beschleunigungssensoren [No07] oder Wearables [NLK20].

Damit dennoch die Privatsphäre der Menschen gewahrt bleibt, können die Videodaten direkt in der Kamera verarbeitet werden, sodass eine Weiterleitung in die Cloud und eine Speicherung der Daten entfällt. Hierfür ist es notwendig, geeignete Hardware zur Verarbeitung in das Sensorsystem zu integrieren, welche leistungsfähig genug ist, um die KI-Algorithmen in Echtzeit auszuführen. Insbesondere die Erkennung zeitlich zusammenhängender Aktivitäten und Vorgänge, etwa der Sturz eines Bewohners, stellen allerdings hohe Ansprüche an die Rechenleistung. Der Energiebedarf der eingebetteten Hardware wie auch der Kamera selbst macht eine ständige Stromversorgung notwendig⁵, wodurch eine Nachrüstung solcher Systeme in bestehendem Wohnraum aufwendig ist. Für einen guten Überblick sollten solche Systeme zudem an der Decke und optimalerweise in der Mitte des Raums installiert werden, wo üblicherweise keine direkt zugängliche Stromversorgung verfügbar ist.

Eine Lösung für diese Problematik, welche den Energiebedarf stark senken kann, ist der Einsatz eines neuromorphen Systems. Neuromorphic Computing bezeichnet Sensorik- und Verarbeitungssysteme, die stark an biologische Vorgänge angelehnt sind, etwa dem menschlichen Auge und Gehirn. Zur Sturzerkennung kann eine Kombination aus neuromorpher Kamera und neuromorpher Verarbeitung basierend auf Spiking Neural Networks (SNN) zum Einsatz kommen. Eine solche Kombination wurde bereits in verschiedenen Szenarien zur Gestenerkennung demonstriert und benötigt dabei lediglich eine Leistungsaufnahme von unter 200mW für das Gesamtsystem [Am17; Ce20]. Ebenso wurde der Einsatz von SNNs bereits für die sog. Human Action Recognition, zu der auch die Sturzerkennung gehört, demonstriert [Kh16]. Die Leistungsaufnahme ist im Vergleich zu einem System mit klassischen Kameras und Artificial Neural Networks (z.B. Convolutional Neural Networks oder Recurrent Neural Networks) um mehrere Größenordnungen geringer. Dadurch werden mobile Lösungen ermöglicht, die mittels Akku mit Energie versorgt werden können, also keine konstante Stromversorgung mehr benötigen. Dies kann die Installation vereinfachen und so die Einstiegshürden für die Anwendung solcher Systeme stark reduzieren. Zwar gibt es auch bei Artificial Neural Networks Bestrebungen, möglichst energieeffiziente Hardware zu entwickeln, allerdings weisen selbst spezielle Beschleuniger wie Googles Coral Edge noch eine Leistungsaufnahme von 1-3W auf [Li20] und sind damit nicht für batteriebetriebene Systeme geeignet.

Da Neuromorphic Computing noch ein vergleichsweise junges Forschungsfeld ist und die Übergangsphase in die praktische Anwendung gerade erst begonnen hat, kann die Anwendung für KMUs trotz allen Potenzials noch eine große Hürde darstellen. In diesem Beitrag soll daher das Pilotprojekt EmbeddedNeuroVision vorgestellt werden. Das Forschungskonsortium, bestehend aus den beiden KMUs HS Analysis und Inferics sowie dem FZI Forschungszentrum Informatik, nimmt im Rahmen des Projekts den Transfer von Neuromorphic Computing in Angriff, um zu zeigen, dass diese Technologien dennoch auch

⁵ Die elektrische Leistungsaufnahme solcher Systeme liegt typischerweise bei ca. 10-20W, wovon ca. 5W auf die Frame-Kamera entfallen und 5-15W auf die Verarbeitung [LA20].

in KMUs relativ unkompliziert angewendet werden können und einen praktischen Mehrwert bieten.

2 Grundlagen

Ein wesentliches Merkmal von Neuromorphic Computing ist die Repräsentation von Informationen als zeitlich diskrete „Events“ bzw. „Spikes“.

Spiking Neural Networks: Die Auswertung von Event-Daten erfolgt mithilfe von Spiking Neural Networks (SNN), deren Synapsen und Neuronen genau auf die Übertragung bzw. Verarbeitung von Events ausgelegt sind [Ma97]. Dies unterscheidet SNNs insbesondere von den heute erfolgreichen Artificial Neural Networks (ANN), welche ihre Ergebnisse basierend auf konstant anliegenden, numerischen Werten berechnen. Die grundsätzliche Funktionsweise eines SNN lässt sich wie folgt beschreiben: Einzelne Synapsen verbinden ein Paar von Neuronen. Die Verbindungen sind gerichtet, d.h. nur das präsynaptische Neuron kann Impulse (Events) an das postsynaptische Neuron übertragen. Die in einem Neuron eingehenden Impulse werden dort integriert bis ein gewisser Schwellwert erreicht ist, woraufhin sich das Neuron entlädt und dadurch seinerseits Impulse an nachgeschaltete Neuronen entsendet. Der Ladezustand eines Neuron wird als Membranpotenzial bezeichnet, angelehnt an das biologische Äquivalent im Gehirn. Eine weitere wichtige Funktion des Neuron ist der Leck-Effekt, durch den das Membranpotenzial während des Integrationsvorgangs stetig abnimmt, solange keine Impulse anliegen. Unabhängig vom Aufkommen neuer Impulse zeigen SNNs damit ein zeitlich dynamisches Verhalten, ähnlich der rekurrenten neuronalen Netzen (RNN). Als einzelne Funktionseinheit betrachtet, implementiert das Neuron im SNN nur eine sehr einfache Funktion. Die Datenverarbeitung mit SNNs wird dementsprechend erst durch eine starke Vernetzung der Neuronen bedeutsam. So kann das Fan-In eines Neuron oftmals hunderte oder tausende Verbindungen umfassen [Bo00].

Neuromorphic Vision: Eine wichtige Komponente für die Integration von SNNs in Edge-Anwendungen sind Sensoren, welche ihre Messwerte durch ihren physikalischen Aufbau nativ als Events ausgeben. Solche Sensoren werden oft als neuromorphe Sensoren bezeichnet und unter ihnen sind die neuromorphen Kameras bereits marktreif [Ga20]. Herkömmliche Videokameras geben ganze Bilder mit einer bestimmten Bildwiederholungsrate aus, z.B. 30 oder 60 Hz. Im Gegensatz dazu wird bei neuromorphen Kameras lediglich die Änderung der Helligkeit im Sichtfeld gemessen. Darüber hinaus werden die Helligkeitsänderungen nicht über einen Kanal als Einzelbild ausgegeben, sondern individuell per Pixel. Die Kodierung der gemessenen Information erfolgt als Event, d.h. jedes Event signalisiert die Helligkeitsänderung eines Pixels um einen festen Wert relativ zum letzten Helligkeitswert, inklusive eines Zeitstempels. Eine weitere Besonderheit ist, dass jeder dieser einzelnen Messkanäle asynchron zu den Restlichen arbeitet. Damit hat der ausgegebene Datenstrom eine variable Datenrate, je nach Aktivität im beobachteten Sichtfeld der neuromorphen Kamera. Die Messung von Helligkeitsänderungen in jedem Pixel basiert auf einer schnellen Schaltung, die neuromorphen Kameras einige vorteilhafte Eigenschaften verleiht [LPD06].

So liegt die maximale Aufnahmezeit der Pixelmesswerte im MHz-Bereich, bei einem hohen Dynamikbereich von ca. 120 dB und einer geringen Energieaufnahme von unter 0,5 Watt. Darüber hinaus sind neuromorphe Kameras direkt mit SNNs kompatibel.

3 Konzept und aktueller Stand

Der Einsatz von Neuromorphic Computing gerade in Edge-Anwendungen ist sehr vielversprechend. Wie die vorherigen Kapitel gezeigt haben, ergeben sich daraus verschiedene Möglichkeiten, neue Marktpotenziale zu erschließen. Im Rahmen von EmbeddedNeuroVision wird dieses Potenzial exemplarisch anhand eines bestehenden Systems zur Sturzerkennung des Partners Inferics gezeigt, das mit verschiedenen neuromorphen Komponenten erweitert wird. Die integrierte klassische Kamera wird durch eine neuromorphe Kamera ersetzt und die zur Verarbeitung eingesetzten rekurrenten neuronalen Netze mit SNNs. Parallel wird eine mögliche Ausführung auf neuromorpher Hardware untersucht werden. Durch die Definition von Schnittstellen und, wo notwendig, Konvertierungsalgorithmen entsteht ein modulares System, bei dem jede Kombination von klassischen und neuromorphen Komponenten möglich ist, siehe Abbildung 1. Nötig ist das bspw. bei der Verwendung einer neuromorphen Kamera mit einer Verarbeitung durch RNNs oder eine klassischen Kamera mit einer Verarbeitung durch SNNs. Dadurch soll ermöglicht werden, den Effekt einzelner Komponenten getrennt zu untersuchen und letztendlich den Mehrwert jeder Komponente für den Einsatz in KMUs abwägen zu können.

Aus den bisherigen Arbeiten haben sich verschiedene Erkenntnisse bzgl. der Anforderungen an neuromorphe Kameras, den Aufbau einer modularen Systemarchitektur und der Implementierung von SNNs ergeben. Für einen raumüberblickenden Einsatz muss insbesondere die Auflösung der Kamera entsprechend groß sein. Eine Anforderungsanalyse hat eine minimale Auflösung von 1280x720 Pixel ergeben, was die Auswahl möglicher neuromorpher Kameras stark eingrenzt [Ga20]. Zum Zeitpunkt der Projektdurchführung bieten nur die CeleX-V von CelePixel sowie die Prophesee Gen4 eine ausreichende Auflösung, welche beide im Rahmen des Projekts evaluiert werden. Tabelle 1 zeigt eine Auswahl verfügbarer neuromorpher Kameras mit ausgewählten Eigenschaften.

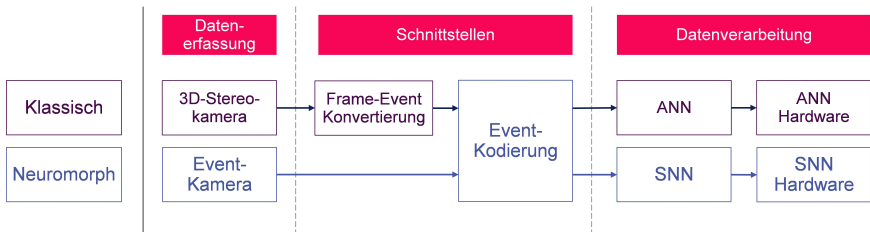


Abb. 1: Modulare Systemarchitektur mit Schnittstellen und notwendiger Konvertierung.

Name	Auflösung (Pixel)	Dynamic Range (dB)
iniVation DAVIS346	346x260	120
iniVation DVXplorer	640x480	110
iniVation DVXplorer lite	320x240	110
Prophesee Gen3	640x480	>120
Prophesee Gen4	1280x720	>124
CelePixel CeleX-V	1280x800	120

Tab. 1: Übersicht von verfügbaren neuromorphen Kameras während des Projekts.

Bei der Verarbeitung der resultierenden Event-Daten durch SNNs scheinen verschiedene Ansätze vielversprechend zu sein. Zum einen ist der Einsatz von Convolutional SNNs (CSNN) möglich, wie es für die Gestenerkennung schon demonstriert wurde [St20]. Als Verarbeitungs-Hardware könnte in diesem Fall Dynap-CNN zum Einsatz kommen, welches speziell für die Ausführung von CSNNs entwickelt wurde [Sy19]. Zum anderen scheint auch eine Kombination aus CSNNs und RSNNs möglich [Wa19; XCS20]. Für die Ausführung der letzten beiden Varianten könnte Intels Loihi als passende Hardware-Plattform dienen [Da18]. Unabhängig von der exakten Lösungsvariante wird eine Leistungsaufnahme von 100mW bis 200mW erwartet [Am17; Ce20]. Daher konzentrieren sich die weiteren Untersuchungen vor allem auf die Genauigkeit und die Einfachheit der Implementierung der jeweiligen Lösungswege.

4 Ausblick

Eine Evaluierung des Gesamtsystems und eine Integration von neuromorpher Hardware werden als nächste Schritte folgen, sobald die laufenden Arbeiten zur Entwicklung eines passenden SNN abgeschlossen sind. Dann wird eine umfassende Aussage über die Möglichkeiten von Neuromorphic Computing zur Sturzerkennung möglich sein. Die bisherige Forschungsarbeit scheint die Erwartungen allerdings zu bestätigen, dass neuromorphe Kameras in Kombination mit Spiking Neural Networks eine mobile Lösung ermöglichen. Die schnellen Fortschritte sowie die kurze Projektlaufzeit von nur einem Jahr zeigen zudem, dass auch der Einsatz neuester Forschung für KMUs ein machbares und gleichzeitig lohnenswertes Unterfangen sein kann.

5 Danksagung

Diese Arbeit ist im Rahmen des Projekts EmbeddedNeuroVision entstanden, welches im Rahmen des KI-Innovationswettbewerbs Baden-Württemberg vom Ministerium für Wirtschaft, Arbeit und Tourismus Baden-Württemberg gefördert wird.

Literatur

- [Am17] Amir, A. et al.: A Low Power, Fully Event-Based Gesture Recognition System. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). S. 7388–7397, 2017.
- [Bo00] Boahen, K.: Point-to-point connectivity between neuromorphic chips using address events. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing* 47/, S. 416–434, 2000.
- [Ce20] Ceolini, E. et al.: Hand-Gesture Recognition Based on EMG and Event-Based Camera Sensor Fusion: A Benchmark in Neuromorphic Computing. *Frontiers in Neuroscience* 14/, 2020.
- [Da18] Davies, M. et al.: Loihi: A Neuromorphic Manycore Processor with On-Chip Learning. *IEEE Micro* 38/1, S. 82–99, 2018.
- [Ga20] Gallego, G. et al.: Event-based Vision: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*/, S. 1–1, 2020.
- [Kh16] Kheradpisheh, S.R. et al.: STDP-based spiking deep neural networks for object recognition. *CoRR abs/1611.01421*/, 2016, arXiv: 1611.01421, URL: <http://arxiv.org/abs/1611.01421>.
- [LA20] Lee, J.; Ahn, B.: Real-Time Human Action Recognition with a Low-Cost RGB Camera and Mobile Robot Platform. *Sensors (Basel)* 20/, 2020.
- [Li20] Libutti, L. et al.: Benchmarking Performance and Power of USB Accelerators for Inference with MLPerf. In: 2020.
- [LPD06] Lichtsteiner, P.; Posch, C.; Delbruck, T.: A 128 128 120 dB 15 s Latency Asynchronous Temporal Contrast Vision Sensor. In: 2006.
- [Ma97] Maas, W.: Networks of spiking neurons: the third generation of neural network models. *Neural Networks* 14/, S. 1659–1671, 1997.
- [NLK20] Nho, Y.-H.; Lim, J. G.; Kwon, D.: Cluster-Analysis-Based User-Adaptive Fall Detection Using Fusion of Heart Rate Sensor and Accelerometer in a Wearable Device. *IEEE Access* 8/, S. 40389–40401, 2020.
- [No07] Noury, N. et al.: Fall detection - Principles and Methods. In: 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. S. 1663–1666, 2007.
- [St20] Stewart, K. et al.: On-chip Few-shot Learning with Surrogate Gradient Descent on a Neuromorphic Processor. 2020 2nd IEEE International Conference on Artificial Intelligence Circuits and Systems (AICAS)/, S. 223–227, 2020.
- [Sy19] SynSense: DYNAP-CNN — the World’s First 1M Neuron, Event-Driven Neuromorphic AI Processor for Vision Processing, 9. Apr. 2019, URL: <https://www.synsense-neuromorphic.com/1380.html>, Stand: 29.04.2021.

- [Wa19] Wang, W. et al.: Temporal Spiking Recurrent Neural Network for Action Recognition. *IEEE Access* 7/, S. 117165–117175, 2019.
- [XCS20] Xing, Y.; Caterina, G. D.; Soraghan, J.: A New Spiking Convolutional Recurrent Neural Network (SCRNN) With Applications to Event-Based Hand Gesture Recognition. *Frontiers in Neuroscience* 14/, 2020.

Towards Designing a User-centric Decision Support System for Predictive Maintenance in SMEs

Domenic Kellner¹, Maximilian Lowin¹, Moritz von Zahn¹, Johannes Chen¹

Abstract: In manufacturing, small and medium-sized enterprises (SMEs) face global competition. In the field of predictive maintenance (PdM), artificial intelligence (AI) helps to prevent machine failures and has the potential to significantly reduce costs and increase process efficiency. Even though PdM has several benefits, it also entails considerable challenges for SMEs, especially when it comes to user interactions. In this short paper, we harness the design science methodology and discuss several problems regarding user interactions with predictive maintenance applications. We incorporate two different literature streams, namely, predictive maintenance and decision support systems. Finally, we present necessary design requirements, principles, features, and propose a research design to further develop and evaluate a user-centric PdM decision support system. Thereby, we contribute to making AI tangible in SMEs.

Keywords: predictive maintenance, machine learning, decision support systems, design science research, explainable artificial intelligence

1 Introduction

In manufacturing, small and medium-sized enterprises (SMEs) face global competition. In order to promote competitiveness, a growing number of SMEs aim to leverage artificial intelligence [HB21]. This can be achieved in the field of maintenance, where artificial intelligence (AI) helps to prevent machine failures and reduces maintenance costs (i.e., predictive maintenance, cf. [Ca19]). Notwithstanding its benefits, predictive maintenance (PdM) entails considerable challenges for SMEs. Many challenges are of technical nature, such as the choice of a suitable algorithm and the availability of training data. However, another key challenge concerns users interacting with PdM applications: Users must be capable of processing, interpreting, and acting upon PdM information [Na16].

In order to overcome user-related challenges in PdM, we built upon design science research (DSR) to infer design requirements for PdM dashboards as decision support systems (DSS) and to answer the following research question: **How to design a user-centric decision support system for predictive maintenance to enhance decision making in SMEs?** To answer our research question, we draw from prior research on predictive maintenance and decision support systems. We present design requirements,

¹ Goethe University Frankfurt, Theodor-W.-Adorno-Platz 4, D-60629 Frankfurt am Main, Germany {kellner, lowin, vzahn, jchen}@wiwi.uni-frankfurt.de

principles, and features as well as a research design to further develop and evaluate a user-centric PdM dashboard. Our proposed research focuses on SMEs, as they tend to lag behind larger organizations in PdM-enabled value creation [Ad20], although they are inherently well-suited to embrace innovation [Qu18]. Hence, we aim to help SMEs to exploit their potential and leverage user-centric PdM in practice.

2 Design Science Methodology

We follow the DSR methodology proposed by [KV08]. The DSR methodology is well suited to address our research question, as it allows us to rigorously design and evaluate our PdM dashboard artifact to obtain new design knowledge for a specific class of IT artifacts [He04]. The overall research design splits into two design cycles. The first design cycle focuses on the design and implementation of a prototype artifact that we plan to evaluate in expert workshops. In addition, we will utilize these expert workshops to further improve our problem definition for the next cycle. The second and final design cycle enables us to iteratively refine the artifact and evaluate it in an online experiment. Subsequently, we reflect on our design and evaluate the results. In this short paper, we present the results of the first two steps of our DSR approach, where we identified PdM problems in the literature. Subsequently, we follow [MMM15] and suggest corresponding design requirements, design principles, and design features. First, design requirements comprise generic requirements that any artifact instantiated from the respective design should meet. Second, design principles are an artifact's generic capabilities regarding the proposed design through which it addresses its requirements. Finally, design features conceptualize technical aspects of the implementation of design principles in an actual artifact.

3 Designing a Decision Support System for Predictive Maintenance in SMEs

3.1 Awareness of Problem

Predictive maintenance (PdM) uses system data to determine a system's condition and to predict when maintenance should be performed [Al14, Mo02]. The system data can originate from various sensors (e.g., thermometers or accelerometers [Ha10]) and can be leveraged by machine learning (ML) as part of AI to identify patterns of unusual device behavior (e.g. anomalies). Once ML techniques identify such anomalies, they can predict a breakdown or a need for service. Commonly used algorithms are so-called black box models, such as artificial neural networks [Ku18]. Thus, PdM can minimize unscheduled breakdowns of equipment, prevent major repairs, save costs, and improve quality by being aware of a system's conditions [Mo02]. However, many current PdM solutions support only dispersed data sources [Ha10] and maintenance cycles are not fully

optimized [Mo02]. Independent of the technical feasibility, PdM solutions must also account for their users, who ultimately decide on maintenance issues. Consequently, a PdM solution should enable users to process the output of PdM solutions with as little cognitive effort as possible [Na16]. Moreover, in contrast to the black box characteristic of many ML applications, PdM solutions should provide reasoning on how system data has driven the prediction. This reasoning promotes user confidence in ML predictions [Sh21] and is thus a prerequisite to realize the cost reduction promised by PdM. Notably, this is particularly relevant for SMEs, for which maintenance costs are of greater magnitude in relative terms.

Decision support systems (DSS) aim to improve decision making by providing users with advice [TAL05]. In order to derive decisional advice, DSS typically leverage either rule-based or data-driven algorithms (e.g., ML techniques). Moreover, DSS build upon various designs to effectively communicate the advice to the user [Mo17]. This is particularly relevant to our study, as a PdM dashboard itself is a data-driven DSS [YV12]. Therefore, a user-centric PdM dashboard should adhere to three major DSS design requirements: high advice quality to improve user decision-making, decision support which reduces cognitive effort of users, and low system restrictiveness which allows users to control the strategy selection [MMM15]. Notably, meeting these requirements is a non-trivial task, as the effects of DSS design on performance, adoption, or cognitive effort depend on the use case at hand [Mo17].

The relevance of our problem has been repeatedly confirmed in discussions with practitioners and decisions-makers from SMEs. In particular, the different understanding of ML among the various participants underlines the need for clear design requirements, principles, and features.

3.2 Suggestion

Our proposed research aims to contribute design requirements, principles, and features for dashboards in the context of maintenance in SMEs (see Fig. 1). The results are particularly relevant for engineers and decision makers who maintain resources in SMEs in order to prevent machine failures and decrease maintenance costs. In addition, we advance the literature by combining PdM and DSS research. Thus, we briefly discuss the rationale and planned evaluation for the four design features derived.

First, PdM dashboards should provide a user with informative guidance (DP1) instead of suggestive guidance, since this has led to superior decision making (DR4) in situations of similar complexity [Mo96]. Moreover, in order to optimize maintenance cycles (DR1), a broad range of data sources (DR2) and granularity (e.g., historical sensor data for ML predictions) have to be integrated (DP3). Hence, we suggest the following design feature: **Integrate historical and real-time data as well as PdM metrics (e.g., machine information, mean time to failure, remaining lifetime information) (DF1)**. In order to evaluate this feature, we plan to conduct an experiment to determine which of the aforementioned metrics are the most important for decision makers in maintenance

tasks. For this, we build upon expert workshops and surveys with SMEs and evaluate the results using a Likert scale questionnaire.

Second, we have to consider the type of end-user to tailor the dashboard solutions appropriately. PdM dashboards should account for various user types (e.g., engineers or managers). To be more precise, we have to account for the motivation and skill level of the end-user to ensure a frequent use of the dashboard and therefore enable different and flexible representation formats (DP4) like natural language, tables, and graphs to ensure customizability (DR5) [YV12]. Hence, we propose to **enable flexible data exploration via different representations (DF2)**. In order to determine whether the choice of representation meets the need of users for customization, we plan to run a series of online experiments to evaluate different forms of representations for PdM tasks.

Third, a PdM dashboard should provide some rationale behind ML predictions to increase user acceptance [Sh21] and improve decision making (DR4) with informative guidance (DP1). Thus, a **PdM dashboard should include predictions with meaningful explanations (DF3)**. Since ML models are usually black boxes, we want to measure the effectiveness of ML model explanations by comparing black box ML with different forms of explanations in experiments. For instance, computer vision and thermal images can be used to detect whether machines are overheating. With black box ML, engineers might only learn which machine is affected. However, with ML explanations (e.g., SHAP, LIME, or saliency maps, see [LAS20] for an overview) engineers learn which machine parts have triggered the detection of overheating. Hence, engineers receive valuable insights to locate the problem more easily, thus promoting efficiency in the maintenance process.

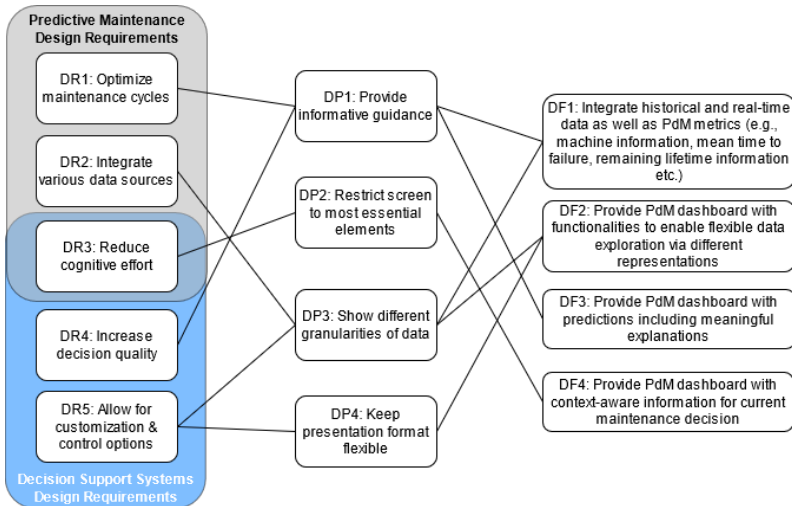


Fig. 1: Design requirements, principles, and features of user-centric DSS for PdM

Last, since the combination of various data sources can easily overwhelm a dashboard user, a dashboard should reduce a user's cognitive effort (DR3) by providing clear information and restrict the screen to the most essential elements (DP2). This context-aware reduction is necessary to support decision makers [Na16] and to promote the adoption of the PdM strategies suggested, since users' ability to process information from DSS is limited [TB99]. Thus, a PdM dashboard should **provide context-aware information for current maintenance decisions (DF4)**. Notably, this is in conflict with providing informative guidance (e.g., through additional ML explanations). Hence, in order to determine a suitable trade-off, we want to investigate how a context-aware reduction mitigates cognitive effort without impairing maintenance decisions by withholding important information from the user.

4 Conclusion and Future Research

In this short paper, we provided the preliminary results of our ongoing design science research on user-centric dashboards for predictive maintenance in SMEs. Following our literature review on PdM problems, we derived several design requirements, principles, and features based on PdM and DSS research. Next, we will prototype a predictive maintenance dashboard for SMEs building on our design principles and features. Based on this prototype, we will conduct expert workshops to analyze whether the design principles enhance decision making. In a second design science cycle, we plan to conduct additional online experiments based on a final artifact for further validation. Future research can implement these insights into a case study with specific SMEs to investigate potential benefits in a real-world scenario.

Acknowledgements

This work was funded by the BMWi through the ForeSight project. The authors state no conflict of interest.

Bibliography

- [Al14] Alestra, S., Brand, C., Erofeev, P., Bordry, C., Burnaev, E., Papanov, A.: Rare Event Prediction Techniques in Application to Predictive Maintenance of Aircraft. In: Proceedings of ITaS conference. pp.32–37, 2014.
- [Ad20] Adu-Amankwa, K., Attia, A., Janardhanan, M., Patel, I.: A predictive maintenance cost model for CNC SMEs in the era of industry 4.0. In: The International Journal of Advanced Manufacturing Technology 104, pp.3567-3587, 2020.
- [Ca19] Carvalho, T.P., Soares, F.A.A.M.N., Vita, R., Francisco, R. da P., Basto, J.P., Alcalá, S.G.S.: A systematic literature review of machine learning methods applied to

- predictive maintenance. *Computers & Industrial Engineering* 137, pp. 106024, 2019.
- [Ha10] Hashemian, H.M.: State-of-the-art predictive maintenance techniques. *IEEE Transactions on Instrumentation and measurement* 60, pp. 226–236, 2010.
- [HB21] Hansen, E.B., Bøgh, S.: Artificial intelligence and internet of things in small and medium-sized enterprises: A survey. *Journal of Manufacturing Systems* 58, pp. 362–372, 2021.
- [He04] Hevner, A.R., March, S.T., Park, J., Ram, S.: Design Science in Information Systems Research. *MIS Quarterly* 28, pp. 75–105, 2004.
- [Ku18] Kulkarni, K., Devi, U., Sirighee, A., Hazra, J., Rao, P.: Predictive maintenance for supermarket refrigeration systems using only case temperature data. In: 2018 Annual American Control Conference (ACC). pp.4640–4645. IEEE, 2018.
- [KV08] Kuechler, B., Vaishnavi, V.: On theory development in design science research: anatomy of a research project. *European Journal of Information Systems* 17, pp. 489–504, 2008.
- [LAS20] Lakkaraju, H., Adebayo, J., Singh, S.: Explaining machine learning predictions: State-of-the-art, challenges, and opportunities. *NeurIPS Tutorial*, 2020.
- [MMM15] Meth, H., Mueller, B., Maedche, A.: Designing a Requirement Mining System. *Journal of the Association for Information Systems* 16, 2015.
- [Mo02] Mobley, R.K.: An introduction to predictive maintenance. Elsevier, 2002.
- [Mo17] Morana, S., Schacht, S., Scherp, A., Maedche, A.: A review of the nature and effects of guidance design features. *Decision Support Systems* 97, pp. 31–42, 2017.
- [Mo96] Montazemi, A.R., Wang, F., Khalid Nainar, S.M., Bart, C.K.: On the effectiveness of decisional guidance. *Decision Support Systems* 18, pp. 181–198, 1996.
- [Na16] Nadj, M., Jegadeesan, H., Maedche, A., Hoffmann, D., Erdmann, P.: A situation awareness driven design for predictive maintenance systems: The case of oil and gas pipeline operations. *ECIS Prototypes* 2016.
- [Qu18] Quinton, S., Canhoto, A., Molinillo, S., Pera, R., Budhathoki, T.: Conceptualising a digital orientation: antecedents of supporting SME performance in the digital economy. *Journal of Strategic Marketing* 26, pp. 427–439, 2018.
- [Sh21] Shin, D.: The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human-Computer Studies* 146, pp. 102551, 2021.
- [TAL05] Turban, E., Aronson, J.E., Liang, T.P.: *Decision Support Systems and Intelligent System, (Sistem Pendukung Keputusan dan Sistem Cerdas)* Ed. 7. Jld. 2. ANDI, 2005.
- [TB99] Todd, P., Benbasat, I.: Evaluating the Impact of DSS, Cognitive Effort, and Incentives on Strategy Selection. *Information Systems Research* 10, pp. 356–374, 1999.
- [YV12] Yigitbasioglu, O.M., Velcu, O.: A review of dashboards in performance management: Implications for design and research. *International Journal of Accounting Information Systems* 13, pp. 41–59, 2012.

Workshop: (Agiles) Enterprise
Architecture Management in
Forschung und Praxis (EAM)

Unternehmensarchitekturmanagement im Zeitalter von Agilität und Künstlicher Intelligenz

Eldar Sultanow¹, Carsten Brockmann², und Christian Czarnecki³

Keywords: Enterprise Architecture Management, Künstliche Intelligenz, Agilität

Künstliche Intelligenz (KI) und Agilität sind entscheidende Faktoren eines modernen Unternehmensarchitekturmanagements (Enterprise Architecture Management, EAM). Die Geschäftswelt ist volatiler, unsicherer, komplexer und schwerer prognostizierbar geworden. In diesem Kontext haben u. a. Entwicklungen im Bereich der KI in den letzten Jahren erhebliche Fortschritte erzielt. Sowohl die zunehmende Technologisierung als auch die Forderung nach Agilität wirken sich aus strategischer, organisatorischer und technischer Sicht auf die Strukturen von Unternehmen aus.

Eine Kernfrage dieses Workshops ist: Wie kann EAM von diesen Entwicklungen profitieren, und umgekehrt, welchen Beitrag kann beziehungsweise muss EAM leisten? Dies wird mit einem anwendungsorientierten Fokus anhand konkreter Fallstudien, praxisorientierter Beispiele und Forschungsarbeiten diskutiert.

KI in Verbindung mit Data Science kann Muster in komplexen Strukturen aufdecken, und damit einen Wertbeitrag für EAM liefern. Beispielsweise, welche IT-Services werden nicht mehr genutzt (und sind obsolet), oder welcher Trend zeichnet sich im Markt ab, der das Unternehmen beeinflusst? Bestimmte Aspekte, die angesichts der Komplexität unvorhersehbar erscheinen, können mithilfe von KI-basierten Analysen doch beherrscht werden.

Andererseits wird sich auch der Wertbeitrag von EAM in der heutigen, agilen Welt verändern. Entwicklungszyklen sind kürzer, der Markt dynamischer sowie intransparenter und daher der Anspruch an organisatorische Strukturen und technische Architekturen höher, auch was Elastizität, Resilienz und Anpassungsfähigkeit angeht.

In diesem Spannungsfeld aus Agilität und zunehmender Technologisierung sind die diesjährigen Workshopbeiträge angesiedelt. So werden praktische Empfehlungen für eine Applikationsarchitektur mit Bezug zur Industrie 4.0, für ein agiles EAM sowie für die Architektur von Netzwerkorganisationen diskutiert. Aus einer datengetriebenen Sicht werden Konzepte für eine performanceorientierte Architekturgestaltung und die automatische Erkennung von Architektur-anomalien vorgeschlagen. Fallstudien demonstrieren den Einsatz von EAM in Kombination mit Chatbots und Robotic Process Automation (RPA).

¹ Universität Potsdam, Karl-Marx-Straße 67, 14482 Potsdam, eldar.sultanow@wi.uni-potsdam.de

² Deloitte Consulting GmbH, Kurfürstendamm 23, 10719 Berlin, cbrockmann@deloitte.de

³ FH Aachen, Eupener Straße 70, 52066 Aachen, czarnecki@fh-aachen.de

A Reference Architecture for On-Premises Chatbots in Banks and Public Institutions

Guidance on Technologies, Information Security and Data Protection

Christian Koch,¹ Benjamin Linnik,² Frank Pelzel,³ Eldar Sultanow,⁴ Sebastian Welter⁵ and Sean Cox⁶

Abstract: Chatbots have the potential to significantly increase the efficiency of banks and public institutions. Both sectors, however, are subject to special regulations and restrictions in areas such as information security and data protection. The policies of these organizations therefore, in some cases, reject the use of cloud and proprietary products because in their view they lack transparency. As a result, the implementation of chatbots in banks and public institutions often focuses on open-source and on-premises solutions; however, there are hardly any scientific guidelines on how to implement these systems. Our paper aims to close this research gap. The article proposes a reference architecture for chatbots in banks and public institutions that are a.) based on open-source software and b.) are hosted on-premises. The framework is validated by case studies at TeamBank AG and the German Federal Employment Agency. Even if our architecture is designed for these specific industries, it may also add value in other sectors – as chatbots are expected to become increasingly important for the practical application of artificial intelligence in enterprises.

Keywords: Chatbots; Machine Learning Architecture; Enterprise Architecture; Reference Architecture; Capability Map; Banking; Public Sector

1 Introduction

In scientific literature there are numerous works on the use of chatbots in enterprises and their potential. On one hand, the technology can be used to improve external communication with customers and business partners. Examples are self-service and automated help [Mc21, p. 13]. On the other hand, chatbots offer the possibility of applying artificial intelligence to facilitate internal processes, such as daily work in customer service, sales teams or in

¹ Technische Hochschule Nürnberg Georg Simon Ohm, Kesslerplatz 12, 90489 Nuremberg, Germany
christian.koch@th-nuernberg.de

² Goethe University Frankfurt, Max-von-Laue-Str. 1, 60438 Frankfurt, Germany, blinnik@uni-frankfurt.de

³ IT Dept. of Germany's Federal Employment Agency, Südwestpark 26, 90449 Nuremberg, Germany
Frank.Pelzel@arbeitsagentur.de

⁴ University of Potsdam, Chair of Business Informatics, Processes and Systems, Karl-Marx Straße 67, 14482, Potsdam, Germany eldar.sultanow@wi.uni-potsdam.de

⁵ External Consultant for Germany's Federal Employment Agency, Germany sebastian.welter@arbeitsagentur.de

⁶ RatPacDune Entertainment, Los Angeles, CA, USA sean.cox@ratpacent.com

administrative workflows [Ko20, p. 129]. The \$11 billion valuation of the German startup Celonis in 2021 impressively demonstrated that process mining and process optimization are capable of adding significant value to businesses [Ko21]. According to Celonis, chatbots often serve as a starting point for so-called Robotic Process Automation (RPA), which leads to a rationalization of company workflows [Sc19].

When implementing chatbots in banks and public institutions, both technological and non-functional challenges need to be addressed. Due to their extremely sensitive data, it is obvious that the two sectors have special requirements regarding information security and data protection. In the European Union, the storage of personal data is governed by the General Data Protection Regulation (GDPR) [Eu16]. With member states acting as role models in implementation, their public agencies pay strong attention to compliance with the regulation. The same is true for banks. As a consequence of the subprime financial crisis in 2008, European banks are subject to close supervision by national as well as international authorities. The European Central Bank (ECB) and the European Banking Authority (EBA) hold key roles in this context. Together with their national counterparts, they oversee compliance with both general and specific regulations to banks such as BCBS 239 [Ba13]. The focus of banking authorities has increasingly expanded into the area of big data and advanced analytics in recent years [Eu20]. At the official level in Germany, the Federal Commissioner for Data Protection (BfDI) and the Federal Office for Information Security (BSI) act as supervisory authorities [De19] [Fe21].

Establishing a chatbot architecture in banks and public institutions is therefore not a purely technology-driven decision. Non-functional requirements such as information security and data protection are equally important. Success depends on both factors. In-house compliance departments have an essential function in this regard to ensure that regulations are being met and have a right of co-determination. Their policies, however, in some cases reject the use of cloud and proprietary products, which in their view lack transparency. As a result, the implementation of chatbots in banks and public institutions often focuses on open-source and on-premises solutions. Yet there are hardly any scientific studies on this special case which is what this article aims to address.

The paper outlines a reference architecture for chatbots in banks and public institutions that are a.) based on open-source software and b.) are hosted on-premises. In the first step, the functional requirements for chatbots are analyzed in the form of a capability map. Secondly, we define system components that cover these requirements. In this context, we describe which frameworks offer the identified elements. Finally, we validate our architecture with real use cases from TeamBank AG and the German Federal Employment Agency. We explain how the functional and non-functional requirements were met in the institutions and which frameworks and technologies were used for the implementation.

2 Related Work

Existing literature categorizes business capabilities enabled by chatbots in various dimensions. Examples are *modes of communication* and *industry vertical*, see Kahn and Das [KD18, pp. 98-105]. The authors distinguish between different dimension attributes. Modes of communication include business-to-business (B2B), business-to-consumer (B2C), consumer-to-consumer (C2C), business-to-employee (B2E), and employee-to-employee (E2E). Examples of industries using chatbots are banking, financial services and insurance (BFSI), travel (especially booking bots), food and restaurants, e-commerce, and utility services – among many others.

Cem Dilmegani [Di21] differentiates between two types of business capabilities. *External capabilities* include patient engagement, drug information, adverse events reporting and social media interaction, as well as collecting customer feedback. Examples for *internal capabilities* are expert lookup, internal help desk, enterprise Q&A as well as virtual sales assistants that integrate with the enterprise's customer relationship management system to notify and guide personnel and thus simplify lead creation and updates. Another internal capability mentioned by Dilmegani is pulse monitoring through chatbots connecting with the company's collaboration environment (e.g., Slack) to monitor satisfaction and productivity.

Hajjar [Ha21] classifies capabilities by *reactive* and *proactive* functions. Moreover, he provides statistics on chatbot usage. Farkash [Fa18] examines chatbots in government. He describes five ways of how bots are capable of personalizing public services.

A study on chatbot applications in government, using the example of public administrations in Latvia, Vienna, and Bonn has been conducted by van Noordt and Misuraca [NM19]. The European Commission published in 2019 an architecture for public service chatbots, authored by PwC [Pw19]. Their basic design serves as a foundation of our component model in section 3.2, adapted for our specific case. Analogously, there are publications with focus on chatbots in the banking sector. For instance, Li et al. [Li20] have analyzed a three-month conversation log of 1,685 users with a task-oriented banking-bot. Considering that information security has become a vital issue for the financial sector, Lai et al. [LLL19] describe a security control procedure for chatbots. Their method aims to protect the security and privacy of user data. Sağlam and Nurse [SN20] address the issue of GDPR compliance and the corresponding aspects that need to be considered when designing agents.

In general, there is plenty of literature on chatbot technology. We find articles on how to build bots and which frameworks and platforms are available in the cloud and on-premises. Kohne et al. [Ko20, pp. 67-79], for example, describe various platform types. Examples of *conversational platforms* are Cognigy.AI, Amelia, kore.ai, and Rulai. *Code platforms* include API.AI/Dialogflow, AWS.AI, Luis.ai, Rasa, recast.ai, reply.ai and IBM Watson. Instances of *klick-chatbot platforms* are MessengerPeople, Botmaker, Spectrm, CharterOn, Chatfuel and Manychat. Finally, *flow platforms* are based on such technologies as Cacao, Creately, Draw.io or Botsociety.

Lokman and Ameen [LA18] conduct a technical review of modern chatbot systems. For this purpose, the authors investigate five technologies, namely DeepProbe, AliMe, SuperAgent, MILABOT, and RubyStar.

Galitsky [Ga19] describes chatbot components and architectures as well as how to implement them using machine learning techniques such as Bayesian networks, neural networks and Markov models. Garga [Ga20] introduces a conversational UI technology maturity model that can be used to classify frameworks and chatbot components along different dimensions. This model serves as a basis of our reference architecture in section 3.3.

Currently, there is limited literature with a special focus on open-source and on-premises solutions. Singh, Ramasubramanian and Shivam [SRS19], for example, provide their insights into the open-source world of chatbots. The authors examine processes in the banking and insurance sectors, which are related to the use cases in our article. Moreover, they identify publicly available data sources that can be used by developers to train bot models. The authors evaluate popular open-source NLP and NLU tools as NLTK, spaCy, CoreNLP, gensim, TextBlob, and fastText. Finally, Singh et al. describe the chatbot frameworks IRIS, Microsoft Bot, Rasa and Google Dialogflow. A great added value of their book is that it guides developers through practical implementation and contains Python code snippets.

Sharma and Joshi [SJ20] focus on Rasa and provide an analytical study and review of this open-source framework. Nayyar [Na19] offers a brief overview on the open-source frameworks Botkit, Botpress, BotMan, VISEO Bot Maker, Rasa and Microsoft Bot. A similar, but more scientific overview of chatbot technologies is provided by Adamopoulou and Moussiades [AM20].

In the following, we build on the discussed literature and further study open-source and on-premises chatbots from a practitioner's perspective. The result will be a draft of a reference architecture for this particular setting.

3 Construct a Reference Architecture

3.1 Capability Map

As a basis for our reference architecture, we analyze the functional requirements for chatbots from a business view in the first step. Therefore, we establish a capability map that outlines use cases in three different application areas, namely Customer Relationship Management, Product & Service Development, and Marketing Services. Figure 1 shows the result. The map is based on the literature reviewed in the previous section and our practical experience.

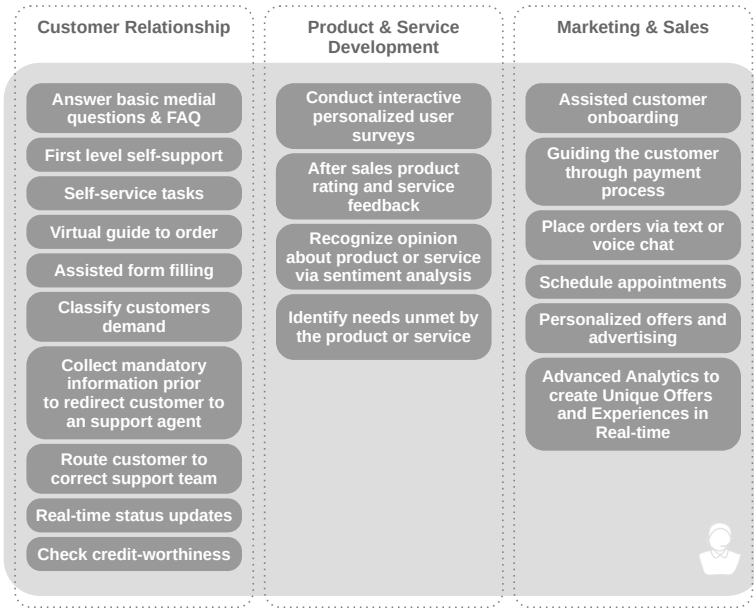


Fig. 1: Business Capability Map for Chatbots

Even though the capability map is primarily oriented towards commercial enterprises, the capabilities can in principle be transferred to public services too. As our map shows, chatbots are being used in a variety of different fields. Hajjar, for example, describes customer support, product suggestion and offering of discounts via conversational interfaces [Ha21]. He highlights the prevention of churn as one of the most important goals for subscription-based industries. According to Hajjar, the top five sectors which benefit from chatbots are real estate, travel, education, healthcare and finance.

3.2 System Components

Let us now analyze the system components needed to implement the capabilities described in section 3.1. For this purpose, we build on an architecture for public service chatbots issued by the European Commission [Pw19]. The high-level design of the study serves as basis for our system components, shown in Figure 2.

According to the above-mentioned study [Pw19], chatbot applications typically consist of three elements: a.) different types of user interfaces such as websites or messengers, b.) the actual digital assistant, and c.) a knowledge base that incorporates, for example, core and analytical systems of an institution. For our specific research question, we extend the model

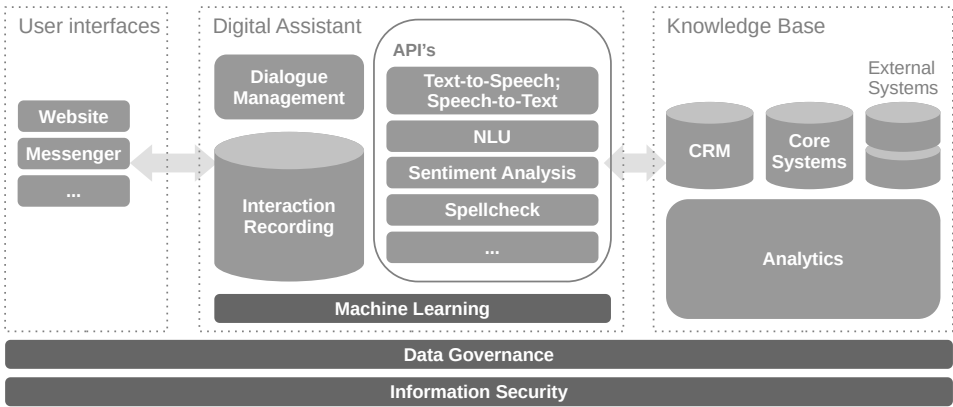


Fig. 2: Chatbot System Components

to include components for data governance and information security, which play a vital role in banking and public administrations. In the next section, we provide an overview of technologies and frameworks that can be used to implement these components.

3.3 Technologies & Frameworks

Our technology overview is based on the categorization of Smriti Garga [Ga20]. She distinguishes chatbots according to a.) their maturity and b.) the mode of interaction, integration and intelligence. In our reference architecture we use Garga's model to structure frameworks along these dimensions as displayed by Figure 3. In line with our research question, the technologies are (free) open-source and can be operated on-premises.

There are various open-source tools and frameworks available which can be used to build or supplement a chatbot architecture. The graphic shown in Figure 3 does not claim to include all open-source frameworks, it is rather an attempt to categorize a selection of technologies and demonstrates a possible grouping method. Similar to our capability map, the choice of frameworks in Figure 3 is based on a literature review and our hands-on experience.

While certain open source frameworks incorporate multiple features into a single package, others specialize in a certain domain. According to the number of GitHub forks and stars, some of the most popular and most comprehensive open-source technologies are Rasa (11.5k GitHub stars) [Bo17], Microsoft Bot Framework (6.5k) and DeepPavlov (5.2k) [Bu18]. These frameworks are designed to support a wide set of chatbot features, including functions for interaction, integration and intelligence. Each of the frameworks is capable of NLU. Their strength is that they can be integrated with other technologies in a variety

of ways when certain functionality is missing. For example, incorporating a chatbot into a website is possible through the customer experience (CX) tools shown in the upper-left corner of Figure 3, if the framework does not support it out of the box. Likewise, any chatbot can be upgraded with voice support to become a voice assistant by converting the user’s audio input to text and the bot’s text response to audio, using the tools depicted in the CX-Voice and Bot-2-Human rectangles.

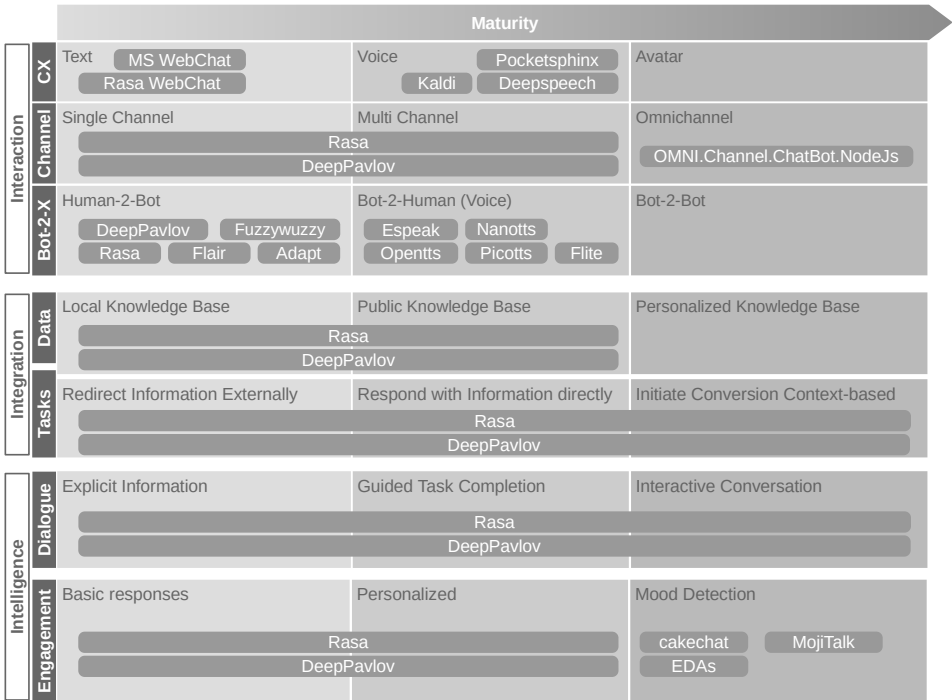


Fig. 3: Chatbot Technology Overview

In order to provide a deeper insight, we discuss technical details of a chatbot implementation with the most popular framework Rasa below. The analysis is based on real use cases from banking and public administration, validating our architecture.

4 Real use cases

4.1 Customer Service at TeamBank AG

TeamBank is the competence center for modern liquidity management in the German Volksbanken Raiffeisenbanken cooperative financial group (Genossenschaftliche FinanzGruppe).

The bank counts about one million clients. Its business activities are carried out in Germany and in Austria by approximately 1,000 employees. Their most popular product, easyCredit, is distributed mainly by cooperative partner banks and online. The processes of the company are highly automated and digitization plays an important role in its business strategy.

In 2020, TeamBank launched a chatbot to enhance workflows in its customer service center. A prior analysis had revealed that the most frequently asked question by clients in the live chat referred to the current processing status of their credit request. In order to receive an answer to this question by an employee, the customer must first legitimize themselves by providing personal data and the corresponding application number. While this process is mandatory from a legal perspective, it is tedious for both the live chatter and the prospect because not every customer has their request number immediately at hand.

Together with the analytics department of the company, TeamBank's customer service center therefore developed a chatbot to streamline the process. The bot guides clients through the legitimation workflow and takes away the pressure of having to find their application number spontaneously. Once customers have gathered the necessary information, they are directed to a live chat with an employee. As a result, the latter has more capacity for actual customer consulting due to the time saved by the bot.

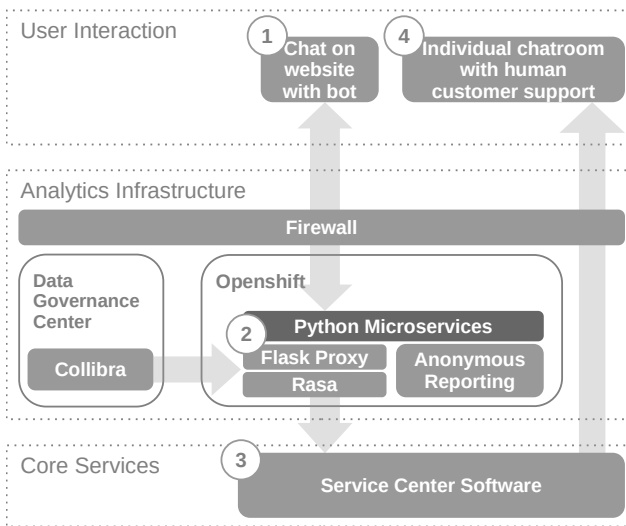


Fig. 4: Customer Service Architecture at TeamBank

In terms of the capability map from section 3.1, the case relates to the area of customer relationship management. Following the technology overview in chapter 3.3, this is an example of human-2-bot communication. Information entered by clients is redirected to human agents. The dialogue flow of the bot is guided and response messages are basic

and not personalized. Consequently, the framework Rasa provided a natural choice as a technology for implementation.

Figure 4 shows the components used, as well as the operational flow of the process. Initially, a visitor to the easyCredit website is offered a chat button (1). After clicking this button and accepting the terms and conditions, a secure connection with the Rasa chatbot service is established (2). The bot asks the customer about their question and collects the information required to answer it. Within this process, the client is guided through a dialogue. All user input is validated to ensure its veracity. As soon as the required information is retrieved, a private chat room with a human support agent is created (3). The user is then redirected to this chat room, where the employee awaits them and has the complete information to take care of the request (4).

The scalability of the architecture is ensured by implementing all functions in the form of Python microservices, as shown in Figure 4 at point (2). The services are stateless per se, the state of conversations is persisted in a database by Rasa. Since highly sensitive data is provided by the user, the infrastructure is hosted on-premises. This way, it can be guaranteed that the data is stored, retrieved and also deleted compliant with company policies. As required by BCBS 239 and GDPR, data lineage information and retention periods are maintained in TeamBank's Data Governance Center (DGC). The DGC is implemented on basis of Collibra and builds the backbone of the bank's operational data governance.

A robust firewall secures the infrastructure and all outside connections are encrypted. To ensure user chatbot sessions cannot be manipulated, a pen-tested Cross-Site-Request-Forgery (CSRF) protection is utilized using a custom Python Flask proxy service. This proxy shields the Rasa chatbot and has various other protection mechanisms in place to block malicious attacks and user requests outside of business hours. Moreover, the proxy service decouples the Rasa chatbot sessions from the server-side browser sessions. The browser sessions are then deleted after 15 minutes.

For quality assurance, a bot usage report can be generated by TeamBank's business departments. This report is anonymous by design, as all statistics are presented in an aggregated view. Consequently, no personal data leaves the platform. This pattern is a viable way to gain insights into user information without violating privacy and/or GDPR.

By having a secure connection, a pen-tested architecture, unique random short-living browser sessions, anonymous reporting, and strict retention policies, TeamBank simultaneously maximizes the safety of user-provided data and offers a user-centered chatbot. The reporting has shown that the bot has already saved a significant amount of time that employees could invest in valuable customer service.

4.2 Chatbot at the German Federal Employment Agency (Bundesagentur für Arbeit)

As stated in the annual report 2019, "the Federal Employment Agency (BA) sees itself as a customer-oriented service provider in the labor market. Its goal is to provide guidance and support for people in their lives and to facilitate their integration into developments in the labor market. In doing so, it places the concerns and expectations of its clients at the center of its work. (...) The BA counsels people on job-related subjects and supports citizens with financial services such as unemployment benefits and child benefits every day. The human being takes center stage for the BA"[Fe19]. The work of the agency's customer service reflects this claim.

In 2019 BA managed 95,000 customer telephone calls daily, conducted 14,000 guidance sessions daily, and granted 8,900 unemployment benefit applications daily [Fe19]. With the onset of the coronavirus pandemic, customer communication has changed massively for the institution. While face-to-face access was more or less impossible, telephony and online channel have significantly increased in importance. This change in communication between BA and its clients is assumed to continue after the pandemic. Considering the changed customer behavior, the BA vigorously pushed ahead with the development of chatbots to process standardized issues, such as special load peaks and providing a 24/7 availability for simple matters.

The chatbot (Digital Guide, "Digitaler Lotse") is a rule-based chatbot, online since 2020 on the portal website of the BA (www.arbeitsagentur.de). It enables customers to obtain information on specific topics, ask questions and receive automatically generated answers. It offers an initial orientation on their respective concerns about basic security benefits, short-time working allowance, unemployment benefits, and child benefits. According to the capability map from section 3.1, the case therefore relates to the field of customer relationship management.

The entire system consists of multiple topic-bots, each one covering a specific topic area ("Lebenslagen") in the BA portfolio. A bot controller enables handovers between the different topic bots; however, this is usually driven by the user by browsing to another topic area on the web portal. The digital assistants are self-learning, continuously developed and technically optimized; additional topic areas will be added over time. In addition, a central coordinating bot for the landing page is being developed, which deals with general questions and redirects to specific topic bots as needed.

Following the technology overview in chapter 3.3, the case is an example of human-2-bot communication. The system interacts with clients in multiple ways:

- Guided dialogues with buttons, navigate the user through scripted processes, e.g. scheduling an appointment.
- Guided dialogues query backend systems to deliver a specific answer to the user, e.g. calculation of short-time working allowances (Kurzarbeitergeld)
- Answers to free text input queries

All bots are implemented with Rasa as core technology. For language understanding, a custom BERT language model for German was trained and optimized to cope with the domain-specific lingua [DC18]. In selecting technologies, frameworks other than Rasa were excluded for the following reasons:

- The bot had to be ready for production in the shortest possible time (first production release in less than four weeks). Consequently, an existing framework solution was needed that offered as much functionality as possible without the necessity of a time-consuming commercial tender.
- The bot had to be operated entirely on-premises and being capable of running in Docker environments (MESOS, Kubernetes).
- The institute's preference was open-source.

The functionality is implemented in the form of microservices. Each bot is a service, with backend integration and logic services (Kurzarbeitergeld) implemented independently. All services are integrated with the web portal frontend via a middleware service that manages the session and coordinates which bot should connect to which user. Figure 5 shows the core services as well as their data stores.

The integration in the BA web portal is implemented using HTML5 components with launch buttons and in-site popups. The middleware service allows for page-specific configuration to be managed centrally. Each site uses the same chatbot web component, and the middleware configuration determines which chatbot to launch for which page.

A Cassandra database is used to manage the session information and the page-specific configuration. Logging of events for performance reporting is accomplished using a centralized ELK stack. In order to ensure GDPR-compliance, user interaction data and feedback get anonymized by removing user identifiable information and omitting user input where sensitive data is being entered. The data is stored for a limited time to allow for reporting and improvement of the bot performance (retraining) in a database.

With the chatbot development, BA is taking an important step on the path to express digitization, which it successfully took last year under the difficult conditions of the

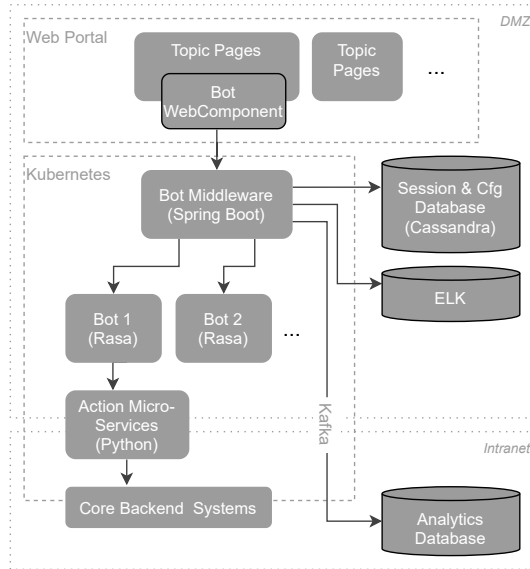


Fig. 5: Bot Architecture "Digitaler Lotse"

coronavirus pandemic. The new technical possibilities that have arisen through the provision of chatbots expand BA's digital service offering. And the customers gratefully accept the new service: About 11,000 unique users had even more unique bot conversations a week (June 2021).

Following the success of the chatbot, subsequent implementation of voice bots for the call center is considered, as well as a re-evaluation of the platform from an editorial point of view: This entails separate interfaces for editors to curate and expand the content independent of the technical chatbot implementation and the introduction of collaboration tooling to foster testing, simulations and enable a systematic handover process between editors and bot technicians.

5 Transferability of our results

Although in this contribution we have focused on on-premises solutions that are open source, the shown procedure is transferable to creating a reference architecture for enterprise chatbots that may be hosted in public clouds and are not limited to be developed using open-source frameworks. Furthermore, the method presented here could be applied to build an enterprise architecture for conversational chatbots incorporating text-to-speech, although this requires a number of adjustments.

While the architecture was designed with specific sectors in mind, it can also add value in other industries. As we have highlighted, chatbots can be a starting point for rationalizing business processes. Aside from finance and public services, they are valuable for many other sectors as well.

6 Limitations

The architecture described here is composed of open-source tools and in contrast to other possibly managed (cloud) solutions, may need a more complex initial setup. The maintenance and development requires expert know-how.

In addition, NLU is an active area of research. Our studied use cases, like most contemporary chatbots, rely on structured training data to follow a conversation. They also depend on user-written plots, so-called stories, to engage in a goal-oriented dialogue. Future generations of chatbots might use artificial intelligence to perform far more generalist tasks with an increased autonomy.

In terms of enterprise integration, the open-source platforms, which have been evaluated in this paper, require manual integration in surrounding systems like artifact management (versioning), monitoring technical and functional/business performance, editorial or CMS systems, databases, other NLP services like text-to-speech, speech-to-text, translation or search (long-tail bots) and channels beyond web (e.g., messenger, telephony). This can be an advantage, as it allows for tailored and best fitting integration but also requires more effort compared to commercial options.

Finally, it is important to emphasize that our reference architecture is a first draft. With two use cases, our sample size is limited. Validating the architecture in a study that includes different cases from several industries could be the topic of a future work.

7 Conclusions and Outlook

The purpose of this paper was to outline a reference architecture for developing and operating chatbots. The requirement was that both training and inference take place on-premises in the form of best-of-breed open-source software.

Based on related work, we provided a technical, methodological, and content structure on required components as well as technologies available from a practitioner's perspective. Subsequently, specific implementations were presented by best practices from the banking and public sector. Rasa served as the framework in both cases, proving the capability to operate high-quality chatbots based on the proposed tooling. However, apart from the technologies and frameworks, the content/stories and (the tuning of) linguistic models constitute a good or bad chatbot. Apart from simple FAQ and decision trees, handling more

complex domain-specific services like in finance or public sector would extend the benefit from quantity to quality. Yet, the fine-tuning of huge linguistic transformer models (e.g. BERT or GPT) is resource-intensive and requires appropriate infrastructure.

The next logical steps in many cases, but not covered by this paper in depth, are the enhancement of bot intelligence and the development of voicebots. For the latter, speech-to-text and text-to-speech components have to be added to the applications. Voice also requires dealing with increased semantic complexity and additional challenges in infrastructure (e.g. callcenter integration, SIP handling, performance and lag).

The handling of data by the major cloud providers in accordance with the governing European and German directives will determine to which extent on-premises solutions remain a valid option. Regardless whether chatbots are run on-premises or in a public cloud, we expect them to become a primary medium for the practical application of artificial intelligence in enterprises.

Literature

- [AM20] Adamopoulou, E.; Moussiades, L.: An Overview of Chatbot Technology. In (Maglogiannis, I.; Iliadis, L.; Pimenidis, E., Hrsg.): Artificial Intelligence Applications and Innovations: 16th IFIP WG 12.5 International Conference, AIAI 2020, Neos Marmaras, Greece, June 5–7, 2020, Proceedings, Part II. Bd. 584. IFIP Advances in Information and Communication Technology, Springer Nature, S. 373–383, 2020.
- [Ba13] Basel Committee on Banking Supervision: Principles for effective risk data aggregation and risk reporting, Techn. Ber., Bank for International Settlements, Jan. 2013, URL: <https://www.bis.org/publ/bcbs239.htm>, Stand: 13. 07. 2021.
- [Bo17] Bocklisch, T.; Faulkner, J.; Pawlowski, N.; Nichol, A.: Rasa: Open Source Language Understanding and Dialogue Management, 2017, arXiv: 1712.05181 [cs.CL].
- [Bu18] Burtsev, M.; Seliverstov, A.; Airapetyan, R.; Arkhipov, M.; Baymurzina, D.; Bushkov, N.; Gureenkova, O.; Khakhulin, T.; Kuratov, Y.; Kuznetsov, D.; Litinsky, A.; Logacheva, V.; Lymar, A.; Malykh, V.; Petrov, M.; Polulyakh, V.; Pugachev, L.; Sorokin, A.; Vikhreva, M.; Zaynutdinov, M.: DeepPavlov: Open-Source Library for Dialogue Systems. In: Proceedings of ACL 2018, System Demonstrations. Association for Computational Linguistics, 2018, URL: <https://doi.org/10.18653/v1/F18-4021>.
- [DC18] Devlin, J.; Chang, M.-W.: Open Sourcing BERT: State-of-the-Art Pre-training for Natural Language Processing, 2018, URL: <https://ai.googleblog.com/2018/11/open-sourcing-bert-state-of-art-pre.html>, Stand: 13. 07. 2021.

- [De19] Der Bundesbeauftragte für den Datenschutz und die Informationsfreiheit (BfDI): Aufgaben und Befugnisse, Techn. Ber., 2019, URL: <https://www.bfdi.bund.de/DE/DerBfDI/Inhalte/DerBfDI/AufgabenBfDI.html>, Stand: 13. 07. 2021.
- [Di21] Dilmegani, C.: 30+ Chatbot Usecases / Applications in Business in 2021, <https://research.aimultiple.com/business-chatbot/>, Mai 2021, Stand: 13. 07. 2021.
- [Eu16] European Parliament: REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), Techn. Ber. 32016R0679, Mai 2016, URL: <https://eur-lex.europa.eu/eli/reg/2016/679/oj>, Stand: 13. 07. 2021.
- [Eu20] European Banking Authority: EBA REPORT ON BIG DATA AND ADVANCED ANALYTICS, Techn. Ber. EBA/REP/2020/01, European Union, Jan. 2020, URL: <https://www.eba.europa.eu/eba-report-identifies-key-challenges-roll-out-big-data-and-advanced-analytics>, Stand: 13. 07. 2021.
- [Fa18] Farkash, Z.: Government Chatbots: 5 Ways Chatbots Can Personalize Public Services, <https://chatbotsmagazine.com/government-chatbots-5-ways-chatbots-can-personalize-public-services-9aaec7f94d84>, Jan. 2018, Stand: 13. 07. 2021.
- [Fe19] Federal Employment Agency (BA): Annual Report of the Federal Employment Agency 2019 (Geschäftsbericht der Bundesagentur für Arbeit, 2019), 2019, URL: <https://www.arbeitsagentur.de/datei/ba146386.pdf>, Stand: 13. 07. 2021.
- [Fe21] Federal Office for Information Security (BSI): Taking advantage of opportunities – avoiding risks, accessed 2021-07-11, 2021, URL: https://www.bsi.bund.de/EN/TheBSI/thebsi_node.html, Stand: 13. 07. 2021.
- [Ga19] Galitsky, B.: Developing Enterprise Chatbots - Learning Linguistic Structures. Springer Nature, 2019, ISBN: 978-3-030-04298-1.
- [Ga20] Garga, S.: A Conversational UI Maturity Model: a guide to take your bot to the next level, <https://chatbotslife.com/a-conversational-ui-maturity-model-a-guide-to-take-your-bot-to-the-next-level-4552d16724a2>, Sep. 2020, Stand: 13. 07. 2021.
- [Ha21] Hajjar, A. J.: Top 36 Chatbot Applications / Usecases in 2021: in-Depth Guide, <https://research.aimultiple.com/chatbot-applications/>, Apr. 2021, Stand: 13. 07. 2021.
- [KD18] Khan, R.; Das, A.: Build Better Chatbots: A Complete Guide to Getting Started with Chatbots. Apress, 2018, ISBN: 978-1-4842-3111-1.
- [Ko20] Kohne, A.; Kleinmanns, P.; Rolf, C.; Beck, M.: Chatbots: Aufbau und Anwendungsmöglichkeiten von autonomen Sprachassistenten. Springer Vieweg, 2020, ISBN: 978-3-658-28849-5.

- [Ko21] Konrad, A.: Celonis Raises \$1 Billion At \$11 Billion Valuation, Making It New York's — And Germany's — Most Valuable Startup, <https://www.forbes.com/sites/alexkonrad/2021/06/02/celonis-process-mining-raises-at-11-billion-valuation/>, Juni 2021, Stand: 13.07.2021.
- [LA18] Lokman, A. S.; Aamedeen, M. A.: Modern Chatbot Systems: A Technical Review. In (Arai, K.; Bhatia, R.; Kapoor, S., Hrsg.): Proceedings of the Future Technologies Conference (FTC) 2018. Bd. 881. Advances in Intelligent Systems and Computing, Springer, S. 1012–1023, 2018, ISBN: 978-3-030-02682-0.
- [Li20] Li, C.-H.; Yeh, S.-F.; Chang, T.-J.; Tsai, M.-H.; Chen, K.; Chang, Y.-J.: A Conversation Analysis of Non-Progress and Coping Strategies with a Banking Task-Oriented Chatbot. In: CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. ACM Press, S. 1–12, 2020.
- [LLL19] Lai, S.-T.; Leu, F.-Y.; Lin, J.-W.: A Banking Chatbot Security Control Procedure for Protecting User Data Security and Privacy. In (Barolli, L.; Leu, F.-Y.; Enokido, T.; Chen, H.-C., Hrsg.): Advances on Broadband and Wireless Computing, Communication and Applications: Proceedings of the 13th International Conference on Broadband and Wireless Computing, Communication and Applications (BWCCA-2018). Lecture Notes on Data Engineering and Communications Technologies, Springer, S. 561–571, 2019, ISBN: 978-3-030-02612-7.
- [Mc21] McTear, M.: Conversational AI: Dialogue Systems, Conversational Agents, and Chatbots. Morgan & Claypool, 2021, ISBN: 9781636390314.
- [Na19] Nayyar, A.: Chatbots and the Open Source Tools You Can Use to Develop Them, <https://www.opensourceforu.com/2019/01/chatbots-and-the-open-source-tools-you-can-use-to-develop-them/>, Jan. 2019, Stand: 13.07.2021.
- [NM19] van Noordt, C.; Misuraca, G.: New Wine in Old Bottles: Chatbots in Government. In (Panagiotopoulos, P., Hrsg.): Electronic Participation: 11th IFIP WG 8.5 International Conference, ePart 2019, San Benedetto Del Tronto, Italy, September 2–4, 2019, Proceedings. Bd. 11686. Lecture Notes in Computer Science, Springer, S. 49–59, 2019, ISBN: 978-3-030-27396-5.
- [Pw19] PwC EU Services: Architecture for public service chatbots, Techn. Ber., European Commission, Sep. 2019, URL: <https://joinup.ec.europa.eu/collection/catalogue-services/document/study-architecture-public-service-chatbots>, Stand: 13.07.2021.
- [Sc19] Scurry, M.: RPA and AI: Business process automation gets smart, <https://www.idginsiderpro.com/article/3398596/rpa-and-ai-business-process-automation-gets-smart.html>, Mai 2019, Stand: 13.07.2021.
- [SJ20] Sharma, R. K.; Joshi, M.: An Analytical Study and Review of open Source Chatbot framework, RASA. International Journal of Engineering Research & Technology (IJERT) 9/6, S. 1011–1014, 2020, ISSN: 2278-0181.

- [SN20] Sağlam, R. B.; Nurse, J. R. C.: Is your chatbot GDPR compliant?: Open issues in agent design. In: CUI '20: Proceedings of the 2nd Conference on Conversational User Interfaces. ACM Press, S. 1–3, 2020.
- [SRS19] Singh, A.; Ramasubramanian, K.; Shivam, S.: Building an Enterprise Chatbot: Work with Protected Enterprise Data Using Open Source Frameworks. Apress, 2019, ISBN: 978-1-4842-5033-4.

An architecture for detecting infrastructure anomalies at Germany's Federal Employment Agency

Johannes Ludsteck,¹ Eldar Sultanow,² Alina Chircu,³ Gebhard Herget,⁴ Matthias Seßler⁵

Abstract: The data centers of Germany's Federal Employment Agency (FEA) provide an information technology (IT) infrastructure that is critical for both external stakeholders and internal processes. For FEA and many other organizations like it, it is essential that any IT infrastructure anomalies - deviations from normal behavior - are detected and their underlying causes are understood and, if appropriate, addressed. In this paper we develop a solution that can help increase the availability of an IT landscape such as FEA's that is characterized by increasing technical complexity and increasing relevance of its applications. The solution detects IT service anomalies based on IT service access logs analyzed with time series methods. The solution also provides visualizations to support further analyses.

Keywords: Machine Learning Architecture; Enterprise Architecture; Time Series; Operations

1 Introduction

Today's organizations rely increasingly on Internet-connected information technology (IT) systems to provide services to external stakeholders as well as support internal operations. The availability, integrity and security of the IT infrastructure and of the associated online services is critical for maintaining the external reputation of the organization, ensuring operational efficiency, and protecting sensitive data. As the complexity of the IT infrastructure increases every day, organizations need to understand and mitigate the risk of IT infrastructure anomalies due to operational failures or malicious attacks, such as Distributed Denial of Service (DDoS) attacks. One challenge, however, is that the resources and expert knowledge that can detect and correct anomalies in IT service provision are limited.

In this paper, we develop a solution for detecting IT service anomalies where the degree of monitoring of IT systems increases, is up-to-date and proactive. The solution focuses

¹ Institute for Employment Research (IAB), Regensburger Straße 104, 90478 Nürnberg, Germany johannes.ludsteck@iab.de

² Capgemini, Bahnhofstraße 30, 90402 Nuremberg, Germany eldar.sultanow@capgemini.com

³ Bentley University, 175 Forest Street, Waltham, MA 02452-4705, USA achircu@bentley.edu

⁴ IT Dept. of Germany's Federal Employment Agency, Tafelhofstraße 4, 90443 Nuremberg, Germany gebhard.herget@arbeitsagentur.de

⁵ IT Dept. of Germany's Federal Employment Agency, Tafelhofstraße 4, 90443 Nuremberg, Germany matthias.sessler@arbeitsagentur.de

on the detection of DDoS and other anomalies for online services at Germany's Federal Employment Agency (FEA) as an example. It relies on predictive models for IT systems telemetry data and key performance indicators (KPIs) for service quality and availability. It is intended to automate assessments (mapping of expert knowledge into intelligent agents") for current and predicted telemetry data. Moreover the solution automates complex monitoring tasks (correlation, cause-effect analysis).

This paper is organized as follows. Section 2 presents a summary of survey studies focused on anomaly detection systems and methods. Section 3 describes the architecture for our proposed solution, the data and methods used, and the results. Section 4 discusses transferability of our results to other contexts. Finally, section 5 discusses conclusions and future research directions.

2 Related Work

Anomalies in IT systems are deviation from normal system behavior that occur due to several major factors: outside threats (initiated by malicious actors who want to disrupt service provision or steal or destroy information), internal events (such as configuration errors, hardware failures, or human error, as well as problems with service quality data collection), and legitimate but abnormal use events (from both internal or external surges that occur when many users try to access IT services at the same time and overload the system) [Fe19]. Anomalies can also be classified based on the nature of the unusual activity patterns as point, collective and contextual anomalies [Fe19; MHS19]. Point anomalies are characterized by the appearance of an outlier, or data point deviation from the usual pattern. Collective anomalies are characterized by a group of data points deviating from the usual pattern (although the individual data points are not indicating point anomalies). Contextual, or conditional anomalies are defined by data deviation from normal expected patterns in a particular context, such as location or point in time [Fe19; MHS19].

While all anomalies are important to detect, outside threats tend to receive the most attention in both research and practice, as well as media, because of their hostile intent and their significant negative consequences regarding the security and availability of IT services. One particularly damaging type of outside threat is a Denial of Service (DOS) attack, which floods an IT service with malicious requests that deplete its processing and connectivity capacity and prevent legitimate users from accessing the service. A DOS attack that is conducted using many computer systems, hacked by the initiator of the attack, is called a Distributed Denial of Service (DDoS) attack [Kh19; KKB17; MHS19; YUS19]. During 2020, there were over 10 million DDoS attacks - an average of almost 840,000 attacks per month and a significant increase over 2019 numbers - with the most affected areas being e-commerce, streaming services, online learning, and healthcare and with the majority of organizations affected reporting service outages as a result of the attack [Hu21].

Detection systems can help automate the monitoring, detecting, analyzing, and even correcting anomalies in complex IT infrastructures [Fe19; MHS19; Sa18]. They can monitor network data, activities on a specific server, or a combination of these data sources [Fe19]. They can detect problems based on known patterns of attacks (signature-based detection, as in the case of already-known external threats coming from specific IP addresses, for example), on deviations from normal or expected behavior (anomaly-based detection), or on a combination of the previous two approaches (hybrid detection) [Fe19; KKB17; MHS19; Sa18].

An anomaly-based detection system uses past data to develop a baseline profile of normal, or expected, behavior. When presented with new data, the detection system analyzes the deviation of this new data set from the baseline profile and flags unusual occurrences [Fe19]. A multitude of approaches, each with its own advantages and disadvantages, have been proposed to implement such a system - including statistical methods, data mining, and artificial intelligence, among others [Fe19; Kh19; KKB17; Sa18; YUS19]. Hybrid methods that combine several different types of approaches, are also being investigated [Fe19]. While none of these methods is superior in all instances, analyses show that statistical methods are very good at detecting both new, unknown anomalies and long-term "low and slow" hostile attacks, despite some disadvantages regarding difficulty in setting parameters, assumptions about normal distributions, and time required to process the data [KKB17]. By comparison, data mining methods are very good at analyzing lengthy continuous patterns but produce high false-positives, and artificial intelligence methods are highly accurate provided they are trained with sufficient data, but also require a lot of processing resources for complex computations [KKB17].

3 Developing the Architecture

To make IT infrastructure anomalies from DDoS or other causes recognizable and visible, all architecture layers need to be involved as shown in Figure 1. Within the infrastructure (in the lowest layer), events that form the basis for anomaly search occur. Various technologies (on the second lowest layer) enable the anomaly detection capabilities (defined in the second layer from the top), that are used by the anomaly detection process (defined in the top most layer).

The infrastructure depicted in the first layer of the architecture comprises (among many other components) of networks, a messaging backbone, and servers that are online accessible. Access protocols (pure infrastructure metrics like number of accesses from an IP address) are gathered by logging and monitoring technologies.

The second layer of the architecture describes the technologies used to collect and process data from the infrastructure. To enable event logs centralization and analysis, we use Elastic Stack, a technology that combines three widely used open source projects, Elasticsearch (based on Lucene), LogStash and Kibana. We use Elastic Stack for collection of the log

files from servers, especially from Internet-facing ones, from ESB (Enterprise Service Bus) that serve as a messaging backbone and from other components that reside in FEA’s network. The technology layer supports business capabilities, as described below. To further support this, we conducted a careful, upfront analysis (before implementing our solution) of technologies that are most established, have a large user community and provide the necessary requirements (such as Mahalanobis Distance, Kalman Filter, etc.) out of the box.

The capabilities described in the third layer of the architecture represent the things that the organization can do with our proposed solution - including perception of the state of the infrastructure, detection of anomalies such as DDoS, recognition of sub-optimal set-ups, and adapting to changes. These business capabilities are important not just for solving FEA’s specific anomaly detection problem, but for any other organization that is interested in monitoring its IT infrastructure and deviations from normal system behavior. In particular, the last capability (adapt to new realities) enables our model to be used in different contexts by different organizational units or even different organizations.

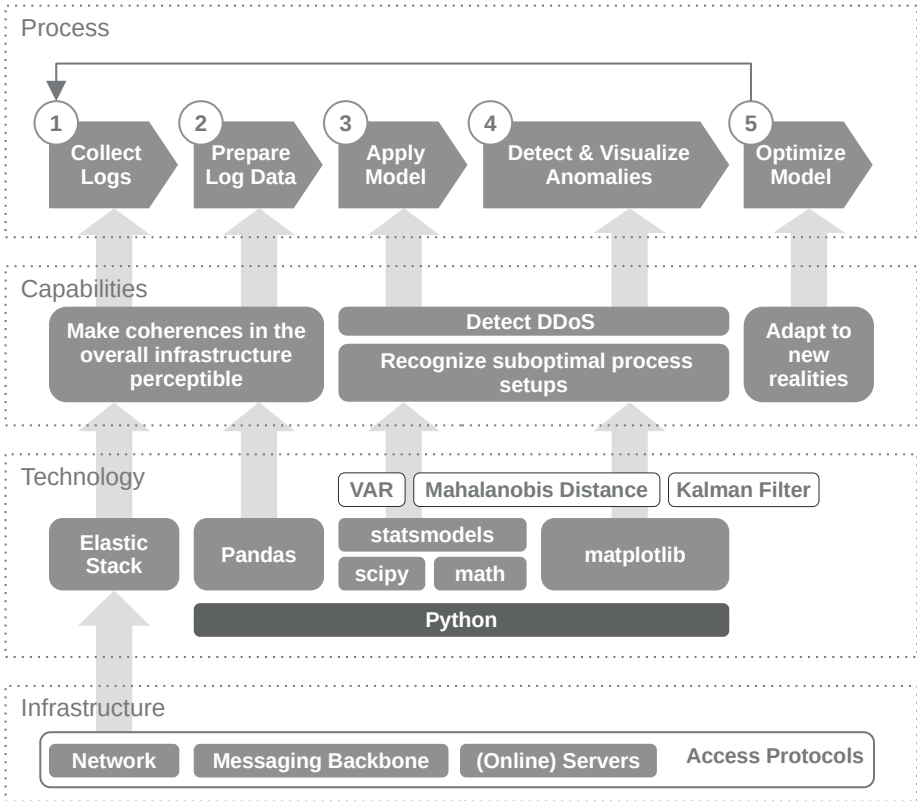


Fig. 1: Architecture for detecting anomalies in German’s Federal Employment Agency infrastructure

The last architectural layer describes the processes that implement the business capabilities - namely, how log data is collected and processed, analyzed through the application of specific models, interpreted and visualized, and how the model is optimized over time. The following sections describe each one of these processes in detail.

3.1 Collecting Logs

Logs are available for determining access to various online services of FEA. The format and content of these logs correspond to Apache's http server logs and are thus standardized. This simplifies the processing considerably and allows for statements regarding the data quality: when is a log entry generated?; under which conditions is the data interpretable?; how can the data be interpreted?; etc. Filebeat is used as the agent for collecting and tagging logs. The challenge is the large amount of data generated per host at peak times (up to 3000 logs/sec). The logs are prepared with Logstash before they are written to an Elasticsearch Database. Essential preparation steps are:

- Extract the timestamp from the log line
- Decompose log line into data fields and assign suitable data types
- Enrich with additional data such as GeoIP and transport times

The IT access logs are collected at a measuring point. Each log entry represents exactly one access. Merging logs or any other form of pre-processing is not required. During real-time processing, three spots in the transmission path that have high latency require special attention: the log mechanism of the source system (1), the Filebeat agent (2), and the preparation by Logstash (3).

Log mechanism of the source system: A log entry receives a record at the completion of the transmission. The time stamp of this record corresponds to the log timestamp, which can be read by the agent. The delay is only determinable with significant effort.

The Filebeat agent: This Filebeat agent is configured so that data is transmitted at intervals of no more than 20 seconds if new data is available.

Preparation by Logstash: No exact time specifications are possible at this point. From practical experience we can say that the latency increases with increasing load. For control purposes, we have included another field that represents the delta between log time and the end of processing by Logstash. Evaluations showed that at peak load times approx. 97% of the data can be retrieved within 60s (storage in Elasticsearch and retrievability via query). This delay is also difficult to measure, but simple manual tests have shown that it is less than 2 seconds.

3.2 Preparing the Log Data

Generally the IT access logs are converted into time series (with an interval of 1 minute). The logs themselves contain information like timestamp, response time, URL and response. In the metadata we include path and name of the log file and the host. Various aggregations are performed and the results stored in a time series database. The Aggregations are:

- Group by timestamp (Intervall 1min), vhost, city, response
- Group by timestamp (Intervall 1min), vhost
- Group by timestamp (Intervall 1min), vhost, response

In addition to the response time (awz) measured as a metric, other statistical parameters are formed for each aggregate using the Elasticsearch database. Listing 1 provides an example.

```
1  "stats-awz": {
2    "count": 149,
3    "min": 0,
4    "max": 1207,
5    "avg": 22.2751677852349,
6    "sum": 3319,
7    "sum_of_squares": 1756401,
8    "variance": 11291.743074636279,
9    "std_deviation": 106.2626137201428,
10   "std_deviation_bounds": {
11     "upper": 234.8003952255205,
12     "lower": -190.25005965505068
13   }
14 }
```

Listing 1: Additional statistical parameters for log aggregates

3.3 Applying the Model

There are several approaches to consider for predicting the time series generated from the IT access log data.

Holt-Winters Method: The Holt-Winters method was suggested by Holt (1957) and Winters (1960), who utilize exponentially weighted moving averages for updating estimates of the seasonally adjusted mean (so-called "level"), slope, and seasonals [CM09, p. 59]. Thus the Holt-Winters method inherently involves the consideration of seasonality and trend (a decomposition of time series data into trend and seasonality is not needed). However this

method is not multivariate and thus not practical for our problem. Moreover the pattern of the past needs to be static, this model requires the weights to be exogenous variables.

Multivariate Temporal Convolutional Network (M-TCN): Wan et al. [Wa19a] propose the M-TCN model for predicting a periodic multivariate time-series. As special features of a M-TCN the authors mention the use of an 1D convolution instead of causal convolutions, the augmentation with two different asymmetric residual blocks, the construction of a sub-model for each feature of input data, and the combination of all sub-models as basis for making predictions.

Wang et al. [Wa19b] used an optimized TCN with stacked auto-encoders to accomplish multivariate time series prediction.

Vector Autoregressive (VAR) Models are widely used, "flexible, and easy to use models for the analysis of multivariate time series"[ZW06, p. 385]. The weights are estimated by the model (and thus not static).

Convolutional Neural Network (CNN) are bio-inspired neuronal networks, which are commonly used for image recognition and classification. They can also applied to multivariate time series prediction, but are more computation-intensive – especially in regard of cross-validation. Another disadvantage, which has been experienced by us in practice, is that with 2% noise everything went well, with more than 5% no useful results were obtained (the behavior became unstable).

Multivariate ARCH and GARCH Models are suitable for time series that are conditionally heteroskedastic and "GARCH models have been proved to be very successful at modeling the serial correlation in the second order moment of the underlying time series"[ZW06, p. 481]. More details on Multivariate ARCH and GARCH Models are also provided for example in [Lü05, p. 557].

3.3.1 SVARX Models: Vector Autoregressions, allowing for Seasonality at multiple Frequencies

Autoregressive (AR) processes explain realizations y_t of a time series $Y = (y_1, y_2, \dots, y_T)'$ as a linear combination of lagged values plus noise (denoted by ϵ_t):

$$y_t = y_{t-1} \alpha_1 + y_{t-2} \alpha_2 + \dots + y_{t-p} \alpha_p + \epsilon_t.$$

They are well fitted as a forecasting tool due to their simplicity, transparency and computational efficiency (since the parameters α_j can be estimated using ordinary least squares). As it is well known, any time series can be approximated with arbitrary precision using autoregressive processes of *sufficient* order: If the number of lags equals the number of observations of the time series, the process yields a perfect fit. Such a model is, however,

useless for forecasting since it fails to separate structure (systematic movements) from noise. This task (separating structure from noise) can be accomplished by selecting lags with high predictive power. The selection can either be based on information criteria (Akaike, Schwarz) or be performed in a data-driven manner by minimizing measures (Mean Squared Error, MSE or Mean Absolute Error, MAE) of the forecast errors.

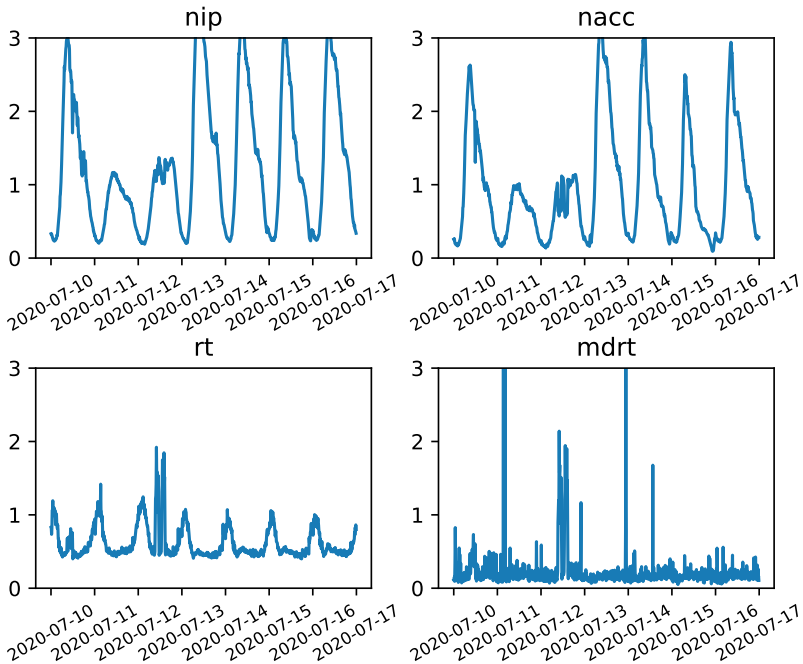


Fig. 2: Time Series Plots

The graphs in Figure 2 show time series (spanning seven days) of the series:

nip	Number of Unique IP Adresses
nacc	Number of Accesses
rt	Response Time
mdrt	Mean Deviation of Response Time

All time series are aggregated to 10 minute granularity by summation (nip, nacc) and averaging (rt, mdrt). Note that all time series are re-scaled by standardization with their standard deviation, i.e. all series have standard deviation 1.

Clearly both information criteria-based as well as data-driven methods become practically infeasible for base models with more than 100 lags, since the predictive power of all possible

subsets of a set containing 100 parameters have to be evaluated in order to find the optimum one. Fortunately, the selection task can be simplified substantially by exploiting the fact that our data show strong and stable seasonal patterns at daily and weekly frequencies (see Figure 2. To understand the numbers correctly, note that all series are re-scaled such that they have standard deviation 1).

This is confirmed by an inspection of the partial autocorrelation function. Filtering all partial autocorrelation coefficients of *nip* exceeding 0.07 in absolute value yields the frequencies 0, 1, 2, 3, 4, 5, 6, 7, 25, **144, 145**, 289, 865, **1008, 1009**. Since our data are aggregated to 10 minute frequency, lag 144 corresponds to $144 \cdot 10/60 = 24$ hours (or one day), and lag 1008 to $1008 \cdot 10/60/24 = 7$ days (or one week).

This simplifies the specification search since it allows us to pre-select lags at daily and weekly frequencies. For example, we restrict our search to all subsets of specifications with

- lags 1, 2, 3, 4, 5 (lags up to one hour)
- lags **144**, 145, 146, 147, 148 (lags around one day)
- **1008**, 1009, 1010, 1011, 1012 (lags around one week).

In principle, each of our four series (*nip*, *nacc*, *rt*, *mdrt*) could be modeled separately. If noteworthy correlations across these series exist, it is, however, preferable to model them together using a vector autoregression. Noteworthy correlations across series exist if lags of one series have predictive power for another series. In the terminology of time series analysis we then say that series *x* Granger-causes series *y*. Our model has the form of equation 1.

$$\begin{aligned}
 y_t = & c + A_1 y_{t-1} + A_2 y_{t-2} + \dots + A_p y_{t-p} \\
 & + \sum_{\tau=0}^{pd} A_{d-\tau} y_{t-d-\tau} \\
 & + \sum_{\tau=0}^{pw} A_{w-\tau} y_{t-w-\tau} \\
 & + B x_t + \epsilon_t
 \end{aligned} \tag{1}$$

Hereby the variable $y_t = (nip_t, nacc_t, rt_t, mdrt_t)'$. The matrix A_1 is defined as follows, where the (scalar) coefficient $a_{nip_{t-1} \rightarrow nip_t}$ represents the partial correlation of nip_{t-1} and nip_t and the coefficient $a_{nacc_{t-1} \rightarrow nip_t}$ represents the partial correlation of $nacc_{t-1}$ and nip_t and so on:

$$A_1 = \begin{pmatrix} a_{nip_{t-1} \rightarrow nip_t} & a_{nacc_{t-1} \rightarrow nip_t} & a_{rt_{t-1} \rightarrow nip_t} & a_{mdrt_{t-1} \rightarrow nip_t} \\ a_{nip_{t-1} \rightarrow nacc_t} & a_{nacc_{t-1} \rightarrow nacc_t} & a_{rt_{t-1} \rightarrow nacc_t} & a_{mdrt_{t-1} \rightarrow nacc_t} \\ a_{nip_{t-1} \rightarrow rt_t} & a_{nacc_{t-1} \rightarrow rt_t} & a_{rt_{t-1} \rightarrow rt_t} & a_{mdrt_{t-1} \rightarrow rt_t} \\ a_{nip_{t-1} \rightarrow mdrt_t} & a_{nacc_{t-1} \rightarrow mdrt_t} & a_{rt_{t-1} \rightarrow mdrt_t} & a_{mdrt_{t-1} \rightarrow mdrt_t} \end{pmatrix}$$

The matrices A_1, \dots, A_p and A_{d-1}, \dots, A_{d-pd} and A_{w-1}, \dots, A_{w-pd} are defined analogously.

The variables $\epsilon_t = (\epsilon_{t,nip}, \epsilon_{t,nacc}, \epsilon_{t,rt}, \epsilon_{t,mdrt})'$ are residuals comprising all other unobservable or omitted effects.

The variable $d = 6 \cdot 24 = 144$ is the number of lags corresponding to one day. The number of lags corresponding to one week is $w = 7 \cdot d = 1008$. The vector of regression constants is $c = (c_{nip}, c_{nacc})'$.

The variables pu, pd, pw represent the maximum lag at unit 10 minutes, daily and weekly frequency.

The vector of exogenous predictors (e.g. holiday dummies, weekday dummies and interactions between lags and weekdays) is x_t . Dummies are indicator variables taking the value 1 if an event occurs and zero otherwise. For example, holiday dummies take on value 1 if t falls into a holiday. Interactions between lags and weekdays allow lag coefficients to vary across weekdays.

In the terminology of standard time series analysis the above described model can be considered as a SVARX (Seasonal Vector Autoregression containing also eXogenous predictors). It deviates from standard SVARX models by including lags at multiple seasonal frequencies and by allowing for gaps in the sequence of lags. A Standard VAR model of order p contains all lags $\in (1, 2, \dots, p)$. In our case, using a base frequency of 10 minutes, including all lags up to one week would require estimation of $6 \cdot 24 \cdot 7 = 1008$ lag coefficients and thereby induce severe multicollinearity problems.

The specification search is conducted by computing a sequence of 144 one-hour forecasts (therefore covering 24 hours) for every combination of parameters $pu \in \{1, 2, 3, 4\}$, $pd \in \{1, 2, \dots, 5\}$, $pw \in \{1, 2, \dots, 5\}$, obtaining the average of the MSE over these 144 forecasts and selecting the model with the minimum MSE. This means that 144 forecasts are computed for each of $3 \cdot 5 \cdot 5 = 72$ models.

The MSE and MAE of a sequence of n_f one-hour-forecast starting at time t are defined as $1/n_f \sum_{\tau=1}^{n_f} (\hat{y}_{t+6+\tau} - y_{t+6+\tau})^2$ and $1/n_f \sum_{\tau=1}^{n_f} |\hat{y}_{t+6+\tau} - y_{t+6+\tau}|$.

At the current exploratory stage of the development, the SVARX is re-estimated by Ordinary Least Squares Regression (OLS) based on a 4-week-sample ranging from $t - 4 \cdot w$ to t in order to obtain the coefficients to compute forecast \hat{y}_{t+6} . We expect considerable efficiency

gains by using recursive least squares (RLS) or the Kalman Filter for the estimation in future versions.

There are further improvements/optimizations to be considered in the further development.

- The choice of the estimation window may have considerable influence on the bias/variance-trade-off: Choosing a long estimation window will increase the bias if the structure of the time series to be forecasted changes quickly. On the other hand, a very short estimation window may shrink bias at the cost of higher variance of parameters and forecasts due to small observation numbers. To tackle this we will include the estimation window length as a further parameter in the optimization.
- Parameter combinations yielding optimum forecasts for one series may be inferior for other series included in the same VAR model. This actually happened in our explorative investigations. The problem can be solved by choosing optimum parameters for each series to be predicted separately.
- The time span of our development data includes one holiday only. Therefore they are unsuitable for a thorough investigation of the relevance of holidays. In the simplest case, holidays can be captured with holiday dummies. If server accesses are shifted from or to adjacent days by the users, it will be necessary, however, to handle this with dummies for the adjacent days. Furthermore, holiday effects may depend on the time distance of the holiday from the weekend. This may require special treatment.

Table 1 shows the first 10 rows of the result generated by the parameter search procedure. It is sorted by the MSE of *nip* (*mse_nip*). Note that the sorting order may (and does vary) depending on the series for which the MSE is computed. If obtaining precise forecasts for *all* four series is important, the parameter search should be performed separately for each parameter. In our case, if the optimization is performed by minimizing the MSE of *nip*, we obtain the following relative decreases for the MSEs of (*nip*, *nacc*, *rt*, *mdrt*): 6.2%, 16.1%, 4.8%, 9.3%. The relative decrease is defined as $(\min_{\theta}(MSE) - \max_{\theta}(MSE)) / \max_{\theta}(MSE)$ where $\min_{\theta}(MSE)$ is the minimum MSE over all parameter combinations contained in the search.

3.3.2 Choosing a frequency

For predicting time series we need to understand which frequency is the best to choose (do we need to consider one measurement per day, hour or minute?). For our problem we decided that one measurement per 10 minutes is sufficient. An important aspect for the choice of the frequency of the data for a forecast model is whether at a lower frequency there are still systematic movements (patterns) or erratics dominate. For a financial income regression, for example, it would be completely unproductive or counterproductive to use

max_lagu	max_lagd	max_lagw	mse_nip	mse_nacc
2.0000	3.0000	5.0000	0.0269	0.1236
2.0000	3.0000	6.0000	0.0269	0.1234
2.0000	3.0000	4.0000	0.0269	0.1241
2.0000	4.0000	5.0000	0.0270	0.1236
2.0000	4.0000	6.0000	0.0270	0.1234
2.0000	5.0000	5.0000	0.0270	0.1237
2.0000	6.0000	5.0000	0.0270	0.1239
2.0000	4.0000	4.0000	0.0270	0.1240
2.0000	5.0000	4.0000	0.0270	0.1241
2.0000	6.0000	4.0000	0.0270	0.1242

Tab. 1: MSEs and MAEs for some Parameter Combinations of the VAR Model

daily data for a medium-term forecast if the bargaining parties only negotiate tariffs once a year. Of course, there can still be changes during the year or even on a monthly basis if there are other influences, such as the economy or special events. The question regarding the existence of a statistical structure (or the lack of it) can be answered by correlation analysis. With financial market data, where something can change every day, every hour and every second, it looks quite different.

3.4 Detect & Visualize Anomalies

What anomalies can be detected in the available data and what reasons can be ascribed to them? In general, it can be observed that users generate stable daily curves with regard to the number of clients, access rates and response time behavior. Strong deviations indicate interaction by automata. These automata can be crawlers, DOS or DDOS attacks, or the preliminary stages of such an attack. This also includes the exploration of limits of security mechanisms.

It is almost impossible to detect all conceivable variants and to classify them without errors. The focus here is on the detection of non-human access, but without classifying it more precisely. The latter is done by an expert. For example, a technical malfunction as the cause must be excluded (avoidance of false alarms).

The deviation from the daily curve can have different forms that should be detected. A visualization represents an easily accessible form of displaying the results of those anomalies. We want to be capable of recognizing frequency or periodicity patterns too.

3.4.1 Spike Detection

Goin and Ahern [GA19] test the performance of several spike detection methods in a simulation parameterized by monthly violence rates in nine California cities. For their research the authors randomly inserted spikes into the series. They tested ARIMA modeling, Kalman filtering and smoothing, wavelet modeling with soft thresholding, and an iterative outlier detection method. Goin and Ahern revealed that the Kalman filter and smoothing procedure had the best overall performance.

In our case, spikes are frequently occurring patterns. Two possible causes are either delay situations in the backend such as garbage collections in the server application (1) or actors that frequently check/scan URLs for accessibility and version (2). It is not easy to identify these spikes within time series data that contain a large amount of noise.

3.4.2 Outlier Detection

Not each spike is an outlier. We assume that outliers do not play a significant role in the data. If we want to do outlier detection, we should unconditionally integrate it into the model. This can be efficiently done in VAR models. A diagnostic statistic for outlier detection is DFBETAS [BKW04]. This tells us, for each estimated coefficient and each individual observation, by what percentage (related to the standard deviation of the coefficient) each coefficient would change when removing the observation out of the sample. And the great thing about this is that we would actually have to estimate 100 times for a sample of 100 observations, so we would always have to delete one observation at a time and then re-estimate. But in fact this is not necessary. Linear algebra makes it possible to calculate all DFBETA values for all observations and the entire sample with one formula [BKW04].

In general, human interaction leads to uniformly distributed metrics (number of clients per time unit, number of requests per time unit, number of cities per time unit). This results in a daily and a weekly cycle. In addition, there are weak seasonal fluctuations or effects of specialized business actions. Outliers clearly indicate automatic actors and require a deeper analysis. We must exclude that for example a technical malfunction or maintenance is the cause.

3.4.3 Change Point Detection

Aminikhanghahi and Cook [AC17] define change points as "abrupt variations in time series data", which may represent "transitions that occur between states". In their survey article "A Survey of Methods for Time Series Change Point Detection" the authors introduce, categorize, evaluate and compare methods including supervised and unsupervised algorithms for detecting such change points in time series. Automatic actors like bots try to operate as

inconspicuously as possible. To avoid spikes, they choose a procedure that only slightly increases the "base load" without leaving the usual metric ranges in the given context. This increase is limited in time. Both at the beginning and at the end of the action, this change can be observed from the change points in the metrics.

3.4.4 Turning Point Detection

A turning-point date is "occurring when a swing in one direction ends and a swing in the other direction begins", see Chin, Geweke, and Miller [CGM00]. With their article "Predicting Turning Points" Chin et al. developed a probit model that estimates the probability of a turning point. Wecker [We79] extends the linear least squares (LLS) technique in order to allow calculating the distribution of the turning points in a time series. Deng [LD07] introduced a Machine Learning approach based on a Radial Basis Function Network (RBFN) for predicting turning points which occur in chaotic financial time series.

Visible Trend alterations (which occur at a turning point) allow conclusions to be drawn about actors. However, the professional evaluation of such a turning point is not possible without extensive further analyses. So far, we identified those trend alterations only a few times in an observation period of several months.

3.4.5 Visualizing the Result

Different options are available for visualizing the results. At the current stage we are not yet able to state which option will be most practical. However, we will list and demonstrate several variants that have turned out to be feasible approaches for our use cases.

Visualizations of time series for telemetry data includes several metrics of an endpoint measured over a period of time. Figure 3 displays three related metrics over a period of approx. 20 hours: requests per minute (green), cities per minute (blue) and unique IP addresses per minute (orange). The timestamp "05-30 22" is read as "30th of May 2020 at 10:00 p.m.". The rapid increase of IP addresses indicates requests from automatic actors. The simultaneous increase in cities shows that the increase in IP addresses is not caused by only one location but is geographically distributed. This points to accesses by a botnet.

This raises the next questions: can the strong increase be assigned locally, e.g. country or city? In our given case, it can be seen that the increase is distributed over approximately 60 locations (cities). Consequently, it is possible that the requests origin from a distributed automatic actor (botnet). A further analysis is required to determine which IP addresses were involved in this access – do they origin from the same provider or from the same region?

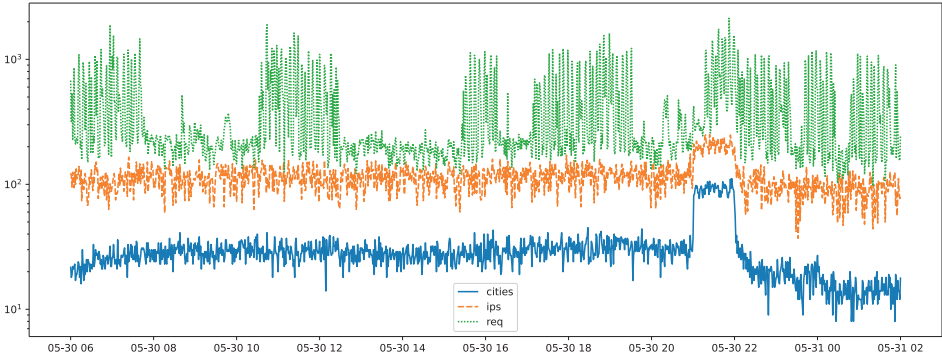


Fig. 3: Three related metrics over a period of approx. 20 hours

Visualization to discover correlations is best done with heat map representations. From a purely analytical point of view, heat maps are easy to read and offer a sufficiently high information density if coded appropriately. Figure 4 provides an example of a heat map for the number of clients per city per interval. This representation can be further supported by a filtered correlation as given by Figure 5.

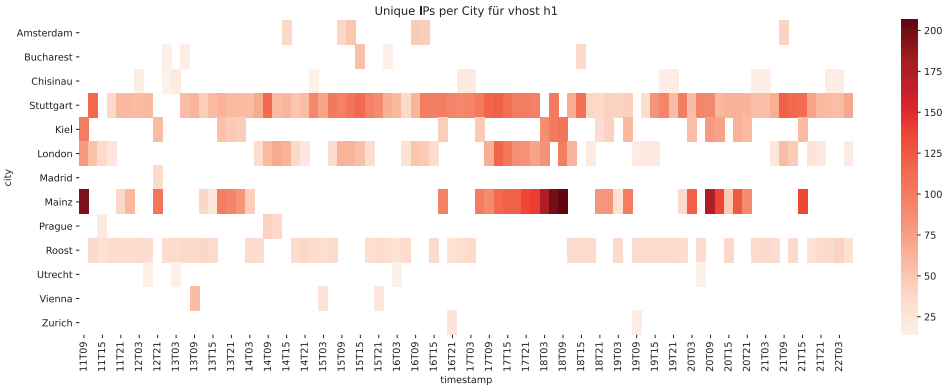


Fig. 4: Heat map with the number of clients per city



Fig. 5: Correlation as filtered visualization of Figure 4

The high density of clients from three cities each in the same time interval provides an indication of where we should focus using subsequent analysis. A malfunction as the root cause can almost be ruled out. Based on the given data, we cannot make a more precise statement as to whether a crawler or a DoS/DDoS attack is the cause. However, legitimate access by users (for example due to training) can be excluded at such a time of day.

3.5 Optimizing the Model

There are many approaches for ensuring that the model remains up-to-date as new data becomes available. The sledgehammer method would be a frequent refitting approach, but this is very computation-intensive. More efficient is a weighting function where younger data has more weight than older data. A best practice is using a Kalman filter that distinguishes between seasonal lags, hourly lags etc.

4 Transferability and Limitations

Although we have developed our solution for detecting IT service anomalies in the specific context of Germany's Federal Employment Agency, we believe our solution is applicable to many organizations around the world. Reports indicate an increasing number of organizations, big and small, and their online services are subject to DDoS attacks every year [Hu21]. As more organizations allow remote working and learning using consumer-grade devices with fewer security protections, such attacks are expected to increase even more in the future

[Hu21]. And because the patterns of DDoS attacks are similar, our method can be applied to other cases as well, irrespective of how IT service access logs are collected and structured. This is because the mathematics behind our approach is independent from the procedure of log collection and log aggregation. Furthermore, our approach is technology-independent - the algorithms presented in this paper can be implemented in any Python libraries, and even with other programming languages.

One limitation of the current solution is that the model has been developed with a data set that includes just one holiday; therefore the treatment of holiday effects and of holiday-proximal weekend combinations have not yet been incorporated in the model. Another limitation is that our solution uses Elastic Stack for log collection and aggregation; even if this is a well-known and widely-used technology, different companies may prefer a different technology. Last, but not least, the model has to be updated as new data become available, and it is possible that patterns of normal behavior and anomalies will change over time; at present, we do not know if the model developed here will be appropriate for future IT service access patterns or if different data and model decisions, or even if different approaches will be required.

5 Conclusions and Future Research

We developed an architecture for detecting IT infrastructure anomalies at Germany's Federal Employment Agency. We used a Vector Autoregressive model allowing for seasonality at multiple frequencies (SVARX), which has turned out to be robust. While this research is still exploratory in nature, the case study organization, FEA, has already benefited from this solution by identifying previously undetected attack patterns. We further plan to extend the solution and connect it to the incident (INC) process using a technical platform that will include log and time series databases with raw data, Python-based implementation (including all necessary libraries for interfaces, data processing (ETL) and analyses), and software deployed in the form of agents as container instances on Kubernetes.

Future research can investigate different time series analysis methods in terms of performance, accuracy and maintainability. For example, in section 3.3 we argued that the Holt-Winters Method is not suitable for our problem, since it is not a multivariate model. By future research this the Holt-Winters Method might be extended to a multivariate model. A straightforward extension of exponential smoothing to multivariate data is, for bivariate time series $(s_{1t}, s_{2t})'$:

$$\begin{pmatrix} s_{1t+1} \\ s_{2t+1} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_{1t} \\ x_{2t} \end{pmatrix} + \begin{pmatrix} 1 - a_{11} & 0 \\ 0 & 1 - a_{22} \end{pmatrix} \begin{pmatrix} s_{1t} \\ s_{2t} \end{pmatrix}$$

or with $s_t = (s_{1t}, s_{2t})'$:

$$s_{t+1} = Ax_t + Bs_t \quad \text{with} \quad B = \begin{pmatrix} 1 - a_{11} & 0 \\ 0 & 1 - a_{22} \end{pmatrix}$$

The stability condition $0 \leq \alpha \leq 1$ applying in the univariate case is to be adapted. We expect that all eigenvalues of A must have an absolute value (a modulus) less than 1. Such questions an a deeper evaluation of this approach form a basis for future research.

Future research can also study the architecture for a detection system capable of handling other types of anomalies - not just DDoS - either other external malicious attacks or internal operational problems. The appropriate frequency of optimizing the model given potential changes in anomaly patterns is an open question as well. A potentially fruitful area of research could be collaborative and distributed detection of anomalies for cloud-based IT systems [YUS19]. Furthermore, as organizations start to integrate a variety of heterogeneous Internet of Things devices into their infrastructure, future research can investigate the types of anomalies affecting these emerging infrastructures and how to identify them using the enormous amounts of data generated in Internet of Things environments [KA21].

Literatur

- [AC17] Aminikhanghahi, S.; Cook, D. J.: A survey of methods for time series change point detection. *Knowledge and Information Systems* 51/2, 2017.
- [BKW04] Belsley, D. A.; Kuh, E.; Welsch, R. E.: *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. John Wiley & Sons, 2004, ISBN: 0-471-69117-8.
- [CGM00] Chin, D.; Geweke, J.; Miller, P.: *Predicting Turning Points*, Techn. Ber. Research Department Staff Report 267, Federal Reserve Bank of Minneapolis und Congressional Budget Office, Juni 2000.
- [CM09] Cowpertwait, P. S.; Metcalfe, A. V.: *Introductory Time Series with R*. Springer, 2009, ISBN: 978-0-387-88697-8.
- [Fe19] Fernandes, G.; Rodrigues, J. J.; Carvalho, L. F.; Al-Muhtadi, J. F.; Proença, M. L.: A comprehensive survey on network anomaly detection. *Telecommunication Systems* 70/3, S. 447–489, 2019.
- [GA19] Goin, D. E.; Ahern, J.: *Identification of Spikes in Time Series*. *Epidemiologic Methods* 8/1, 2019.
- [Hu21] Hummel, R.: *DDoS in a Time of Pandemic*, Techn. Ber. NETSCOUT Threat Intelligence Report, Issue 6, NETSCOUT, Apr. 2021.
- [KA21] Khraisat, A.; Alazab, A.: A critical review of intrusion detection systems in the internet of things: techniques, deployment strategy, validation strategy, attacks, public datasets and challenges. *Cybersecurity* 4/1, S. 1–27, 2021.

- [Kh19] Khalaf, B. A.; Mostafa, S. A.; Mustapha, A.; Mohammed, M. A.; Abdualah, W. M.: Comprehensive review of artificial intelligence and statistical approaches in distributed denial of service attack and defense methods. *IEEE Access* 7/, S. 51691–51713, 2019.
- [KKB17] Kaur, P.; Kumar, M.; Bhandari, A.: A review of detection approaches for distributed denial of service attacks. *Systems Science & Control Engineering* 5/1, S. 301–320, 2017.
- [LD07] Li, X.; Deng, Z.: A Machine Learning Approach to Predict Turning Points for Chaotic Financial Time Series. In: 19th IEEE International Conference on Tools with Artificial Intelligence. IEEE Computer Society, 2007.
- [Lü05] Lütkepohl, H.: *New Introduction to Multiple Time Series Analysis*. Springer, 2005, ISBN: 3-540-40172-5.
- [MHS19] Moustafa, N.; Hu, J.; Slay, J.: A holistic review of network anomaly detection systems: A comprehensive survey. *Journal of Network and Computer Applications* 128/, S. 33–55, 2019.
- [Sa18] Salo, F.; Injadat, M.; Nassif, A. B.; Shami, A.; Essex, A.: Data mining techniques in intrusion detection systems: A systematic literature review. *IEEE Access* 6/, S. 56046–56058, 2018.
- [Wa19a] Wan, R.; Mei, S.; Wang, J.; Liu, M.; Yang, F.: Multivariate Temporal Convolutional Network: A Deep Neural Networks Approach for Multivariate Time Series Forecasting. *Electronics* 8/8, 2019, URL: <https://www.mdpi.com/2079-9292/8/8/876>.
- [Wa19b] Wang, Y.; Liu, Z.; Hu, D.; Zhang, M.: Multivariate Time Series Prediction Based on Optimized Temporal Convolutional Networks with Stacked Auto-encoders. In (Lee, W. S.; Suzuki, T., Hrsg.): *Proceedings of The 11th Asian Conference on Machine Learning, ACML 2019, 17-19 November 2019, Nagoya, Japan*. Bd. 101. *Proceedings of Machine Learning Research*, PMLR, S. 157–172, 2019, URL: <http://proceedings.mlr.press/v101/wang19c.html>.
- [We79] Wecker, W. E.: Predicting the Turning Points of a Time Series. *The Journal of Business* 52/1, S. 35–50, 1979, URL: <https://www.jstor.org/stable/2352662>.
- [YUS19] Yusof, A. R.; Udzir, N. I.; Selamat, A.: Systematic literature review and taxonomy for DDoS attack detection and prediction. *International Journal of Digital Enterprise Technology* 1/3, S. 292–315, 2019.
- [ZW06] Zivot, E.; Wang, J.: *Modeling Financial Time Series with S-PLUS*. Springer, 2006, ISBN: 978-0387-27965-7.

Realizing Robotic Process Automation Potentials: An Architectural Perspective on a Real-life Implementation Case

Konstantin Ritschel¹, Adina Stenzel², Christian Czarnecki³ and Chin-Gi Hong⁴

Abstract: The initial idea of Robotic Process Automation (RPA) is the automation of business processes through a simple emulation of user input and output by software robots. Hence, it can be assumed that no changes of the used software systems and existing Enterprise Architecture (EA) is required. In this short, practical paper we discuss this assumption based on a real-life implementation project. We show that a successful RPA implementation might require architectural work during analysis, implementation, and migration. As practical paper we focus on exemplary lessons-learned and new questions related to RPA and EA.

Keywords: Robotic Process Automation, Enterprise Architecture, Implementation Case.

1 Introduction

Robotic Process Automation (RPA) comprises a broad range of concepts to automatically execute processes by software robots [CF21]. The main ideas are the 1:1 automation of human activities and the handling of existing application systems through their presentation layer, e.g., [ABH18], [SDC19]. The implementation of RPA is dominated by standard software systems [En+20]. The ease of implementation combined with the usage of existing, unchanged application landscapes are essential characteristics of the value proposition. Manual activities are learned and automated by software robots that emulate formerly human data entries on the application's presentation layer. Software vendors even suggest a pure implementation by the business side while technical knowledge (e.g., script programming) is not required [WLC15]. The successful implementation of RPA is documented in various practical cases, e.g., [Cz+21], [He+20], [HHF19]. Furthermore, the combination of RPA with cognitive capabilities is proposed to realize an automation of complex processes with decision points [AMM19].

An essential question is related to the matter on how those RPA potentials are realized in practice, e.g., [CF21], [He+20]. It is common sense that (1) the selection of an appropriate RPA standard software, and (2) the selection of the right processes to be

¹ Detecon International GmbH, Sternengasse 14-16, 50676 Köln, konstantin.ritschel@detecon.com

² Detecon International GmbH, Winterfeldtstraße 21, 10781 Berlin, adina.stenzel@detecon.com

³ FH Aachen, Eupener Straße 70, 52066 Aachen, czarnecki@fh-aachen.de

⁴ Detecon International GmbH, Hahnstraße 43, 60528 Frankfurt/Main, chin-gi.hong@detecon.com

automated are crucial points. In addition, the management and governance of RPA projects contain some specifics, wherefore Herm et al. [He+20] have proposed an RPA project implementation framework. From an architectural perspective, the initial idea of RPA assumes no changes of the existing enterprise architecture (EA) which might lead to the conclusion that no architectural work is required. This assumption has already been discussed by Auth et al. [ACB19] who propose a differentiated view based on the complexity of the RPA solution.

In this short, practical paper we continue the discussion of the required architectural work for a successful RPA implementation. Based on a real-life implementation case, we derive lessons-learned from an architectural perspective and propose required EA elements mapped to different stages of an RPA implementation project. As this paper at hand has a practical focus, we discuss exemplary insights based on an actual real-life implementation with the purpose of identifying new questions and research opportunities related to the field of EA and RPA. Our research is based on our own case studies and observations in practice.

2 RPA from an Architectural Perspective

The development, optimization, or implementation of information systems is a complex task that influences various parts of an enterprise. The modelling of those parts is supported by EA concepts providing structures, templates, and methods for the overall design and description of an enterprise [WS07]. Typically, those concepts are generic and independent from concrete industries. From a methodological perspective, EA can be seen as a structural element that supports the overall solution design. The scope of EA is normally broad, and often covers the whole or major parts of the enterprise [WF07]. Therefore, there is not one single representation of a concrete architecture, but various different views according to purpose and target audience.

The general idea of RPA is the automation of activities by emulating the manual usage of application systems through a software robot that acts exactly like the formerly human user, e.g., [ABH18], [SDC19]. The whole interaction between the RPA system and the application system(s) is realized through existing user interfaces. Hence, RPA can be understood as a lightweight solution that does not change the existing application architecture. It distinguishes itself from traditional approaches to automation, which usually require technical and/or organizational changes such as integration. With respect to an existing EA, RPA impacts the processes and applications layer (cf. Fig. 1) while the level of change might be limited to a pure automation and usage of existing elements. However, Auth et al. [ACB19] already propose that further impact on the EA should be considered, e. g., the implementation of new processes by RPA and the integration of different RPA systems.

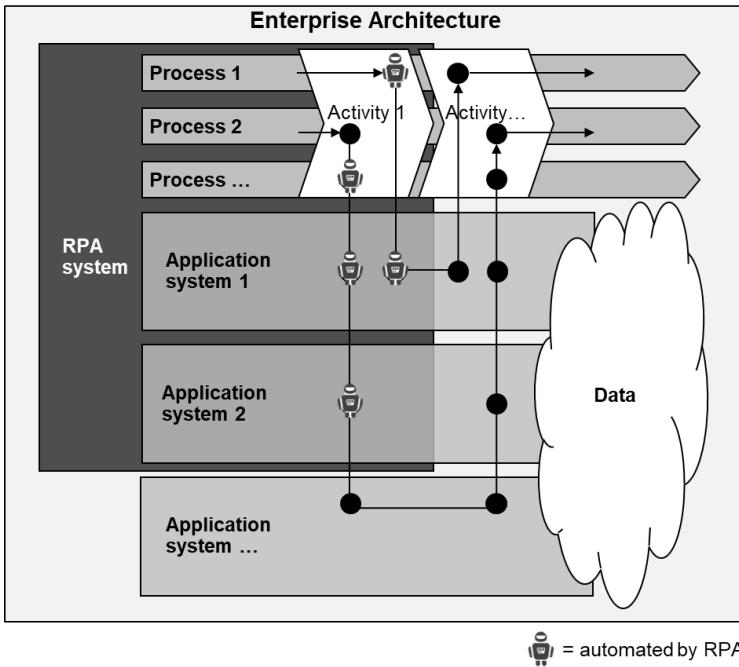


Figure 1: Impact of RPA on EA layers (according to [ACB19])

3 Practical Implementation Case

The following application case has been implemented at *Deutsche Telekom Service Europe (DTSE)*. The description is based on the authors' observations as well as official project documents. The case has been already discussed in [Cz+21] as part of broader questions of cognitive RPA and process digitalization. In this paper, the discussion is extended by bringing in the architectural perspective.

Use Case	Process	Description	Application	Frequency
1 Electricity consumption data extraction	finance	<ul style="list-style-type: none"> Download of relevant invoices from SAP Extraction of consumption data per consumption point through text field recognition Booking of data in SAP 	<ul style="list-style-type: none"> free text, SAP MM 	<ul style="list-style-type: none"> daily
2 Transfer of offer data in SAP	purchasing	<ul style="list-style-type: none"> Extraction of data from offer documents AI-based text recognition Start of the SAP transaction and entry of fields with information from offer documents 	<ul style="list-style-type: none"> PPT, PDF, DOC, SAP MM 	<ul style="list-style-type: none"> daily
3 Shopping bot	ordering	<ul style="list-style-type: none"> Bot creates shopping cart based on Excel lists for the procurement of smartphones for employees 10,000 of 25,000 shopping carts created within 9 months 	<ul style="list-style-type: none"> MS Outlook, SAP MM 	<ul style="list-style-type: none"> daily

Figure 2: RPA implementation cases (excerpt according to [Cz+21])

DTSE has started an initiative to automate existing processes using RPA. The major objectives are efficiency increase and cost reduction. In a first step, a broad variety of different operational use cases were identified (cf. Fig. 2). For each case, the corresponding processes and applications were considered. Furthermore, the frequency of the process execution as well as the related effort were examined. As of now approximately 170 software robots were implemented successfully.

It follows a detailed description of the implementation of case 1 “electricity consumption data extraction”. In the as-is situation the process is highly manual [Cz+21]: (1) The data is provided by the energy provider as paper invoice, which is (2) scanned and uploaded to SAP by an accountant. (3) A dedicated expert downloads the invoice from SAP, extracts the consumption values per consumption point, and (4) books the information in SAP. Those four manual process steps were automated by using RPA systems (cf. Fig. 3). First, the download of invoices from SAP is handled by a software robot using *Blue Prism*. Second, the optical character recognition (OCR), the data extraction, and the text field recognition are performed by a software robot using *ABBYY Flexicapture*, which is a document workflow platform that captures and classifies unstructured data.

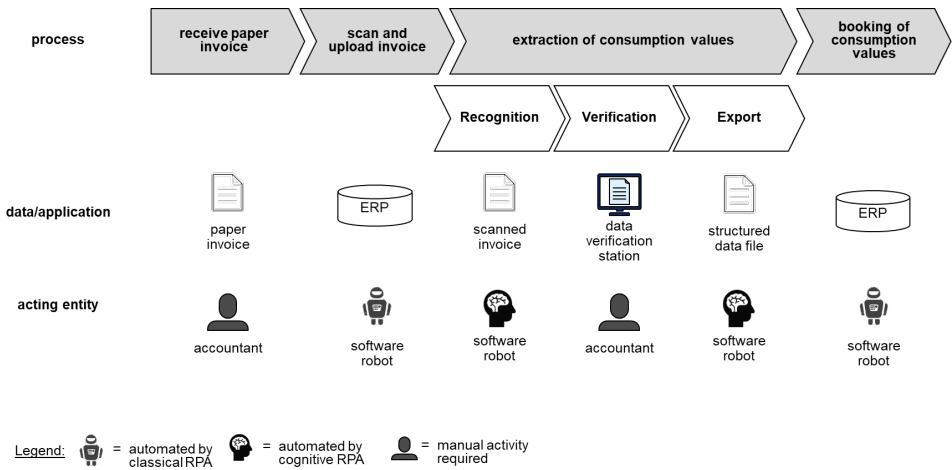


Figure 3: Real-life case example

On a high level, it seems that the existing manual activities are 1:1 transferred to a software robot. However, on a more detailed level the formerly manual extraction of consumption values is newly structured into (1) recognition, (2) verification, and (3) export. All three detailed activities were not existent in the as-is situation. For some invoices a manual verification is still required. A new data verification station was implemented to support this step. The existing ERP system remained unchanged.

The importance of this new verification step is shown in Table 1. Without the verification only 21% of all invoices would have been processed correctly which obviously would not justify the automation. Each invoice includes on average 16 line items and 114 data fields. On the detailed level of data field, only 11% of verification is required. Hence, the proper realization of the verification step is a major success factor of the whole case. As a result, the manual effort as well as the processing time was decreased. The handling of an invoice was on average reduced to 19.5 seconds, which results in a total effort of 76 hours per year and a time reduction of 94%.

Data entity	Miss rate	Fully automated recognition	Verification required
Document	0%	21%	79%
Line item	0%	48%	52%
Data field	1%	88%	11%

Table 1: Accuracy of case example

4 Lessons Learned from an Architectural Perspective

The initial thesis of RPA is that the existing application architecture remains unchanged. Software vendors communicate the ease of implementation as a major advantage of RPA. They even regard an implementation without IT expertise. In this context, self-learning software robots implemented by the business side are a common vision. In literature, the selection of the right processes is discussed [Ho+21], [RO21]. Some authors propose an automatic process selection, for example, based on process mining [Aa21]. Furthermore, the management and governance of RPA projects is expected to require further guidance [He+21].

The practical insights described here provide first indications that a successful RPA implementation requires detailed conceptual and architectural work:

1. *Identification of relevant processes* should be based on an as-is analysis including the corresponding applications, the relevant data and the frequency of process execution.
2. *Development of RPA cases* is comparable to the use cases found in software engineering projects. They include a high-level description of a possible target solution considering the as-is situation, improvement potentials as well as technical restrictions.
3. *Evaluation of RPA cases* requires performance indicators and quantities of the as-is process as well as a valid estimation of target figures which should be based on technical capabilities.
4. *Development of RPA solution* includes the design of target processes as well as the technical solution design which might include new process steps (cf. the verification step explained in Section 3).
5. *Technical realization and roll-out* might require implementation work that goes beyond a pure RPA standard software system (cf. the data verification station described in Section 3).
6. *Operations and management of RPA system(s)* should include a continuous documentation of implemented cases, the corresponding applications and the relevant data objects.

Hence, typical EA elements are required for an RPA project. The upper part of Fig. 4 shows the RPA project implementation framework according to [He+20]. The lower part contains a first proposal of EA elements that are mapped to the different stages of an RPA project.

The identification and evaluation of RPA cases require a documentation of the as-is processes, application landscape as well as the corresponding data models which are all

part of typical EA models. Furthermore, RPA use cases could be based on existing approaches of use case design, and the performance indicators should be the result of process or EA management initiatives. Even though the RPA implementation can be understood as a lightweight concept, still target design activities – including processes, application landscape, performance estimations, and solution design – are required, and do not differ from typical conceptual EA design. Furthermore, the migration should be planned, and due to the common number of software robots and different cases, this part might be quite complex. Typical for RPA is a high number of different cases on an operational level that are continuously adapted to changed requirements. Hence, an ongoing documentation of implemented RPA cases is required. Furthermore, in each RPA case different software robots handle different application systems using their user interfaces which requires a careful documentation of the technical restrictions and assumptions including a continuous change log [SE21].

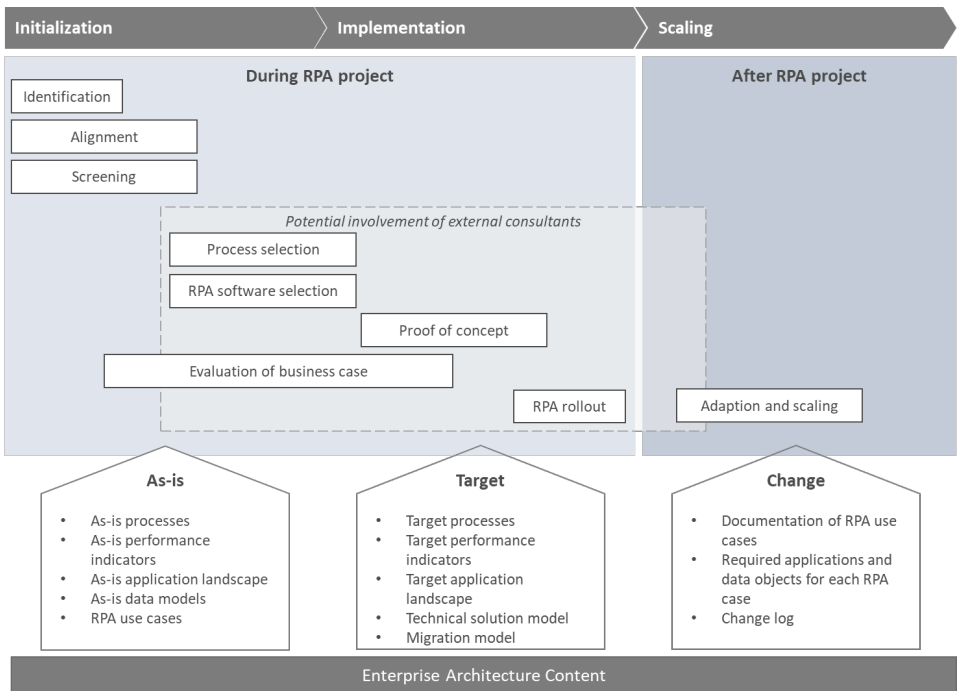


Figure 4: EA elements during RPA implementation (upper part from [He+20])

5 Conclusion and Outlook

The ease of implementation and unchanged existing application systems are major value propositions of RPA. However, this does not mean that no information systems

modelling, or architectural design are required. This paper illustrates a real-life RPA case that automates the extraction of data and booking in SAP. The case demonstrates that on a detailed level processes were changed, and a new data verification station was implemented. In this concrete example, the success was highly related to the newly developed verification steps. The further discussion of the lessons learned provides indications that various EA elements might be required during all stages of an RPA project, e.g., the as-is application landscape during the identification, and a technical solution model during the roll-out. It could be assumed that expertise in the analysis and design of processes and applications is helpful for a successful RPA implementation.

This paper contributes practical insights to the discussion of the required architectural work for a successful RPA implementation. Further work is necessary to validate the results by comparing them with lessons learned from other project cases and to derive general recommendations. Furthermore, the increasing complexity of RPA might require further guidance on general methods and models for a successful implementation.


Bibliography

- [ABH18] van der Aalst, W. M. P.; Bichler, M.; Heinzl, A.: Robotic Process Automation. *Bus Inf Syst Eng* 60(4), pp. 269-272, 2018.
- [Aa21] van der Aalst, W. M. P.: Process mining and RPA: How to pick your automation battles? In: Czarnecki, C., Fettke, P. (eds): *Robotic Process Automation: Management, Technology, Applications*. De Gruyter, pp. 223-239, 2021.
- [AMM19] Agostinelli S, Marrella A, Mecella M: Research Challenges for Intelligent Robotic Process Automation. In: Di Fran-cescomarino C, Dijkman R, Zdun U (eds) *Business Process Management Workshops*. Springer International Publishing, Cham, pp 12–18, 2019.
- [ACB19] Auth, G., Czarnecki, C., Bensberg, F.: Impact of robotic process automation on enterprise architectures. In: *50 Jahre Gesellschaft für Informatik –Informatik für Gesellschaft (Workshop-Beiträge)*. Gesellschaft für Informatik e.V., Bonn, pp. 59–65.
- [CF21] Czarnecki, C., Fettke, P.: Robotic process automation: Positioning, structuring, and framing the work, In: Czarnecki, C., Fettke, P. (eds): *Robotic Process Automation: Management, Technology, Applications*. De Gruyter, pp. 3-24, 2021.
- [Cz+21] Czarnecki, C., Hong, C.-G., Schmitz, M., Dietze, C.: Enabling Digital Transformation through Cognitive Robotic Process Automation at Deutsche Telekom Service Europe. In: Urbach, N.; Röglinger, M. (eds): *Digitalization Cases*. 2nd Edition. Springer International, 2021.
- [En+20] Enriquez JG, Jimenez-Ramirez A, Dominguez-Mayo FJ, Garcia-Garcia JA: Robotic Process Automation: A Scientific and Industrial Systematic Mapping Study. *IEEE Access* 8:39113–39129. 2020.
- [He+20] Herm, L.V., Janiesch, C., Helm, A., Imgrund, F., Fuchs, K., Hofmann, A.,

- Winkelmann, A.: A consolidated framework for implementing robotic process automation projects, in Proceedings of the 18th International Conference on Business Process Management, Lecture Notes in Computer Science Vol. 12168, (Sevilla), pp. 471–488, Springer, 2020.
- [He+21] Herm, L.V., Janiesch, C., Steinbach, T., Wüllner, D.: Managing RPA implementation projects: A framework applied at SYSTHEMIS AG. In: Czarnecki, C., Fettke, P. (eds): Robotic Process Automation: Management, Technology, Applications. De Gruyter, pp. 27-46, 2021.
- [Ho+21] Hofmann, A., Prători, T., Seubert, F., Wanner, J., Fischer, M., Winkelmann, A.: Process selection fo RPA projects: A holistic approach. In: Czarnecki, C., Fettke, P. (eds): Robotic Process Automation: Management, Technology, Applications. De Gruyter, pp. 77-90, 2021.
- [HHF19] Houy, C; Hamberg, M.; Fettke, P.: Robotic Process Automation in Public Administrations. Digitalisierung von Staat und Verwaltung. Fachtagung Verwaltungsinformatik (FTVI) und Fachtagung Rechtsinformatik (FTRI), LNI 291, pp. 62-74, 2019.
- [Ji+19] Jimenez-Ramirez A., Reijers H.A., Barba I., Del Valle C. (2019) A Method to Improve the Early Stages of the Robotic Process Automation Lifecycle. In: Giorgini P., Weber B. (eds) Advanced Information Systems Engineering. CAiSE 2019. Lecture Notes in Computer Science, vol 11483. Springer, Cham.
- [ReOp21] Rechberger, S., Oppl, S.: Selecting processes for RPA: A study of relevant key process indicators in the finance industry. In: Czarnecki, C., Fettke, P. (eds): Robotic Process Automation: Management, Technology, Applications. De Gruyter, pp. 91-110, 2021.
- [SE21] Sultanow, E., Chircu, A., Plath, R., Friedmann, D., Merscheid, T. & Sharma, K.: AI evolves IA, In: Czarnecki, C., Fettke, P. (eds): Robotic Process Automation: Management, Technology, Applications. De Gruyter, pp. 349-376, 2021.
- [WF07] Winter, R.; Fischer, R.: Essential layers, artifacts, and dependencies of enterprise architecture. Journal of Enterprise Architecture, 2, pp. 7-18, 2007.
- [WLC15] Willcocks, L.; Lacity, M.; Craig, A.: The IT Function and Robotic Process Automation. The Outsourcing Unit Working Research Paper Series, Paper 15/05, The London School of Economics and Political Science, 2015, http://eprints.lse.ac.uk/64519/1/OUWRPS_15_05_published.pdf, accessed: 28/05/2019.
- [WS07] Winter, R.; Sinz, E. J.: Enterprise architecture. Information Systems and e-Business Management, 5, pp. 357-358. DOI: 10.1007/s10257-007-0054-0. 2007.

Application-Architecture of an ERP-System for the era of Industry 4.0

The architecture of a new ERP-System for the industry 4.0 based the concept of a Digital Control Twin

Wilmjakob Johannes Herlyn  ¹


Abstract: In this paper we design an architecture of a future ERP-System that bases on the concept of a Digital Control Twin as its kernel. ‘Architecture’ in this paper means how the different applications of an ERP-System work together in combination with the Operation Control System of the Computer. The architecture is designed in the way that a lot of ERP-Applications are integrated in only one System. Nowadays the ERP-applications are separated modules like Master Production Scheduling and Manufacturing Execution or Material Requirement Planning, Warehouse Management, Transportation Control or JIT-Call-Offs which causes a loss of data consistency and accuracy and a lack of response-time and adjustment consequently. Classical ERP-Systems bases on the Waterfall-Model with no direct recursion of results to a preceding module. This architecture is no more suitable for the capabilities of industry 4.0 e. g. to exploit the huge amount of permanent available data and to react instantly on disturbances in production, transportation, storing or handling. The proposed architecture is a kind of Design Science Research approach that bases on the knowledge and experience in one of the biggest automotive OEM in the departments of IT, Purchasing, Production, Logistic and Sales.

Keywords: Enterprise Resource Planning, Digital Control Twin, Production and Material Flow, Data Gravity Centre, Data Grid, Counting Point, Digital Trigger, Digital Shadow

1 Introduction

The innovations of industry 4.0 create new opportunities to change the classical ERP-System and its applications modules. Existent ERP-Systems are not able the use the high amount of data, which are gathered by new techniques and devices like RFID, QR-Code, and cyber-physical objects (CPO). The extremely high-speed of computing and data transmission (5G and expected 6G) requires a new kind of architecture of the ERP-System itself. This approach bases on the concept of a Digital Control Twin (DCT) which was inspiring by the DT concept of Grieves/Vickers [Gr14] which does not mean a ‘simple’ digital representation of a physical object or process [Di17]. It includes the interaction between the real and virtual twin whereby the virtual twin leads the real twin which is a

¹ Otto-von-Guericke University Magdeburg, Institute of Logistics and Material Handling Systems,

Universitätsplatz 2, Magdeburg, 39106, wilm.herlyn@ovgu.de  <https://orcid.org/0000-0003-4522-5417>

paradigm shift in business informatics. For this we need a common Repository for all ERP-Applications which are nowadays separated and connected by interfaces and data-exchange processes [LL07, He11, KU13, L013, Sc16, KT18].

2 Concept of the Digital Control Twin – fundament of ERP-System

Most of the exiting Digital Twin in logistics has the focus on technical aspects of production and material flow control [Di19]. The concept of Digital Control Twin which is not engineering but order-oriented and use the process data of the technical objects [SAP]. There is still lack in research and development of DT-Concepts for business process and especially for the value-added chains [Mi18, ZH20, ZH21, P120, P121].

2.1 The Application Architecture of the Digital Control Twin

The application architecture of the DCT bases on the three main pillars: 1 Reality 2. Repository 3. Regulation and in addition, and as a fourth pillar an Application System Controller (ASC) is needed to control all ERP-applications (fig. 2.1).

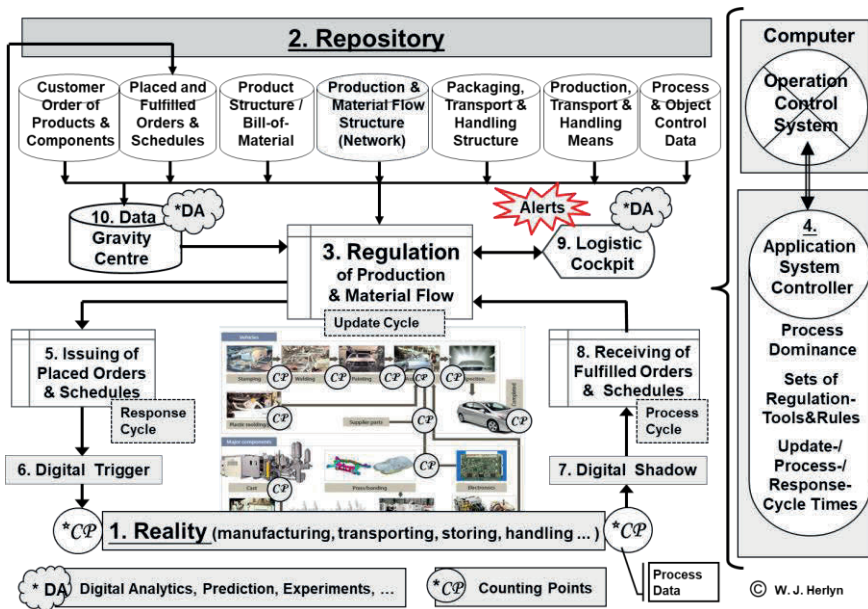


Fig. 2.1 Architecture of an ERP-System for industry 4.0

As first we explain the constituting elements in short, thereafter we focus on regulation as the brain of the ERP-System including a Logistic Cockpit for supervising and a Data Gravity Centre for Data Analytics, process prediction and experimental configurations.

1. Reality: Reality is the actual process of production and material flow (PMF) that must be planned, controlled, and monitored by an ERP-System and covers all activities of manufacturing, transportation, storing, handling etc. of final products, and required components. The reality is mapped by ***Counting Points** that are used for the order-oriented regulation of the process. Objects of regulation are at first final products and required components, but could be all kinds of material flow items (MFI) of transporting and loading means like trucks, trains, bins, container etc. that use material handling equipment and infrastructure like streets and railways or assembly lines and warehouses.

7. Digital Shadow: Digital Shadows (DS) are actual acquired data of real processes that reflect the state, time, locations, and quantity of an identified MFI. The data can be acquired in different ways, by different devices and techniques, it can be entered by hand or scanned by Bar-Code or QR-Code, gathered by an active or passive RFID-Chip, or transferred by a Cyber-Physical Objects which can store, process, share and transmit data itself. For the DCT it does not matter which kind of data acquisition technique or device is used, its decisive only that the acquired process data is available at the time when it is needed for the specific purpose and task of the application. Process data are generated by physical devices or humans, but Digital Shadows are only those process data which are acquired at defined Counting Points for a predefined process cycle time.

8. Receiving Fulfilled Orders and Schedules: Digital Shadows comes from different processes and devices and are often generated extremely fast and in a huge amount. For an ERP-Application it makes no sense to process the acquired data immediately because the processing depends on specific application task and response time. Therefore, a kind of 'filter' is needed to condense the actual data and to pass it to the ERP-Application in an appropriate manner which is controlled by the specific **Process Cycle-Time**.

2. Repository: Repository means the digital mapping of the reality and covers all kinds of Transaction-Data, Master-Data and Control-Data for all ERP-applications. Transaction-Data means all kinds of order-oriented data like customer order or manufacturing orders and schedules like Master Production Schedule (MPS) or Transportation Schedules. Master-Data means all order-independent data and covers the structure of product and material flow, products, packaging, transport, and handling. Control-Data means all kinds of parameters to control the process and activities of production and material flow. The Repository covers not only material flow items (MFI) like final products and its component, but also means for packaging, transportation and handling means like container, bins, trucks, tigger-trains, shelves etc.

3. Regulation: Regulation covers all kinds of planning, scheduling, monitoring, and balancing of MFI especially generating of schedules and calculation of orders. The main task of regulation is to ensure that the right MFI are available in the right quantity, at the right time and the right place. The exact order quantities of products and components must be determined for the entire network. The main regulation method bases on the principal of Closed-Loop-Control; whereby **Placed Orders and Schedules** are the target values and **Fulfilled Orders and Schedules** are the actual values. Deviations of target and actual values are evaluated and balanced by the regulation software. The execution of an

applications is controlled by the Application System Controller, where rules and instructions are explicitly defined and stored which is a completely new ERP-function. The execution of each application is controlled by a specific **Update Cycle-Time** which depends on the specific purposes and task and the level of PMF-Structure.

5. Issuing Placed Orders and Schedules: Results of an ERP-Application are placed orders and planned schedules, but it is not always possible or make no sense to transmit the results immediately to the devices, places or partners, where they are needed. Orders and schedules must be issued in different ways and at different times and are controlled by the specific **Response Cycle-Time** of the specific process and technical and organisational conditions. Therefore a specific tool controls the issuing of the results to avoid problems for physical processes and involved partners. If the orders are issued too early the process and partners can overreact and if orders are issued too late than the process can be delayed, and partners react too late and in an inappropriate manner.

6. Digital Trigger: Digital Triggers are orders and schedules, that are really transmitted to a specific device or partner. Digital Triggers control the real process of manufacturing, transportation, storing and handling of MFIs. At the end Digital Triggers must be executed by physical devices or humans and they are the input for an engineering oriented Digital Twins.

4 Application System Controller: Because all ERP-Applications runs in only one System an Application System Controller (ASC) is needed to control the execution of the application-software by explicitly defined sets of regulation rules. For each application we need specific rules because each application depends on the specific task and process environment, the defined dominance hierarchy of the process and the defined cycle times for updating the regulation and processing Digital Triggers and Digital Shadows. The ASC is therefore a necessary bridge between the ERP-System and the Operation Control System (OCS) of the Computer.

9. Logistic Cockpit: The Logistic Cockpit is like a Control-Tower for supervising the entire process of the real and the virtual world of logistics, on the one side to visualize the state of the process and on the other side to analyze process deviations and to initiate additional activities or counter measures if needed. Depending on the grade of deviations the Cockpit issues "Warnings" or "Alerts" and detailed information about. In case of an 'Alert' the Cockpit-Operator can zoom deeper into the 'problem' e. g. by changing the level of PMF-Structure to get more detailed information about the planned and actual situation or to highlight and monitor the process with additional tools. The operator can initiate an application to predict what will happen in the nearest future if nothing changes or he can start an experiment with a different set of regulation rules to evaluate alternative situations or initiate a simulation algorithm to solve the problem in a different way or to initiate operational counter measures.

10. Data-Gravity-Centre: The Data-Gravity-Centre (DGC) is like a 'Data Lake' that contains the complete data history of the Reality and the Repository and the Results of Regulation, that means all Digital Triggers and Digital Shadows and the detected

deviations. The DGC also includes all data about the state of the System itself and it can be enhanced and enriched by external big data for data analytic tools, to make the processes and the order regulation of an enterprise more resilient in the future.

***Data Analytics:** Data Analytics (DA) covers all kinds of tools to analyse and exploit data of the DGC. We can use the DGC also for experimental scenarios and future **projection** using different sets of Control Data to see in advance what will or can happen in case of accidents or unusual events. Also, optimization tools or simulation can be used to solve certain underlying tasks in a shorter time than planning tool or is not able to solve, e. g. to calculate the actual sequence of transportation means for a Jit-Process for far distant suppliers. A very new function is the **prediction** of future process behaviour with a predictive logistic we can detect problems in advance so the process and/or Master-Data and Control-Data can be maintained and adapted if needed.

3 Repository for all operations and logistics ERP-Applications

The repository is the mapping of reality and covers all data and information for the ERP-Application to schedule, control, and balance PMF-processes by issuing orders and schedules for all involved process partners for the right items, time and places.

3.1 Databases of the Repository – the virtual world of the ERP-System

The Repository consists of a lot of databases which relate to each other; the term database does not mean exactly only one specific database, it's more a synonym for a certain content of information and can be a couple of connected databases or lists.

On the left side there is a database for the original **Customer or Dealer Orders** for final products; it is like an order book in the sales market whereby the orders are not finally scheduled. There can be also another database for industrial orders and orders for spare parts or other demands, which must be fulfilled. Beneath is the database for **Placed and Fulfilled Orders**. These orders are on the one side already scheduled for a certain plant or manufacturing shop for a certain time and on the other side the database consist the data of the actual state of the order and about the final order fulfilment. Both databases are known as transaction-data because these orders are not constant over the time, these data drive the whole production process and figures are changing along the process. The other category are the master-data which are not changed by orders.

The database of **Production and Material Flow** structure is the central database of the Repository and covers the mapping of the entire network of the supplying, production, and distribution of discrete products. This is a closed system, all MFI can flow thru therein only, and it can work accurately only if the real world of logistics is mapped completely and consistently. In classical ERP-Systems we do not have such an explicitly defined network. here information about production and material flow are normally used as a simple identification number of a plant or a supplier, and also inside a plant for

manufacturing shops, cost centres, assembly lines, warehouses, storing places or buffers. Only in some applications on shop floor level we find some information about the material flow structure for executional reasons. So, what is still missing in classical ERP-Systems is an explicitly defined PMF-Structure for all processes, routes, locations etc. which can be used by different ERP-Applications.

The PMF-Structure is closely related to the **Product Structure** (database at the left) and to the **Packaging, Transportation and Handling Structure** (database at the right). Together, these three databases form the skeleton of the ERP-System. The Product Structure contains all components that are required to manufacture the final products that are stored as a **Bill-of-Material** (BoM) normally. For different manufacturing areas we can use different kinds and formats of BoMs, which we cannot discuss here in detail (for more information see [He20]). For different shipping areas we need information about how an object must be handled, as a singular item or in a lot, and how it must be transported, in a special container or as a loose item and how it can be shipped, by a truck or ship etc.

Beside there is the database with information about the **Means and Assets for Production, Transportation, and Handling**. This database contains information about engineering parameters likes dimension and weight of an object and process like the speed of transportation means or takt rates of machinery asset or assembly line. It also contains data about the capacity and constraints of assets that must be taken in account.

Control-data of Objects and Processes: This database contains all disposition parameters for quantitative regulation of PMF and especially for calculation of schedules and orders for all MFIs. The exact setting of control data is an important influencing variable and must be carried out very carefully for each interval and each MFI. Here we focus on lead-time (LT) only which is the most important parameter of regulation, but we know that order calculation and scheduling can't be performed without lot-sizes and batch-sizes and other control data.

3.2 PMF-Structure – ordinated Data-Grid for all ERP-Regulation-Software

For the different applications and tasks, we need a common data structure to integrate the different regulation software functions. The common Data-Grid is the PMF-Structure that relates to the Product Structure for manufacturing processes and to the Packaging, Transportation and Handling Structure, together they build the backbone to integrate all application and functions. For mapping of Production and Material Flow we use the mathematical theory of Boolean Algebra for special classes like Boolean Intervals and Boolean Trees [Ko89]. We assume that PMF is a linear ordered chain which is closed open defined by a left and right boundary and meet the requirements of an ideal Boolean Interval lattice. There are no jumps or overlaps between the next following intervals, no interval in the chain is missing or lies outside the entire interval. The beginning is defined by an 'Entry Point' (EP) which lays inside the interval and the ending is defined by another EP which lays outside the interval and must be the beginning of the next-following interval.

The complete algebraic definition of intervals and subintervals is omitted here for reasons of space, detailed information can be found in [He12, He20]. We will explain this in short for the automotive industry and define an interval 'Int-PP' for the entire manufacturing and shipping process which defines the virtual space of production and material flow (fig. 3.4). This interval starts at the entry-point 'PE' that signalizes the beginning of the entire observed process and ends at the entry-point 'UE' which signalizes the end of the process. 'UE' must be the beginning of the next-following interval which stands for the usage of final products.

$$\text{Int-PP} = [\text{PE} \prec \text{UE}); \text{PE and UE are EPs of the entire production process} \quad (1)$$

We split this upper interval into four strictly ordered subintervals for 'single part production 'Int-SP', for assembly group production 'Int-AP', for module production 'Int-MP' and for final product production 'Int-FP'. These subintervals are disjunctive and complementary and ideal embedded into the upper interval 'Int-PP'. Each of these subintervals represents a typical section of the production process of complex products like in the automotive industry. We define the ideal PMF-Structure in the following way:

$$\text{Int-PP} = \{\text{Int-SP, Int-AP, Int-MP, Int-FP}\} \text{ Int-PP is the entire Interval} \quad (2)$$

Each of the subintervals is defined by two Entry-Points (EP) whereby the ending EP of an interval **must be** the beginning of the next-following interval. We focus only the definition and declaration of the subinterval of final products production.

$$\text{Int-FP} = [\text{FP} \prec \text{FU}); \text{FP and FU are the two bounding Counting Points} \quad (3)$$

$$\text{Int-FP} = \{\text{Int-FF, Int-FS}\}; \text{Int-FA: Final Assembling, Int-FS: Final Shipping} \quad (4)$$

$$\text{Int-FF} = [\text{MH} \prec \text{FF}); \text{MH and FF are the two bounding CPs of the interval} \quad (5)$$

$$\text{Int-FF} = \{\text{Int-FA, Int-FI}\}; \text{Int-FA: Final Assembly; Int-FI: Final Inspection} \quad (6)$$

$$\text{Int-FA} = \{\text{takt-1, \dots, takt-y}\}; \text{takt-1 etc. all intervals of Final Assembly Line} \quad (7)$$

$$\text{Int-FA} = [\text{t1} \prec \text{ti} \prec \text{tn} \prec \text{tz}) \text{ all next-following CPs of Final Assembly Line} \quad (8)$$

If an EP is referenced by an existing Data Acquisition Point (DAP) we call this a Counting Point (CP). This referencing procedure is very new but mandatory otherwise the results of calculation will be not correct. By this the entire process of production of final products and its components is mapped on a consistent mathematical fundament. This mapping procedure is especially important for the DCT. The larger and deeper the PMF-Structure is defined, the better and more accurate can the ERP-application be performed, because the granularity of the hierarchy and the implemented Counting Points are used for regulation of all process. Each CP represents a node in an ordinated Data-Grid that is used for alle ERP-Applications. The calculation starts at CP 'FF' with a breakdown of customer orders into material requirement and runs backwards from one CP to the next preceding CP and ends at the CP 'SP', normally the location of goods receipt of a plant (fig. 3.4).

3.3 Equidistant metric of the PMF-Structure

The ideal PMF-Structure is only a logical mapping without any quantitative meaning and not suitable for arithmetic calculation. But for scheduling and monitoring of processes we need a quantitative metric which is valid for all intervals. It seems that the lengthwise distance between two consecutive CP's is a suitable measure, but the same physical distance between two CP's can be traversed by different freight carrier with different speeds, transportation depends on the type of transport means and the used route. For our purposes Lead-Time (LT) is a more appropriate metric which we can be used for every interval and process, not only for manufacturing and transportation, but also for storing or handling etc. The metric unit must be equal and applicable and countable for all intervals in other words, we need an equidistant metric for all intervals and processes otherwise the result of calculations would be wrong. And we need a specific 'calendar' to 'normalize' the interval by an equidistant metric unit. An example, why is this mandatory: If the manufacturing process of a car takes in a physical meaning three shifts it can be finished in only one day or in three days, if there is only one shift per day in a plant. And if there is a weekend with two idle days in addition, the manufacturing process takes five days. This example shows that we must normalize a normal Gregorian Calendar into a Process Calendar with equidistant units for LT calculation. This can be realized by a Gregorian Calendar wherein idle time units are marked so they could be skipped in calculation because the given LT parameters represent only the pure processing time. The smallest time-unit could be a second, all other time-units must refer to this metric like hours, shifts, days, weeks, decades, months or years. Without such a normalized Process Calendar we cannot calculate the correct time for order values.

3.4 Control of ERP-applications on different PMF-Levels (exemplary)

In this subchapter we will explain exemplary the control of some typical ERP-applications like Sales Planning, Master Production Scheduling, Assembly-Line-Balancing and JiT/JiS-Call-offs using the automotive industry for illustration (fig. 3.4). ERP-Application starts with regulation of sales planning for final products (cars) at the highest PMF-Level, the entire Process Interval (FP → UE), the referred CP for the sales plan is therefore 'FU' for different sales markets. The time unit for sales planning are normally years divided into quarters and/or month and the cycle time to planning resp. replanning is a new year and a revolving update can be made for each quarter for the latest year. Based on the results of sales planning the MPS is generated e. g. for 3 months of the latest year for final products plant. The MPS is divided into weeks and days and the referring CP is 'FF' which represents the finish of final products production. A revolving MPS-update is made every new month and the last two weeks are updated every day. For further regulation we can use the main sub-intervals of 'Press-Parts Production', 'Chassis- Production', 'Car Body Production' and 'Final Product Production', whereby the days of the last four weeks are divided into shifts. Here we focus on the 'Final Assembly Line' where the sequence of final products is determined for every single assembly line. For this we use the next PMF-Level of Interval Int-FA (MH → FF). Assembly Line Balancing is updated every day or

shift of the two latest weeks whereby a frozen horizon can be defined for some days. The Int-FA (takt-1 → takt-y) is used for JiT-/JiS-Applications which are generated in advanced by the concept of pearl-chain and process-data are transmitted in real-time or in the process- cycle- time of the final assembly line.

Knowing the relation between the processes, the PMF-Structure and the application we can control the process by a defined cycle-time for each interval which is crucial for the update and data processing and therefore for the Operation Control System of the Hardware. It makes no sense to calculate every second all data for all applications. The same also applies to the processing of the collected data. At the lowest PMF-level we need the latest actual data, but we should process the data not real-time but only in the process cycle time. The reason is the lead-time of material flow in an interval and the response time of the process to react on a new event. If the defined cycle-time of processing is too slow than the reaction resp. digital trigger comes too late. If the cycle-time is too fast than the reaction resp. digital trigger comes too early, and the process gets nervous and is over-controlled.

In connection with the LT the reaction or response time (RT) of MFI and the cycle-time (CT) for new targets values plays an important role. On the one hand CT depends on the required LT and RT is additively influenced by factors such as batch sizes and departure times. If the CT is (far) below the RT, then the real process cannot keep up with the virtual process so the system will react too quickly which will lead to hectic and incorrect reactions. If the CT is (far) above LT then DCT-System reacts too slowly, which can lead to material shortages and provokes additional avoidable costs.

The ideal PMF-Structure has some useful properties for ERP-Applications. One of these is the fact that actual **Process Data** must be acquired on the lowest PMF-level only. The actual data of all upper PMF-Levels are only a sum of actual Process Data of the lowest PMF-Level, so all ERP-Applications use the same data! This is because the very first CP of a 'lower' interval has the same meaning of the very first CP of the 'upper' Interval: If an MFI passes a very first CP of a 'lower' interval then it passes at the same time the CP of all corresponding 'upper' intervals, e. g. if 't1' is passed then 'FE', 'MH' and 'FP' is passed at the same time (fig. 3.4), so the actual Process-Data on different PMF-Level are at these CP's consistent, but the process- and response-time of applications are different. The placed orders (target values) for a certain CP of a higher PMF-Level is equal for all corresponding CP's of the lower PMF-Level. These ideal properties in the data-grid accelerate the adjustment of different ERP-applications on different PMF-levels because no data transmission is needed and ensures thereby a direct recursion of data between the different ERP-applications.

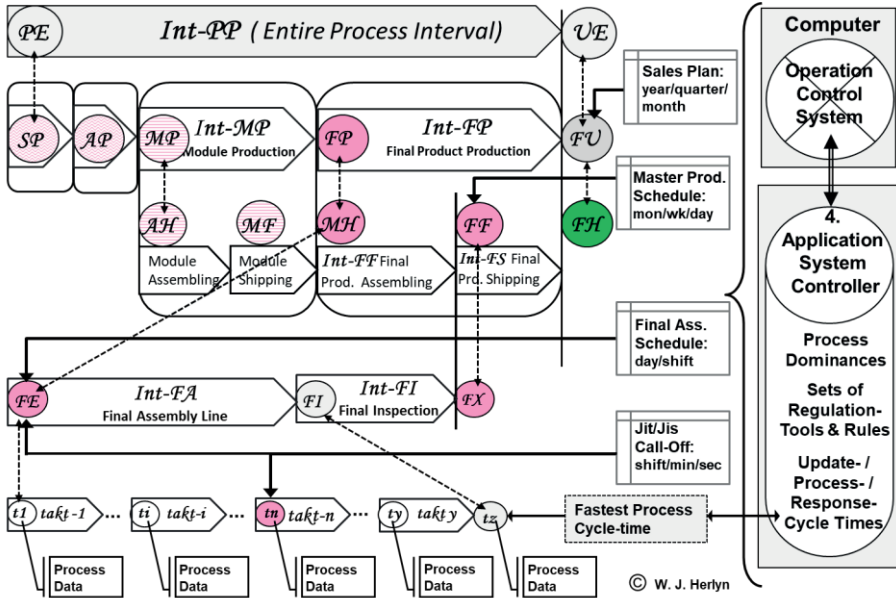


Fig. 3.4 PMF-Structure and control of different ERP-Applications

Because the ERP-System works on the DCT a very new kind of communication bridge between the ERP-Applications and the Computer ‘Operation System Controller’ (OSC) of the computer is required. Because the ERP-System is always ‘on air’ we need a special incorporated ‘tool’ to control permanently the different ERP-Application and to harmonize the data and tasks. The bridge is an ‘Application System Controller’ (ASC) of the ERP-system, it’s a very new Software-tool that contains explicitly defined regulation rules, which are implicitly incorporated in the classical ERP-System. In classical ERP-Systems the different Applications run ‘Job’ by ‘Job’ following the water-fall-concept. The jobs are linked by ‘Job-Control-Procedures’ and data are transmitted thru interfaces. But if the ERP-Applications run in the same IT-System, we must describe explicitly all regulation rules and cycle times in a specific Application Control System (fig. 2.1).

4 Conclusion and short outlook

The designed Application Architecture of the ERP-System uses the capabilities of industry 4.0 and the data resources that are permanent available for an enterprise. This is achieved by the concept of ‘Digital Control Twin’. The DCT is founded on a common Repository with an explicitly well-defined Production and Material Flow Structure related to the Product Structure and Transport Structure. Each Counting Point in the PMF-Structure is a node of an ordinated Data-Grid which is used to calculate Digital Triggers for process control and to get Digital Shadows of the process. The regulation itself is based on the

principal of closed-loop-control for control and balancing the processes. Now, there exists only a small ‘Demonstrator’ for further research and proof of concept which must be developed to a professional IT-System before implementation.

Based on the target and actual values of the ordained Data-Grid all ERP-Applications are executed in One System. Therefore, no interfaces and data transmission between different applications are needed and the loss of data accuracy and a time-lack is eliminated too. And in addition, the regulation rules and tools for each ERP-Application must be defined explicitly in a specific separate Application Control System, that also contains the cycle-times for application update, for issuing Digital Triggers and for receiving Digital Shadows. For supervising the entire process, a Logistic Cockpit is installed by which not only the real and virtual processes are monitored. It is used by experts to initiate countermeasure if needed or to start predictive activities or to launch changes in the process or regulation supported by Data Analytic and Smart Logistic tools to exploit the Data Gravity Centre for process prediction or projections.

And at the end it seems that the future ERP-System has something in common with the of regulation of an autonomously driven vehicle, the vehicle is now the plant which is driven by the ERP-Application-Software and gets the process-data form the technical devices of the factory and the transportation and handling means.

5 References and Bibliography

- [BBG21] Bender B, Bertheau C, Gronau N, Future ERP Systems: A Research Agenda, in: Proceedings of the 23rd International Conference on Enterprise Information Systems (ICEIS 2021) - Volume 2, pages 776-78; DOI: 10.5220/001047730776078
- [De19] Deutsche Post DHL Group, Digital Twins in Logistics - A DHL perspective on the impact of digital twins on the logistics industry, 2019, accessed 24/4/2021; <https://www.dhl.com/content/dam/dhl/global/core/documents/pdf/glo-core-digital-twins-in-logistics.pdf>
- [Di17] Gesellschaft für Informatik, Digitaler Zwilling, in: Informatik-Lexikon, 2017
- [Di19] Deutsche Post DHL Group, Digital Twins in Logistics - A DHL perspective on the impact of digital twins on the logistics industry, 2019, accessed 24/4/2021; <https://www.dhl.com/content/dam/dhl/global/core/documents/pdf/glo-core-digital-twins-in-logistics.pdf>
- [Di20] Wissenschaftliche Gesellschaft für Produktion: Digitaler Zwilling, 2020, accessed: 20/4/2021; http://www.wigep.de/fileadmin/Positions-und_Impulspapiere/Positionspapier_Gesamt_20200401_V11_final.pdf
- [Gr14] Grieves M., Digital Twin-Manufacturing Excellence through Virtual Factory Replication, Whitepaper, LLC, 2014; accessed: 20/4/2021, https://research.fit.edu/media/site-specific/researchfitedu/camid/documents/1411.0_Digital_Twin_White_Paper_Dr_Grieves.pdf

- [He11] Heizer et, al, Operations Management - Sustainability and Supply Chain Management, 2017 12. Ed., Pearson Education Ltd., Essex (England)
- [He12] Herlyn, W., PPS im Automobilbau - Produktionsprogrammplanung und -steuerung von Fahrzeugen und Aggregaten“, Hanser Verlag, München, 2012
- [He19] Herlyn W., ERP und Industrie 4.0, 2019, in: Industriemanagement 15 (2019), 4, S. 26 – 29
- [HE20] Herlyn, W. Comprehensive MRP-concept for complex products in global production and supplying networks: 2020, DOI: 10.13140/RG.2.2.31478.14404
- [Ko89] Koppelberg S., Special Classes of Boolean Algebra, in: Monk J. D., Bonnet R. ‘Handbook of Boolean Algebra’, Vol. 1, North-Holland, Amsterdam, 1989
- [KT18] Khan A., Turowski K., A Perspective on Industry 4.0: From Challenges to Opportunities in Production Systems, in: Proceedings of the International Conference on Internet of Things and Big Data (IoTBD 2016), pages 441-448
- [Ku13] Kurbel K. E., Enterprise Resource Planning and Supply Chain Management Functions, Business Processes and Software for Manufacturing Companies, 2013, Springer, Berlin
- [LL07] Laudon K.C., Laudon J.P.: Management Information Systems, 10. Ed., 2007; Pearson International Edition, New Jersey
- [Lh13] Lödding H., Handbook of Manufacturing -fundamentals, Description, Configuration, 2013, Springer, Berlin Heidelberg,
- [Mi18] SAP: Mit der SAP Digital Supply Chain zum intelligenten Unternehmen“, 2018, <https://www.sap.com/content/dam/site/events/emea/germany/assets/2018-11-06-de-email-sap-digital-supply-chain.pdf>
- [PI20] Bundesministerium für Bildung und Forschung: Plattform Industrie 4.0 - Innovationen im Zeitalter der Digitalisierung, Berlin, 2020, accessed: 23/4/2021; https://www.bmbf.de/upload_filestore/pub/Industrie_4.0.pdf
- [PI21] Bundesministerium für Wirtschaft und Energie; Plattform Industrie 4.0 Industrie 4.0 gestalten. Souverän. Interoperabel. Nachhaltig -Fortschrittsbericht 2020, Berlin 2020, accessed 23/4/2021; https://www.de.digital/DIGITAL/Redaktion/DE/Digital-Gipfel/Download/2020/2020-fortschrittsbericht.pdf?__blob=publicationFile&v=5
- [Sc16] Schönsleben P., Integral Logistics Management: Operations and Supply Chain Management Within and Across Companies, 5. Ed. 2016 CRC Press, Boca Raton
- [SKF20] Schryen G, Kliwer N, Fink M, High Performance Business Computing, 2017, accessed 20/4/2021, <https://link.springer.com/article/10.1007/s12599-018-0549-x>
- [ZH20] Zadek H., Herlyn W., 2020, Der Digitale Steuerungs-Zwilling; in: Zeitschrift für wirtschaftliche Fertigung (ZwF), 2020 115, special „Digitaler Zwilling“, 70-73, <https://doi.org/10.3139/104.112338/>
- [ZH21] Zadek H., Herlyn W., Mastering the supply chain by a concept of a digital control-twin, in: Proceeding of Hamburg International Conference of Logistics (HICL) 29: p. 661-697 (2020), accessed: 20/4/2021 <https://doi.org/10.15480/882.3120>

Towards an agile and adaptive Enterprise

Holger Cermann¹, Sebastian Kern², Florian Bertram³, Martin Sander⁴, Thomas Heimann⁵, Rachel Proctor⁶

Abstract: This paper contributes to the challenge of enterprises responding to factors that require a change and transformation. It proposes a concept that can be adopted by enterprises' organizational functions to empower agility and adaptability as a response to this challenge. The concept fills the gap of transformation challenges towards agility and adaptiveness. This is of particular relevance for classic and hierarchical-oriented businesses in a frequently faster-changing world. The main element of the concept describes a stepwise thinking pattern intended to guide a function's adaption process. It builds on industry-wide accepted principles of adaptive enterprises and applies the concept of reference architectures that shows a potential application of the stepwise approach in more detail. This paper builds around industry experiences that have the function of Enterprise Architecture Management (EAM). This is because the EAM function plays a central role in driving its enterprise with transparency and decision support in the process of transformation.

Keywords: Adaptive Enterprise, Agility, Agile Transformation, Change Management, Principles

1 Introduction

The speed at which industries are disrupted and reshaped is increasing. Since 2000 more than half of the companies in the Fortune 500 have been acquired or ceased to exist [Bo15]. A key capability to overcome the challenge of these complex and fast-moving times is the use of digital technologies. There are many cases in which they have already proven to be essential enablers of value. One example is the use of cloud technologies to adapt an organization's infrastructure and become more competitive [CA16]. By maximizing the utilization of servers and requesting infrastructure resources on-demand, enterprises are empowered to reduce their total cost of ownership and build their infrastructure flexibly around the needs of changing business models. In addition to technological enhancements, enterprises are adopting lean and agile workflow and organization models to increase the adaptiveness of the organization itself. These agile models emerged from technology companies that built their strategies and capabilities around digital technologies to facilitate agility, adaptiveness, and a shared way of thinking between business and IT resulting in fast market response. To overcome the challenges of these fast-moving times, political and economic institutions are promoting the development and adoption of digital

¹ Capgemini, holger.cermann@capgemini.com

² Capgemini, sebastian.b.kern@capgemini.com

³ Capgemini, florian.bertram@capgemini.com

⁴ Capgemini, martin.sander@capgemini.com

⁵ Capgemini, thomas.heimann@capgemini.com

⁶ Capgemini, rachel.proctor@capgemini.com

technologies and agile models [Sa03]. However, practical experience and the available body of research show that the adoption of these approaches is challenging for organizations. This is particularly true for classical business models and comes with great differences in risks and expectations among the stakeholders [Su18]. Consequently, it becomes important for enterprises to mitigate the risks when transforming an enterprise towards adaptiveness and agility.

This paper contributes to the digitalization research by proposing a stepwise approach that allows an enterprise to move its business structure towards adaptiveness by going through six phases. The phases allow an enterprise to reduce the risk that is associated with change and transformation efforts. They were developed together with industry experts who are engaging in transformation projects of enterprises in their day-to-day work.

The proposed concept in this paper adds to the growing body of literature on transformation management and digitalization models. It adopts the two rather static concepts of principles and reference architectures and adds the required activity and processing view that enables the desired state of transformation.

2 Implications for Enterprises

The world is changing more quickly than ever, and enterprises must be adaptive to remain viable and retain their competitiveness. This is mainly driven by advancing digital solutions and new flexible business models. Additionally, existing hierarchical companies are facing new challenges and increased competitive pressure from companies that build their business model around digital solutions. To retain a competitive edge, they are required to adapt their business models and integrate digital technologies to enhance their value creation. The classic build enterprise must understand the characteristics of the required change and consider various perspectives cross-cutting multiple domains, businesses, and IT components. This challenge should not be seen as a one-time project but rather an evolutionary approach that can handle emerging requirements in a fast manner to remain competitive while adapting the enterprise. Among others, this includes the adoption of a business strategy that is considering internal and external factors that are driving and supporting this change. For example, the internal structures in an enterprise must be adapted in a way to encourage change. From a business perspective, this includes the development and empowerment of capabilities, e.g. competencies for the execution of agility.

2.1 Principles for an adaptive enterprise

Due to given challenges enterprises need to increase performance, value creation, and time-to-market to remain competitive. In general, enterprises are faced with the challenge

of doing more and better things in less time and at the same time, the context of the enterprise is evolving at a top speed.

To meet these requirements the concept of principles has evolved and matured through various research and industry applications. Principles are general rules and guidelines to serve the fulfillment of an enterprise's strategy and mission. They may be established at various abstraction levels, both horizontally and vertically, and may be linked in a hierarchical set that forms a structured composition towards a common and pre-defined goal. For example, from generic to specific architectures, business to technology, or strategy to operations. Because of their strategic nature, which is spanning wider time horizons, they are intended to be long-living and have a holistic and general intention. The instance of a principle is a statement that is tied and applied to an action or thinking process – it can be seen as the frame around an action that has a secondary implication to the processing view of an enterprise on how to implement an action. As a result, the output of the executed action or thinking is directly guided towards the overall goal. A principle can be described in a textual form containing a name, statement, rationale, implication, and link to other principles to build up a hierarchical structure.

Principles are long-living and last for several years. Various research and industry applications have proven the effectiveness of principles for enterprises. This concept has also been adopted by enterprises in the view of transformation and adaptiveness. However, little research has been made around how these principles can be applied by a specific organization considering the current increasing demand for adaptiveness.

3 Implications for Enterprise Architecture

Due to the increasing use of digital technologies and agile models to handle transformation challenges, organizational functions that are integrated at the intersection of business/IT and strategy/operation become more important to align these elements. One of these interdisciplinary functions is the Enterprise Architecture Management (EAM) function. EAM functions are used by enterprises to enable goals like transparency standardization, business-IT alignment, and management support on multiple levels such as the strategy, portfolio, project, or operational level. Additionally, various frameworks, methodologies, and working practices that were developed in research-industry collaborations, and have gained industry-wide acceptance, are supporting the execution of the EAM function. Agile methodologies are already adopted by the Enterprise Architecture function to deliver increments according to their business value [Cb18]. The following are examples of how these goals are enabled through the capabilities of the EAM function: First, transparency is enabled by modeling representations of, e.g., the enterprise's as-is application landscape cross-cutting business-process-views. Secondly, standardization is enabled, e.g., by using architectural principles across the IT organization that are in alignment with the strategic direction of the enterprise. Thirdly, business-IT alignment is enabled, e.g., by communicating an enterprise-wide and context-specific glossary that allows for a common

language across relevant stakeholders. Finally, management support is enabled e.g., by creating a reporting and measurement system that provides input for pre-defined quality gates.

Enterprises usually expect a faster time to market, increased ability to innovate, increased quality, effective risk mitigation, and a reduction of costs. The following picture depicts the described sequence and effect through the implementation of the EAM capability in an enterprise:

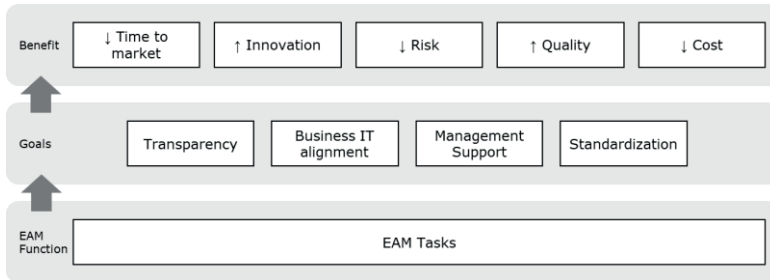


Figure 1: Desired effect by implementing an EAM function

Given the EA effect, capabilities, toolset, deliverables, processes, and industry-wide acceptance of the EAM function, an enterprise might use it to support transformation challenges with IT business participation. Further, the concept of principles, which is adopted by enterprises in the light of transformation, is already embodied in the capabilities of the EAM function. Therefore, the EAM function may support transformation efforts in multiple ways. EAM can enable transparency to find the right level of adaptiveness and agility for the different parts of an enterprise. And it can enable a standardized and stepwise procedure to approach an enterprise's adaptiveness. Furthermore, it can provide insights that are relevant for decision-making and alignment of business and IT in the light of transformation.

This paper will provide a concrete example of how the stepwise approach will be rooted in and executed by the EAM function. It depicts the transformative concept of AI and shows how the capabilities of EAM can put the use of AI into an enterprise-wide context, hence enabling desired business goals. The capabilities of EAM are related to multiple layers of an enterprise and can be tied to multi-domain entities. Therefore, it can provide solution approaches for the use of AI in an enterprise-wide context. AI has already proven its effectiveness in various business domains. Concepts are already used in different areas to increase adaptivity. These include, for example, using AI Ops to improve operations and security, or monitoring networks combined with predictive maintenance to reduce costs and mitigate risks. The concept for AI aligns with the requirement of an EA function to address the transparency of an enterprise holistically. This means that an enterprise is sliced and diced into horizontal and vertical layers that allow stakeholder-specific views to address their given concerns. The IAF Framework is an industry-wide accepted

framework that addresses the representation of architectural blocks and their relation to each other to achieve transparency in an enterprise. In the following section, the paper makes use of Aspect Areas and Domain-specific views from the IAF. Aspect areas are represented horizontally, and Domain-specific views are represented vertically in the capability map depicted as Figure 2.

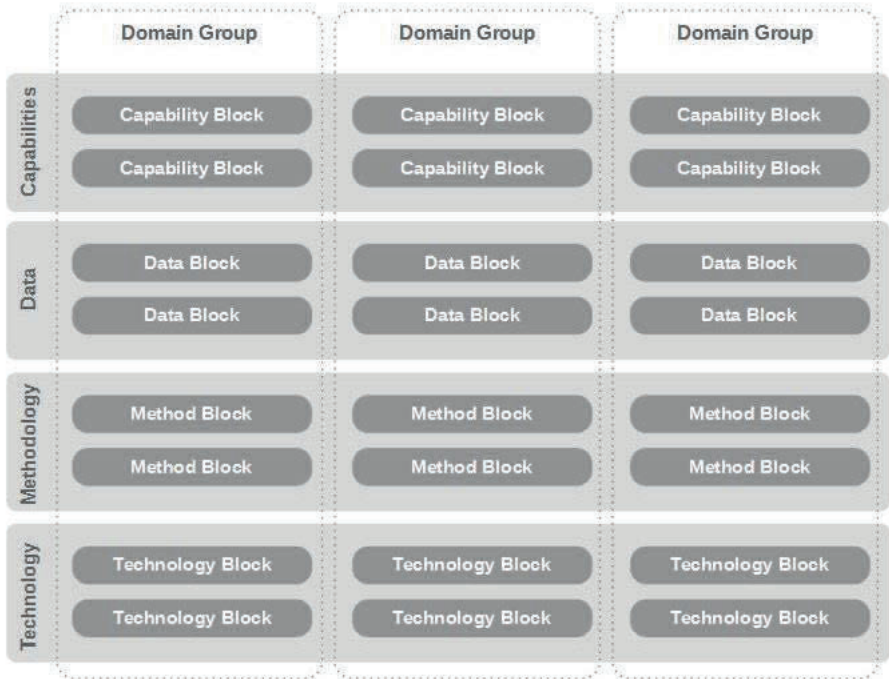


Figure 2: Domain map of architecture blocks adopting IAF Aspect Areas

4 Six phases to transform an Enterprise towards adaptiveness

Enterprises are facing transformation challenges and demand their organizations to be more adaptive towards change. Principles have already been used by enterprises to give direction to the creation of adaptive capabilities. The EAM function provides capabilities at the intersection of IT/business and strategy/operation. Principles are general rules and guidelines with the goal of serving the fulfillment of an enterprise's strategy and mission. Figure 3 depicts the general process, which will be detailed in the next sections.

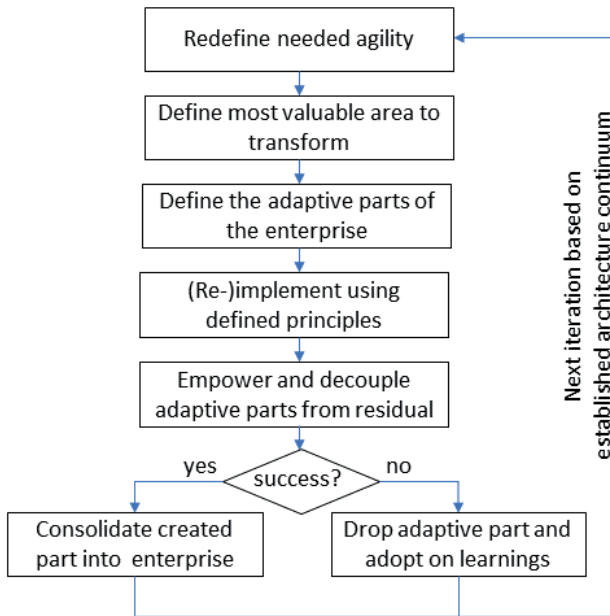


Figure 3: Six phases approach towards an enterprise's adaptiveness

4.1 Redefine needed agility (Phase 1)

Agility comes with a price, and productivity reduces in the first stages. But on the other hand, agility comes with value. Over time it increases productivity, enables the organization to be innovative, and establishes a culture in which work is experienced as positive, highly purpose-driven, and providing pleasure [Ri12].

Levers in positive as well as negative directions will occur in the enterprise's transformation towards agility. Depending on need and context, the levers are to be analyzed independently. Negative resulted levers can be, for example:

- Less long-term planning capabilities in the context of highly regulated and externally embedded processes.
- Needed changes also on interfacing company parts (people and IT)
- Different (and case-wise not optimal) planning methods in adaptive enterprise parts embedded in traditional enterprises.
- Amount of recurring adaptations/changes needed due to adopted target picture.

This evaluation will change over time – with upcoming adaptiveness in the overall enterprise and interconnections. In this phase, the enterprise defines the scope of agility.

Concrete business cases with coupled value propositions can help to understand and redefine the degree of the required agility. Although the transition towards adaptiveness and agility will be worth it in most of the identified business cases [MC09], different parts of the organization will have different potential to move towards adaptiveness. Therefore, the enterprise needs to understand that moving towards an adaptive enterprise comes with a price and a definition of the needed agility helps to prioritize the use of resources that is required for the move towards adaptiveness.

4.2 Identify where the enterprise needs to be most adaptive (Phase 2)

For architectural partitioning, TOGAF proposes to use time, breadth, and level as dimensions [Og18]. In the context of this publication, we focus on Subject Matter (breadth) and organizational structure. The rationale for this lies in the people dimension, which is embedded in the line organization: Adaptive enterprises require adaptive and agile mindsets that are built around people. We propose to make the “people dimension” a key in the transformation approach. And a stepwise transformation must always include a stepwise transformation of the people behind it (team training / upskilling, team rotation, team change). One reason is that most successful changes must include multi-perspective change elements rather than being delivered in a top-down or bottom-up approach itself with an extensive pre-defined plan [Mc09]. And especially for people management, a stepwise and adaptable approach is important to anchor new approaches deeply into the organization’s culture [Jc96]. Furthermore, organizational units of an enterprise vary in their scope and relation to the enterprise, thus requiring different levels of transformation. Proposed levels of transformation might be (ordered in decreasing size):

- Partially decoupled sub-organizations (i.e. small startups) in the enterprise.
- Divisions/Departments with already existing interfaces to the organization.
- Existing or upcoming project, program, or line management teams/departments.
- Agile parts of an existing department focused to take over business processes from the embedded department in an evolutionary process.

In this phase, the enterprise identifies the parts of the organization and its boundaries which shall be in the focus of moving towards adaptiveness. Furthermore, a special emphasis must be given to the breadth and organizational structure with particular attention to the influence coming from the dimension of people. After this phase, the enterprise should have created an awareness and understanding about where and to what extend the move towards adaptiveness is necessary for the respective boundaries and parts of the organization.

4.3 Define the adaptive parts of the enterprise (Phase 3)

Systemic approaches help enterprises to mitigate risks and be more successful with managing changes [Gc04]. As described in the previous chapter, the move towards

adaptiveness requires various considerations regarding the change. To increase the chances of success and value creation when applying this concept, a more detailed structure to go through two critical dimensions in a two-step process is proposed in this phase of the concept:

- Evaluation process on needed adaptiveness: Organisation view
- Evaluation process on needed adaptiveness: Process view

As described in Phase 2, organizational boundaries (and given Conway's law also other structural areas including IT) are proposed to be a key delimiter in this approach. To address the different levels explained in Phase 2 (Suborganisation, Division, Department, Program, Process) a tree-based scanning method can be applied, where – starting at the sub-organization level – the evaluation can be applied. For the evaluation itself, a target company-specific criteria list needs to be applied. Ubiquitous criteria candidates could be:

- Change frequency of the underlying business domain and IT landscape
- Need for differentiation in given domain (in relation to competitors)
- Current strategic business fit
- IT backlog size in terms of demanded business change
- Potential for IT-driven change in business domain (including AI opportunities)

In section three, this paper describes the capability of the EAM function to create transparency in an enterprise by modelling Aspect Areas and Domain-specific views by applying the IAF framework. Picking up the example of AI, which is a transformation-supportive digital technology, this phase can iterate through the EAM function to elicit information from the enterprise that supports decision-making for the definition of the adaptive parts. As an example of how the EAM function may model the use of AI to support decision making, the following picture shows a model instance and domain view of potentially relevant architectural AI blocks cross-cutting the IAF Aspect Areas in the context of adaptiveness.

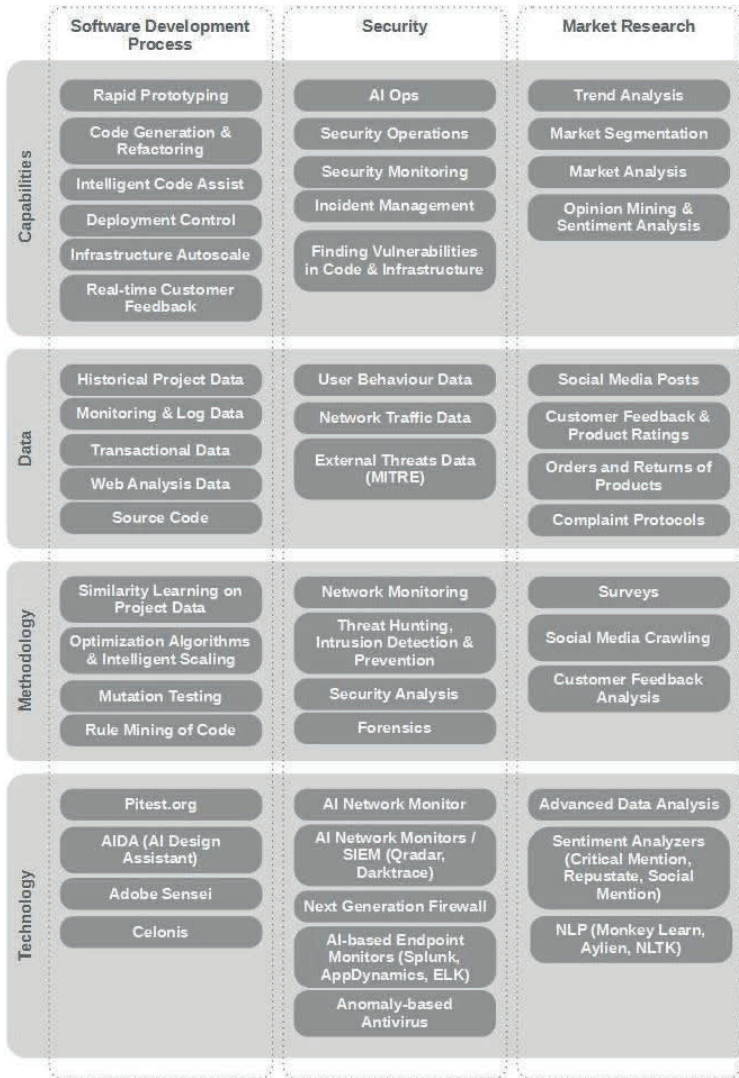


Figure 4: General view of architectural AI Blocks to support decision making

This view enlists AI-based capabilities that support the move of an enterprise towards adaptability. In this step, the paper identifies which of these capabilities can be implemented and how, making use of the EAM function and AI for increasing the adaptability of an enterprise. In alignment with agile principles and in a similar fashion how enterprises can iterate through the EAM function to acquire information for decision

making, the paper proposes to iterate frequently through the EAM function to make informed decisions and adjustments/re-implementations of respective models.

4.4 Use defined principles to (re-)implement the adaptive part (Phase 4)

An enterprise can choose from a plethora of different approaches on their change journey [Tm03, As15]. For the specific change case of adaptiveness, section two highlighted that adaptive enterprises are using principles to support their adaptiveness and achievement of long-term goals. These existing principles, however, are built around the behavior of an already adaptive enterprise. An extract from principals of adaptive enterprises that have shown practical application is listed as follows [Bv05]:

Be innovative: An enterprise must constantly be aware of possibilities that create new values and secure the viability of the enterprise. Also, it must monitor the progress of innovative initiatives within the enterprise and scale the developed solution into an enterprise wide context.

Holistic thinking: An enterprise must have an overall view of its organizations and ongoing initiatives. This principle is strongly related to transparency so that decision-making is supported in adequate ways. Having an overall picture allows the enterprise to pave the way for future states.

Interdependence: This principle describes the importance of the relation of distinct principles. As described in section 2 principles must be linked to each other to provide the relevant systemic structure to support the overarching goal or business case.

Although this list gives non-adaptive enterprises an idea of how a potential adaptive state might look like, there is little guidance on how to move towards adaptiveness. Therefore the paper developed four guiding “principles of procedure” that can be used by an enterprise when approaching or adjusting its adaptiveness:

Revolution: A “revolution” is fast but also risk-prone and only works in specific cases. The paper proposes to use the revolution approach when setting up agile sub-companies/departments or carve out existing - and capable - teams/sub-companies. This is because these parts of the enterprise are more or less isolated, and therefore the impact of an occurring risk can be managed more easily. It must be mentioned that this approach is not always successful and highly depends on the current mindset of the involved people. A mindset change and preparation phase before “flipping the coin” is a mandatory success factor.

Evolution: In most cases where a fast change is not possible, “evolution” can be possible. This is because of multi-level structures and dependencies within and across enterprises that build up a complex system. The drawback is that this approach and the shifting mindset of the people can be very slow, which can be a reason for failure itself.

Furthermore, top-down steering, organizational silos, or deep-rooted processes and behaviors may work against an evolutionary approach.

Submarines/Speedboats: Submarines/Speedboats can help you figure out the best approach, also to generate visible success stories for the part of the enterprise which is still to be transformed. They need to be continued, e.g. by a follow-up revolution.

Recreation: While not an appropriate approach for bigger and already future-oriented and well-performing parts of the organization, this transformation approach can be used to substitute existing structures (business as well as IT) with new capabilities created in parallel. While the initial transformation effort and also the information system transition architecture efforts exceed the variants above, this approach can unlock possibilities with a controllable risk. To succeed in any of the proposed approaches, invariant enterprise architecture capabilities are needed in the part to be transformed.

It is recommended that the principles of procedure described above (this also applies to principles in general) are not an isolated approach and should rather be used in combination and hierarchical linkage with other principles during the transformation journey. Furthermore, they will not fit the enterprise as a whole, and therefore an individual adaption and fit for respective parts of the enterprise should be considered.

As described in section 3, the EAM function plays an important role in the alignment of transformation goals within and beyond the enterprise. To support the move towards transformation and in addition to the enterprise principles above, the paper proposes an underlying principle for the EAM function itself:

EAM empowerment of the basis: When the basis of the enterprise is accelerating and enhancing its capabilities for faster change, the EAM function itself must increase in speed to keep up with the basis of the enterprise. One approach is to increase the agility of the EAM function. This means that the function is required to work more “bottom-up” and less in the more traditional top-down approach. Established (and empowered) EAM capabilities help in a strategic focused transformation while not risking the needed velocity. Introduction and promotion of large-scale agile execution frameworks (e.g., SAFe) can support the EAM as well as the enterprise itself.

The role of the enterprise architect must bridge different views and stakeholders: The empowered basis, the market, and the company strategy. Therefore, EAM behaves like the “active connector” in the enterprise. With these capabilities, the work of an enterprise architect changes: “Think like a CEO and act like a CDO” [Su20], [Ro20], which means that the enterprise architect should not solely feel like part of the IT organization. Also, the architect should not lose the connection to the basis. More enterprise architects and need to work collaboratively. The definition of contracts on architectural high-level targets (as proposed in TOGAF) can still be a part of enterprise architecture governance, but due to the high frequency of changes and empowerment on the basis, this tool is not sufficient alone.

In this phase, the enterprise is using principles to move towards a desired state of adaptiveness. Various principles of an adaptive enterprise are found in literature and can be used in a specific state of adaptiveness. The principles presented in this section will help an enterprise to define its journey towards adaptiveness.

4.5 Decouple the adaptive parts and its team from the company and empower them (Phase 5)

Depending on the granularity chosen in steps two and three, partially different strategies can be applied. While developing this paper we found frequent occurring aspects that showed similar patterns across various transformation projects in different industries. Therefore, this paper proposes five invariant aspects that are to be considered when moving towards adaptiveness:

- Empowerment of the transformed part.
- Organizational decoupling of line organization or of people involved.
- Strong sponsorship on all levels: CxO Sponsor and Sponsors in the chain.
- Adaptive ground rules on interconnections to non-transformed enterprise parts.
- Evolvement tasks in the IT landscape and common systems (Systems as boundaries).

In the following paragraphs, these aspects will be detailed.

Empowerment means: Giving each adaptive part of the enterprise the ability to decide on their own and optimize themselves based on their business demands. Analogous to "fully empowered product teams", agile management methodologies can apply on all management levels. Examples are capability-based budgeting (including line/run and projects). The **organizational decoupling** can define a key element in this empowerment action plan, as in many companies, budget, reporting, and target derivation are directly bound to the organizational structures. Especially within the first adoption phases, **Sponsorship** on all levels is needed to achieve the needed change – also in terms of mindset and culture. Without sponsorship, a transformation project will not have enough supportive resources and will potentially phase-out before achieving the desired output. When carving out adaptive parts from the enterprise, it is important to define the boundaries of interaction between the adaptive and the non-transformed parts of the enterprise. Mindset and culture within the enterprise will also be impacted by these changes to entities and their connections **adaptive ground rules** (or even loose principles) can help an enterprise to set the framework and standardize the interactions of these new connections, avoiding, e.g., clash of these new cultures. Finally, **Evolvement tasks in the IT landscape and common systems** must be built around the (re-)implemented adaptive parts of the system and the move towards adaptiveness to give the decoupled and empowered parts sufficient and long-term competitive ability within the enterprise. A special focus must be given to interfaces – processual and IT-based. A high maturity in classical EAM helps here to a) identify those correctly and b) foster change and adoption.

Dedicated interconnections between the new adaptive part and the residual enterprise with speed-embracing interface methodologies and technologies define a key enabler to empower the carved-out adaptive part.

In this phase, the enterprise is carving out and decoupling adaptive parts of the enterprise and provides the necessary funding from various perspectives. Common patterns and obstacles are considered to mitigate the risk of failure and increase the chances of success. The new or adapted entities of the enterprise are put into a systemic and holistic perspective and managed accordingly.

4.6 Consolidate successful transformations into the enterprise (Phase 6)

While the authors intentionally describe a decision gateway for “success” or “failure” to describe the result of the transformation efforts (cf. Figure 3), a more comprehensive definition must be given. As described earlier in „Agility comes with a price”, levers in positive as well as negative directions will occur in agile enterprise transformation. Considering this, and to judge for success or failure, an enterprise must reflect upon the overall resulting business case. If it is positive, we consider it as a success and worth keeping (maybe subject to later further transformations). If it is negative, a drop/rollback is proposed (or a rework with changed scope).

For the consolidation, a switchover must be made in all included dimensions, like organizational structure, processes, IT integration and IT processes, customer experience, etc. As mentioned before, to safeguard the successful transition, long-lasting empowerment must be ensured, also in terms of the established budgeting processes.

In case of cancelled transformation or with the realization of a negative business case outcome, the paper proposes to ***drop these adaptive parts as quickly as possible***. Accepting the “sunken cost fallacy”, management strength and boldness is needed in these situations to avoid suboptimal “pseudo agile enterprises” in the long run.

Nevertheless, just doing a partial rollback is an option. Also, the definition of a smaller organizational sub-scope and a retry of this sub-scope in the next iteration is possible. Efforts in mindset change can be reused, and therein motivated employees may be leveraged.

5 Conclusion

The concept described in this paper provides an approach to an evolutionary transformation based on an iterative model. While the authors present a single sequential iteration for simplification, with growing maturity and increasing capability in the organization the process can be parallelized, and multiple transformation activities may take place at the same time. Nevertheless, this dimension of parallelization needs to be

managed actively to balance change capabilities, a feedback loop (e.g. lessons learned), and required adoption speed. The EAM function of the enterprise must play a key role in the transformation process.

Independent from optimizations: Transformations to agility must also be executed with help of agile toolsets. There is no longer "one technologically driven" strategic EA target design, which used to be common in many enterprises.

6 References

- [As15] Andreas, Schutkin: Das Geheimnis des Neuen: Wie innovationen entstehen, 2015
- [Bv05] Bas, Verbuggen: The adaptive enterprise: Defining architecture principles for an adaptive enterprise, 2005
- [Bo15] Bonnet, D.; Buvat, J.; KVJ, S.: When Digital Disruption Strikes: How Can Incumbents Respond? 2015, Capgemini Consulting, URL: https://www.capgemini.com/consulting/wp-content/uploads/sites/30/2017/07/digital_disruption_1.pdf
- [Ca16] Cancila, M.; Toombs, D.; Waite, A.; Khnaser, E.: Gartner 2017 Planning Guide for Cloud Computing. Gartner, 2016.
- [Cb18] Carsten, Brockmann; Eldar, Sultanow; Christian Czarnecki: Enterprise Architectures between agility and traditional methodologies, Lecture Notes in Informatics, 2018.
- [Gc04] Guangming, Cao; Steve, Clarke, Lehaney: The Need for a Systemic Approach to Change Management, 2004.
- [JC96] John, Kotterer: Leading Change. Harvard Business School Press, 1996
- [Mc09] Mike, COHN: Addison-Wesley Professional: Succeeding with Agile, 2009.
- [Og18] The Open Group: The TOGAF® Standard, Version 9.2, 2018
- [Ri12] Ries, E.: Lean Startup: Schnell, risikolos und erfolgreich Unternehmen gründen. Redline Verlag, ISBN: 978-3-86881-333-3.

- [Ro20] Roth, S.; Heimann, T.; Sultanow, E.: Innovation ist Architektursache: Stellschrauben für die Innovationsfähigkeit. *Objektspektrum*, 06/2020.
- [Sa03] Sambamurthy, Bharadwaj, and Grover. 2003. "Shaping Agility through Digital Options: Reconceptualizing the Role of Information Technology in Contemporary Firms," *MIS Quarterly* (27:2), p. 237, DOI: <https://doi.org/10.2307/30036530>.
- [Su18] Sultanow, E.; Chircu, A; Schröder, K.; Kern, S.: A Reference Architecture for Pharma, Healthcare & Life Science, Workshop der Informatik, 2018
- [Su20] Sultanow, E.; Duane, J.-N.; Chircu, A.: Skills for Sustainable Enterprise Architectures in a VUCA World, WI 2020 – 15. internationale Tagung Wirtschaftsinformatik vom 9.-11.3.2020 in Potsdam. DOI: https://doi.org/10.30844/wi_2020_x2-sultanow
- [Tm03] Thomas, Malone; Robert Laubacher; *Inventing the organizations of the 21st century*, 2003

Enterprise Architecture für Netzwerkorganisationen

Systematische Koordination kollektiver und dezentral getriebener Aktivitäten

Andreas Becker¹

Abstract: Projekte im Kontext von Unternehmensarchitektur sind anspruchsvoll: die zunehmende, inhaltliche Komplexität angesichts von Cloud-Architekturen und damit verbunden die steigende Zahl an Beteiligten bei der Erstellung von IT-Werkzeugen, sowie die Vielschichtigkeit des Umfelds besonders in größeren Unternehmen stellen die EA-Modelle vor Herausforderungen. Dieser Beitrag schlägt ein Modell vor, welches durch rigorose Orientierung an den herzustellenden Leistungen und durch Abgrenzung der rein technischen Architektur Abhilfe schaffen soll. Das Modell stellt eine konsistente Sicht auf der strategischen Ebene her und integriert sich nahtlos mit den operativen Werkzeugen im Geschäftsprozessmanagement sowie im IT-Service-Management.

Keywords: Unternehmensarchitektur, Enterprise Architecture, Netzwerkorganisationen

1 Unternehmensarchitektur kondensiert

Ausgehend von Informationssystem-Architekturen entstanden zu Ende des 20. Jahrhunderts konzeptionelle Modelle, deren Absicht primär darin lag, die zunehmende, IT-technische Komplexität zu bewältigen. Die wesentlichen Ziele waren (und sind):

- das gemeinsame Funktionieren voneinander unabhängiger Komponenten zu gewährleisten,
- die Wiederverwendung von Werkzeugen oder Leistungen möglich zu machen,
- eine konsistente Weiterentwicklung auch komplexer Systeme zu ermöglichen.

In den vergangenen 30 Jahren haben diese Modelle unter der Bezeichnung „Enterprise Architecture“ oder „EA“ weite Verbreitung gefunden. Eine Zusammenfassung der ersten 20 Jahre findet sich bei [LöLe13, S.102] und bei [Brau07, S.48 ff.]), Übersichten über die verfügbaren Modellierungsnotationen und Rahmenmodelle liefern unter vielen anderen [Matt11], [BeRC16], [AiRW08], [Leis06].

In der Einsicht, dass neben der Beschreibung von Organisationen und ihrer IT-Werkzeuge auch eine gemeinsame Methodik zur Erstellung und Verwendung der konzeptionellen Modelle nötig ist, wurden sie um entsprechende Vorgehensmodelle erweitert. So beschreibt der verbreitete Standard TOGAF® ([Open00]), den systematischen Aufbau und die Weiterentwicklung eines „Architecture Repository“ nach

¹ syde consulting GmbH, CH-4103 Bottmingen, andreas.becker@syde-consulting.ch

einem standardisierten Vorgehen. Das „Scaled Agile Framework“ ([Scal00]) stellt Rollen und Vorgehen stark in den Vordergrund, und berücksichtigt neuere Methoden im Bereich der Erstellung von IT-Werkzeugen (z. B. DevOps, agiles Vorgehen).

In der Praxis werden jedoch nicht wenige Enterprise Architecture-Initiativen nach kurzer Zeit eingestellt oder bleiben Papiertiger. Einer der wesentlichen Gründe für diesen Misserfolg sei gemäß [GoJa20] die „Inability to Handle Complexity by EA“. Den Bedarf, der durch die Modellierung abgedeckt werden soll, sieht [Fran16] jedoch ungebrochen, da ja die Leistungserstellungsketten angesichts weiter zunehmender Spezialisierung und durch Cloud-Architekturen immer mehr Parteien einbinden müssen.

So wird in diesem Beitrag der Versuch unternommen, die erwähnten, komplexen Modelle zu einem einfachen und flexibel einsetzbaren Kernmodell zu kondensieren. Das hier vorgestellte Modell versteht sich nicht als Konkurrenz zu den genannten Modellierungsnotationen und Rahmenwerken. Im Gegenteil ist die Absicht, diese so zu adaptieren, dass eine solch systematische Arbeitsweise auch für kleinere Organisationen möglich wird.

2 Netzwerke zur Leistungserstellung statt „Unternehmen“

Sehr generisch formuliert verfolgen Unternehmen, sowie Organisationen allgemein, den Zweck, Leistungen zu erstellen, die den Bedarf ihrer Nutzenden („Kunden“) erfüllen. Inspiriert vom Begriff des „hybriden Leistungsbündel“ bei [BeBK08, S.9 ff.] wird in diesem Text angenommen, dass solche Leistungen in aller Regel

- aus mehreren Bestandteilen bestehen
- Sachleistungen und Dienstleistungen kombinieren
- gegen Entgelt oder gratis abgegeben werden
- innerhalb oder außerhalb der produzierenden Organisation genutzt werden

Eine Leistung könnte etwa darin bestehen, eine Software mit definierten Funktionen wochentags von 8-18 Uhr für z. B. 20 Nutzerinnen und Nutzer zur Verfügung zu stellen, in dieser Zeit Telefon- und Störungssupport zu bieten und das System bei Unterbrüchen innerhalb 2min wieder herzustellen. Die Leistungen werden oft zur Erstellung weiterer Leistungen genutzt, und so entstehen Wertschöpfungsketten oder gar –Netzwerke.

Das vorgestellte Modell besteht aus einem Netzwerk von „Systemen zur Leistungserstellung“: jeder Knoten beinhaltet die Prozesse, inklusive der dazu benötigten Personen, Ressourcen, und Werkzeugen, die zur Erstellung oder Erbringung der Leistung(en) benötigt werden. Ein solcher Knoten besteht niemals ausschließlich aus technischen Werkzeugen, da ja alle technischen Werkzeuge – sowohl IT-Werkzeuge, als

auch Maschinen, Anlagen o. ä. – menschliche „Betreiber“ brauchen. In Abbildung 1 stehen die runden Elemente für solche „Leistungserstellungssysteme“.

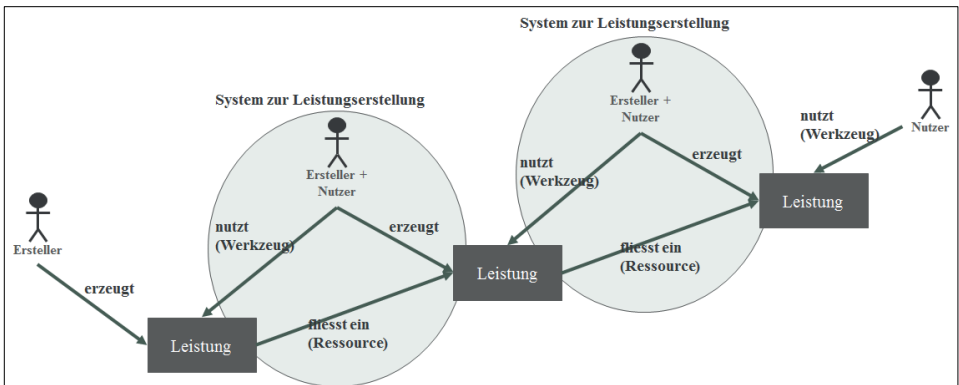


Abbildung 1: Einfache Wertschöpfungskette

Allerdings ist die Differenzierung zwischen dedizierten und geteilten Werkzeugen wichtig: Solange ein Werkzeug nur innerhalb eines Leistungserstellungssystems genutzt wird, können Herstellung und Betrieb innerhalb des entsprechenden Knoten „versteckt“ werden. Sobald jedoch etwa eine Maschine zur Herstellung unterschiedlicher Leistungen (als „shared service“) verwendet wird, wird diese Maschine mit dem Betriebspersonal in ein eigenes, separates Leistungserstellungssystem ausgelagert. Das kann auch innerhalb eines Unternehmens geschehen.

Die Verbindungen zwischen den Knoten stellen Kunden-Lieferanten-Beziehungen dar.

- Der Lieferant stellt die Leistung für den Kunden bereit.
- Der Kunde stellt Anforderungen an seine Lieferanten.

Solche Kunden-Lieferanten-Beziehungen können auch innerhalb eines Unternehmens oder innerhalb eines Konzerns entstehen. Die Abläufe im Kontext des Anforderungsmanagement und der Leistungsvereinbarung wären so für „externe“ und „interne“ Kunden identisch: Lediglich die Art der Vergütung und damit verbunden die Abläufe bzgl. Fakturierung und Inkasso unterscheiden sich diesen beiden Situationen.

Aus der Perspektive der Modellierung betrachtet entspricht ein solches Netzwerk in vielen Aspekten einem Netzwerk von Microservices (vgl. [FoLe14]). Die Architekturprinzipien sind dieselben. Wie Abbildung 2 zeigt, lassen sich die Knoten in unterschiedlicher Detailtiefe modellieren. So entstehen – abhängig von Kontext und den konkreten Anforderungen – unterschiedliche Sichten auf dasselbe Netzwerk.

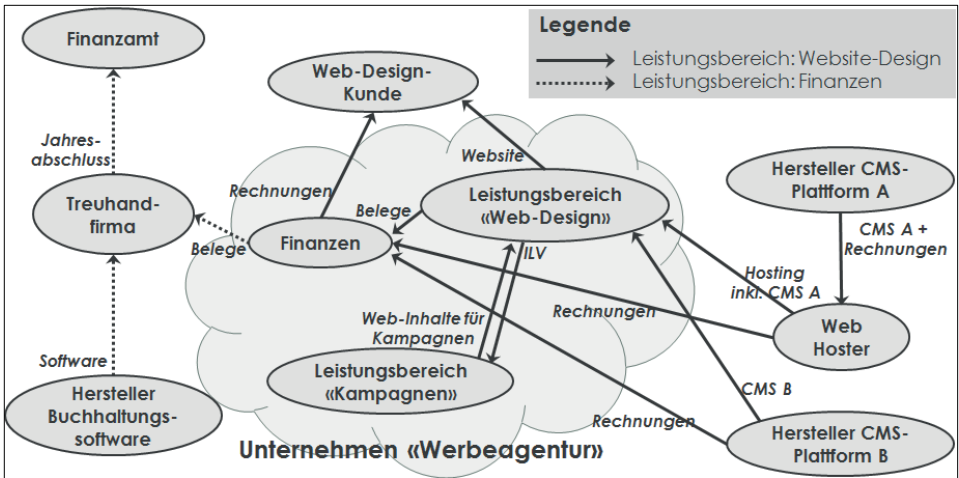


Abbildung 2: Beispiel eines Leistungserstellungnetzwerkes mit zwei Leistungsbereichen

Eine wichtige Fragestellung in diesem Zusammenhang ist die Frage des Zuschnittes der Knoten im Netzwerk. Sie ist auf organisatorischer Ebene vergleichbar mit der Problematik der Dekomposition von Services, vgl. Diskussion bei [GKGZ16]. Das Modell selbst erlaubt, ein und dieselbe Realität in unterschiedlichen Detaillierungstiefen abzubilden: Etwa könnte in Abbildung 2 in einer Darstellung zur Betrachtung der Web-Hosting-Firma umgekehrt die Werbeagentur als „Black Box“ erscheinen, während die Aktivitäten des Web-Hosters entsprechend der dortigen, internen Zuständigkeiten detaillierter dargestellt sind.

Im Grundsatz sollte die Gesamtverantwortung für die Leistungen eines Knoten in jeder Detaillierung eindeutig bestimmbar sein. Neben dieser Gesamtverantwortung können die Zuständigkeiten und Rollen innerhalb eines Knotens granularer aufgeteilt werden. Das geschieht etwa entlang der im folgenden Kapitel vorgestellten Betrachtungsperspektiven.

3 Internes Modell der Leistungserstellungssysteme

In diesem Kapitel werden die internen Strukturelemente der Knoten in den Leistungserstellungnetzwerken betrachtet. Der hier vorgestellte Ansatz strebt v. a. eine Modularisierung der Unternehmensarchitektur entlang der Leistungserstellung an und vermeidet damit Modelle, die allein ihres Umfangs wegen schwierig zu pflegen und zu verwenden sind. Daneben wird das Ziel verfolgt, Redundanzen zwischen einem EAM-Tool und operativ genutzten Werkzeugen wie Configuration Management Database [CMDB] oder Prozessmanagement zu vermeiden.

Die in Kapitel 1 erwähnten Modelle der Unternehmensarchitektur strukturieren Architekturmodelle oft in Schichten: neben der Geschäfts-, Organisations- oder Prozessebene, findet sich häufig eine Schicht, in der die Informationssysteme logisch abgebildet sind („Applikationen“ oder „IT-Services“), sowie eine Modellierung von IT-technischen und IT-Infrastruktur-Elementen. Die Modellierungselemente auf den verschiedenen Ebenen stehen durchgehend und konsistent miteinander in Beziehung. Eine Diskussion hierzu mit zahlreichen Literaturangaben findet sich z. B. bei [Brau07, S.89 ff.].

Diese Schichtung, die die Verbindung von der Geschäftsarchitektur über die logische Modellierung der Werkzeuge bis zur Technologie herstellt, ist notwendig und sinnvoll. Oft führt sie jedoch zu umfangreichen Modellierungen, die in der Praxis schwierig handhabbar sind. Der hier vorgestellte Ansatz schlägt ergänzend eine Gliederung in zwei Handlungsebenen vor:

- Die Handlungsebene „**Strategie, Architektur, Zuständigkeiten**“ enthält das logisch-abstrakte Modell für die Architektur- und Strategiearbeit.
- Die Handlungsebene „**operative Steuerung und Implementierung**“ enthält die real implementierten Objekte, die zur operativen Leistungserstellung benötigt werden.

Die vorher erwähnte Schichtung findet sich auf beiden Handlungsebenen – die Elemente sind jeweils unterschiedlich ausdetailliert. Das Modell arbeitet dazu mit drei Arten von Entitäten: Prozessen, Werkzeugen und Informationen (vgl. dazu Abbildung 3).

Die Fach- und Geschäftsperspektive wird durch die Modellierung der **Prozesse** abgebildet. Prozessmodelle sind hierbei allerdings auf einen der Netzwerkknoten begrenzt und bleiben so überschaubar. Für knotenübergreifende Leistungserstellung sind die beteiligten Knoten durch Kunden-Lieferanten-Beziehungen miteinander verbunden, vgl. dazu Kapitel 2. Auf der strategischen Handlungsebene enthält das Modell eine Leistungs- und Prozessübersicht. Auf der operativen Ebene werden detaillierte Workflows und Abläufe modelliert, bis hin zu konkreten Aufgabenbeschreibungen.

Die **Werkzeuge** umfassen alle Produktionsmittel, inklusive der Informatik. Auf der strategischen Handlungsebene wird z. B. ein IT Servicekatalog mit den fachlogischen Schnittstellen modelliert. Für Nicht-IT-Werkzeuge gibt es analog ein Inventar von Maschinen und sonstige Anlagen. Auf der operativen Ebene werden sämtliche technischen Komponenten der Werkzeuge modelliert, die innerhalb des Knotens betrieben werden, beispielsweise Software, Geräte, Datenbanken, usw. mit ihren Beziehungen untereinander. Für IT-Werkzeuge findet das üblicherweise in einer CMDB statt. Auch Elemente der technischen Infrastruktur (z. B. Virtualisierung) werden dabei sowohl „strategisch“ als IT-Service, als auch „operativ“ mit sämtlichen Elementen modelliert. Die Erstellung der Werkzeuge findet in den meisten Fällen in anderen Knoten des Netzwerks statt.

Mit Blick auf weiter zunehmende Bedeutung von **Informationen** und die wachsenden, technischen Möglichkeiten ist es wichtig, auch diese in die Modelle zu integrieren. Die strategische Architekturebene enthält dabei die Geschäftsobjekte oder Informationsentitäten, auf der operativen Ebene werden die implementierten Datenmodelle verwendet. Das strategische Informationsmodell dient als Grundlage etwa für Datenschutzaktivitäten und hilft bei der strategischen Steuerung des IT Service-Portfolio.

„Rollen“ stellen keine eigenständige Modellierungsentität dar, . Sie ergeben sich im Gegenteil aus den anderen Perspektiven: für jedes Element der genannten Perspektiven wird eine zuständige Rolle mit entsprechenden Aufgaben definiert. Im „Job-Profil“ einer Person können jedoch mehrere dieser Rollen kombiniert werden. So kann eine Person für einen oder mehrere Prozesse zuständig sein, oder aber für Prozesse, Werkzeuge und Informationen in einem Leistungsbereich. Dabei meint der Begriff „Zuständigkeit“ die Verantwortung für die Qualität der erbrachten Leistung, nicht zwangsläufig hierarchische „Führungsverantwortung“.

Die beiden Handlungsebenen sind eng miteinander verbunden. In Unternehmen ist oft zu beobachten, dass die drei genannten Entitäten operativ in eigenständigen Lösungen abgebildet sind: die Prozesse in einer BPM-Lösung, die Werkzeuge in einem ITSM-System, die Informationen in einer Software für Master-Data-Management. Das ist in dieser Form sinnvoll und ermöglicht den operativen Teams größtmögliche Autonomie in der Wahl der Arbeitsweisen und Werkzeuge.

Damit die Ziele der Architekturarbeit erreicht werden können, erfolgt dann zusätzlich die Modellierung aller Entitäten der strategischen Ebene in einem einzigen Werkzeug. Die festgelegten „Verbindungselemente“ zur operativen Ebene müssen ggf. redundant geführt werden.

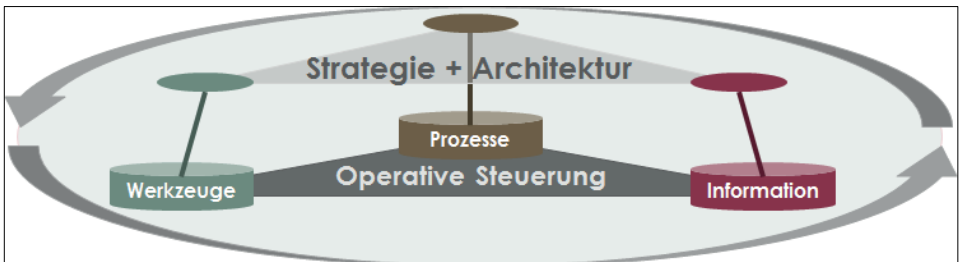


Abbildung 3: Modellierung eines Leistungserstellungssystems

Ein solches System zur Leistungserstellung lässt sich zu einem definierten Zeitpunkt als Klassenmodell darstellen. Ein entsprechendes Metamodell findet sich in Abbildung 4.

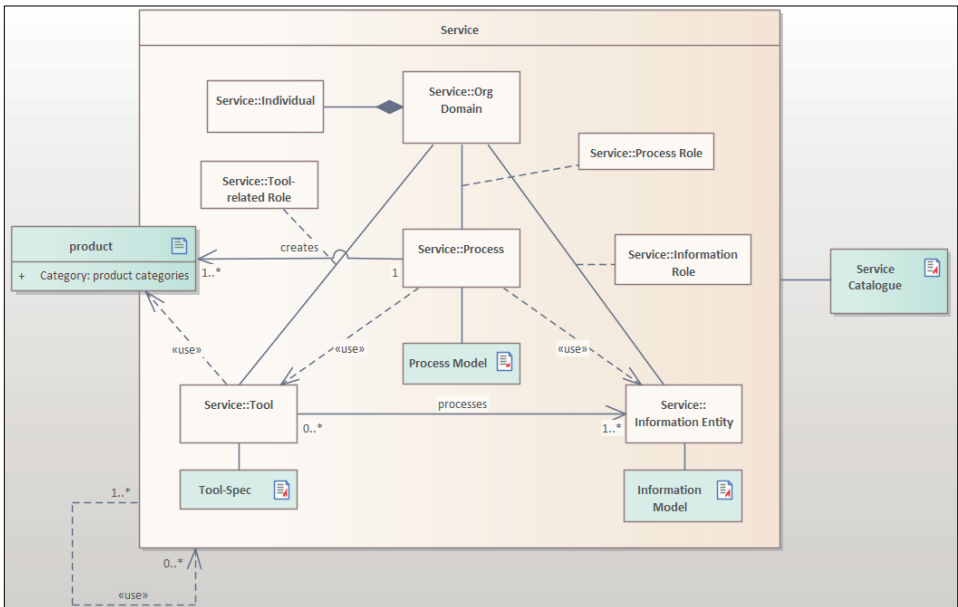


Abbildung 4: Versuch eines Metamodell eines Systems zur Leistungserstellung

Darüber hinaus zeigt Abbildung 5 das Gesamtbild dieser „statischen“ Welt, ergänzt um ihre Dynamik der Veränderung. Zu einem späteren Zeitpunkt sieht dasselbe System stellenweise anders aus, verwendet veränderte Werkzeuge und liefert veränderte Leistungen. In dieser Logik lässt sich die gesamte Organisation, inklusive der Werkzeuge sowie der Informationen, systematisch entwickeln. Die in der IT-Systementwicklung erprobten Methoden lassen sich auch hier einsetzen. So werden Änderungswünsche und Anforderungen festgehalten, Lösungsszenarien beschrieben, schließlich eines der Szenarien ausgewählt und umgesetzt. Änderungen, die Auswirkungen auf andere Knoten haben, müssen als Projekte eingespeist werden. Solche Änderungen erfolgen idealerweise im Rahmen von zyklischer und kontinuierlicher Verbesserung (vgl. beispielsweise [Bert20, Kap.11]) z. B. im Rahmen der Steuerung des Projektportfolio.

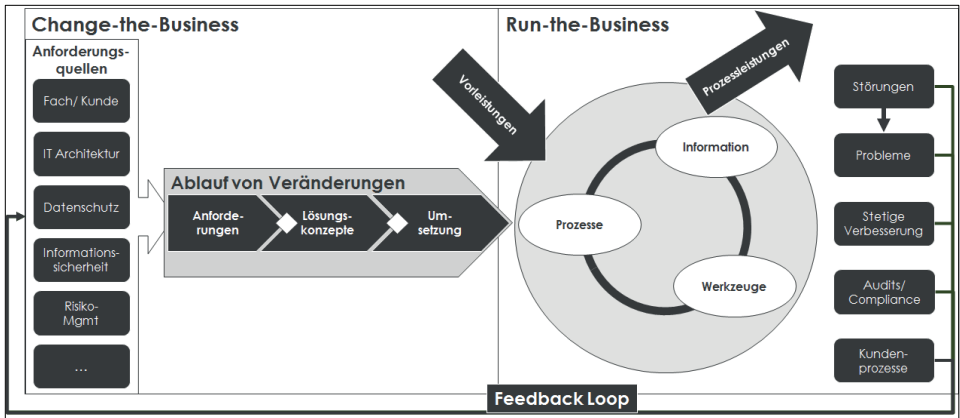


Abbildung 5: Dynamik eines Systems zur Leistungserstellung

4 Schluss und Ausblick

Das hier vorgestellte Modell, mit Fokus auf die Erstellung von Leistungen, verschlankt die Unternehmensarchitektur v. a. auf der Umsetzungsebene und ermöglicht dort eine größere Autonomie. Die Verknüpfung von Werkzeugen zur Unterstützung des operativen Betriebs ist vorgesehen. Daneben ermöglicht es auf der strategischen Ebene eine vergleichsweise einfach herzustellende Transparenz, und erlaubt die Umsetzung kontinuierlicher Verbesserung, mit der Perspektive, robuste Organisationen zu gestalten.

Zur Umsetzung von umfassenderen Veränderungen über mehrere Knoten hinweg ist eine entsprechende Koordination nötig. Dieser Bedarf entsteht jedoch nicht durch das Modell selbst, sondern ist primär durch die stärkere Arbeitsteilung in den Leistungserstellungsprozessen ausgelöst. Die Abbildung und die Steuerung unternehmensübergreifender und kollaborativer Prozesse sind vergleichsweise einfach möglich.

Eine solche, systematische Weiterentwicklung über Unternehmensgrenzen hinaus setzt eine Veränderungsbereitschaft voraus, die nicht mit den hier vorgestellten Modell allein erreicht werden kann. Sie herzustellen erfordert adäquate Steuerung und eine Kultur, in der die verantwortlichen Personen auf der operativen Ebene sich an den relevanten Schnittstellen abstimmen und eigenverantwortlich agieren können. Neben anderen ist ein dazu geeignetes Organisationsmodell [Holo00], zu den Schlagworten „Arbeitswelt 4.0“ und „Digital Leadership“ ist umfassende Literatur verfügbar.

5 Literatur

- [AiRW08] AIER, STEPHAN ; RIEGE, CHRISTIAN ; WINTER, ROBERT: Unternehmensarchitektur – Literaturüberblick und Stand der Praxis. In: *WIRTSCHAFTSINFORMATIK* Bd. 50 (2008), Nr. 4, S. 292–304
- [BeBK08] BECKER, JÖRG ; BEVERUNGEN, DANIEL ; KNACKSTEDT, RALF: Wertschöpfungsnetzwerke von Produzenten und Dienstleistern als Option zur Organisation der Erstellung hybrider Leistungsbündel. In: *Wertschöpfungsnetzwerke*, 2008, S. 3–31
- [BeRC16] BECKER, JÖRG ; RIEHLE, DENNIS M. ; CLEVER, NICO: Ansätze zur Unternehmensmodellierung – Eine Einordnung. In: BENKER, T. ; JÜRCK, C. ; WOLF, M. (Hrsg.): *Geschäftsprozessorientierte Systementwicklung. Von der Unternehmensarchitektur zum IT-System*. Bd. 37. Wiesbaden : Springer Fachmedien Wiesbaden, 2016 — ISBN 978-3-658-14825-6, S. 415–425
- [Bert20] BERTAGNOLLI, FRANK: *Lean Management*. Wiesbaden : Springer Gabler, 2020 — ISBN 978-3-658-31239-8
- [Brau07] BRAUN, CHRISTIAN: *Modellierung der Unternehmensarchitektur*. St. Gallen, Universität St. Gallen, Dissertation, 2007
- [FoLe14] FOWLER, MARTIN ; LEWIS, JAMES: *Microservices*. URL <https://martinfowler.com/articles/microservices.html>. - abgerufen am 2018-03-21. — a definition of this new architectural term
- [Fran16] FRANK, ULRICH: Konzeptuelle Modellierung: Obsoleter Kostentreiber oder zentraler Erfolgsfaktor der digitalen Transformation? In: BENKER, T. ; JÜRCK, C. ; WOLF, M. (Hrsg.): *Geschäftsprozessorientierte Systementwicklung. Von der Unternehmensarchitektur zum IT-System*. Bd. 3. Wiesbaden : Springer Fachmedien Wiesbaden, 2016 — ISBN 978-3-658-14825-6, S. 33–49
- [GKGZ16] GYSEL, MICHAEL ; KÖLBENER, LUKAS ; GIERSCHE, WOLFGANG ; ZIMMERMANN, OLAF: Service Cutter: A Systematic Approach to Service Decomposition. In: AIELLO, M. ; JOHNSEN, E. B. ; DUSTDAR, S. ; GEORGIEVSKI, I. (Hrsg.): *Service-Oriented and Cloud Computing. 5th IFIP WG 2.14 European Conference SOCC 2016 Vienna Austria September 5-7 2016 Proceedings, Lecture Notes in Computer Science*. Bd. 9846. Cham : Springer International Publishing, 2016 — ISBN 978-3-319-44481-9, S. 185–200
- [GoJa20] GONG, YIWEI ; JANSSEN, MARIJN: Exploring Causal Factors Influencing Enterprise Architecture Failure. In: SHARMA, S. K. ; DWIVEDI, Y. K. ; METRI, B. ; RANA, N. P. (Hrsg.): *Re-imagining Diffusion and Adoption of Information Technology and Systems: A Continuing Conversation*, 2020

- [Holo00] *Holacracy - Evolve Your Organization*. URL <https://www.holacracy.org>. - abgerufen am 2021-05-28. — Holacracy
- [Leis06] LEIST-GALANOS, SUSANNE: *Methoden zur Unternehmensmodellierung - Vergleich, Anwendungen und Integrationspotenziale*. Berlin : Logos, 2006 — ISBN 3-8325-1041-9
- [LöLe13] LÖHE, JENS ; LEGNER, CHRISTINE: Overcoming implementation challenges in enterprise architecture management: a design theory for architecture-driven IT Management (ADRIIMA). In: *Information Systems and e-Business Management* (2013), Nr. 1/2014, S. 101–137
- [Matt11] MATTHES, DIRK: *Enterprise Architecture Frameworks Kompendium*, 2011
- [Open00] THE OPEN GROUP: *TOGAF®*. URL <https://www.opengroup.org/togaf>. - abgerufen am 2021-05-18. — The TOGAF® Standard, Version 9.2
- [Scal00] SCALED AGILE INC.: *SAFe® 5.0 Framework*. URL <https://www.scaledagileframework.com/>. - abgerufen am 2021-05-18. — Scaled Agile Framework

Organizational Performance Management – From Planning to Execution and Re-Alignment

A holistic approach

Philipp Klingmann¹, Albrecht Altenhoener², Dr. Carsten Brockmann³, Selina Silbernagel⁴

Abstract: The organizational performance management frameworks presented in this contribution cover the whole lifecycle from planning to monitoring up to intervening and re-aligning. The business architecture shall always be aligned with the technological architecture through the supporting elements. In addition, Governance and people play a crucial role in defining and achieving performance goals, even allowing organizations to exceed their goals.

Keywords: Organizational performance management, plan, monitor, intervene, CFO Services, strategic cost transformation, business finance, controlling.

1 Introduction

Our increasingly volatile economic environment is providing enormous challenges for today's businesses. Factors such as globalization and demographic change as well as disruptions in business continuity have become massive turning points that are also redefining the role of the Chief Finance Officer. In addition to pushing for value-adding growth, CFOs and their teams need to take on several roles at once instead of focusing solely on monthly/quarterly/yearly results. Business partnering is likely to gain a more dominant position within Finance, moving the CFO role more towards that of value architect. Capable of balancing major business issues in a volatile and complex environment, value architects identify competitive cost advantages and design digital structures that are fit for purpose. This also includes leveraging specific growth potential and differentiating the business from competitors. Transparency as well as legal and regulatory compliance are also becoming an increasingly important part of the Finance function's portfolio. From identifying and implementing compliance policies to establishing governance structures and issuing sustainability and environmental impact reports, today's CFO is called to deliver a cross-functional set of skills.

¹ Deloitte Consulting GmbH, Finance & Performance, pklingmann@deloitte.de

² Deloitte Consulting GmbH, Finance & Performance, aaltenhoener@deloitte.de

³ Deloitte Consulting GmbH, Finance & Performance, cbrockmann@deloitte.de

⁴ Deloitte Consulting GmbH, Finance & Performance, sisilbernagel@deloitte.de

On one side, CFOs can make their own actions more efficient; on the other side, they are in a position to establish an Organizational Performance Framework designed to leverage gains across the organization as a whole.

The Finance function has made enormous efficiency gains over the last few years. Back in 2010, the average company spent 1.5 percent of its revenue running Finance; today it only spends 1.0 percent. Top performers even spend as little as 0.56 percent of their revenues to run the Finance department [1]. An increase in process efficiency, systems and people have been key drivers for this change.

By documenting, standardizing and streamlining processes, they paved the way to develop a mature governance model that not only covers process specifics but also challenges the existing data basis. Digitalization has become a significant agent of change on the systems side, where the finance function can lead the way [2, 3]. New digital technologies, from cloud solutions to robotic process automation, are replacing manual tasks and reducing both the effort involved in transactional processing as well as headcount in Finance. Although digitalization is high on the CFO agenda, plans are far from being completed: estimates show that the Finance workforce is likely to decrease by 15-24 percent in the future [4] – an even stronger indicator for efficiency gains in Finance.

Finance departments are responsible for delivering the results that support the business and bolster the company's growth in a competitive market. They can use a structured approach and accelerators such as architectural templates [5], scenario analyses, analytics-based solutions and business-modelling capacities to differentiate their business from the competition. Establishing a corporate culture that supports this mission is a critical success factor. Finance professionals can make a decisive impact by implementing a well-designed Organizational Performance Management framework.

Analytics, driver-based forecasting and scenario planning generate progressive insights once deployed. For the Finance function, focusing future decision-making on creating value, driving operational excellence and providing multi-dimensional insights is key in providing holistic performance management for the entire organization based on transparency and reliability.

2 Our organizational performance framework

Our performance management framework consists of three major process steps – setting the target, monitoring the execution and intervening if required. People, IT and organizations are all enablers that allow the Finance function to grow and thrive.

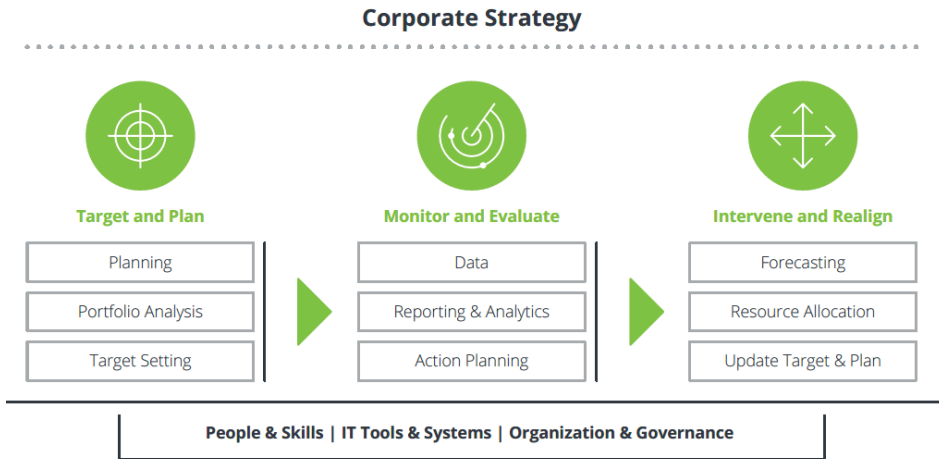


Figure 1: Organizational Performance management framework

2.1 Phase Target & Plan

In this phase, organizations set targets using a structured portfolio analysis and the corporate strategy to identify and operationalize overarching goals.

The planning cycle starts by updating strategically relevant long-term goals covering a timespan of three to five years. In a next step, the focus shifts to the upcoming financial year by setting financial and non-financial targets. This requires the Finance team to identify key drivers of business value and operationalize strategic goals accordingly by taking both current and future drivers into account, from an internal as well as external perspective.

A portfolio analysis allows us to operationalize high-level targets for growth. For each growth target, we identify and group opportunities by product, customer and market, and then identify and quantify both vertical and horizontal growth opportunities. Market opportunities may also highlight potential shifts in the offering portfolio, which could change the overall business portfolio in terms of short and long-term value creation.

During the target setting phase, the opportunities we identified are cascaded into an action plan, resulting in goals measured with financial and non-financial KPIs. Our experience has shown that it can be beneficial to choose a setting physically separate from the usual office environment to define your overarching goals. We find formats like conclaves, CFO labs or other events very productive, as they allow participants to focus on goal-setting for a full day or longer. Strategic goals are broken down into concrete long and short-term targets, shifting the focus from past performance to a more forward-looking view. Once the targets are set, it is time to cascade them to all levels of the organization and place the

focus clearly on the KPIs we need to measure progress and make adjustments when required.

2.2 Phase monitor and evaluate

In the Monitor & Evaluate phase, the focus is on the continuous and close monitoring of the targets and the action plan to improve the current state.

Finance departments need to acquire data from multiple sources to determine the overall progress with their goals, focusing on consistency in data structures. Developing a central governance and quality management structure helps ensure that a single view on the truth can be generated. Where no central governance structure exists, individuals at the business unit level are tasked with setting up and maintaining data collection systems as well as the continuous provision for interconnected areas.

The reporting and analytics step analyzes all of the data collected to determine where there are discrepancies and the potential actions to remedy them. The reporting feature uses an event-driven model to issue reports outlining actions that will improve on the current situation. In order to make reports more actionable, the system sets and updates thresholds for key values and the resulting tasks. Structured real-time data is used to identify performance discrepancies.

The action-planning step to improve the current situation relies on data as well as reports. These individualized reports with detailed visualizations can generate insights faster and ease the decision-making process. Precise measures are defined for each target, while forecasting as well as action plans become more flexible to adapt to the outcome of business events.

2.3 Phase Intervene and Re-Align

During the planning process and subsequent evaluation of the measures taken, managers may have to make interventions if the established plan changes. They develop corrective measures during this phase and implement them in response to the situation as it now stands.

In the forecasting step, Finance departments update data from the original plan in line with the latest developments, providing a new forecast as an initial indicator for a possible mismatch between the current state and the established plan. In addition to analyzing the plan, they can elaborate multiple scenarios that consider external variables by applying their knowledge on intrinsic and extrinsic motivations of actors.

Overall, resources should be allocated where they add the most value. If plans change, the Finance function needs to make a holistic evaluation of any adjustments in resource allocation that result from the intervention. On the one hand, this pertains to financial

resources designed to create value. On the other hand, the allocation of non-financial resources, such as human resources, is also a key element to success. After all, the most value is created when only the right resources are deployed at the right location.

In this last part of our framework, Finance departments update overall targets and plans with any changes that have occurred in the respective systems. They communicate any modifications made to the resource allocation across the hierarchy and adjust individual goals/targets where necessary.

2.4 Enabling Elements

Enablers foster value creation throughout the planning cycle by creating links between the process steps. They also support process execution and record all changes that result from execution or re-planning.

In terms of personnel management, we know well that competences and capabilities vary between employees. Simply deploying resources does not take their specific background into account. Each employee has his or her own personal perspective on performance management and the value created by individual actions. That is why individualized communication with clear objective-setting between employees and their superiors is so important for good performance, along with clear guidance on how to achieve those goals. Providing open dialogue on performance and the right incentives, such as aligned and structured targets, short and long-term horizons or relative ambitions, can boost performance for individual staff members. Teams that introduce modern approaches like lean or agile management may increase employee satisfaction by involving staff more closely with the value creation process.

For all elements of the performance management methodology, Finance departments need robust IT systems to store and monitor the results of each process step. The best tools will provide an analytical environment as well as the possibility to run predictive forecasts and build custom reports.

In terms of corporate structure, Finance should promote its position as a partner to business entities. Defining clear roles and responsibilities will enable Finance to fulfill that mission, while still focusing on improving current transactional processes.

3 Our recommendations

One of the primary goals of Finance is to reduce costs and it is precisely for this reason that shared services are once again coming to the fore. 70 percent of German CFOs are looking to reduce costs by deploying Shared Service Centers or Global Business Services [6]. This is quite a drastic development: in a previous survey conducted just one year

earlier, they had described Shared Services as a means to add decisive business value for the company as a whole [7].

We recommend focusing on the value being created, while letting go of activities that do not create value. Zero Based Budgeting (ZBB) has proven to be one of the leading practices for companies looking for a different approach to allocating funds. With this approach, the budget is planned from scratch each period, i.e. starting with a “zero base”. This ensures that the budget is based exclusively on necessary expenses. Stakeholders need to explain and justify each spending item, which guarantees that the focus is placed solely on essentials.

Our approach to Strategic Cost Transformation recognizes the general trend moving away from strictly tactical cost reductions towards a more strategic approach. It supports and enables clients to optimize and align their costs strategically by implementing traditional cost approaches, a targeted redesign of the operating model and/or next-generation cost solutions that leverage state-of-the-art technologies [8]. The overall objective is to make clients adopt a save-to-transform mindset, i.e., a sustainable change in their business and operations [9], which will finally pave the way for integrated organizational performance management.




4 References

1. APQC. *Key Finance Organization Benchmarks – Cross Industry*. 2020; Available from: <https://www.apqc.org/resource-library/resource-listing/key-finance-organization-benchmarks-cross-industry>. Last access: 31st of May 2021
2. Epstein, R., et al., *Die Finanzfunktion als Accelerator einer agilen Organisation*. CFO Insights, 2019. **November**: p. 1 - 7.
3. Brockmann, C., et al., *Stepping stones to an agile enterprise*. Deloitte Review, 2019(25): p. 36 - 47.
4. Deloitte. *CFO Survey - Konjunktur, Klima, Talente*. 2019; Available from: https://www2.deloitte.com/content/dam/Deloitte/de/Documents/finance-transformation/CFO_Survey_Herbst_2019.pdf. Last access: 31st of May 2021
5. Brockmann, C., et al., *Die IT-Finanzarchitektur im Cloudumfeld*, in *INFORMATIK 2020*, R.H. Reussner, A. Koziol, and R. Heinrich, Editors. 2020, Gesellschaft für Informatik, Bonn: Bonn. p. 197 - 203.
6. Deloitte. *CFO Survey - Sustainable Finance & Service Delivery – Transformation als Chance*. 2020; Available from: <https://www2.deloitte.com/content/dam/Deloitte/de/Documents/finance-transformation/CFO-Survey-2020-Sustainable-Finance.pdf>. Last access: 31st of May 2021
7. Deloitte. *Global Shared Services Survey Report - 11th biennial edition*. 2019; Available from: <https://www2.deloitte.com/content/dam/Deloitte/us/Documents/process-and-operations/2019-global-shared-services-survey-results.pdf>. Last access: 31st of May 2021
8. Deloitte. *Solutions - Strategic Cost Transformation*. 2021; Available from: <https://www2.deloitte.com/de/de/pages/operations/solutions/strategic-cost-transformation.html>. Last access: 31st of May 2021
9. Deloitte. *Deloitte's second biennial global cost survey - Save-to-transform as a catalyst for embracing digital disruption*. 2019; Available from: <https://www2.deloitte.com/content/dam/Deloitte/global/Documents/Process-and-Operations/gx-consulting-global-cost-survey-2019.pdf>. Last access: 31st of May 2021

Workshop: Datengetriebene
Geschäftsmodelle für nachhaltige
Wertschöpfungssysteme
(NaWerSys)


Workshop: Datengetriebene Geschäftsmodelle für nachhaltige Wertschöpfungssysteme (NaWerSys)


Eine Einführung


Friedemann Kammler ¹, Paul Christoph Gembarski ², Thorsten Schoormann ³
Simon Hagen⁴

1 Ausgangssituation

Produkte und Dienstleistungen, die neue Features durch den Einsatz von Informations- und Kommunikationssystemen (IKT) ermöglichen, stellen heute kein Novum mehr dar. In der Anwendung für innovative Geschäftsmodelle bilden „Smarte Produkte“ [PH15] und deren systematische Vernetzung in Cyber-Physischen Systemen (CPS) bereits heute einen wichtigen Mechanismus zur Erbringung innovativer Leistungen. Treibstoff dieses Mechanismus⁴ sind Daten, die als virtuelles Wirtschaftsgut zwischen Herstellern und ihren Kunden ausgetauscht werden und auf diese Weise Kreisläufe fundamental verändern. Dabei treten Kunden aus Unternehmensperspektive nicht länger ausschließlich als Konsumierende auf, sondern werden für die Hersteller zu wichtigen Anbietern von Daten und darin enthaltener Information. Hierdurch entwickeln sich neue Geschäftsmöglichkeiten. Offensichtlich erscheint die Erfassung von Marktwissen, für die Hersteller bereits heute automatisch Betriebsdaten zum Gebrauch ihrer Produkte beziehen und zukünftige Entwicklungen passgenau auf Kundenbedürfnisse ausrichten [St19]. Gleichermäßen verspricht der „Treibstoff“ Daten aber auch kundenseitiges Nutzenpotenzial: Datenaustausch zwischen vernetzten Produkten trägt zum Beispiel zu auf Künstlicher Intelligenz basierenden Fähigkeiten, wie der Perzeption und situativen Schlussfolgerung des Systems bei und erlaubt Systemen auf diese Weise das „kontinuierliche Erlernen“ komplexer Fähigkeiten, wie der Fähigkeit zur Adaption in einem dynamischen Umfeld (u.a. spezifiziert als Kontextadaption in „Smart Systems“ [RO20]). Derartige Technologien implizieren einen großen Schritt in Richtung kundenzentrierter Lösungen. Durch virtuelle Konfiguration (i.S.v. Mass Customization) werden Produkte und Dienstleistungen unterschiedlicher Anbieter integriert sowie automatisch und kontinuierlich an die Bedürfnisse und Gewohnheiten ihrer Nutzer

¹ Deutsches Forschungszentrum für Künstliche Intelligenz GmbH (DFKI), Smart Enterprise Engineering, Parkstraße 40, 49080 Osnabrück, friedemann.kammler@dfki.de,  <https://orcid.org/0000-0002-2418-4882>

² Leibniz Universität Hannover, Institut für Produktentwicklung und Gerätebau (IPEG), An der Universität 1, 30823 Garbsen, gembarski@ipeg.uni-hannover.de,  <https://orcid.org/0000-0002-2642-3445>

³ Universität Hildesheim, Institut für Betriebswirtschaft und Wirtschaftsinformatik, Universitätsplatz 1, 31141 Hildesheim, schoormann@uni-hildesheim.de  <https://orcid.org/0000-0002-3831-1395>

⁴ Deutsches Forschungszentrum für Künstliche Intelligenz GmbH (DFKI), Smart Enterprise Engineering, Parkstraße 40, 49080 Osnabrück, simon.hagen@dfki.de

angepasst [KKG21]. Auf der anderen Seite setzt die umfassende Gestaltung derartiger Leistungen hohe Entwicklungsaufwände und Koordination im späteren Betrieb voraus.

Der Workshop NaWerSys verbindet drei erforderliche Betrachtungswinkel. Zunächst erweitern wir den gestalterischen Fokus von der Einzelleistung auf deren Beitrag zu ganzen Wertschöpfungssystemen. Hierfür ist im Sinne der ökonomischen, aber auch ökologischen und sozialen Nachhaltigkeit eine Vielzahl von Wechselwirkungen zu berücksichtigen, die harmonisiert werden müssen, um langfristige Wirksamkeit sicherzustellen. Neben der reinen Erklärbarkeit setzt die praktische Anwendung letztlich auch die Entwicklung von neuen Konstruktionstechniken voraus, die komplexe, interdisziplinäre Funktionszusammenhänge in plausible Teilprobleme zerlegen, die fachkundig gelöst werden können.

2 Von innovativen Leistungen zu datengetriebenen Wertschöpfungssystemen

Produkte und Dienstleistungen, die innovative Funktionen durch Daten erschließen, entstehen gegenwärtig in nahezu jeder Wertschöpfungskette. Im Kern substituieren viele Leistungen manuelle Arbeitsschritte durch automatisierte Funktionen, z.B. Recommender Systeme als Ersatz von wissensintensiver Beratung im Verkauf [FE19], oder optimieren Zeit- und Kostenaufwand sowie Qualität, z.B. über Bilderkennung in der Qualitätsprüfung oder Bauteilidentifikation [LSK20]. Traditionell als Endprodukt betrachtete Güter erzeugen mit den erfassten Daten darüber hinaus aber oftmals auch einen Rohstoff für komplementäre Leistungen: Maschinen, die Betriebsdaten zur Diagnose von Stillständen bereitstellen, schaffen so etwa einen Mehrwert für die Optimierung eines universellen Instandhaltungsprozesses auf tatsächlich benötigte Arbeitsschritte („Tailored Services“) [Ka19]. In der Sache bereits in Beziehung stehende Güter können durch die Erweiterung mit IKT also auch faktisch als Datenerzeuger und -verbraucher in einem Ökosystem verknüpft werden. Die Entwicklung der hierfür erforderlichen Grundstruktur wird bspw. durch die Initiative GAIA-X verfolgt [GA20].

Am Anwendungsfall betrachtet wird diese Veränderung noch deutlicher: Vereinfacht kann jedes Geschäft als abgeschlossener Leistungsaustausch zwischen einem Anbieter und seinem Kunden verstanden werden. Dies ist im Kern unmittelbar auf „High-Tech“-Produkte und -Dienstleistungen anwendbar. Erfolgt stufenweise Kauf, Veredelung und Verkauf einer Leistung, werden diese Schritte logisch als Wertschöpfungskette zusammengefasst: Auf diese Weise tragen Stahlproduzenten, Teilefertiger und Automobilbauer in Stufen zum gleichen Produkt – einem Auto – bei. Wird dieser lineare Zusammenhang aufgebrochen und unterschiedliche Wertschöpfungsketten ineinander verschränkt, um eine Reihe komplementärer Leistungen zu realisieren, sprechen wir von einem Wertschöpfungssystem. In der Natur von Wertschöpfungssystemen liegt deshalb bereits eine grundsätzliche *Komplexität*, die gerade dann entsteht, wenn einzelne Teilnehmer eine Vielzahl von Rollen und Beziehungen zueinander einnehmen und sogar

„Rückflüsse“ entgegen der ursprünglichen Veredelungsfolge der „Kernleistung“ verursachen. Zum Vergleich: In einem Wertschöpfungssystem wird der Leasingnehmer eines Fahrzeugs auch gleichzeitig zum Anbieter von Verkehrsdaten auf den gefahrenen Strecken. Dabei folgen die ineinander verschränkten Leistungen keinesfalls synchronen Lebenszyklen, sondern überlappen einander und erzeugen auf diese Weise *Kontinuität*. Wertschöpfungssysteme dienen in diesem Sinne der Verstetigung von Kundenbeziehungen. In ihrer Anlage können derartige Systeme außerdem *offen* – d.h. jederzeit durch komplementäre Angebote anderer Anbieter erweiterbar – oder *geschlossen* gestaltet werden. Während geschlossenen Systemen die Herausforderung innewohnt, dass ein gesamtes Ökosystem aus eigener Hand bewirtschaftet werden muss, bringen offene Systeme *Dynamik* mit sich. Denn Kunden und Anbieter können das System verlassen (oder herausgedrängt werden), während andererseits neue hinzukommen und so eine Prognose über Bedarfe und Absatzzahlen langfristig mindestens mit großen Unsicherheiten verbinden.

Um die vielschichtigen Zusammenhänge eines Wertschöpfungssystems zu entwirren, erscheinen Geschäftsmodelle als zweckdienliche Abstraktion einzelner Interessen. Deren Modellierung (z.B. mittels Software [Sz20]) umfasst u.a. die Benennung zentraler Ressourcen, Kundengruppen, Partner sowie Kosten- und Ertragsstrukturen.

3 Nachhaltigkeit als Schnittpunkt komplementärer Interessen

Durch die passgenaue Optimierung des Ressourceneinsatzes auf die individuellen Bedürfnisse des einzelnen Kunden bieten datengetriebene Wertschöpfungssysteme das Potenzial, einen Beitrag zur Erreichung von Nachhaltigkeitszielen zu leisten. Aus *ökologischer Perspektive* betrifft dies etwa die Virtualisierung von Einzelleistungen, indem einfache physische Produkte, wie ein Wecker, in universellere und standardisierte Produkte, wie Smartphones, integriert und durch datengetriebene Funktionen, wie die Ableitung von typischen Weckzeiten, angereichert werden. Analog werden auch Dienstleistungen, die Wege- und Zeitaufwand verursachen, durch digitale Services ersetzt (i.S.v. Variantenreduktion). Mit Perspektive auf den langfristigen Betrieb von Kundenlösungen erscheint die virtuelle Konfiguration von Leistungsbündeln (i.S.v. Product-Service Systemen) als Ziel, um die kontinuierliche Anpassung an Veränderungen der Kundenbedürfnisse und des technischen Stands (i.S.v. Retrofit und Refurbishment) zu ermöglichen [GL18]. Letztendlich bieten sich auch für „ausgemusterte“ Leistungen Potenziale, indem bspw. alternative Verwendungen und Recycling-Pfade (i.S.v. Circular Economy) abgeleitet werden.

Der Datenbedarf für Betrieb und Verbesserung derartiger Funktionen begründet neue Märkte. Hierdurch sind den ökologischen Potenzialen auch neue *ökonomische Fragestellungen* gegenüberzustellen, die beantwortet werden müssen. Kurze Innovationszyklen und das schnelle Eintreten der Obsoleszenz einer Technologie erzeugen in solchen Märkten eine kontinuierliche Dynamik, die die Verfügbarkeit und Kompatibilität von

angebotenen und benötigten Daten negativ beeinflussen kann. Um diese Gefahr zu entschärfen, gilt es, auch die ökonomische Nachhaltigkeit ins Blickfeld zu rücken und auch den automatisierten Datenaustausch zwischen Produkten und Dienstleistungen als eigenständige Leistungen im Wertschöpfungssystem zu betrachten. Hierfür ist die automatisierte Preisbildung erforderlich, die in vielen Fällen von der Informationsqualität eines Datensatzes (Enthaltene Fakten, Richtigkeit, Aktualität, Aufbereitung) abhängig ist. Mit der Implementierung von Abrechnungsmechaniken, wird die Bedeutung von Produkten und Dienstleistungen als „systemrelevanter“ Datenanbieter hervorgehoben und in kontinuierliche Wertschöpfungsprozesse eingewoben (direkt, oder z.B. durch Tausch- und Kompensationsgeschäfte).

Letztlich gilt es auch, die *soziale Nachhaltigkeit* von datengetriebenen Wertschöpfungssystemen genauer zu untersuchen [SK20]. Dabei ist abzusehen, dass individuelle Kundenlösungen, die innerhalb eines Systems virtuell gebündelt werden, einzelne Produkte und Dienstleistungen obsolet machen können und so einen zentrierten und verschärften Wettbewerb auslösen können. Beherrschen einzelne Anbieter ganze Systeme, besteht außerdem die Gefahr von schwer reversiblen Marktstrukturen, die letztlich weitere Anbieter verdrängen („Winner-takes-all-Effekt“ [HA17]). Dem gegenüber steht das Potenzial offener Wertschöpfungssysteme, in denen neue Anbieter zur Erweiterung des Lösungsraums beitragen und langfristige wechselseitige Kooperationen mit ihren Kunden etablieren.

Der Workshop NaWerSys möchte diese ökonomischen, ökologischen und soziale Potenziale durch datengetriebene Wertschöpfungssysteme und Geschäftsmodelle aufgreifen und bietet eine Plattform für Forschung und Praxis, um sowohl methodische Aspekte, als auch (praktische) Beispiele zu diskutieren.

4 Integriertes Engineering für datengetriebene Wertschöpfungssysteme

Es scheint absehbar, dass auch zukünftig Geschäftsmodelle die Potenziale von Daten und deren Analyse stärker betonen werden, sei es durch deren Bepreisung als Rohstoff oder durch die funktionale Steigerung des Kundennutzens im Hinblick auf Zeitersparnis, Prozessvereinfachung, Integration von Einzelleistungen oder generellen Qualitätssteigerung [ASB16]. Um datengetriebene Geschäftsmodelle erfolgreich umzusetzen, müssen materielle und virtuelle Entwicklungstätigkeiten integriert betrachtet werden. Eine Kerntechnologie für diese Aufgabe ist der digitale Zwilling, der Produkte und Dienstleistungen virtuell repräsentiert und so Tests und Optimierung neuer Betriebsparameter, die Analyse von geänderten Umweltbedingungen oder die digitale Fehleranalyse ermöglicht. Sinnbildlich gesprochen erzeugen derartige Technologien eine „virtuelle Parallelwelt“, in der die informationellen Anteile der Leistung abgebildet werden. Der Nutzen liegt auf der Hand: Eine Änderung an der Maschine, egal ob bei Wartung, Instandsetzung oder betrieblicher Optimierung, kann im Vorfeld präzise getestet

werden, die Stillstände und auch das Risiko sind deutlich geringer [AGC19]. Bei der Erbringung von Dienstleistungen steigern digitale Assistenten ebenfalls die Sicherheit und Qualität und ermöglichen auf diese Art und Weise im After-Market neue, Self-Service-basierte Wartungsmodelle. In diesem Sinn könnten in Zukunft ganze Geschäftsmodelle (und damit Anteile eines Wertschöpfungssystems) vollständig durch Information abgebildet werden. Dieses Potenzial wird in der betrieblichen Praxis jedoch bislang häufig noch nicht gehoben, weil das „konsequente Mitdenken“ virtueller Leistungsbestandteile und alternativer Verwendungsformen der eigenen Produkte und Dienstleistungen ausbleibt. Mehr noch, eine Reihe von praktischen Beispielen zeigt, dass Produkte ergebnisoffen mit Technologie zur Datenerfassung und -verarbeitung ergänzt werden, ohne weiterführende Nutzungsszenarien zu spezifizieren und in deren Ergebnisse in den Konstruktionsprozess einfließen zu lassen. Im Ergebnis lassen sich aus den aufgenommenen Daten keine, oder nicht die erforderlichen Informationen ableiten.

Für Entwickler ergibt sich daraus zukünftig der klare Auftrag, auch Fragen an das Produkt zu formulieren, die auf informationeller Ebene beantwortet werden können, und auf dieser Basis zu integrierende IKT (z.B. integrierte Wissensspeicher, aber auch Sensorik) auszuwählen [De16]. Bekannte erfolgreiche Implementierungen der virtuellen Gestaltungsebene verfolgen zumeist einem top-down-Ansatz. Sie werden also als abgeschlossenes System durch einen einzelnen Anbieter mit ausreichenden Kapazitäten und einer strategischen Geschäftsplanung entwickelt. Im Betrieb des Systems erschließt der Anbieter dann neue Geschäftsfelder, indem zusätzlicher Kundennutzen durch Leistungsbausteine auf der virtuellen Gestaltungsebene erzeugt wird. Ein vielbeachtetes Beispiel ist die nachträgliche Integration der autonomen Steuerung in Tesla Fahrzeugen. Entwicklungen setzen dafür einerseits auf einem strategischen Programmmanagement für Funktionen auf, müssen andererseits aber auch einen „virtuellen Freiheitsgrad“ implementieren, der Folgeinnovationen zulässt. Aus dieser Perspektive darf es keine Rolle spielen, ob die spätere Verwendung erfasster Daten in der Zustandsüberwachung einer im Betrieb befindlichen Maschine liegt, oder die Ableitung von Anforderungen für die Folgegeneration des Produkts ermöglicht [Ge20]. Ein analoger Gestaltungsfreiheitsgrad muss auch für Dienstleistungen betrieben werden, um eingesetzte IKT und den damit deutlich gesteigerten Entwicklungs- und Integrationsaufwand für kontinuierliche Dienstleistungsinnovationen operationalisieren zu können.

5 Anlass und Anliegen des NaWerSys-Workshops

Die nachhaltige Konzeption, Entwicklung und der Betrieb von datengetriebenen Wertschöpfungssysteme erfordern interdisziplinäre Kompetenzen, die eine Herausforderung für alle beteiligten Gestaltungsdisziplinen darstellen. Das Anliegen des NaWerSys-Workshops liegt deshalb darin, Entwickler übergreifend zusammenzubringen, erfolgreiche Anwendungen zu diskutieren und gemeinsame Konzepte, Modelle und Methoden zu prägen, die zur systemischen Gestaltung beitragen. Artefakte werden so im interdisziplinären Austausch aus einer Vielzahl von Perspektiven validiert und können –

insofern sie sich als zweckdienlich erweisen – als akzeptierte Standards zur Lösung von Schnittstellenproblemen beitragen. Der erste NaWerSys-Workshop auf der INFORMATIK 2021 nimmt diese Diskussion in fünf Beiträgen auf.

Der Beitrag von *Christoph Heinbach* und *Oliver Thomas* adressiert die Herausforderung der Dekarbonisierung von Lieferketten. Mit dem „Green Service Canvas“ entwickeln die Autoren ein visuelles Werkzeug, mit dem insbesondere kleine und mittelständische Unternehmen in der Entwicklung neuer ökologisch nachhaltiger Servicemodelle unterstützt werden sollen.

Jonas Rebstadt, *Henrik Kortum*, *Simon Hagen* und *Oliver Thomas* betrachten das Wertschöpfungssystem Smart Living und zeigen in Anlehnung an die europäische Initiative GAIA-X mit der „Service Registry“ einen prototypischen Ansatz zur Koordination unterschiedlicher Dienstleistungsangebote.

Hannes Jahn und *Hans-Jürgen Pfisterer* entwickeln und evaluieren eine prototypische Lösung zur dynamischen Energieversorgung, indem sie einen mobilen Speicher um eine Datenerfassungstechnologie erweitern.

Im vierten Beitrag erforschen *Philipp Fukas* und *Oliver Thomas* die Schnittstelle zwischen Verfahren der Künstlichen Intelligenz und „Progressive Web Apps“ und zeigen deren Potenzial, Beiträge zur Nachhaltigkeit zu leisten.

Michael Ganske, *Katherina Meißner* und *Julia Rieck* gehen der Frage nach, inwieweit Computerspiele aus dem Bereich der Wirtschaftssimulationen für das Erlernen von Supply Chain Management-Fähigkeiten im Bereich der universitären Lehre angewendet und hinsichtlich Nachhaltigkeitsfragestellungen erweitert werden können.

6 NaWerSys-Programmkomitee und Danksagung

Der Workshop wäre nicht möglich gewesen ohne die Unterstützung durch das Organisationsteam der Jahrestagung INFORMATIK 2021. Ein besonderer Dank geht darüber hinaus auch an das NaWerSys-Programmkomitee, das diesen Workshop mit sorgfältig erstellten Gutachten unterstützt hat:

- Dennis Behrens (Data Scientist)
- Helge Fischer (FOM Essen)
- Lucas Hüer (HS Osnabrück)
- Marie Kammler (HWG Ludwigshafen)
- Robert Keller (Hochschule Kempten)
- Babett Kühne (Universität Hamburg)

- Sven Jannaber (LVM Versicherung)
- Kai Lingemann (DFKI Osnabrück, Planbasierte Robotersteuerung)
- Cristina Mihale-Wilson (Goethe-Universität Frankfurt)
- Frederik Möller (TU Dortmund)
- Sebastian Sieloff (Startup Center HI-Cube GmbH)
- Maren Stadtländer (Universität Hildesheim)
- Frederic Stahl (DFKI Oldenburg, Marine Perception)
- Andreas Varwig (Consultant Data Science)
- Sebastian Werning (IU Internationale Hochschule)
- Jannis Vogel (Universität Osnabrück)

Wir wünschen allen Teilnehmerinnen und Teilnehmern des NaWerSys-Workshops viele neuen Erkenntnisse, spannende Diskussionen und interessante Kontakte.

September 2021

Friedemann Kammler, Paul Gembarski, Simon Hagen und Thorsten Schoormann

Literaturverzeichnis

- [AGC19] Aivaliotis, P.; Georgoulas, K.; Chryssolouris, G.: The use of Digital Twin for predictive maintenance in manufacturing. *International Journal of Computer Integrated Manufacturing* 32/11, S. 1067-1080, 2019.
- [ASB16] Almquist, E.; Senior, J.; Bloch, N.: The elements of value. *Harvard Business Review* 94/9, S. 47-53, 2016.
- [De16] Demminger, C.; Mozgova, I.; Quirico, M.; Uhlich, F.; Denkena, B.; Lachmayer, R.; Nyhuis, P.: The concept of technical inheritance in operation: analysis of the information flow in the life cycle of smart products. In: *Procedia Technology*, S. 26, 79-88, 2016.
- [FE19] Felfernig, A.; Polat-Erdeniz, S.; Uran, C.; Reiterer, S.; Atas, M.; Tran, T. N. T.; Azzoni, P.; Kiraly, C.; Dolui, K.: An overview of recommender systems in the internet of things. *Journal of Intelligent Information Systems* 52/2, S. 285-309, 2019.
- [GA20] GAIA-X: A Pitch Towards Europe - Status Report on User Ecosystems and Requirements. Federal Ministry for Economic Affairs and Energy, Berlin, 2020.

- [Ge20] Gembarski P.C.: The Meaning of Solution Space Modelling and Knowledge-Based Product Configurators for Smart Service Systems. In Borzemski L.; Świątek J.; Wilimowska Z. (Hrsg.): Proceedings of 40th Anniversary International Conference on Information Systems Architecture and Technology. Advances in Intelligent Systems and Computing 1050. Springer, Cham, 2020.
- [GL18] Gembarski, P.C.; Lachmayer, R.: Product-Service-Systems – What and why Developers can learn from Mass Customization. Enterprise Modelling and Information Systems Architectures 13, S. 16-1, 2018.
- [HA17] Hagi, A.; Altman, E. J.: Finding the platform in your product. Harvard Business Review 95/4, S. 94-100, 2017.
- [Ka19] Kammler, F.; Hagen, S.; Brinker, J.; Thomas, O.: Leveraging the value of data-driven service systems in manufacturing: a graph-based approach. In: Proceedings of the European Conference on Information Systems. Stockholm-Uppsala, Schweden, 2019.
- [KKG21] Kammler, F.; Kortum, H.; Gembarski, P.C.: Leveraging the Value of Data in the Continuum of Products and Services: Business Types in the Function-Oriented Offerings Model. In: Proceedings of the 10th World Mass Customization and Personalization Conference. Aalborg, Denmark, 2021.
- [LSK20] Lehr, J.; Schlüter, M.; Krüger, J.: Decentralised identification of used exchange parts with a mobile application. International Journal of Sustainable Manufacturing 4/2-4, S. 150-164, 2020.
- [PH15] Porter, M. E. & Heppelmann, E.: How smart, connected products are transforming companies. Harvard Business Review 93/10, 2015.
- [RO20] Romero, M.; Guédria, W.; Panetto, H.; Barafort, B.: Towards a characterisation of smart systems: A systematic literature review. Computers in industry 120, 2020.
- [SK20] Schoormann, T.; Kutzner, K.: Towards Understanding Social Sustainability: An Information Systems Research-Perspective. In: Proceedings of the International Conference on Information Systems. Hyderabad, India, 2020.
- [Sz20] Szopinski, D.; Schoormann, T.; John, T.; Knackstedt, R.; Kundisch, D.: Software tools for business model innovation: current state and future challenges. Electronic Markets 30/3, S. 469-494, 2020.
- [St19] Stark, R.; Fresemann, C.; Lindow, K.: Development and operation of Digital Twins for technical systems and services. CIRP Annals 68/1, S. 129-132, 2019.

Innovation by Information Technology Recombination: How Artificial Intelligence Progressive Web Apps Foster Sustainable Development

Philipp Fukas ¹, Oliver Thomas²


Abstract: Artificial Intelligence (AI) and Progressive Web Apps (PWAs) represent two major trends in today's development of modern information systems. AI aims to automate intelligent behaviour whereas PWAs aim to provide fast, reliable, and engaging applications. The influence of these two key technologies on organisations and sustainable development on their own has already been explored. However, there is no research that merges these technologies in terms of recombinant innovation to show their joint potential. By conducting a systematic literature review this article reveals the positive impact of "Artificial Intelligence Progressive Web Apps" (AI-PWAs) on sustainability. It is shown that AI-PWAs can realise economic, environmental and social benefits and thus can support the achievement of the United Nations' Sustainable Development Goals.

Keywords: Artificial Intelligence, Progressive Web Apps, Recombinant Innovation, Sustainability, Systematic Literature Review.

1 Introduction

One of the most important challenges in modern economy and society are the United Nations' Sustainable Development Goals (SDGs) [Vi20]. To achieve these SDGs, the development of innovative business models with reinforcement by increasing digitization and data availability is necessary. This paper discusses the potential of information technology (IT) innovation for sustainable development by recombining Artificial Intelligence (AI) and Progressive Web App (PWA) technologies, which are two major trends in today's development of modern information systems.

On the one hand, AI represent a relatively mature research area that has already been extensively explored since the 1950s [MP43], [RN20], [Tu50]. Nevertheless, AI has been recently granted a huge potential to transform the present business landscape as its core due to increased computational power and new algorithmic approaches [So20]. The amount of publications, projects and funding regarding AI has risen sharply in past years [Pe2019]. While most AI research until now attempted to obtain state-of-the-art results in

¹ Osnabrück University, Information Management and Information Systems (IMWI), Parkstraße 40, Osnabrück, 49080, philipp.fukas@uni-osnabrueck.de,  <https://orcid.org/0000-0002-2169-5505>

² Osnabrück University, Information Management and Information Systems (IMWI), Parkstraße 40, Osnabrück, 49080, oliver.thomas@uni-osnabrueck.de

accuracy through the use of massive computational power without respecting sustainability goals [Sc20], the relationship between AI technologies and sustainability currently gains increasing attention [Fe21], [GT20]. On the other hand, PWAs represent a relative new concept that combines the advantages of traditional web and native applications to provide a better user experience while limiting the development costs [Mo20]. PWAs have also been granted a huge potential to transform the present business landscape as current companies can engage their customers while bypassing app stores [Cl21]. Due to its novelty, the relationship between PWA technologies and sustainability is in its infancy, but first research was already conducted.

The impact of AI on sustainable development has already been reviewed, analysed and discussed [Di20], [Vi20]. PWA technologies have also been reviewed and analysed [Mo20]. However, there is no research, which combine AI and PWAs to show their joint potential towards sustainable development. Based on the idea of recombinant innovation, the fusion of these known technologies can create a new innovation called Artificial Intelligence Progressive Web App (AI-PWA). Therefore, this paper aims to identify current use cases and implementations, where PWAs and AI are combined and analyses these cases towards their impact on sustainable development. The following research question (RQ) will be answered:

How can Artificial Intelligence Progressive Web Apps foster sustainable development?

To answer this RQ, the paper is structured as follows: In section 2, the foundations of IT innovation, sustainable development, AI and PWAs are elaborated as an introduction into relevant concepts. In section 3, the applied research method of the systematic literature review (SLR) is described. In section 4, the findings of the SLR are explained in detail including a discussion towards their impact on sustainable development. Section 5 concludes the paper and provides an outlook on further research activities.

2 Theoretical Foundations

2.1 The Role of Information Technology Innovation for Sustainable Development

A common view on innovation is to see it as a crucial factor for long-term successful business management [SMA17]. However, innovation can also be seen as a key driver for sustainability, because sustainable development is a pressing issue that requires immediate action from governments, businesses, and society [ST19]. To analyse the influence of innovations on sustainable development, it can be subdivided an economic, an environmental and a social dimension [Vi20]. Moreover, many recent innovations were enabled through IT [Le10]. Innovations through IT are commonly referred to as IT innovations. Research in this field explores organisational and technological factors (e.g. firm size, scope, technological competency and expected benefits) that determine IT adoption and diffusion [MR08]. Maturity models for concrete IT types including AI have

been developed and used to assess these factors [Fu21a], [We12]. To reach a higher organisational maturity towards a certain IT type and therefore discover new value creation opportunities, IT innovations are needed [GLS17]. But innovators rarely come up with completely new ideas and instead often recombine known ideas in order to innovate [La93]. This concept is known as recombinant innovation in literature [BM14], [FIZ12]. Therefore, innovation by recombining known IT components like AI and PWAs seem to be a promising way to foster sustainable development.

2.2 Artificial Intelligence and Progressive Web Apps

AI is an umbrella term for a type of IT, which has the capacity “to act intelligently or to imitate cognitive and intelligent behaviour” [Fu21a]. AI covers every aspect to automate intelligent behaviour and can be subdivided into various subdisciplines [RN20]. To innovate entire organisations with AI, different dimensions must be considered. AI maturity models can assess these dimension to foster the adoption and diffusion of AI within entire organisations [Fu21a]. Besides this management-oriented view, the relation between AI and sustainability gains increasing awareness in the research and business community [Fe21], [GT20]. One common way to assess this relationship is by researching how the use of AI can help to achieve sustainability goals [NKC20], [Vi20]. Another way is to find approaches to design the application of AI-based technologies themselves in a sustainable way [Dh20], [Sc20]. PWAs are responsive web applications that leverage the capabilities of browsers for continuous improvements based on the principle of progressive enhancement [Pa10]. In 2015, Alex Russell coined the term “Progressive Web App” to describe apps taking advantage of new features supported by modern browsers, that let users upgrade traditional web apps to PWAs in their native operating system [Ru15]. PWAs ensure a reliable, motivating, and native user experience through installation, offline functionality and push notifications [TJ18]. Once retrieved online, content can be very quickly displayed from the cache even without an internet connection (offline functionality) [BMG18]. PWAs can reduce the environmental thread by reducing its energy consumption with the help of cache and offline capabilities [Bu19]. With a well-designed service worker unnecessary network requests can be limited and this reduces the energy consumption [Ro20]. PWAs positively contribute to social sustainability as they can provide digital services with native app-like features to all end devices without eluding parts of society only limited to one type of device [We18a].

3 Research Method

To answer the RQ a complete overview over current combinations of AI and PWA technologies as well as an analysis towards their technical components and sustainability impacts is necessary. To identify and analyse all available research relevant to a particular research question or topic area the method of a SLR can be used [KC07]. In general, the SLR follows an multi-step process [Br09]. In the beginning, the review scope was defined

by using the six characteristics of the taxonomy of Cooper [Co88]. After defining the review scope, the first fundamental literature was identified and the topics were conceptualised via the method of concept mapping to derive relevant key terms for the search strings (cf. Fig. 1) [RS04], [SNB82]. The search was intentionally not restricted with any terms related to sustainability in order to identify as many publications as possible and to subsequently analyse all of them regarding their motivations and benefits achieved by the fusion of AI and PWAs. From these motivations and benefits the ones that have an impact on sustainability were identified to answer the RQ. The search strings were subsequently tested in the identified databases to guarantee their functionality.

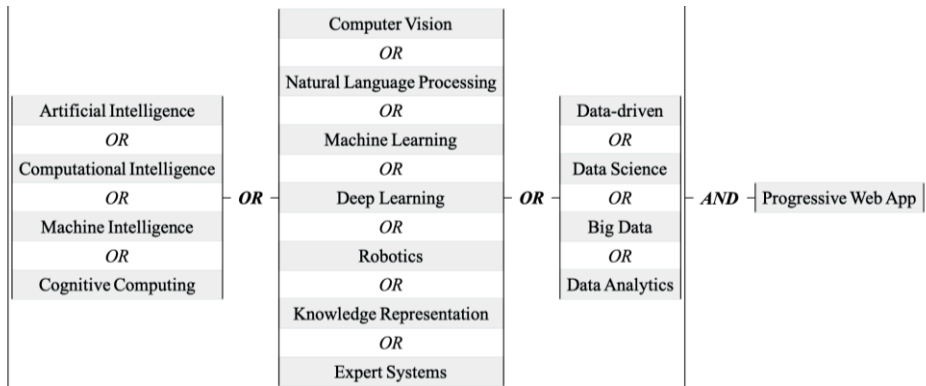


Fig. 1: Search terms for the systematic literature review

In total, the literature was searched via ten different scholarly databases providing access to leading journals. Since the topic of combining AI with PWA is relatively new and immature, we also included patents, whitepaper and technical reports and thesis instead of relying only on scientifically accepted journals. This ensured that scientific as well as practical top-tier sources from various domains were considered. On the contrary, this led to a number of 603 search results including irrelevant articles (cf. Fig. 2). After eliminating all duplicates and results that were not primarily written in English or German 251 distinct articles remain. To exclude all irrelevant articles, the results were scanned based on their titles [Dy08]. In line with the inclusion and exclusion criteria, many articles, whose titles were not referring to AI or PWA concepts could be sorted out. In the next step, abstracts and complete texts of the remaining articles were intensively analysed. The final backward and a forward search based on the resulting articles resulted in no additional relevant article [WW02]. All articles were analysed and synthesized using two concept matrices [SNB82], [WW02]. For transparency, supplementary material including all 62 sources and the complete concept matrices is provided in Fukas and Thomas (2021) [Fu21b]. Finally, a definition of AI-PWAs including their technical components (cf. section 4.1) and their implications towards sustainable development (cf. section 4.2) were elaborated.

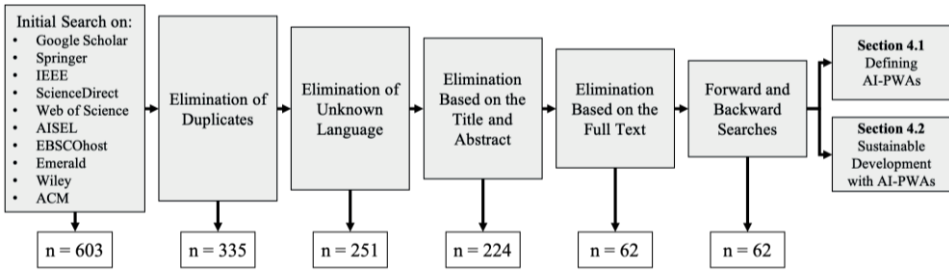


Fig. 2: Search, elimination and analysis steps of the systematic literature review

4 Results

4.1 Defining Artificial Intelligence Progressive Web Apps

To fully understand and define the innovation of AI-PWAs, 62 publications were identified and analysed towards their combination of AI and PWA technologies (cf. Fig. 3). The majority of these publications are journal articles and conference proceedings (39 out of 62 articles). A significant amount of theses and dissertations were also included (14 out of 62 articles). The minority of articles are book chapters, patents and technical reports/whitepapers (9 out of 62 articles). An increasing trend from 2017 until today can be observed (excluding the number of articles from 2021 for the trend analysis, since the SLR was conducted in the first quarter of 2021). Thus, it can be assumed that AI-PWAs already represent an emerging innovation.

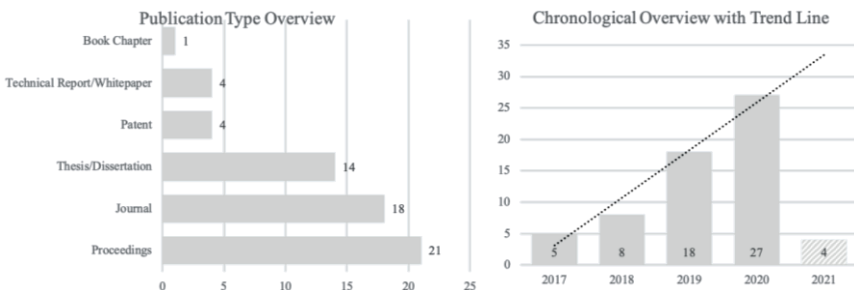


Fig. 3: Publication type and chronological overview with trend line of retrieved articles

As already mentioned earlier, AI can be divided into different subcomponents (cf. section 2.2). In the identified literature, the authors combined one or more of these subcomponents with PWA functionalities in order to provide an innovative solution to their respective research problem. Tab. 1 summarises these AI components including the number of sources mentioning the corresponding AI component (No.) and a selection of reference publications:

AI Component	Short Description	No.	Reference Publications
Computer Vision (CV)	CV describes the automatic processing and recognition of visual inputs.	15	[Bi19], [Hu19], [Wi19]
Deduction, Reasoning, Problem Solving (DRP)	DRP imitates step-by-step reasoning that humans use when solving puzzles or drawing logical conclusions.	6	[EKK19], [Gr20], [Kh17], [Sa19a], [Sa19b]
Knowledge Representation (KR)	KR model intelligence in form of knowledge-based systems.	13	[EKK19], [Gr20], [HZ18], [Kh17]
Machine Learning (ML)	ML describes the automatic generation of knowledge from experience.	36	[Al17], [Ch20], [Me19], [VGG20]
Natural Language Processing (NLP)	NLP describes the processing of natural language by a machine.	14	[HB19], [Ka20], [Ly18], [Pa17]
Planning (P)	Planning describes the automatic decision making in achieving goals.	11	[Ea19], [La18], [VGG20]
Robotics (R)	R creates systems for a sovereign interaction with natural environment.	2	[Da21], [Ma18]
Social Intelligence (SI)	SI aims to recognise, interpret, process or simulate human affects.	3	[Ai18], [Ga19], [Sp19]

Tab. 1: Overview of AI components in AI-PWAs

Further, the SLR showed successful examples how AI-PWAs can be implemented. The development follows mostly two recombination strategies. In the first strategy, AI components are developed independently of the PWA. Then, the PWA was developed aiming on a cost efficient, intuitive and cross-platform deployment as well as a central user interface (UI), where users can interact with AI components [HZ18], [Pa17]. In the second strategy, first the PWA is developed consisting of non-AI functionality to engage users. Then, the existing PWA was enhanced by AI components aiming on better UI, efficiency gains or higher service quality [Kh17], [La18]. Finally, the following definition and technical overview of AI-PWAs (cf. Fig. 4) was derived from the findings of the SLR:

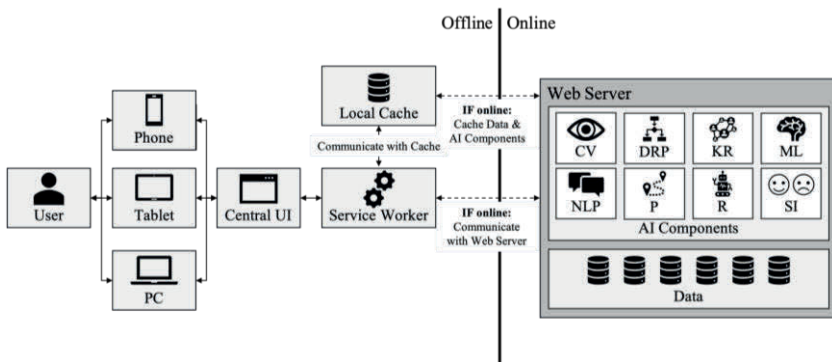


Fig. 4: Basic structure of Artificial Intelligence Progressive Web Apps

Definition: *Artificial Intelligence Progressive Web Apps are information systems that combine one or more Artificial Intelligence components, e.g. computer vision, natural language processing or machine learning, with one or more Progressive Web App functionalities, e.g. offline ability, push notifications or native app experience.*

4.2 Sustainable Development with Artificial Intelligence Progressive Web Apps

Based on our definition, the impact of AI-PWAs on sustainable development were examined. A common way to analyse the influence of innovations on sustainable development is by structuring it into an economic, an environmental and a social dimension [Vi20]. Therefore, the motivations and benefits of AI-PWAs were inductively analysed in the identified literature of the SLR and then related to the three sustainability dimensions. We found, that the sustainability impacts of AI-PWAs can be described by six characteristics, which are further explained in Tab. 2:

Sustainability Impact	Short Description	No.	Reference Publications
Process Efficiency	Process Efficiency describes the productivity gain of AI-PWA users compared to users of other IT systems.	11	[EKK19], [PSM20a], [PSM20b]
Cost Efficiency	Cost Efficiency describes the lower financial resources that have to be invested to develop and operate AI-PWAs compared to other IT systems.	2	[Bi19], [Me19]
Consumption Optimisation	Consumption Optimisation describes the optimisation of resource consumption achieved through the use of AI-PWAs.	3	[Al17], [Ch20], [VGG20]
Usage Optimisation	Usage Optimisation describes the lower resource usage of AI-PWAs themselves compared to other IT systems.	9	[Ea19], [HZ18], [La18], [Ly18]
Enabled Society	Enabled Society describes the assistance provided to users in achieving their desired goals, which would not be possible without the use of AI-PWAs.	20	[Gr20], [Hu19], [PSM20a], [PSM20b], [Wi19]
Engaged Society	Engaged Society describes the ease of involving a large number of AI-PWA users compared to the more difficult user engagement in other IT systems.	17	[Ai18], [Ga19], [Kh17], [Sa19a], [Sa19b], [Sp19]

Tab. 2: Overview of sustainability impacts of AI-PWAs

These six characteristics directly relate to sustainable development and the economic, environmental and social dimensions. The sustainability impacts are illustrated in Fig. 5 and explained in detail including examples from the SLR in the following:

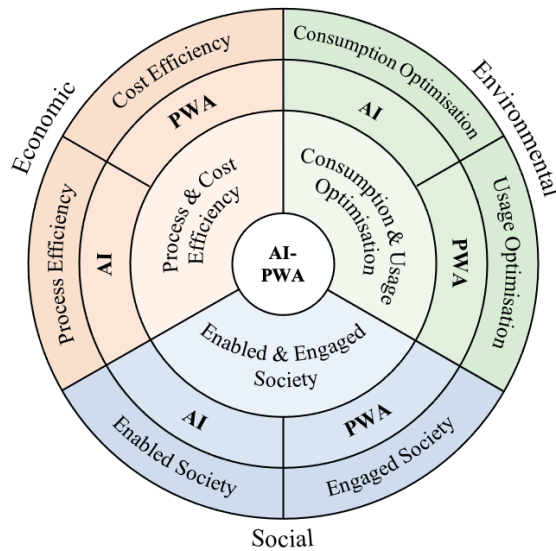


Fig. 5: Sustainability impacts of Artificial Intelligence Progressive Web Apps

Economic sustainability is supported by an AI-PWA mainly on two ways. On the one hand, most articles of the SLR considered the use of AI to reach a higher *Process Efficiency*. AI is linked to a net positive impact associated to increased productivity [Vi20]. For example, incorporation of CV to automatically recognize image features and thus automate manual, error-prone actions directly can improve process efficiency [PSM20a], [PSM20b]. On the other hand, the deployment as a PWA was often chosen due to avoiding the cost of developing multiple native applications by still ensuring the same usability [Bi19], [Me19]. By using PWA technologies, only one single web-based application needs to be developed, which can be used on every device with native-like features instead of developing an own, separate native applications for every operating system. In some instances, this can drastically reduce the investment by gaining the same value proposition and result in a better *Cost Efficiency*. The development of an AI-PWA can combine these two advantages to enhance economic sustainability.

Environmental sustainability of AI-PWAs can be considered in two ways as well. The AI components of AI-PWAs can be used to optimise the consumption of resources (*Consumption Optimisation*), e.g. electricity, water or food [Vi20]. It can also underpin low-carbon systems, for instance, by supporting the creation of circular economies and smart cities that efficiently use their resources [Ne19]. For example, concrete AI-PWAs were already developed for optimising waste management or water usage [Ch20], [VGG20]. The PWA components of AI-PWAs can be used to optimise the usage of the app itself (*Usage Optimisation*). Through PWA technologies, a better usability and optimised program code can ensure a more energy efficient use of the app itself [Ea19]. PWAs allow caching and offline functionalities, which reduce the unnecessary data

transfer [Bu19]. With a well-designed service worker unnecessary network requests can be limited and this reduces the energy consumption [Ro20]. AI-PWAs therefore can realise environmental advantages regarding their operation as well as their implementation purpose.

Social sustainability can be achieved by enabling and engaging the society. AI components can be used to enhance humans by providing automated assistance, reducing error-prone, manual tasks and teaching new skills (*Enabled Society*) [Hu19], [PSM20a], [PSM20b], [Wi19]. Moreover, AI components can help by supporting the provision of food, health, water, and energy services to the population by autonomously predicting the corresponding needs [Vi20]. PWA components are able to engage humans by a better usability, cross-platform capability and lower entrance barriers (*Engaged Society*) [Kh17], [Sa19a], [Sa19b]. They positively contribute to social sustainability as they can provide digital services with native app-like features to all end devices without eluding parts of society that are only limited to one type of device. Moreover, the attempts to provide AI-based social intelligence through a PWA UI [Ai18], [Ga19], [Sp19] or to engage social user behaviour through AI-based chatbots directly count towards the social dimension of sustainability [HB19], [Pa17]. Therefore, AI-PWAs can induce an enabled and engaged society.

5 Conclusion

The SLR showed successful examples how AI-PWAs can be developed and used. Different AI subdisciplines like CV, DRP, KR, ML, NLP, P, R and SI were deployed using PWA functionalities to achieve synergies and combine the benefits of both technologies resulting in a new type of IT innovation, namely AI-PWAs. AI-PWAs impact sustainable development in a positive manner. Therefore, the answer to the RQ is the following: Regarding the economic, environmental and social dimensions of sustainability, AI-PWAs can realise process and cost efficiency, consumption and usage optimisation as well as an enabled and engaged society. Moreover, negative impacts on sustainability by AI can be diminished by combining it with PWA functionalities. The social drawback of AI being unevenly distributed and not everyone being able to access AI functionalities since it requires specialised end devices is reduced by the cross-platform capabilities of PWAs [We18b]. The cost-intensive deployment of AI functionalities in native apps negatively impacting economic sustainability can be reduced by implementing AI-PWAs, as only one application needs to be developed for all end devices. Finally, the energy intensive use of AI components can be reduced by caching and offline features as well as better usability reducing negative environmental impacts of AI.

This paper directly contributes to the computer science and information systems body of knowledge by introducing the concept of AI-PWAs based on concrete evidence in literature. Moreover, novel insights to foster sustainable development by recombinant innovation are provided that not have been explored before. Our findings support the

theory of recombinant innovation as AI-PWAs represent a perfect example how innovations are achieved by recombining known ideas. In addition to the theoretical insights, practitioners can profit of the findings by gaining knowledge about a new technique to implement innovative business cases more effectively in a sustainable way. Moreover, by systematic structuring the relating literature, practitioners can learn from the experiences others made in implementing AI-PWAs. In relation to AI maturity models, AI-PWAs have the potential to increase the AI maturity of an organisation and to further diffuse AI technologies within it.

The SLR revealed further research opportunities regarding AI-PWAs. Only a few attempts concretely show how PWA caching mechanisms can be used to cache AI components locally and thus ensure a full offline capability of AI-PWAs. By intelligently caching actual AI models rather than just caching responses or data, AI functionality could be transferred from a computationally intensive server to a computationally light end device. This concept is known as edge AI. Edge AI does not yet cover PWA technologies although AI-PWAs show some promising advantages in terms of sustainability. In the future, caching of AI-functionalities through PWA technologies should be explored as this could contribute crucial insights to the field of edge AI. Moreover, this paper represents one of the first research items to elaborate the sustainability of AI-PWAs. In the future, sustainability impacts could be further explored by implementing concrete technical analyses and prototypical sustainability measurements in AI-PWAs. As this paper has presented promising AI-PWA approaches that can benefit various businesses to foster sustainable development, the transfer from AI-PWAs to concrete business model innovations is also worth exploring.

Bibliography

- [Ai18] Aikaterini, T.: Affective system monitoring personal expenses, helping the user to stay on budget. International Hellenic University, 2018.
- [Al17] Alurkar, A. A. et al.: A Proposed Data Science Approach for Email Spam Classification using Machine Learning Techniques. *Internet of Things Business Models, Users, and Networks*, pp. 1–5, 2017.
- [Bi19] Bilsten, J.: Generating a website from digital image metadata. U.S. Patent Pub. No. US20190065613A1, 2019.
- [BM14] Brynjolfsson, E.; McAfee, A.: *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. 2014.
- [BMG18] Biørn-Hansen, A.; Majchrzak, T. A.; Grønli, T. M.: Progressive Web Apps for the Unified Development of Mobile Applications. In (Majchrzak, T. A. et al., eds.): *Lecture Notes in Business Information Processing Volume 322*, pp. 64–86.
- [Br09] vom Brocke, J.: Reconstructing the Giant : On the Importance of Rigour in Documenting the Literature Search Process. In: *Proceedings of the European Conference on Information Systems, Verona, Italy*, pp. 2206–2217, 2009.

- [Bu19] Butterworth, C.: How creating a Progressive Web App has made our website better for people and planet, 2019. <https://www.wholegraindigital.com/blog/progressive-web-app-better-for-people-and-planet/>, accessed: 20/06/2021.
- [Ch20] Chandrashekar Murthy, B. N. et al.: Prediction of Water Demand for Domestic Purpose Using Multiple Linear Regression. In (Smys, S. et al., eds.): *New Trends in Computational Vision and Bio-inspired Computing: Selected works presented at the ICCVBIC 2018*, pp. 811–817, 2020.
- [Cl21] Cloud Four Inc.: PWA Stats, 2021. <https://www.pwastats.com/>, accessed: 20/06/2021.
- [Co88] Cooper, H. M.: Organizing knowledge syntheses: A taxonomy of literature reviews. *Knowledge in Society* 01/01, pp. 104–126, 1988.
- [Da21] Dandabathula, G. et al.: Design and Development of Aquayaan: An IoT based Robotic Boat for Inland Water Surveys, 2021.
- [Dh20] Dhar, P.: The carbon impact of artificial intelligence. *Nature Machine Intelligence* 02/08, pp. 423–425, 2020.
- [Di20] Di Vaio, A. et al.: Artificial intelligence and business models in the sustainable development goals perspective: A systematic literature review. *Journal of Business Research* 121, pp. 283–314, 2020.
- [Dy08] Dybå, T.; Dingsøy, T.: Empirical studies of agile software development: A systematic review. *Information and Software Technology* 50/09, pp. 833–859, 2008.
- [Ea19] East, A. R.: Timetable Scheduling via Genetic Algorithm. National University of Ireland, 2019.
- [EKK19] Espenakk, E.; Knalstad, M. J.; Kofod-Petersen, A.: Lazy Learned Screening for Efficient Recruitment. In (Bach, K., ed.): *Case-Based Reasoning Research and Development*, pp. 64–78, 2019.
- [Fe21] Fernandez-Aller, C. et al.: An Inclusive and Sustainable Artificial Intelligence Strategy for Europe Based on Human Rights. *IEEE Technology and Society Magazine* 40/01, pp. 46–54, 2021.
- [FIZ12] Frenken, K.; Izquierdo, L. R.; Zeppini, P.: Branching innovation, recombinant innovation, and endogenous technological transitions. *Environmental Innovation and Societal Transitions* 4, pp. 25–35, 2012.
- [Fu21a] Fukas, P. et al.: Developing an Artificial Intelligence Maturity Model for Auditing. In: *29th European Conference on Information Systems Research Papers*, 2021.
- [Fu21b] Fukas, P.; Thomas, O.: Supplementary Material for the Publication “Innovation by Information Technology Recombination: How Artificial Intelligence Progressive Web Apps Foster Sustainable Development”. <https://doi.org/10.5281/zenodo.4994772>, accessed: 20/06/2021.
- [Ga19] Garg, S.: Group Emotion Recognition. University of Manchester, 2019.
- [GLS17] Griffith, R.; Lee, S.; Straathof, B.: Recombinant innovation and the boundaries of the firm. *International Journal of Industrial Organization* 50, pp. 34–56, 2017.
- [Gr20] Gray, F. M. et al.: Occupant Feedback and Context Awareness: On the Application of

- Building Information Modeling and Semantic Technologies for Improved Complaint Management in Commercial Buildings. In: 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), pp. 101–108, 2020.
- [GT20] Goralski, M. A.; Tan, T. K.: Artificial intelligence and sustainable development. *International Journal of Management Education* 18/01, 2020.
- [HB19] Hobert, S.; Berens, F.: Small Talk Conversations and the Long-Term Use of Chatbots in Educational Settings—Experiences from a Field Study. In (Følstad, A. et al., eds.): *Chatbot Research and Design*, pp. 260–272, 2019.
- [Hu19] Huynh, S. T.: *Developing App Features for Figure Drawing Practice*. Worcester Polytechnic Institute, 2019.
- [HZ18] Hu, K.; Zhu, J.: A Progressive Web Application on Ancient Roman Empire Coins and Relevant Historical Figures with Graph Database. In (Ioannides, M. et al, eds.): *Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection*, pp. 235–241, 2018.
- [Ka20] Kavanagh, C. S.: *Künstliche Intelligenz in der Kultur: Design und Entwicklung eines Chatbots für kulturfinder.sh..* Kiel University of Applied Sciences, 2020.
- [KC07] Kitchenham, B. A.; Charters, S. M.: *Guidelines for performing Systematic Literature Reviews in Software Engineering*, 2007.
- [Kh17] Khan, Z. et al.: Developing Knowledge-Based Citizen Participation Platform to Support Smart City Decision Making: The Smarticipate Case Study. *Information* 08/02, no. 47, 2017.
- [La18] Lakatosova, P.: *Taxi system front-end*. Charles University, 2018.
- [La93] Larsen, T. J.: Middle managers' contribution to implemented information technology innovation. In: *Journal of Management Information Systems* 10/02, pp. 155–176, 1983.
- [Le10] Lemke, C.: Rethinking How Students Learn. In (Bellanca, J.; Brandt, R., eds.): *21st century skills*. pp. 243–274, 2010.
- [Ly18] Lyndon-James, P.: *Networked System and Method for Increasing Brand Awareness and Data Collection Using Gamification*. U.S. Patent Pub. No. US20180060901A1, 2018.
- [Ma18] Mahi, S. M. Al: Multi-Modal Multi sensor Interaction between Human and Heterogeneous Multi-Robot System. In: *Proceedings of the 20th ACM International Conference on Multimodal Interaction*, pp. 524–528, 2018.
- [Me19] Mendes, F. M. et al.: EJ: A Free Software Platform for Social Participation. In (Bordeleau, F. et al., eds.): *Open Source Systems*, pp. 27–37, 2019.
- [Mo20] Mole, P. V.: *Progressive Web Apps: A Novel Way for Cross-Platform Development*, 2020.
- [MP43] McCulloch, W. S.; Pitts, W.: A Logical Calculus of Ideas Immanent in Nervous Activity. *Bulletin of Mathematical Biophysics* 5, pp. 115–133, 1943.
- [MR08] Melville, N.; Ramirez, R.: Information technology innovation diffusion: An information requirements paradigm. *Information Systems Journal* 18/03, pp. 247–273, 2008.

- [Ne19] Nerini, F. F. et al.: A research and innovation agenda for zero-emission European cities. *Sustainability* 11/06, pp. 1–13, 2019.
- [NKC20] Nishant, R.; Kennedy, M.; Corbett, J.: Artificial intelligence for sustainability: Challenges, opportunities, and a research agenda. *International Journal of Information Management* 53, no. 102104, 2020.
- [Pa10] Parker, T. et al.: *Designing with Progressive Enhancement*. 2010.
- [Pa17] Pathak, N.: Designing a Multilingual Virtual Agent Capable of Interacting with Uneducated People for Automated Data Collection. In: *IEEE Symposium Series on Computational Intelligence (SSCI)*, pp. 1–7, 2017.
- [Pe2019] Perrault, R. et al.: *The AI Index 2019 Annual Report*. Stanford University, 2019.
- [PSM20a] Prasad, R. B.; Singh, A.; Marathe, J.: Ionic4 and Pay-Pal Merging for Apps and Progressive Web Apps. *International Research Journal of Modernization in Engineering Technology and Science (IRJMETS)* 02/09, pp. 243–246, 2020.
- [PSM20b] Prasad, R. B.; Singh, A.; Marathe, J.: Ionic 4 Paypal Payment Integration for Apps and Progressive Web Apps. *International Research Journal of Engineering and Technology (IRJET)* 07/09, pp. 3234–3235, 2020.
- [RN20] Russell, S. J.; Norvig, P.: *Artificial Intelligence: A Modern Approach*. 4th ed., 2020.
- [Ro20] Rousset, D.: Green Energy Efficient Progressive Web Apps, 2020. <https://devblogs.microsoft.com/sustainable-software/green-energy-efficient-progressive-web-apps/>, accessed: 20/06/2020.
- [RS04] Rowley, J.; Slack, F.: Conducting a literature review. *Management Research News* 27/06, pp. 31–39, 2004.
- [Ru15] Russell, A.: Progressive Web Apps: Escaping Tabs Without Losing Our Soul, 2015. <https://infrequently.org/2015/06/progressive-apps-escaping-tabs-without-losing-our-soul/>, accessed: 20/06/2021.
- [Sa19a] Santos-Gago, J. M. et al.: SportsWoman: A semantic-based recommender platform for women who practice sports. *Journal on Advances in Theoretical and Applied Informatics* 05, no. 1, 2019.
- [Sa19b] Santos-Gago, J. M. et al.: Towards a Personalised Recommender Platform for Sportswomen. In (Rocha, Á. et al., eds.): *New Knowledge in Information Systems and Technologies*, pp. 504–514, 2019.
- [Sc20] Schwartz, R. et al.: Green AI. *Communications of the ACM* 63/12, pp. 54–63, 2020.
- [SMA17] Savino, T.; Messeni Petruzzelli, A.; Albino, V.: Search and Recombination Process to Innovate: A Review of the Empirical Evidence and a Research Agenda. *International Journal of Management Reviews* 19/01, pp. 54–75, 2017.
- [SNB82] Salipante, P.; Notz, W.; Bigelow, J.: A Matrix Approach to Literature Reviews. *Research in Organizational Behavior* 04, pp. 321–348, 1982.
- [So20] Soni, N. et al.: Artificial Intelligence in Business: From Research and Innovation to Market Deployment. *Procedia Computer Science* 167, pp. 2200–2210, 2020.

- [Sp19] von Sperling, O. K.: UnB Sense: a web application to probe for signs of depression from user profiles on social media. University of Brasília, 2019.
- [ST19] Silvestre, B. S.; Țircă, D. M.: Innovations for sustainable development: Moving toward a sustainable future. *Journal of Cleaner Production* 208, pp. 325–332, 2019.
- [TJ18] Tandel, S.; Jamadar, A.: Impact of Progressive Web Apps on Web App. *International Journal of Innovative Research in Science, Engineering and Technology* 07/09, pp. 9439–9444, 2018.
- [Tu50] Turing, A. M.: Computing Machinery and Intelligence. *Mind* 49, pp. 433–460, 1950.
- [VGG20] Vohra, D.; Gunjan; Gandhi, D. A.: Incentivized Techbin: a Step Towards Sustainable Environment. *International Journal of Engineering Applied Sciences and Technology (IJEAST)* 05/02, pp. 141–148, 2020.
- [Vi20] Vinuesa, R. et al.: The role of artificial intelligence in achieving the Sustainable Development Goals. *Nature Communications* 11/01, pp. 1–10, 2020.
- [We12] Wender, R.: The maturity of maturity model research: A systematic mapping study. *Information and Software Technology* 54/12, pp. 1317–1339, 2012.
- [We18a] Weller, J.: The Business Perspective — Progressive Web Apps, 2018. <https://medium.com/@jedihacks/progressive-web-apps-the-business-perspective-9ea411ddee90#:~:text=A Progressive Web Application is,while bypassing the app store,accessed: 20/06/2021>.
- [We18b] Wegren, S. K.: The “left behind”: Smallholders in contemporary Russian agriculture. *Journal of Agrarian Change* 18/04, pp. 913–925, 2018.
- [Wi19] Williford, B. et al.: DrawMyPhoto: Assisting Novices in Drawing from Photographs. In: *Proceedings of the 2019 on Creativity and Cognition*, pp. 198–209, 2019.
- [WW02] Webster, J.; Watson, R. T.: Analyzing the Past to Prepare for the Future: Writing a Literature Review. *Management Information Systems Quarterly* 26/02, pp. xiii–xxiii, 2002.

Game-based Learning: Acquiring Supply Chain Management Skills Through Playing Computer Games

Michael Ganske,¹ Katherina Meißner,¹ Julia Rieck¹

Abstract: Computer games can support university teaching by promoting strategic, planning, and networked thinking. Especially in supply chain management, many processes require coordination, effects of actions need to be predicted, and decisions are made under time pressure. A structured and analytical approach is therefore of great importance, which is particularly promoted by playing computer games. We demonstrate the use of Rise of Industry to deepen students' understanding of supply chain management, covering logistics and production control as well as interactions and interdependencies between the components of supply chains. Using Constructive Alignment, we define intended learning outcomes and activities of a computer game-based learning course and derive requirements for selecting and adapting a computer game. Evaluating a three-day seminar, we find that students were able to apply prior knowledge to the game world and improve their decision making while experiencing the immediate economic and ecological impacts of their actions through the game.

Keywords: Game-based Learning; Computer Games; Supply Chain Management; Constructive Alignment; Design Science Research; Sustainability

1 Problem Identification and Motivation

Sustainable supply chain management (SCM) is a diverse subject area with many interconnections. Thus, decision-making is complex and challenging. In particular, interdependencies such as the trade-off between economic success and ecological sustainability are difficult to grasp at first glance. Yet it is precisely this interplay that is of great importance in today's society [TGRW15]. Students learn this, e. g., through lectures and discussions, but also in seminars, where *business simulation games* are used.

Game-based learning (GBL) as an educational approach has become very popular in recent years. Using games, students are encouraged to think critically and solve problems themselves [HKS14]. Furthermore, games have a positive impact on developing leadership skills [SR19] and enable a deep understanding of relationships and interactions in corporate processes [PGPCC16]. Many SCM-oriented business games focus on the management of a fictitious company. Students play in groups, taking on the role of different actors in a company and make decisions. These decisions are then entered into a system that simulates their effects for a given time interval. This process is repeated until the game ends after a predefined time [MAA17], e. g., three years. However, how the decisions ultimately affect

¹ University of Hildesheim, Institute for Business Administration and Information Systems, Universitätsplatz 1, 31141 Hildesheim, Germany, {ganske,meissner,rieck}@uni-hildesheim.de

the environment during the simulation is not apparent. Therefore, students cannot reflect on their decisions and consequences in detail, but only on a strategical level.

Since business simulations often lack visual elements, i. e., only offer text and tables, complex subjects within SCM, e. g., environmental impacts of production and logistics, are difficult to represent. Complementary, more visual, and interactive approaches are considered promising to make interrelationships and sustainable actions in supply chains tangible and thus easier to grasp. *Computer games* are designed to offer a fun, visual, and interactive experience. This leads to players remaining engaged and immersed in the game world. The increased intrinsic motivation can be used to improve the learning transfer of complex topics, as the players learn through the game without being aware of it [MDM08]. As the effects of the player’s actions are directly perceptible within the game world, computer games enable experiential learning [Zu19]. This real-time feedback sets computer games apart from business simulations and makes it easier to see and react to decisions’ effects. In this way, dependencies and interactions are better learned and understood [Li17, So19], which is especially relevant in teaching SCM skills.

This paper presents the design and implementation of a GBL course using computer games, as opposed to digital or analog business simulations currently found in educational practice, e. g., MIT beer game (<https://beergame.org/>). To promote sustainability-oriented behavior in SCM decisions, we look at computer games that were “not intentionally developed for educational purposes, but which nonetheless have a positive effect on students’ learning or development” [CC06]. In order to adequately address the educational need for implementation, the learning objectives as well as the teaching and learning activities of the course are formulated using Biggs’ Constructive Alignment (CA) [Bi96]. Since the new course has to be integrated into the existing curriculum, design and development processes are necessary, which we realize in a structured way within the framework of Design Science Research (DSR) [Pe07]. In Sec. 2, we define the educational objectives for the course before we go into the details of the design and development in Sec. 3. Section 4 contains the demonstration of a game-based SCM course in a university setting. The evaluation is presented in Sec. 5 before concluding the paper with a discussion and an outlook on future research in Sec. 6. The graphical abstract in Fig. 1 presents the phases of the development process of a course based on CA and DSR.

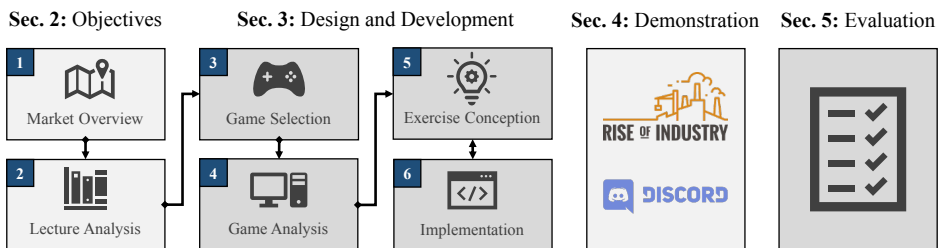


Fig. 1: Graphical Abstract

2 Definition of Educational Objectives

In order to define the objectives for the design of the university course from a competence-based perspective, we apply the theory of Constructive Alignment. The objectives, or rather the targeted learning outcomes, are used to guide the development of teaching and learning activities as well as the assessment of the course, but also to derive specific requirements for the computer game (cf. Sec 3.1) to be implemented in the course. The Structure of Observed Learning Outcomes (SOLO) taxonomy [Bi96] divides students' cognitive processes into five levels (SOLO 1, . . . , SOLO 5) with an increasing order of complexity. These range from missing the point of an exercise entirely to being able to generalize a structure beyond what is given. We structure our Intended Learning Outcomes (ILO) according to these levels (cf. [Br08]), assuming that the students can already make obvious connections and look at individual aspects (SOLO 1 and 2 can be omitted).

After the course, students are expected to be able to . . .

- (a) *create* and maintain an economic and sustainable supply chain for a given product by *applying* the theories from former lectures (SOLO 4);
- (b) *act* as supply chain managers under timely pressure, *relate* external events to their supply chain, and *apply* sensible solutions for them (SOLO 5);
- (c) *formulate* aims and strategies for a company or supply chain department (SOLO 5);
- (d) *discuss* their strategy and actions with a team partner (SOLO 3);
- (e) *present*, evaluate, and defend their business decisions in front of the class (SOLO 3);
- (f) *reflect* their strategy, decisions, and actions (SOLO 5).

In the first phase of the course development process (cf. Fig. 1), we prepared a *market overview* of business simulation games focusing on supply chains. From the overview, we captured the typical implementation of a business simulation game, which is reflected in the teaching and learning activities (TLAs) listed below. The TLAs, incorporating adaptations for computer games, are designed to promote achievement of the ILOs.

- (1) **Founding of a virtual company** including the *formulation* of aims and strategies to engage students in the computer game and activate their creativity.
- (2) **Playing the computer game** in teams of two to enhance *peer*-activities (learning partnerships) in multiple (short) exercises with given focus on certain aspects of SCM. Here, instructors are responsible for game-related and technical assistance.
- (3) **In-class discussions** of decisions to consolidate learning outcomes and to provide space for presentation and discussion skills.
- (4) **Writing a group report** based on activating questions in a (longer) reflection phase to assess the learning outcome.

In a *lecture analysis* (phase 2, Fig. 1), we examined the content of students' previous lectures (e.g., business administration) to identify prior knowledge that should be referenced and extended in the course. This analysis also identified appropriate content for the exercises in TLA (2). Structurally, the course iterates through TLAs (1) and (2) multiple times. This gives students space to repeat the process of *creating* and *maintaining* a supply chain at various occasions and allows for insights and feedback from previous assignments to be considered. The course grade is then *assessed* based on performance in the presentations TLA (3) and the group report TLA (4). Gaming skills were not graded. Since the learning activities reflect the time spent in the game, they inherently align with all specified ILOs.

3 Design and Development

In this section, we aim to provide a preliminary framework for how a computer game can be purposefully integrated into a university learning environment. Therefore, the phases 3 to 6, numbered in boxes in Fig. 1, are described.

3.1 Game Selection and Analysis

The intended learning outcomes from Sec. 2 are used as a basis for the computer *game selection* (phase 3, Fig. 1). We deduct the following requirements from ILOs (a) and (b) for the selection process. A suitable game must represent an economy, include sustainability aspects, provide for the establishment and monitoring of supply chains, allow for the production and sale of various products, and provide the necessary data to enable the application of theories from the lectures (e.g., amortization calculation). In order to support the requirements derived from ILO (a), a game must offer a sufficiently interactive visualization of the game world and thus a systemic perspective of the supply chains (i.e., coordination of procurement, production, distribution, and retail). Furthermore, it should be possible to depict various events that influence the supply chain (ILO (b)), so that the players can deal with them, i.e., demand fluctuation. In order for students to achieve the competencies from ILOs (a) and (b), the game should depict a continuous, i.e., close-to-reality, time progression. Therefore, fast-forwarding must be prevented. External timing of the game, i.e., definition of exercise start and end, would also facilitate instructors' management of the course.

For the selection of a suitable game, we suggest an approach similar to a literature search. In a first stage, computer game repositories, e.g., Steam (<https://store.steampowered.com>), were browsed with appropriate genres (e.g., construction and management simulations with a strategy focus) and keywords (e.g., production, logistics, management and business simulation). Exclusion criteria were specified in a second stage (e.g., the presence of competitors and an economy system), to decide whether a game is suitable for the intended purpose. Unsuitable games were sorted out by reading the games' description and by reviewing screenshots, trailers, and gameplay videos. In a third stage, the final game was selected through play-testing.

During testing, particular attention was paid to a balanced coverage of all areas of the supply chain within the games and that players could interact with all processes in the game in order to encourage strategic, analytical, and interconnected thinking. Due to its variety of available products, the realistic representation of supply chains, and the possibility to create customized scenarios, we identified “Rise of Industry” (RoI) (<https://riseofindustry.com>) as the most suitable game for the course. In particular, the need to make strategic decisions, the ability to build a complex manufacturing environment in which production times can be optimized, and the presence of realistic demands made RoI stand out from other games (i. e., especially the Anno Series, Cities: Skylines, Factorio, Industry Giant 2, and Transport Fever). In RoI, players take on the role of an entrepreneur within the manufacturing industry being in full control over the supply chain and all inherent manufacturing steps, from the procurement of raw materials to the sale of finished goods.

Once a computer game is selected, it must be thoroughly tested in a *game analysis* (phase 4, Fig. 1). In order to best align the course to the game, we identified how the achievement of the ILOs can be supported through RoI. This included taking a closer look at the player’s perspective on the supply chain, and the way that real-time feedback on participants’ decisions and actions manifested in the game world. To better guide reflection and application of knowledge in the course, we reviewed, among other things, how the conflict of interest between economic success and environmental sustainability is addressed and visualized in RoI, e. g., a factory’s pollution impact on its immediate environment.

In the game, environmental impacts become evident through visual indicators. In the case of increased air or water pollution, the landscape turns gray, trees wither, and land for farming and livestock becomes unusable. If pollution is not dealt with, buildings are abandoned and the settlement’s population decreases, which in turn negatively affects the demand for products. By researching, building, and continuously maintaining filtration systems, these effects can be contained by the player.

3.2 Exercise Conception, Scenario Creation, and Game Modification

Playing computer games leads to learning. However, “the question remains what is learned, why is it learned, and which parts of this knowledge are transferred from the game to the real world” [MDM08]. In order to avoid solely relying on the knowledge imparted by simply playing the game, TLA (2) can be implemented through *creating exercises* (Fig. 1, phase 5) that allow addressing individual learning outcomes in the game world, in particular ILOs (a) and (b). Splitting up content over multiple exercises can also reduce overstimulation and confusion resulting from playing an unfamiliar game. The time between exercises can be used to integrate pauses for discussion, presentation, and reflection (ILOs (d) – (f)). To separate exercises, predefined game states, hereafter referred to as *scenarios*, can be created and then distributed to the students. Scenarios can be used to incorporate game elements into a course that would otherwise only be accessible at an advanced stage of the game. This approach is particularly useful to counteract the limited time available in a course. To achieve a certain part of the game in free play, potentially much more time is needed.

The strategic perspective of gameplay can be enhanced through additional exercises that let students utilize in-game data. This way, competencies from previous lectures can be applied to the game, e. g., corporate strategy definition and amortization calculation. In addition to the exercise conception and scenario creation, exercises should be linked to tell a coherent story, i. e., through task formulation. This story allows for a deeper immersion in the game world, which helps students to stay focused during an exercise [MDM08]. Considering the characteristics of RoI and the ILOs, exercises had to be designed and associated scenarios had to be created. Section 4.2 depicts a sample exercise from the course.

From the observations made in the game analysis (Fig. 1, phase 4), games need to be tailored to the intended learning outcomes. If the computer game does not natively align with all ILOs, *implementation of changes* (Fig. 1, phase 6) through modifications of the games' source code and with so-called *mods* can be considered. Since RoI did not support external timing of the game (cf. Sec. 3.1), the game developers adjusted the source code to allow the speed of the game to be affected via a mod. This change is available through the build (a so-called *beta*) named 'university' in the Steam library. To further align the game-based course with the ILOs outlined in Sec. 2, a mod was created that extracted in-game metrics, e. g., profit and pollution. With this data, students receive information beyond that provided by the game which can be used for decision-making. The data also supports compliance with ILOs (d) – (f).

3.3 Company Score Calculation and Web Interface Development

The data extracted from the mod was displayed on a web interface (i. e., a dashboard) to give students a better overview of the impact of their decisions in the game world. From this data, a *company score* was calculated to allow comparison of the teams' performance against each other and to foster competition among the teams. The considered key figures were weighted in such a way that the interaction between economic success indicators and ecological effects could be read from the company score. Equation (1) shows the calculation of the company score together with the weightings used.

$$\text{Company Score} = \frac{3}{1,000,000} \cdot \text{CV} + \frac{1}{100,000} \cdot \text{PRO} - 100 \cdot \text{MD} - 2,500 \cdot \text{POL} \quad (1)$$

The *company value* (CV) includes the value of the regions owned by a player and the investment costs of factories built by the player. The *machine downtime* (MD) is determined based on the machine running time of all owned factories in the previous month. Sustainability aspects are considered through the inclusion of the players' regions average *pollution* (POL). The *profit* (PRO) is composed of product sales in the respective shops, operating expenses, and other costs. Machine downtime and pollution are entered as percentage values and have a negative effect on the company score. Players started with a company score of 100.

Through the extracted and calculated data, aspects such as waste of resources, utilization of production capacities and effects of production on the environment can be observed. This helps to make statements about the economic and environmental sustainability of the players' actions taken in RoI.

3.4 Communication Server Setup and Virtual Machine Configuration

As the course was designed for an online setting, communication between course participants and teachers had to be ensured. Therefore, a server with multiple text and voice channels was set up using Discord (<https://discord.com/>). Text channels were used to distribute assignments and additional information. Voice channels allowed for easy communication during and between exercises. Additionally, application programs (e. g., computer games) could be screen-shared through the program. We set up virtual machines with Paperspace (<https://paperspace.com/>) using the Parsec Template which is made for cloud gaming to ensure accessibility of the game for all participants.

4 Demonstration of a Computer Game-based University Course

We integrated a modified version of Rise of Industry into a seminar course at the University of Hildesheim in the summer semester of 2020. With the objectives stated in Sec. 2 in mind, we created exercises that are adapted to the game and learning outcomes. In a three-day course, 24 students from various IT degree programs completed five exercises in RoI. During the course, the players and instructors communicated via Discord. It was decided to play in teams of two, a team size that has proven appropriate during internal play-testing to promote communication and teamwork. In the course, the measures defined in Sec. 3.3 represent an operationalization of ILOs (c) – (f); thus the experience of sustainability with all its facets is possible in the context of SCM. The company score (cf. Eq. (1)) can be seen as the enabler for reflection on the part of the players. The players' behavior, i. e., the decisions made during the game, show whether sustainable reflective action was taken in relation to the computer game. During the exercises, the company score was displayed on the web-based dashboard to evoke a sense of competition which strengthens learning effects and increasingly provides for a reflection of one's own decisions [B115]. Teams without adequate hardware played the game in the cloud (cf. Sec. 3.4). For more information on the implementation of the course, which is beyond the scope of this article, please visit <https://github.com/joscha999/PlanspielWeb/>. We now describe the used course structure, the exercise structure and example, and the observed results (cf. Sec. 4.1 – 4.3).

4.1 Course Structure

In the online course, we first explained the relevant functionalities as well as important overviews, e. g., a factory's GUI, and essential tools in RoI to the students. This was done since it cannot be assumed that students know a computer game in advance and to avoid that a lack of computer game affinity might affect the attendance of the course. Following the *game tutorial*, students performed a simple exercise in the game world to create an initial, shared understanding. Afterwards, the teams were tasked with founding a fictitious company and make up a name, logo, slogan, and strategy (TLA (1)).

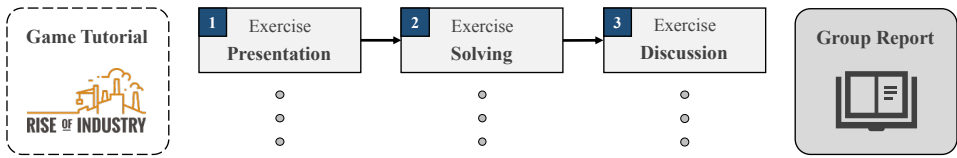


Fig. 2: Structure of the course

As seen in Fig. 2, three steps were repeated for each exercise. The exercises were presented by the instructors (step 1), then solved by the students in pairs (step 2, TLA (2)), and finally the results were discussed and reflected in class (step 3, TLA (3)). The in-class discussion was also intended for students to learn from the experiences of the other teams for following exercises. Each exercise had to be solved within a given time limit. Depending on the complexity, this was either one or two in-game years, i. e., 45 or 90 minutes. Once a scenario was loaded, the time within RoI unpaused and automatically stopped at the predefined end in-game date through the mod. During the exercise, team members communicated in their individual voice channels. The game was played by one person at a time, but the player role could be switched between exercises. While one student played the game, the teammate watched the live stream and helped with decisions and calculations. Upon completion of the course, students wrote a group report on the decisions made in the game and their implications to further reflect on what they had learned (TLA (4)).

4.2 Exercise Structure and Example

In the seminar, students performed five exercises, aimed at building, managing, and expanding their company's supply chain. The introductory exercise focused on a product with low vertical manufacturing depth. In subsequent exercises, the supply chain was extended to include products with a more complex bill of materials (BOM).

At the beginning of each exercise, the BOM of the demanded products must be determined so that the depth of the supply chain can be estimated (vertical dimension). To build the required number of factories at each level of the supply chain (horizontal dimension), the product demand can be referred to. From the demand it can be determined which and how many raw materials are needed for production and whether they are available in the players' region. Unavailable materials must be imported or extracted in another region. Once resources are available, factories and associated transportation infrastructure can be built and configured. The supply chain should be designed from the 'bottom up' to avoid production downtime and idle costs. Besides the construction and maintenance of supply chains, students had to calculate production times and quantities, time-to-market (TTM), and amortization times for various products. Furthermore, suitable production locations and sustainability aspects in the supply chain had to be identified and taken into account. In what follows, an exercise from the course is presented that covers topics such as TTM, investment cost analysis, and the assessment of the BOM. After setting up a production for wooden toys, the students had to produce marbles and deliver them to the toy store.

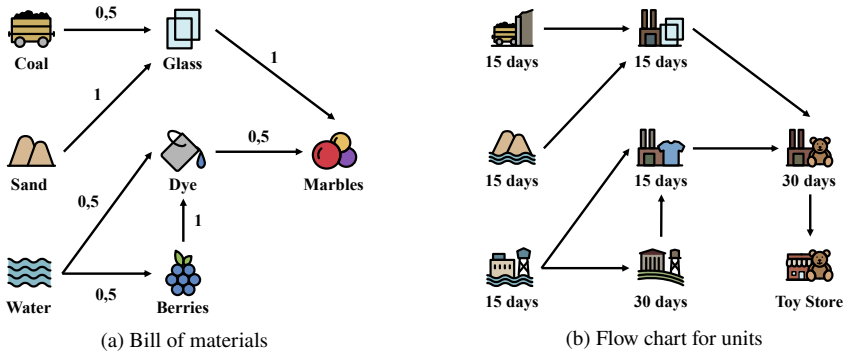


Fig. 3: BOM and flow chart for marbles production

In RoI, marbles are made of various raw materials (coal, sand, and water) and intermediate products (glass, dye, and berries). The BOM is shown in Fig. 3a. Initially, six units of marbles were demanded monthly at the local toy shop (primary demand $x_{Marbles} = 6$). The total demand r of raw materials and (intermediate) products can be calculated as follows:

$$\begin{aligned}
 r_{Coal} &= 0.5 \cdot r_{Glass} + x_{Coal} &&= 0.5 \cdot 6 + 0 &&= 3 \\
 r_{Sand} &= 1 \cdot r_{Glass} + x_{Sand} &&= 1 \cdot 6 + 0 &&= 6 \\
 r_{Glass} &= 1 \cdot r_{Marbles} + x_{Glass} &&= 1 \cdot 6 + 0 &&= 6 \\
 r_{Water} &= 0.5 \cdot r_{Dye} + 0.5 \cdot r_{Berries} + x_{Water} &&= 0.5 \cdot 3 + 0.5 \cdot 3 + 2 &&= 5 \\
 r_{Berries} &= 1 \cdot r_{Dye} + x_{Berries} &&= 1 \cdot 3 + 0 &&= 3 \\
 r_{Dye} &= 0.5 \cdot r_{Marbles} + x_{Dye} &&= 0.5 \cdot 6 + 0 &&= 3 \\
 r_{Marbles} &= x_{Marbles} && &&= 6.
 \end{aligned}$$

Tables 1 and 2 show the determination of investment costs for two types of factories, those without and those with harvesters (e. g., mines, pumps, and orchards). By default, factories can have a maximum of three harvesters each with a monthly production capacity of two units. Demand fluctuations can be responded to by adjusting the factory efficiency (e. g., 75% or 125%). In the game, an efficiency adjustment leads to a shortened production time with a simultaneous increase in air pollution. Since sustainability aspects are considered in our course, air purifiers that combat air pollution should be built by the students. For each cleaner, costs of \$125,000 and a monthly demand of two units of water ($x_{Water} = 2$) are to be considered. Summing up the building costs including a purifier, results in an investment cost of \$5,175,000.

Tab. 1: Investment costs of factories without harvesters

Product	Demand	Building Costs (\$)	Capacity	Buildings needed	Total (\$)
Glass	6	600,000	4 / month	1.5(2)	1,200,000
Dye	3	450,000	4 / month	0.75(1)	450,000
Marbles	6	600,000	2 / month	3	1,800,000

Tab. 2: Investment costs of factories

Product	Demand	Building Costs (\$)		Buildings needed		Total (\$)
		Harvester	Factory	Harvester	Factory	
Coal	3	175,000	225,000	1.5(2)	1	575,000
Sand	6	75,000	125,000	3	1	350,000
Water	5	75,000	150,000	2.5(3)	1	375,000
Berries	3	125,000	175,000	1.5(2)	1	425,000

The TTM of the marble production can be derived from the longest path in the flow chart in Fig. 3b. At least 90 days pass before the first marbles are delivered. The time starts with the collection of water for the berry plantation. Besides the production time, transportation times must also be considered. These are not included in the figure, because they depend on the road and vehicle type as well as the positioning of the factories in RoI.

4.3 Observed Results

During the exercises, student decisions and their subsequent impact, represented by performance metrics, e. g., company score, were collected in Rise of Industry on a monthly basis for all teams (cf. Sec. 3.3). Figure 4 presents the data weighted according to Eq. (1) over the course of the exercises played. The x-axis shows the months played by students. Vertical lines mark the beginning of the exercises, i. e., at months 0, 12, 30, 54, and 78, where jumps in the data are observable due to loading new scenarios. The remaining demand plot describes the absolute deviation of the products sold by the player and the demand for these products. This value is not included in Eq. (1), but helps to evaluate how well the exercises were solved by the teams. In the last exercise, the winning team, visualized with the green line, almost completely covered all demands, which indirectly had a positive effect on the profit and thus on the company value.

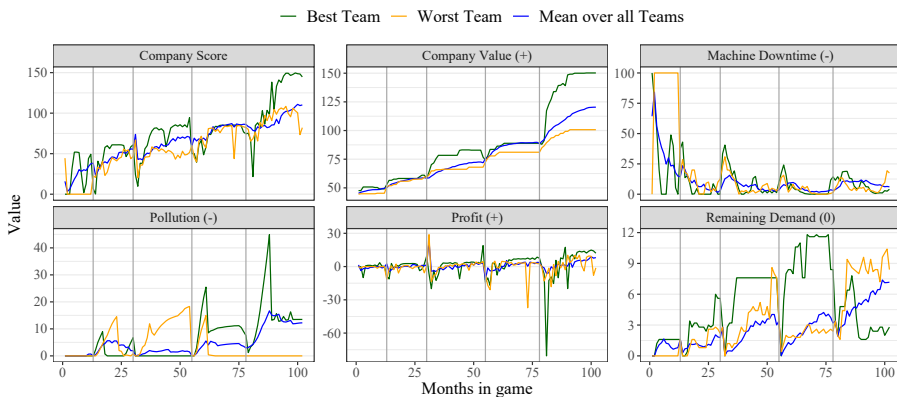


Fig. 4: Progression of Company Score

In Fig. 4, the local extrema of the plots indicate various student actions within the game world. Between months 75 and 85, the profit of the best team decreases due to the purchase of new factories and the expansion into a new region. At the same time, the company value increases. A subsequent delayed start of production due to late raw material deliveries temporarily increases machine downtime. After the start of production, pollution increases significantly, because an air purifier has not yet been set up. Therefore, there is a short-term drop in the company score, which stabilizes after the construction of the air purifier. At the end of the seminar, student feedback was gathered through a follow-up group discussion and an anonymous online form that included questions about the game-based and online nature of the course. The feedback showed that the students enjoyed participating in the course. In addition to the achievement of educational objectives (cf. Sec. 2), we evaluated whether the fun of playing the game in the course led to increased perceived learning transfer.

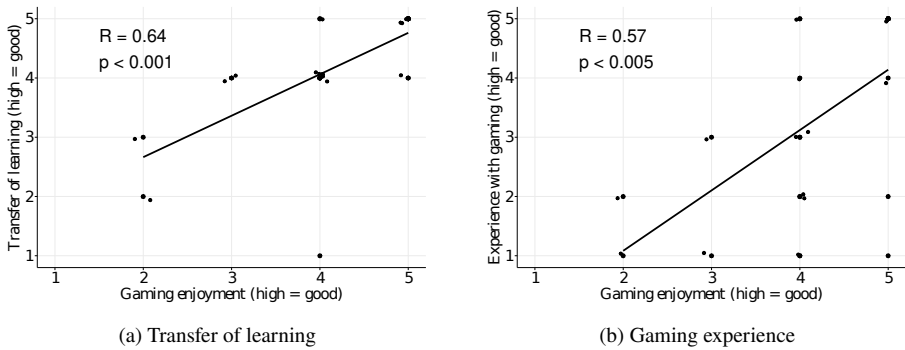


Fig. 5: Correlations of gaming enjoyment

Figure 5a shows that there is a strong correlation between the enjoyment of playing and the learning transfer. The more fun the students had in the game, the higher they rated their learning transfer ($R = 0.64$). It should be noted that in our target group the fun of playing the game depended on the previous experience with games (cf. Fig. 5b). The correlation is slightly weaker here ($R = 0.57$). However, students with very little experience also indicated that they sensed a high level of enjoyment and thus a high level of learning transfer. The p-values show that in both cases a significant correlation can be assumed ($p < 0.05$).

5 Descriptive Evaluation

In this section, we descriptively evaluate the first iteration of the computer game-based university course using Rise of Industry. The goal is to use the feedback to draw conclusions to further improve the quality of the course design and contents for future semesters. In particular, the evaluation investigates how well the TLAs supported the achievement of the educational objectives and thus the growth in competency among students.

In our study's game selection, RoI proved to be the most appropriate computer game for supporting ILOs (a) and (b). The remaining ILOs were corroborated through the course design. Through playing the game (TLA (2)) and observing the dashboard, students received direct feedback (in-game real time) on the consequences of their actions in the game world, were able to integrate different points of view, and thus better prepare and think through decisions. During the course, the exercises served as a frame of reference to ensure compliance with the given objectives. Although the exercises allowed us to specifically address individual aspects of the SCM, the compulsory loading of predefined scenarios prevented the students from experiencing the influence of their long-term strategic decisions. This meant that additional efforts of the students, e. g., the investment in an additional production line, were neglected by the new scenario. Moreover, with a new exercise the students received an optimal supply chain solution for prior exercises and were therefore not challenged to sufficiently maintain their own supply chain.

In order to be able to investigate students' competence development, students were observed during gameplay (TLA (2)). To solve the exercises, students applied prior knowledge and competencies received from lectures. In the online evaluation, students assessed their individual perceived learning transfer through playing RoI. It was noticed that the student assessment corresponded with the observations during the game, the discussions between the exercises (TLA (3)), and the reflective group reports (TLA (4)). From Fig. 4, it can be seen that students applied sensible solutions to external events in the game world (ILO (b)). In doing so, they made informed and sustainable decisions such as building an air filter or adjusting production in response to a demand fluctuation.

Application and reflection of knowledge (ILOs (a) and (f)) was achieved by combining the game with exercises on business calculations. From the group reports, it could be seen that students thought strategically, analytically, and in a networked manner. The reports also provided insight into how students thought about the impact of their decisions on the game. Overall, the seminar was characterized by teamwork among the participants, but also by the exchange between the teams during in-class discussions (TLA (3), ILOs (c) – (e)). As the course progressed, students' strategies and decisions improved as they worked in teams and utilized the data provided. By making the students aware of the consequences of their decisions in the game world, they worked in a feedback-oriented way. Whether the increase in competence experienced in the course can be applied in real-life situations remains to be seen. To this end, further iterations and research are necessary.

6 Conclusion and Outlook

In this work, drawing on Constructive Alignment, we defined intended learning outcomes and activities of a GBL course and derived requirements for the selection and adaption of a computer game. To guide the course design and development, we used DSR as an underpinning framework. In the university course, the computer game Rise of Industry was used to empower students to acquire and consolidate SCM-competencies.

Our results indicate that complex learning objectives can be covered not only in traditional

lectures-based, but in computer game-based formats as well. We found that visualizing the supply chain through the computer game and incorporating in-game metrics for decision-making worked well in the presented course. The modified version of RoI along with the exercises and the dashboard provided students with a concrete reflection tool for sustainable SCM-decisions. In the game, students experienced both the economic and ecological impacts of their actions. Although no generalizations are possible based on one iteration, our approach and results indicate that the embedding of a computer game in the university context is a promising approach in the area of SCM. According to our observation, it seems useful to conduct a learning status survey of the students before the seminar and (longer) time after the seminar in order to verify the knowledge growth as well as knowledge constancy and persistency.

Further iterations will follow to improve the combination of RoI and the dashboard and to provide a quantitative examination of the achievement of the goals. In order to raise the students' awareness of the long-term effects of their decisions and strategies, a structural change of the exercises is necessary. Students should play on a single game state for the entire duration of the seminar. Additionally, we plan to adjust the dashboard and the calculation of the company score to further encourage reflection and more sustainable decision-making, e. g., through the use of nudging elements. A graphical representation of the data through charts could help students in making more informed decisions as well. Moreover, it can be investigated how a higher temporal resolution of the submitted data, i. e., weekly or daily in-game, affects student decisions.

Bibliography

- [Bi96] Biggs, John: Enhancing teaching through constructive alignment. *Higher Education*, 32(3):347–364, 1996.
- [BI15] Blötz, Ulrich: *Planspiele und Serious Games in der beruflichen Bildung*. Bundesinstitut für Berufsbildung, Bonn, 2015.
- [Br08] Brabrand, Claus: Constructive Alignment for Teaching Model-Based Design for Concurrency. In: *Transactions on Petri Nets and other models of concurrency I*, volume 5100 of LNCS, pp. 1–18. Springer, Berlin, 2008.
- [CC06] Can, Gulfidan; Cagiltay, Kursat: Turkish Prospective Teachers' Perceptions Regarding the Use of Computer Games with Educational Features. *J. Educ. Techno. Soc.*, 9:308–321, 2006.
- [HKS14] Hamari, Juho; Koivisto, Jonna; Sarsa, Harri: Does gamification work?—a literature review of empirical studies on gamification. In: *2014 47th Hawaii international conference on system sciences*. IEEE, pp. 3025–3034, 2014.
- [Li17] Liu, Chiung-Lin: Using a video game to teach supply chain and logistics management. *Interact. Learn. Environ.*, 25(8):1009–1024, 2017.

- [MAA17] Mustata, Cristian; Alexe, C.E.; Alexe, C.M.: Developing Competencies with the General Management II Business Simulation Game. *Int. J. Simul. Model.*, 16(3):412–421, 2017.
- [MDM08] Martens, Alke; Diener, Holger; Malo, Steffen: Game-Based Learning with Computers – Learning, Simulations, and Games. *T. Edutainment*, 1:172–190, 2008.
- [Pe07] Peffers, Ken; Tuunanen, Tuure; Rothenberger, Marcus; Chatterjee, Samir: A Design Science Research Methodology for Information Systems Research. *J. Manage. Inf. Syst.*, 24(3):45–77, 2007.
- [PGPCC16] Pando-García, Julián; Periañez-Cañadillas, Iñaki; Charterina, Jon: Business simulation games with and without supervision: An analysis based on the TAM model. *J. Bus. Res.*, 69(5):1731–1736, 2016.
- [So19] de Souza, Bruno Campello; Roazzi, Antonio; de Souza, Jaélison Rodrigues; de Paula, Sílvia Luiz: Commercial Videogames and their Potential for Higher Education: A Study with 347 Brazilian Undergraduate Business Management Students. *Acta Scientiae*, 21(1):116–138, 2019.
- [SR19] Sousa, Maria José; Rocha, Álvaro: Leadership styles and skills developed through game-based learning. *J. Bus. Res.*, 94:360–366, 2019.
- [TGRW15] Tognetti, Alice; Grosse-Ruyken, Pan Theo; Wagner, Stephan: Green supply chain network optimization and the trade-off between environmental and economic objectives. *Int. J. Prod. Econ.*, 170:385–392, 2015.
- [Zu19] Zulfiqar, Salman; Sarwar, Binesh; Aziz, Saira; Ejaz Chandia, Khurram; Khan, Muhammad Kaleem: An analysis of influence of business simulation games on business school students’ attitude and intention toward entrepreneurial activities. *J. Educ. Comput. Res.*, 57(1):106–130, 2019.

Decarbonization in the Digital Platform Age – A Green Service Canvas for Small and Medium-sized Road Carriers


Christoph Heinbach ¹, Oliver Thomas^{1,2}

Abstract: Decarbonization initiatives are increasingly promoted by the emergence of digital platforms (DPs) and small and medium-sized road carriers (SMRCs) are urged to implement green services from DPs to ensure responsible management of their carbon footprint leading to sustainable efficiency gains, new revenue streams, and competitive advantages. However, platform-based decarbonization services are challenging for SMRCs due to scarce technology capabilities and it is necessary to identify green service offerings from existing DPs and to capture the dimensions for an adapted business model. This paper develops a Green Services Canvas (GSC) and provides a tool for SMRCs to use services from DPs for decarbonization. The conceptualized GSC is demonstrated by three use cases and evaluated by 10 digital experts. From the insights of this study, the authors explore (a) digital platforms for decarbonization in road transportation, and (b) the relevance of green service business models for SMRCs through DPs.

Keywords: Decarbonization, Digital Platform, Small and Medium-sized Road Carriers, Green Service Canvas, Use Case

1 Introduction

Today, the total amount of CO₂ emissions related to the transport sector are 8.26 Gigatons [IE21] and due to its vital function in our growing economies within intertwined supply chains, global freight transport accounts for 20% of the emissions generated, and this share is expected to rise by 2050 to 24% [In19]. More specifically, road freight activities account for 20% of global tonnes-km (tkm) while more than 70% of the energy in freight modes is consumed by road freight [Te18]. Thus, the need to cut CO₂ emissions and promote decarbonization in road freight transport is increasingly realized by freight transport stakeholders (e.g., shippers, carriers). Recently, a study of road carriers in Europe has analyzed decarbonization of *Small and Medium-sized Road Carriers (SMRCs)* to understand their environmental efforts [TM21]. The authors conclude that the decarbonization is scarcely promoted in the transport industry and missing information for carriers impedes proper actions to cut emissions while freight buyers do not sufficiently consider decarbonization in the procurement activities.

¹ University Osnabrück, Informationsmanagement und Wirtschaftsinformatik, Parkstr. 40, 49080 Osnabrück, {christoph.heinbach, oliver.thomas}@uni-osnabrueck.de,  <https://orcid.org/0000-0003-1602-5192>

² Deutsches Forschungszentrum für Künstliche Intelligenz GmbH (DFKI), Smart Enterprise Engineering, Parkstraße 40, 49080 Osnabrück, oliver.thomas@dfki.de

At the same time, *digital platforms (DPs)* are disrupting the road freight industry [OK21], and a transformation into green logistics platforms reveals green service opportunities addressing logistics efficiency, decarbonization, and costs [Kw20]. DPs appear in the forwarding industry as cost-efficient and flexible IT resources for SMRCs. Beyond, green service initiatives emerge from DP service providers by increased *utilization* of fleet capacities and empty run savings, the *aggregation* of vehicle data for fuel-efficient management, and the *tracking* of greenhouse gas (GHG) emissions generated for shipments according to carbon taxation [He21], [HO17], [Uc19]. Thus, SMRCs are required to screen DP services, assess the service capabilities, and decide on their business model adaption to remain competitive in a complex business environment.

However, it is difficult for SMRCs to identify and introduce green services from existing DPs and to adapt the business model due to limited human resources, the lack of IT abilities, and monetary constraints [SM16]. Against this backdrop, the promotion of guidelines for SMRCs to identify and develop platform-based green services is suggested to address the above-mentioned aspects. A structured approach for the green service innovation of DPs provides new opportunities for SMRCs to establish sustainable business models, meet customer requirements, and cope with increasing environmental pressure towards zero-emission freight transportation. Hence, this paper proposes a **Green Service Canvas (GSC)** for SMRCs in the road freight industry to navigate transport operators in the DP age towards green services for more sustainable transportation. While a traditional *Business Model Canvas* presents a useful instrument to visualize and structure business models, a GSC is considered as a *process model* which starts with the identification of green service based on the firm's initial situation, supports the evaluation of capabilities to succeed with the services, and arrives at an adapted *green service business model* systematically. To develop the GSC, existing models need to be critically reviewed in terms of the individual steps to successfully apply the derived concept. Furthermore, the conceptualized GSC is applied by use cases with carriers in practice to demonstrate their applicability in the real-world business environments.

2 Fundamentals

2.1 Digital Platforms for Road Freight Transportation

DPs in road freight operate as digital marketplaces [Bi19], digital freight forwarders [EG19], and DP integrators while the services provided are based on "*as a Service*" concepts [He21]. By the use of data as the only resource to offer customer services, DPs remain flexible and cost-competitive resources for SMRCs and "logistics platforms" [Ro20] appear in the road freight industry related to three different types [EG16]: *transaction platforms*, *data-focusing platforms*, and *integration platforms*. For instance, the improvement of freight efficiency by match-making platforms of freight shipments and load capacities [Gr17] refers to transaction-based marketplaces for shippers and

carriers. Similarly, the connectivity of IT systems through integration platforms supports the facilitation of end-to-end visibility [Ri18] for transport chain stakeholders. In essence, DPs operate in the road freight in the form of multisided platforms (MSPs) [HA15] enabling direct interactions between shippers and carriers to support transport management [Ki18], [Wu20] (Figure 1).

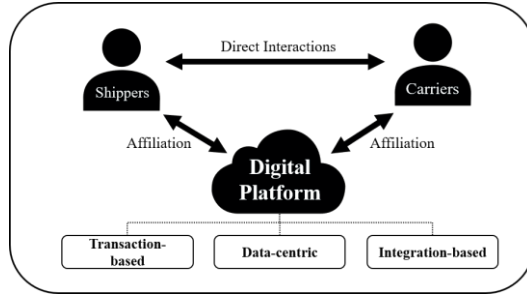


Fig. 1: Concept of digital platforms in the Road Freight Industry based on [EG19]

2.2 Platform-based Green Services Related to Decarbonization

Green logistics platforms aim to manage freight supply and truck resources between carriers and shippers efficiently and can be specified as “*shared logistics platforms*” [Kw20]. However, DP services based on interactions and related to efficiency gains that result from improved operations are recognized as “*green services*” due to their environmental impact. For instance, an advanced vehicle routing from DP services leads to shorter distances and supports the reduction of emissions from truck operations [He19]. Based on the analysis of DP service providers in the freight forwarding sector conducted by [He21], three green services were derived aligned to transport management for carriers which aim at ecological improvements as illustrated in Figure 2.

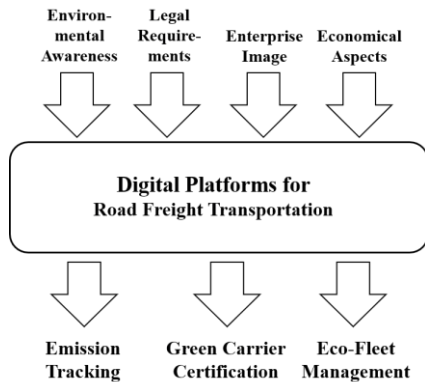


Fig. 2: Drivers and green services for carriers from digital platforms derived from [He21]

“*Emission Tracking*” refers to the calculation of GHG and carbon emissions associated with operational performance of trucks and the costs from CO₂ taxations to be charged to customers according to the load utilization. This implies also the provision of a decarbonization Key Performance Indicator (KPI) to support procurement activities of freight buyers. A “*Green Carrier Certification*” describes initiatives from DPs to provide shippers with a status on carrier’s responsibility and their ecological management. Therefore, the green certificate may be received by DP providers after performing a green audit program. Furthermore, “*Eco-Fleet Management*” refers to vehicle diagnostics of road freight trucks based on telematics systems that are widely in use for improved fleet management focusing on monitoring of driver and vehicle performance. Beyond, vehicle diagnostics for predicted maintenance services are available by connecting telematics systems with other information systems (IS) for improved vehicle management (e.g., Transport Management System).

3 Related Work

The exploration and introduction of green services from DPs indicates a close connection to the process of *digital transformation* for carriers in logistics [PS18]. In addition, DP service providers are regarded as an enabler for SMRCs to leverage digital transformation successfully by making use of the platform-based service offerings [Li18]. To support the process of digital technologies in *Small and Medium-sized Enterprises (SMEs)* different process models have been developed recently.

A Digital-Transformation-Tool³ was developed by emcra GmbH [em19] within project initiative "DIGITALISATION: Strategy Development Tool for the Digitalisation of SMEs" funded by the European Commission which facilitates the process of digital transformation in *SMEs* using a step by step approach. SMEs are guided through the process of digital transformation in the context of project management described by four phases: *Transformation Basics*, *Current Status Analysis*, *Manage the Process Change*, and *Implementation*. To manage a successful transformation, different analysis techniques are suggested to support the individual steps. The tool starts without a technology focus and SMEs risk missing DP innovations since a technology-induced perspective is not given to the beginning of the process.

Garzoni et al. [Ga20] developed a four-level approach for SMEs based on case studies comprising *digital awareness*, *digital enquirement*, *digital collaboration*, and *digital transformation*. The different levels focus on the process of digital transformation resulting from case analysis and decision-makers can evaluate the readiness level for processes of SMEs to make use of digital technologies. However, the approach also starts from the point of digital capabilities of SMEs without concise technology focus while the management of digital transformation is explored within an ecosystem boundary.

³ <https://digital-transformation-tool.eu>, accessed on 30th March 2021.

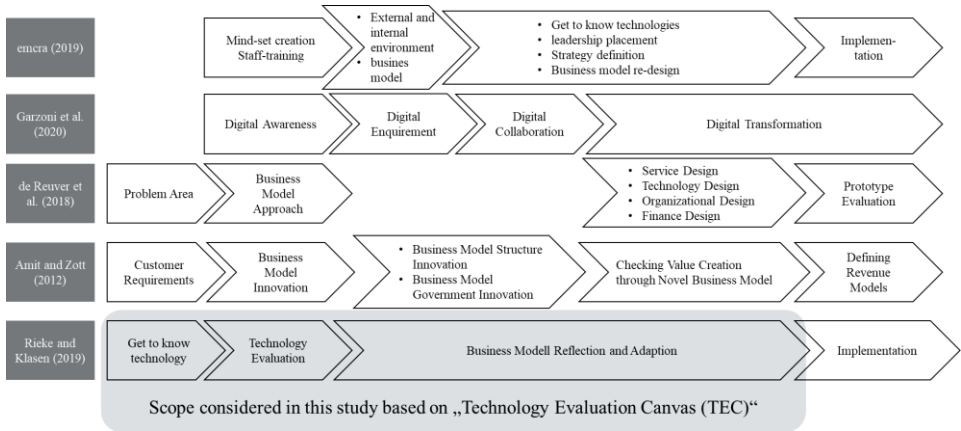


Fig. 3: Selected process models for technology introductions and digital transformation in enterprises based on [AZ12], [em19], [Ga20], [Re18], [RK19]

The developed process model proposed by **Rieke and Klasen** [RK19] does specifically consider digital technologies from a bottom-up approach and suggests a *Technology Evaluation Canvas (TEC)* implying the stages of *technology identification*, *assessment*, and *business model adaption* followed by *implementation*, *introduction*, and *monitoring & support*. While general process models and digital transformation roadmaps (Figure 3) are following a top-down approach without a technology focus in the beginning, this paper does follow the presented canvas concept of Rieke and Klasen [RK19] addressing the *introduction of technologies* which is consequently transferred into an adapted business model. Moreover, the utilization of digital technologies (e.g., cloud computing) is a key for organizations to establish innovative business models and enables service engineering to design new value propositions. From this starting point, the development of a **Green Service Canvas (GSC)** in this study is focusing platform-based services in the freight transport industry and presents the first approach to transform green services into an adapted business model in the road freight industry for carriers. To achieve the GSC we obtain design guidance for visual innovation canvas from [Av20] and [Th19] to facilitate a common understanding for SMRCs and to consolidate our knowledge visually. The GSC considered in this paper starts with a *service-induced identification* from DPs and decides on an adapted *green service business model*. The strength of this *bottom-up approach* relates to the initial exploration of relevant green services for SMRCs focusing on business needs and specifications. Besides, green services are assessed in terms of their capabilities for the carrier organization and from the market perspective in the next stage. Subsequently, the current business model, the business opportunities, and the adapted business model are addressed according to the green service impacts. This proposal aims at promoting decarbonization by SMRCs from DP services. To this end, a GSC is designed and evaluated throughout this paper to provide an instrument that guides SME carriers in the road freight industry for transforming platform-based services into a *green service business model*.

4 Development of Green Service Canvas for SME Road Carriers

In this section, we establish a *Green Service Canvas (GSC)* for the development of a *green service business model* based on available services of DPs comprising three phases: *green service identification* (phase 1), *assessment of capabilities* (phase 2), and *business model adaption* (phase 3). While DPs indicate a *data-driven* or *smart service* context in the road freight industry due to their disruptive service potential, carriers are required to explore green service innovations for the adaption of their current business models to establish new revenue streams and to sustain their competitive edge.

4.1 Green Service Identification of Digital Platforms

A GSC starts in the first phase with the exploration of existing DP technologies and the identification of relevant green services for SMRCs. From the networking and scalability effects of DPs in logistics and Supply Chain Management [PA19] an impact of cloud-based services on business models and their value propositions exist for road freight transport operators [He21]. Therefore, a systematic green service identification is suggested to be aligned to business model aspects to match services with the business scope of transport operators. To achieve the alignment, a conceptualization of relevant dimensions can be guided by four questions from Amit and Zott [AZ12] to develop business models presented in the following.

(1.) Business Cases (asking “Why” does the green service generate value?)

The value propositions from green services of DPs refer to *external and internal business cases* for SMRCs within an inter-organizational context. For instance, customers have an interest in their *CO₂ performance for reporting* purposes while internal cases are found in *transport lane-related fuel-management* to improve truck utilization. These green services lead to cost efficiencies as well as the legal requirements from the Quality and Environmental Department, i.e. to report results. The specified green services achieve new value propositions by the realization of *environmental awareness* and *key performance indicators* for benchmark purposes.

(2.) Service Screening (asking “What” digital platform services are relevant?)

The screening of technologies presents a core task for SMRCs to identify the most appropriate DP service for decarbonization. Relevant green services are represented by *Emission-Tracking Services in addition to transport management system (TMS) applications* and *Eco-Fleet Management for trucks and transport subcontractors*. Relevant state-of-the-art technologies are presented in *technology trend radar* which emerge specifically by institutions linked to logistics such as *BVL.digital*⁴. Transport operators can establish a screening process and frequently apply the tools for observing relevant green services of DPs.

⁴ <https://technologieradar.de/>, accessed 9th April 2021.

(3.) IS Alignment (asking “How” can the green service be connected and integrated to existing IS?)

Once a service has been observed, transport operators need to discover the options to connect the data from services with other IS in their organization to reach integrated service operations. For instance, an emission tracking service can be incorporated in the TMS through *API interfaces for shipment-based CO2 calculations* for customers based on the *transport-order-to-payment* process. Furthermore, the *remote Fleet Manage System Standard (rFMS)* represents a standardized interface within trucks that enables the integration of telematics data through cloud-supported integration. However, these steps require an understanding of technical specifications to anticipate barriers or limitations.

(4.) Networking Effects (asking “Who” are the general stakeholders for green services?)

Transport value chains are complex and composed of different stakeholders while networking effects arise from green services for *immediate consumption* in vertical and horizontal structures within end-to-end transport chains. A horizontal perspective is related to the use of green services by *shippers and consignees* to support their (green) *compliance* and to support *procurement processes* by KPIs. Vertical dimensions comprise the collaboration with other subcontracted carriers which apply a green service to deliver environmental data or information to partners or final customers. From the resulting networking effects by using green services, DP service providers can develop the platform into an advanced digital *green ecosystem* to further extend the service offerings to *insurance companies* to benefit from CO2 performance based on operational activities

Figure 4 summarizes the four dimensions for the identification phase of green services from DPs including the questions and proposed actions for SMRCs.

<p>① Business Cases</p> <p>“Why”: <i>Why does the green service generate value?</i></p> <ul style="list-style-type: none"> ▪ CO2 performance for shippers to support reports ▪ Fuel-management for transport lane and truck specific benchmarks 	<p>② Service Screening</p> <p>“What”: <i>What digital platform services are relevant?</i></p> <ul style="list-style-type: none"> ▪ Emission-Tracking Service in addition to TMS by modular ▪ Eco-Fleet Management of trucks and subcontractors
<p>③ IS Alignment</p> <p>“How”: <i>How can the green service be connected and integrated to existing IS?</i></p> <ul style="list-style-type: none"> ▪ API interfaces for shipment-based CO2 calculations ▪ Remote FMS-Standard (rFMS) to obtain truck data 	<p>④ Network Effects</p> <p>“Who”: <i>What are the general stakeholders?</i></p> <ul style="list-style-type: none"> ▪ Shippers, consignees an SMRCs as immediate service consumers ▪ Insurance companies within green ecosystems

Fig. 4: Steps for green service identification phase adapted from [RK19]

4.2 Evaluation of Green Service Capabilities

During the second phase, an assessment of digital competencies is an important task to understand the technology fit of SMRCs. Beyond, green services capabilities are required to be evaluated to benchmark a service output and analyze service risks and their opportunities. Since green services from DPs are operated within a cloud-based infrastructure, the digital capabilities of SMRCs are essentially reflected by the competence of employees to apply green service aligned to operations successfully. The concept of *Digital Competence Screening* developed by Bayrle [Ba17] provides a useful instrument that helps visualizing data-driven service concepts and to obtain details on competence requirements by users to manage green services. In addition, the application of *digital readiness checks* can support the analysis of the digital status quo and observe risks impeding successful green service deployment.

Moreover, a *Digital Framework* [Ba17] can be useful for SMRCs to investigate current and targeted service conditions and to reveal necessary actions resulting from a *Gap Analysis* to address future competencies for training and employee's qualification purpose. The objective is to understand if certain gaps exist within the carriers' organization that may inhibit a business model adaption according to the green service offerings to customers. To assess green services in this context, a *SWOT analysis* is recommended which can reveal service capabilities and their limitations. SMRCs can discuss the results with stakeholders and build an *evaluation catalog* that helps to evaluate the green service idea according to different business parameters. Finally, these parameters can be scored employing a *Scoring Model* and the green service implying the highest score is decided for business model implementation.

4.3 Green Service Business Model Adaption

In the third phase of the GSC the business model is reviewed and re-designed according to the decided green services. The aim is to compare the existing business model against the capabilities of the identified green services and to transform the green services into a *green service business model*. Green services of DPs are driven by data and to achieve a *data-driven business model* Bayrle et al. [Ba18] suggest applying a *Business Model Puzzle* related to three dimensions: *distant environment*, *close environment*, and *internal environment*. The distant environment comprises ecological and legal aspects, i.e. environmental data protection regulations. A close environment encompasses factors such as customer value, costs, partners, and channels. For the internal environment the processes, capabilities, and value propositions are considered. The "puzzle" provides visual assistance within interdisciplinary teams for the definition of the current business model which can be supported in addition by a *PESTLE analysis*.

From the existing business model, business opportunities are derived by matching results from an earlier assessment of service capabilities with requirements of internal (e.g., dispatchers, fleet managers) and external (e.g., shippers, consignee) stakeholders. Hence, the creation of *User Stories* is helpful to explore service specifications for process

implementation within real business environments. The knowledge can be consolidated in *green service profiles* to help to identify business opportunities in practice. Subsequently, *Best Practices* [RSK19] may support SMRCs to obtain practical orientation specifically for the service implementation at a later stage.

Against the existing business model, an evaluation of business opportunities in terms of their *Digital Fit* [RSK19] is suggested to ensure successful business model adaption. Reviews of green services are recommended in combination with *SWOT analysis* that allows the critical reflection of defined opportunities while a *Scoring Model* allows the comprehensive justification of green services related to business parameters. The results from the evaluation may lead to the conclusion that the identified green services are not addressing the stakeholder needs appropriately. Hence, the green services can be either rejected or further be developed in case of a positive decision. To this end, a *green service business model* is achieved by the identification of green services, the assessment of capabilities, and the adaption of the business model.

The findings of our conceptualization are summarized in the Green Service Canvas (Figure 5) comprising *three phases* and *nine steps*. Following the idea of a Technology Evaluation Canvas [RSK19], the model is to be maintained continuously by SMRCs leading to a transformed *green service business model*. Based on the GSC, carriers operating in the road freight transportation industry can systematically develop a digital service and promote decarbonization to take a sustainable stance. Since the concept presents a new approach for SMRCs in a complex business environment this paper aims at validating the model in the real world to understand its applicability.

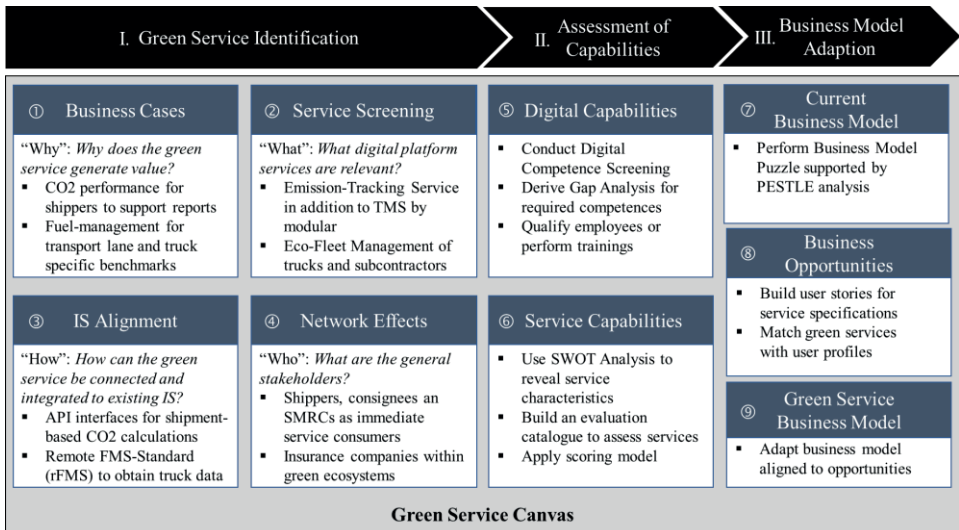


Fig. 5: Green Service Canvas for small and medium-sized road carriers adapted from [RK19]

5 Use Case Evaluation

5.1 Methodology and Use Cases

How well does the conceptualized GSC reflect the advancements of decarbonization based on DPs in road freight enterprises? To answer this question, we asked transport operators in the road freight industry about real-world green service offerings following the established process model using an online survey. The objective was to receive carriers' feedback on the different phases and individual steps systematically derived in this paper which allows them to develop a *green service business model* based on relevant cloud-service offerings. For the survey, green service use cases were obtained from [GS19] in combination with a google website search by querying the different green services options presented in section 2.2. Respective DP service providers were selected as use cases according to their green service offerings for SMRCs (Table 1).

Use Case ⁵	Description
EcoTransIT	EcoTransIT provides an automatic emission calculation tool that offers a flexible online solution for transport services. Customers can specify transport characteristics to enable emission calculations based on data from TMS while the tool can be incorporated in third-party websites.
Saloodo!	The digital freight platform "Saloodo!" offers a "Green Carrier" certification program that aims to certify freight forwarders according to their sustainable activities. The certificate can be used by SMRCs to demonstrate environmental responsibility during freight procurement.
RIO Cloud	Based on truck telematics systems RIO Cloud represents a digital platform operating as an integrator that offers specific services for SMRCs such as vehicle performance and diagnostics. Carriers can manage their mixed eco fleet efficiency and make improvements visible.

Tab. 1: Identified use cases for Green Service Canvas evaluation

Due to the exclusive access to SMRCs, 13 responsible digital managers were invited from individual carrier organizations and 10 surveys were finally completed. The experts had work experiences between six to 23 years with the freight forwarding business. To evaluate the nine GSC steps against the use cases, the respondents were asked to decide on the relevance of every single step towards a *green service business model*, and detailed information on the steps and the use cases were provided to avoid any bias or misunderstanding. The steps were measured on a five-point Likert scale, ranging from 'strongly disagree' (1) to 'strongly agree' (5). Additional room was left in the online survey to receive personal comments and suggestions for improvement beyond the scope of the GSC steps. By this approach, it was possible to engage practitioners from a real business environment to the proposed process model and to explore the level of GSC

⁵ <https://www.ecotransit.org>, accessed 13th April 2021; <https://www.saloodo.com>, accessed 15th April 2021; <https://rio.cloud/>, accessed 18th April 2021.

applicability for different green services of DPs. The results were gathered in a table and computed by a score for each of the GSC steps. Specifically, the score is defined as the median score of all SMRCs for the individual GSC steps of the use cases.

5.2 Results and Discussion

In Figure 6 the use cases are illustrated according to the individual steps of the GSC phases evaluated by SMRCs in a spider chart. The line representing green services that show high applicability of the GSC steps is grey (RIO Cloud) with an overall median score of 4.5, followed by blue (EcoTransIT) with an overall median score of 4.3 both ahead of orange (Saloodo!) with an overall median score of 3.8. The results give the confidence to confirm that the developed GSC steps are appropriate for application to develop green services business models systematically since the lowest median score of 3.2 from the use case “Saloodo!” (item “5. Digital Competences”).

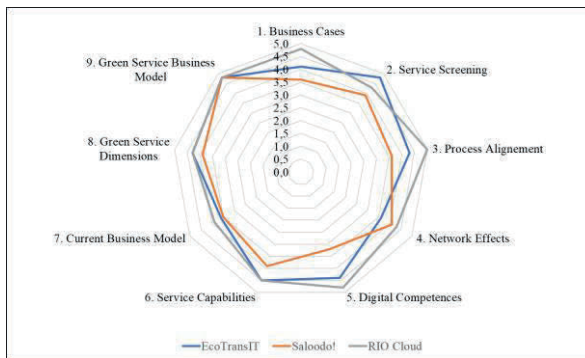


Fig. 6: Results of evaluated use cases related to Green Service Canvas

The green service with a higher score from “1. Business Cases” (e.g., RIO Cloud, score of 4.8) than green services with lower score show higher scores of “3. Process Alignment” (score of 5.0) and “5. Digital Competences” (score of 4.8) in addition. This may be explained by the fact, that green services related to integration platforms rely more on data connection and integration capabilities of enterprises, i.e. by a standardized interface to their TMS. Considering the use case overall, the item “9. Green Service Business Model” is the highest median score of 4.8 while “7. Current Business Model” shows the lowest median score of 3.7. This may indicate growing green service business expectations by SMRCs through DPs towards a transformed business model related to increasing business benefits. Surprisingly, all other GSC steps show a consistent median score ranging between 4.0 (“4. Network Effects”) and 4.3 (“3. Process Alignment”).

Since DPs arise as a flexible and cost-efficient IS resource, SMRCs seem to be motivated to make use of cloud-services fostering green services linked to the use of their operational data. The picture that arrives from the median score results is that green services aiming at decarbonization are increasingly of interest for SMRCs. The results

are coherent with the trend of decarbonization which is growingly regarded as a competitive factor to position SMRCs ahead in a fragmented market [TM21]. Although freight procurement is still primarily driven by freight rates, freight buyers are expected to request green service capabilities from carriers and integrate a decarbonization performance of freight suppliers into their procurement process for road freight transportation in the future. Higher environmental awareness and the increasing visibility of carrier performance through DPs may lead to the implementation of a “Decarbonization KPI” in digital freight platforms business supporting a green freight procurement process. Thus, future development of platform-based green services is likely to merge into digital green transport ecosystems addressing shared services [Kw20] and vertical green collaboration to further promote decarbonization.

6 Conclusions and Outlook

Decarbonization is strongly discussed across all business sectors and road freight transportation must contribute urgently to mitigate the severe GHG effects on our world. This contribution reveals new directions for SMRCs to take actions by a *Green Service Canvas (GSC)* which is developed to foster decarbonization from a platform-based “*as a Service*” perspective. The final GSC established in this paper is composed of three phases comprising *Green Service Identification*, *Assessment of Capabilities*, and *Business Model Adaption* based on the proposed concept of *Technology Evaluation Canvas (TEC)* from Rieke and Klasen [RK19]. Accordingly, nine steps were adapted to address a *platform-induced* approach for *green service development*. Based on the GSC concept, a service developing tool is provided as an artifact that supports road freight carriers to explore DPs and systematically transform the offered green services by platform providers into a *green service business model*. From the evaluation of the use cases by experts, the conceptualized GSC steps have shown validity for the application in reality and provide practical guidance for green service innovations of SMRCs.

However, subsequent studies should include a more diverse use case distribution to enhance the generalizability of results. Although the small sample of SMRCs has given confidence on the validity of the presented GSC steps, further empirical research is required to confirm the developed concept on an international scale comprising different road freight supplier market. Beyond, additional aspects related the *data-driven service engineering* is to be implemented in the concept to design *digital green ecosystems* and to realize *smart services* towards *green transport systems* to succeed in decarbonization collaboratively. For instance, shippers can start collecting “*green miles*” with carriers to benchmark decarbonization actively against other shipper-carrier constellations while collecting emission data along manufacturers’ product lifecycles. Hence, *smart transport decarbonization* may also be achieved in the future through the establishment of a *transport-related carbon footprint of products* with effects on the final buying decision of consumers. These technology-driven efforts are elementary not only because of their direct impact but also due to their inspiring effects for the entire freight transportation

sector to adopt a more sustainable business mindset.

Bibliography

- [AZ12] Amit, R., Zott, C.: Creating value through business model innovation, 2012.
- [Av20] Avdiji, H. et al.: A Design Theory for Visual Inquiry Tools. JAIS. 21, 695–734, 2020.
- [Ba17] Bayrle, C.: Digital Competence Screening: Kompetenzen für datenbasierte Dienstleistungen identifizieren-Ein Handlungsleitfaden. IPRI-International Performance Research Institute, 2017.
- [Ba18] Bayrle, C. et al.: Service-Business-Innovation-Lab – Ein Vorschlag für ein praxisorientiertes Vorgehen zur Entwicklung datenbasierter Service-Geschäftsmodelle. In (Bruhn, M. and Hadwich, K. eds.): Service Business Development, pp. 497–517 Springer Fachmedien, Wiesbaden, 2018.
- [Bi19] Bierwirth, C. et al.: Logistics Management. Springer, Lecture Notes in Logistics, 2019.
- [EG19]. Elbert, R., Gleser, M.: Digital Forwarders: A Market Oriented Taxonomy. In: Bierwirth, C. et al. (eds.) Logistics Management. pp. 19–31 Springer, Cham (2019).
- [em19] emcra GmbH: Process Model for Digital Transformation of SMEs, https://digital-transformation-tool.eu/wp-content/uploads/2020/01/DIGIT_Process-model_O1_FINAL.pdf, 2019.
- [EG16] Evans, P.C., Gawer, A.: The rise of the platform enterprise: a global survey, 2016.
- [Ga20] Garzoni, A. et al.: Fostering digital transformation of SMEs: a four levels approach. MD. 58, 8, pp. 1543–1562, 2020.
- [GS19] Göpfert, I., Seeßle, P.: Innovative Startups in der Logistikbranche–Eine Betrachtung der neuen Marktteilnehmer und empirische Erkenntnisse einer Fragebogenstudie. In: Logistik der Zukunft-Logistics for the Future, Springer, pp. 253–280, 2019.
- [Gr17] Graser, K. et al.: Digital Disruption in Freight and Logistics. Accenture, 2017.
- [HA15] Hagi, A., Wright, J.: Multi-sided platforms. International Journal of Industrial Organization. 43, pp. 162–174, 2015.
- [He21] Heinbach, C. et al.: Freight-Logistics-as-a-Service – Innovative Geschäftsmodelle für ein datengetriebenes Transportmanagement. HMD Praxis Wirtschaftsinformatik, 2021.
- [He19] Helmke, B.: Digitalisierung in der Logistik. In (Hartel, D.H. ed.): Projektmanagement in Logistik und Supply Chain Management. Springer Fachmedien, Wiesbaden, pp. 183–207, 2019.
- [HO17] Hofmann, E., Osterwalder, F.: Third-Party Logistics Providers in the Digital Age: Towards a New Competitive Arena? Logistics, 1, 2, 9, 2017.
- [IE21] IEA: CO2 emissions statistics, <https://www.iea.org/statistics/co2emissions/>, accessed 2021/03/18.
- [In19] International Transport Forum: ITF Transport Outlook 2019. OECD, 2019.

- [Ki18] Kille, C.: Digital Logistics. In (Fend, L. and Hofmann, J. eds.): Digitalisierung in Industrie-, Handels- und Dienstleistungsunternehmen. Springer Fachmedien, Wiesbaden, pp. 125–137, 2018.
- [Kw20] Kwak, S.-Y. et al.: Intention to Use Sustainable Green Logistics Platforms. 17, 2020.
- [Li18] Li, L. et al.: Digital transformation by SME entrepreneurs: A capability perspective. *Info Systems J.* 28, 6, pp. 1129–1157, 2018.
- [OK21] Ortwein, P., Kuchinke, J.: Digital Freight Forwarders Disrupt Road Freight Space. In Wurst, C. and Graf, L. eds.): *Disrupting Logistics*. Springer, Cham, pp. 163–175, 2021.
- [PA19] Pflaum, A., Fischer, C.: Von der Pipeline zur Plattform: Denkanstöße und Handlungsempfehlungen für das Supply Chain Management. In (Ulrich, P. and Baltzer, B. eds.): *Wertschöpfung in der Betriebswirtschaftslehre*. Springer Fachmedien, Wiesbaden, pp. 409–429, 2019.
- [PS18] Pflaum, A., Schulz, E.: Auf dem Weg zum digitalen Geschäftsmodell: „Tour de Force“ von der Vision des digitalisierten Unternehmens zum disruptiven Potenzial digitaler Plattformen. *HMD Praxis Wirtschaftsinformatik* 55, 2, pp. 234–251, 2018.
- [Re18] de Reuver, M. et al.: The Digital Platform: A Research Agenda. *Journal of Information Technology*. 33, 2, pp. 124–135, 2018.
- [Ri18] Riedl, J. et al.: *The Digital Imperative in Freight Forwarding*. Boston Consulting, 2018.
- [RK19] Rieke, T., Klasen, A.S.: Einführung von digitalen Technologien in KMU – Vorgehensmodell und Technology Evaluation Canvas. In (Linszen, O. et al. eds.): *Projektmanagement und Vorgehensmodelle*. Gesellschaft für Informatik e.V., Bonn, pp. 97–109, 2019.
- [Ro20] Roy, D. et al.: Logistikplattformen als Treiber für smarte Ökosysteme – Supply Chain Visibility als initialer Schritt für Transparenz und Steuerung von integrierten Real-time Supply Chains. *I40M* 5, pp. 63–66, 2020.
- [SM16] Stölzle, W., Martin, J.: *Logistikinnovation–Auf dem Weg zu Business Innovation in der Logistikdienstleistung*. Business Innovation: Das St. Galler Modell. Springer, pp. 499–529, 2016.
- [Te18] Teter, J.: The future of trucks—Implications for energy & the environment. In Stakeholder meeting on the Impact Assessment on HDV CO2 emission standards., https://ec.europa.eu/clima/sites/clima/files/events/docs/0121/iea_en.pdf, 2018.
- [Th19] Thoring, K. et al.: Exploring the Design Space of Innovation Canvases. In *Conference Proceedings of ADIM*. 2, 1, 2019.
- [TM21] Tölke, M., McKinnon, A.: Decarbonizing the operations of small and medium-sized road carriers in Europe. *Smart Freight Centre and Kühne Logistics*, 2021.
- [Uc19] Uckelmann, D. et al.: Grüne IT für eine grüne Logistik – Umweltorientierter Einsatz von Informationstechnologien für eine nachhaltige Logistik. In (Lochmahr, A. et al. eds.): *Digitalen Wandel gestalten*. Springer, Wiesbaden, pp. 141–150, 2019.
- [Wu20] Wurst, C.: Chancen von Logistik 4.0 nutzen. *Control Manag Rev.* 64, 2, 34–39, 2020.

Entwicklung eines Datenerfassungssystems zur Verwendung mobiler Speicher in lokalen Netzen, um erneuerbare Energien stärker zu nutzen

Hannes Jahn¹, Hans-Jürgen Pfisterer²

Abstract: Für eine nachhaltige elektrische Energieversorgung eignet sich die Nutzung verfügbarer Strom-Speicher von Elektrofahrzeugen. Damit kann Strom aus regenerativen Energiequellen erzeugungsunabhängig zur Verfügung gestellt werden, ohne den Ausbau von kostenintensiver Speicherinfrastruktur zu fokussieren. Für die Nutzung dieser Speicher wird eine bidirektionale Ladeinfrastruktur sowie eine gezielte Datenerfassung benötigt, um die Verfügbarkeit von Speicherkapazitäten zu prognostizieren und zu nutzen. Für bestehende Versuchsanlagen wurde ein umfangreiches Datenerfassungssystem entwickelt und eingesetzt. Diese erfassten Daten liefern neben den Energieflüssen und Positionsdaten der Fahrzeuge die Basis für eine Verfügbarkeitsprognose zur optimalen Auslegung und effizienten Bewirtschaftung solcher lokalen Netze.

Keywords: Datenerfassung, Daten-Management, GPS, Datenlogger, Smart Grids, lokale Netze, regenerative Energie, mobile Speicher, Stationäre Speicher, Bidirektionale Ladeinfrastruktur

1 Einleitung

Mit dem neuen „Erneuerbaren Energien Gesetz 2021“ möchte die Bundesregierung die Treibhausgasneutralität des Stroms in Deutschland bis 2050 erreichen [ME21]. Um dieses Ziel zu erreichen, muss der Ausbau von erneuerbaren Energien stark vorangetrieben werden. Schon heute werden 2,8% der regenerativ erzeugten Energien mit steigender Tendenz gedrosselt. Das entspricht einer Ausfallarbeit von 6,5 TWh, die geschätzte Entschädigungsansprüche von 709,5 Mio Euro verursacht [BU19] [Wi14]. Damit ist es von essenzieller Bedeutung die erzeugte Energie nicht nur effizient zu nutzen, sondern auch die Lasten zu verschieben (Demand Side Management) [SC14] [LO12] und Überschüsse zu speichern. Die gesonderte Speicherung von elektrischem Strom ist verlustbehaftet und setzt für einen wirtschaftlichen Betrieb, eine hohe Nutzungsfrequenz und eine kostengünstige Technologie voraus [KO21] [WA18]. Zusätzlich basiert die Verschiebung von Lasten auf speicherbasierenden oder zeitlich flexiblen Prozessen, wobei die zeitliche Flexibilität stark vom Nutzer und dessen Akzeptanz abhängig ist. Ein Ansatz, um elektrische Speichermöglichkeiten und verschiebbare Lasten zu kombinieren, bietet die Elektromobilität [HE14] [DÜ13]. Für den Betrieb von elektrischen Fahrzeugen,

¹ Hochschule Osnabrück, Fakultät Ingenieurwissenschaften und Informatik, Albrechtstr. 30, 49076 Osnabrück, hannes.jahn@hs-osnabrueck.de

² Hochschule Osnabrück, Fakultät Ingenieurwissenschaften und Informatik, Albrechtstr. 30, 49076 Osnabrück, j.pfisterer@hs-osnabrueck.de

müssen diese geladen werden, was besonders bei langen Standzeiten eine hohe Flexibilität bei der Lastverschiebung mit sich bringt. In diesen Zeiträumen können die Fahrzeuge nicht nur zeitlich flexibel geladen, sondern auch im moderaten Rahmen entladen und damit als elektrischer Speicher verwendet werden [SE14]. Das bidirektionale Laden ist bei Elektrofahrzeugen technisch möglich und bei der Gleichstromladung mit geringem Aufwand zu realisieren. Die Belastung der lithiumbasierten Speicher von Elektrofahrzeugen ist abhängig von der Lade/Entladeleistung wie auch von der Ladezyklentiefe und damit maßgeblich vom Fahrverhalten [HE10]. Mit der Einbindung solcher mobilen Speicher in Energienetze und einer moderaten Nutzung verringert sich die Lebensdauer der Fahrzeugspeicher im Vergleich nur geringfügig.

Mit dem Forschungsprojekt „LokSmart Jetzt!“ wurden seit dem Jahr 2012 in mehreren Phasen Versuchsanlagen entwickelt und angepasst. Mit regenerativen Energiequellen und elektrischen Speichern wurden lokale intelligente Netze aufgebaut und erweitert, die nicht nur überschüssige regenerative Energien nutzen, sondern auch die Verwendung eines Strompreissignals ermöglichen. In diesem Projekt lag der Fokus neben der Nutzung von stationären Speichern auch in der Nutzung von mobilen Speichern in Form von Elektrofahrzeugen [LI12]. Diese mobilen Speicher werden durch eine bidirektionale Ladeinfrastruktur in die lokalen Netze eingebunden [LO21]. Aus normativen Gründen bietet sich für das bidirektionale Laden das CCS- oder Chademo-Ladesystem an [CH21] [RU21]. Der Chademo-Standard wurde in Japan entwickelt und bietet als erste normierte Ladeschnittstelle eine bidirektionale Gleichstromladung [RU20]. Der CCS Standard bietet die Gleichstromladung als Erweiterung des europäischen Typ 2 Ladestandards an. In den letzten Jahren wurde, mit der Beteiligung des Forschungsprojektkonsortiums am Normungsgremium, die bidirektionale Gleichstromladung in die Norm mit aufgenommen. Mit der Auslagerung der Lade-Umrichter in den Ladepunkt spielt die Baugröße sowie das Gewicht nur noch eine untergeordnete Rolle und die höhere Nutzungsfrequenz erlaubt die Verwendung von hochwertigeren Ladesystemen, die ein größeres Leistungsspektrum und höheren Wirkungsgrade bereitstellen können. In diesem Forschungsprojekt mussten auf eigens entwickelte Fahrzeuge und Ladeinfrastruktur zurückgegriffen werden, um einen bidirektionalen Ladebetrieb zu ermöglichen, da es zu diesem Zeitpunkt noch keine kommerziellen Produkte zur bidirektionalen Ladung gab. Systeme für die Erfassung von positionsbezogenen Daten werden viele angeboten. Gerade im Bereich der Navigation und im Bereich der Leistungsüberwachung von Personen und Fahrzeugen sind diverse Systeme verfügbar. Auch im Bereich der Energienetze gibt es Systeme, die Energieflüsse erfassen. Jedoch sind diese Systeme häufig in sich geschlossen und von den Schnittstellen und Möglichkeiten begrenzt. Ein umfassendes Datenerfassungssystem für die Positionserfassung von mobilen Speichern, wie auch für die flexible Datenerfassung von Energiesystemen zur Bewirtschaftung der Speicher und zur Optimierung von Energiesystemen, gibt es in dieser Form noch nicht. Damit ist dieses Forschungsprojekt eines der ersten, das den CCS-Ladestandard zur bidirektionalen Einbindung von mobilen Speichern nutzt und die Optimierung der Energiesysteme mittels einer umfangreichen Datenanalyse fokussiert [SC17] [LI12].

Für eine optimale Auslegung und eine intelligente Bewirtschaftung der lokalen Netze ist es maßgeblich das Betriebsverhalten der Energiesysteme zu erfassen und auszuwerten. Das Verhalten der Energiequellen, der Speicher, die Energiesenken werden erfasst, um Korrelationen der Energiesysteme zu analysieren und zu prognostizieren. Damit ist es möglich, die Energieflüsse der Erzeuger und Verbraucher zu prognostizieren und die stationären, wie auch die mobilen, Speicher sinnvoll einzusetzen und zu bewirtschaften [JI13]. Weiterhin können diese Netze in mathematischen Modellen [PA14] [NA14] abgebildet und der Betrieb der Energiesysteme mit den erfassten Daten simuliert und analysiert werden. Das Verhalten dieser lokalen Energienetze wird optimiert und die Teilsysteme passend skaliert. Für den idealen Betrieb und die optimale Auslegung solcher Systeme ist neben einer umfangreichen Datenerhebung, auch die Datenhaltung von essenzieller Bedeutung. Daraus ergibt sich die hier behandelte Forschungsfrage: „Wie weit ist es in der Praxis möglich, die Energieflüsse und Einflussfaktoren von Energiesystemen und die Bewegungsprofile von mobilen Speichern zu erfassen und auszuwerten, um eine effizientere Nutzung der Ressourcen und einen optimalen Betrieb von lokalen Netzen zu gewährleisten“.

Eine wichtige Aufgabe in dem Forschungsprojekt war es, für die historisch gewachsenen und in sich geschlossenen Versuchsanlagen, geeignete Möglichkeit zu finden, um mit minimalem Eingriff in den Betrieb die benötigten Daten zu erfassen und auszuwerten. Die relevanten Energieflüsse in den lokalen Netzen werden in Quellen, Senken und Speicher unterteilt und dahingehend weiter differenziert. Die Stromquellen bestehen zum einen aus flexiblen Quellen, aus denen jeder Zeit Strom entnommen werden kann, und zum anderen aus volatilen Quellen. Für die Prognose dieser volatilen Quellen werden zusätzliche Parameter benötigt, um Rückschlüsse auf das Verhalten der Quellen zu bekommen. Daher spielt die Erfassung der Wetterdaten und deren Vorhersage eine wichtige Rolle für die Beschreibung und Analyse der volatilen Energiequellen. Für die Prognose ist es auch von Interesse, Abhängigkeiten von vermeintlich flexiblen Quellen mit einzubeziehen, um neue Geschäftsmodelle und optimale Nutzungsverhalten zu identifizieren. Beispielsweise spielen die volatilen Strompreise an der Börse oder der CO₂-Ausstoß des erzeugten Stroms für den Endkunden nur eine untergeordnete Rolle, da diese Schwankungen an den Endkunden nicht weitergereicht werden und dieser darauf üblicherweise nicht reagieren kann. Je stärker die Stromerzeugung und -speicherung mit dem „Prosumer“ [AR16] dezentralisiert wird, desto relevanter werden Geschäftsmodelle zur flexiblen Stromnutzung, um Versorgungsnetze, regenerative Energien und Kraftwerke besser auszulasten, wie auch optimierter zu betreiben. Daher ist es wichtig, auch Werte wie Strompreise oder CO₂-Preise mit zu erfassen und in die Analyse mit einzubeziehen. [HÜ19]

Die Fahrzeuge bzw. mobilen Speicher, sowie die Senken sind stark nutzerabhängig und die Verhaltensprognosen dementsprechend komplex. Bei Unternehmen wie auch bei Haushalten kann das Nutzerverhalten anhand von Arbeitsprozessen abgeleitet und damit gut prognostiziert werden. Bei energieintensiven Anlagen, die regelmäßig oder zyklisch betrieben werden, können hilfreiche Nutzungsprofile entwickelt werden. Dies gilt ebenso bei Pendlern im Schichtdienst oder bei Lieferfahrzeugen, die regelmäßig die gleichen

Routen fahren. Bei individuellen Fahrprofilen, bei denen variable Ziele fokussiert werden oder verkehrsbedingten Einschränkungen unterliegen, ist es sinnvoll diese Profile aufzuschlüsseln und mit Stau- und Wetterprognosen abzugleichen oder die Terminkalender in die Analyse mit einzubeziehen.

2 Ergebnisse

2.1 Methodische Konzeptentwicklung des Datenerfassungssystems

Um die Forschungsfrage zu betrachten, ist die Entwicklung und Nutzung von Prototypen in den Versuchsanlagen essenziell. Für die Entwicklung des Datenerfassungssystems wurde eine Anforderungsanalyse durchgeführt. Aufbauend auf einer Literaturrecherche nach Webster und Watson [WE02] konnte ein erster Entwurf der Anforderungsanalyse entwickelt werden. Dabei wurde Suchbegriffe definiert und Wissenschaftsdatenbanken systematisch nach Veröffentlichungen durchsucht. Die Ergebnisse wurden reduziert, analysiert und daraus ein Anforderungskatalog entwickelt. Diese Anforderungen wurden in Expertengesprächen und Gruppendiskussionen erweitert und überarbeitet. Mit einer auf den Ergebnissen aufbauenden Marktanalyse wurde eine Übersicht von Angeboten zu verfügbaren Datenerfassungssystemen erstellt. Zusätzlich zu dieser Vorarbeit wurde eine Beobachtung der Versuchsanlagen sowie deren Betriebsstrukturen durchgeführt mit der Schlussfolgerung, dass die verfügbaren Produkte den Anforderungen nicht gerecht werden. Auf Basis des Anforderungskatalogs wurde ein Prototyp des Datenloggers entwickelt und aufgebaut. Des Weiteren wurde auf Basis der Anforderungsanalyse eine Datenstruktur entwickelt und mittels einer Datenbankarchitektur auf einem Server in Betrieb genommen. Mit diesem Prototypensystem wurden Versuche durchgeführt und Optimierungen in die Entwicklung übertragen. Im letzten Schritt wurden die Datenerfassungssysteme in die Versuchsanlagen integriert. Anhand der Versuchsanlagen wurden Simulationsmodelle konzipiert und mit den erfassten Daten verifiziert.

2.2 Datenlogger

Die wichtigsten Punkte der Anforderungsanalyse waren die hohe Flexibilität des Datenerfassungssystems und der wartungsarme und autonome Betrieb. Die Versuchsanlagen waren weit entfernt, so dass ein regelmäßiges Auslesen der Daten und ein Wechseln/Laden der Batterie von Hand schwer zu realisieren waren. Der Datenlogger sollte darüber hinaus ein breites Spektrum an Schnittstellen zur Verfügung stellen, um mit einer Vielzahl von Sensoren und Energiemanagementsystemen kommunizieren zu können. Die personenbezogenen Daten mussten aus Gründen des Datenschutzes sicher gespeichert und verwaltet werden. Daher wurde ein eigener Server bereitgestellt, um die Daten zu schützen.

Um eine detaillierte Routenanalyse zu gewährleisten, muss eine hochauflösende Positionserfassung realisiert werden, die sekundengenau die Koordinaten, wie auch die Höhe und Geschwindigkeitswerte erfasst. Der Datenlogger wird in einem internen sicheren Netzwerk registriert und sendet die erfassten Daten nach der Fahrt über einen WLAN-Hotspot des Unternehmensnetzwerks an die Datenbank. Der Datenlogger wird flexibel über den Zigarettenanzünder an das 12 V Netz des Fahrzeugs angeschlossen und so im Betrieb versorgt. Beim Parken und Verlassen des Fahrzeuges wird auch die Spannungsversorgung unterbrochen. In diesem Fall muss der Datenlogger über eine interne unterbrechungsfreie Stromversorgung (USV) weiter mit Strom versorgt werden, um die Daten über den WLAN-Hotspot auf den Server hochzuladen. Sobald die Daten übertragen worden sind, bringt der Datenlogger sich in einen Ruhezustand und wartet bis er erneut mit der Fahrzeugspannung versorgt wird. Bei der nächsten Fahrt erfasst er wieder die Positionsdaten und speichert sie ab, um bei dem nächsten verfügbaren Netz die Daten wieder hochzuladen. Allerdings müssen bereits bei der Fahrzeugzulassung alle mit der Fahrzeugsteuerung agierenden Systeme berücksichtigt werden, was in diesem Fall nachträglich nicht möglich war. Daher mussten auf die Messdaten vom Speicher, wie auch vom Antriebsstrang verzichtet werden und die Informationen konnten nur über die Positionsdaten abgeleitet werden. Bei den folgenden Neuzulassungen können diese Datenlogger mitberücksichtigt und so weitere wichtige Informationen erfasst werden.

In der stationären Anwendung stehen eine kontinuierliche Stromversorgung und ein dauerhaft verfügbares Netz zur Verfügung, weshalb der Datenlogger bei Verbindungsausfällen nur als Datenspeicher und ansonsten als Schnittstelle zur Datenbank fungiert. Der Datenlogger verfügt neben den Ethernet-Schnittstellen über diverse Kommunikationsschnittstellen, wie USB, Controller Area Network (CAN), Serial Peripheral Interface (SPI), Inter-Integrated Circuit (I²C). Über diese kann der Logger an Geräte und Bussysteme angeschlossen und die kommunizierten Daten an die Datenbank weitergeleitet werden.

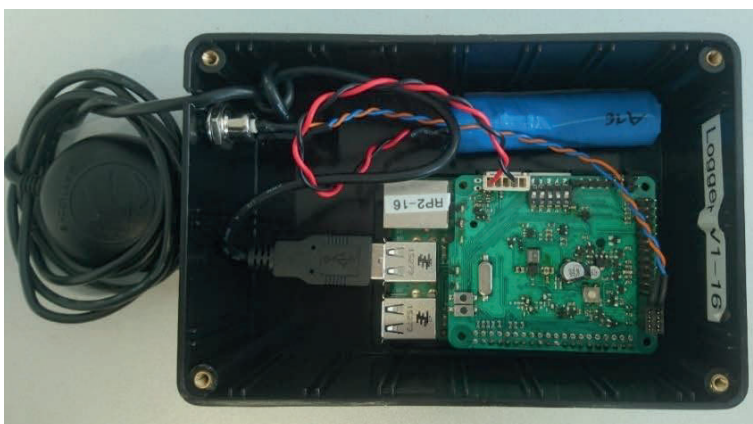


Abb. 1: Geöffneter Datenlogger für den Mobilen Einsatz

Für den flexiblen Einsatz bietet sich ein einfaches Betriebssystem auf Basis des Linux-Kernels an, das auf einem kompakten Einplatinencomputer betrieben wird. Um die Funktionalität der Hardware mit einer unterbrechungsfreien Stromversorgung, einem Energiemanagementsystem und den benötigten Schnittstellen zu erweitern, wurde eine zusätzliche Hardware entwickelt. Darüber hinaus sind weitere Funktionen in der Hardware implementiert, wie eine Real-Time-Clock (RTC) oder Signal-LEDs, wie auch Taster und Schalter, um den Datenlogger zu konfigurieren, zu überwachen und zu steuern. Für die Verwendung in Fahrzeugen wurde der Datenlogger in einem robusten Gehäuse integriert und mit einem GPS-Sensor über die USB-Schnittstelle erweitert, siehe Abb. 1. Mit einer standardisierten Steckverbindung kann das System über den Zigarettenanschluss an das Fahrzeug angeschlossen und mit Strom versorgt werden.

Weiterhin gibt es die Möglichkeit, internetfähige Steuergeräte und Energiesysteme direkt oder indirekt an die Datenbank anzubinden.

2.3 Datenstruktur

Wie zuvor beschrieben, werden nicht nur Informationen über die elektrischen Energieflüsse der lokalen Netze benötigt, sondern auch weitere Daten, wie Temperaturen, Sonneneinstrahlung, Fahrprofile, etc. Für den Einsatz der Datenbank in unterschiedlichsten lokalen Netzen wird eine entsprechende Struktur benötigt, die eine hohe Form von Flexibilität zulässt.

Im Zentrum der Datenstruktur steht der Sensor im Sinne der Datenquelle. Dieser hat Attribute und Beziehungen, die ihn detaillierter beschreiben und Zuordnungen zulassen. Der Sensor ist auf einen frei wählbaren Wertetyp reduziert und ist dadurch vielseitig einsetzbar, siehe Abb. 2. Durch die Zuordnung des Sensors zu einem Datenbesitzer (Data Owner), einem Standort (Location), einer Einheit (Unit), einer Sensorart (Kind), einem Gerät (Ressource) und weiteren Werten (Data), ist es möglich, jeden gesammelten Wert in der Datenbank anzulegen und abzuspeichern. Diese Werte können auf vielen Ebenen gruppiert und so ideal für die Auswertung zur Verfügung gestellt werden. Zusätzlich ist für die Nutzung der Daten eine Beziehung zwischen den Datenbesitzern und den Datenbanknutzern vorgesehen, um Berechtigungen für die Nutzung der einzelnen Daten zu bestimmen. Um Korrelationen zwischen den erfassten Werten, wie beispielsweise den Positionsdaten und den einzelnen Nutzern (Fahrern) zu erkennen, ist es möglich, jeden Wert einem bestimmten Nutzer zuzuordnen.

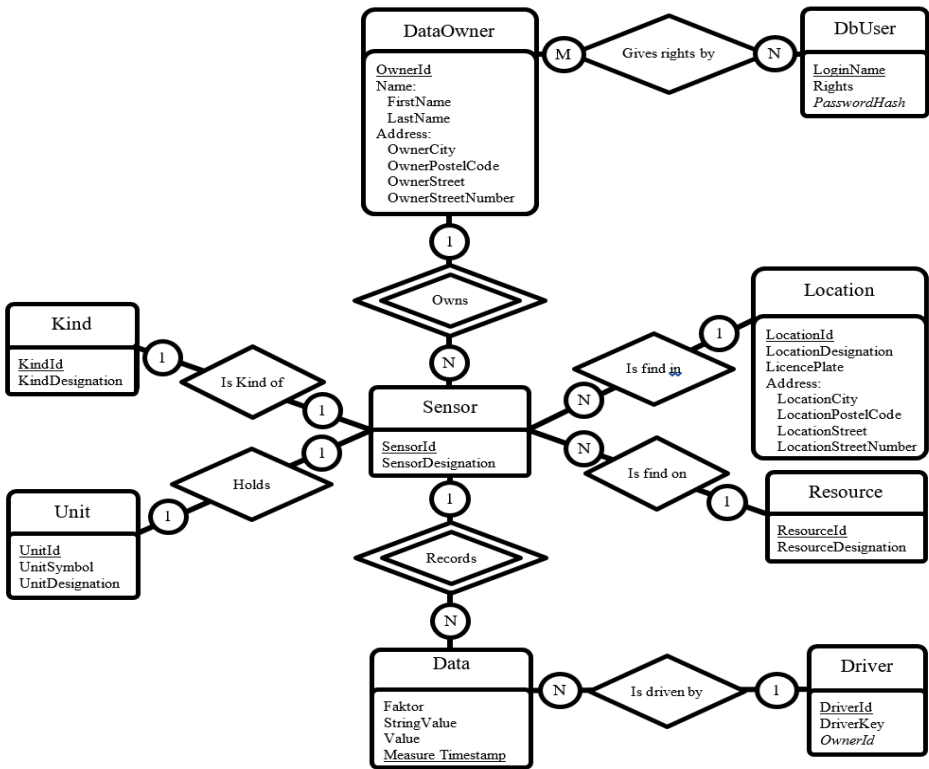


Abb. 2: Entwickelte Datenbankstruktur

2.4 Schnittstelle zur Datenauswertung

Für Zugriffe auf die Datenbank und die Nutzung der Daten wurden unterschiedliche Anwendungen geschrieben. Eine Anwendung greift direkt aus Matlab [MA21] auf die Datenbank zu und ruft die zuvor ausgewählten Daten ab. Damit ist es möglich, die Daten zu analysieren und in die Modelle von lokalen Netzen einzuspeisen. So können die Modelle der existierenden Anlage durchgeführt und prognostiziert werden. Mit der Analyse solcher lokalen Netze können die Auslegung der Speicher, die erneuerbaren Energieerzeuger und die Verbraucher besser aufeinander abgestimmt und die Regelungen optimiert werden, um beispielsweise den Autarkiegrad, wie auch den Eigenverbrauchswert zu steigern. Mit dem prognostizierten Verhalten der Energiesysteme können die entwickelten Regelalgorithmen für die Energieflusssteuerung optimiert werden und effizienter arbeiten. Dadurch ist es möglich die Speichereigenschaften und die Lastverschiebung dieser Systeme voll auszuschöpfen und bei entsprechenden Anreizprogrammen Gewinne zu erwirtschaften.

Eine weitere Anwendung wird verwendet, um ausgewählte Werte für die flexible und einfache Handhabung in CSV-Dateien abzulegen. Die so erzeugten Daten können einfach genutzt werden. Diese Datengrundlage kann für die simple Darstellung, Gegenüberstellungen oder die Erzeugung von Diagrammen, Bildern, etc. in Berichten und Präsentationen verwendet werden. Für diese Nutzung ist eine gute Vorverarbeitung essenziell. Daher ist das Programm zum Abrufen der Daten so aufgebaut, dass der Zeitraum, die Auswahl, wie auch die Reihenfolge der einzelnen Datenquellen gewählt werden können. Die Zeitintervalle der einzelnen Sensordaten können sich unterscheiden, weshalb jede Datenquelle auf ein einheitliches Zeitintervall angepasst werden kann. Dafür gibt es unterschiedliche Funktionen zur Reduktion oder Expansion der Datenreihen. Damit ist es möglich, eine einheitliche Schrittweite zu wählen und somit eine synchronisierte Datenbasis zu erhalten. Die Daten werden mit dem Zeitstempel und einem gewünschten Präfix/Suffix abgelegt. Zusätzlich gibt es die Möglichkeit, die Konfigurationen abzuspeichern, um die konfigurierten Datentransformationen zu reproduzieren. Die aufbereiteten Datensätze können mit unterschiedlichsten Programmen verarbeitet und ausgewertet werden, wie beispielsweise mit Microsoft Excel oder GIS-Programmen im Abschnitt 2.5.

2.5 Möglichkeiten zur Datenauswertung

Zur Veranschaulichung wurden die GPS-Daten von einzelnen Fahrzeugen grafisch ausgewertet. Die einzelnen GPS-Punkte wurden, farblich mit einer Geschwindigkeitsinformation versehen, als Route dargestellt und mit dem Satellitenbild hinterlegt, siehe Abb. 3. Weiterhin konnten Fahrverhaltensanalysen durchgeführt werden, bei denen wiederholende Fahrten (Pendelrouten) lokalisiert und somit Zeiträume prognostiziert werden, in denen der mobile Speicher zur Verfügung steht. Über die Zeit- und Geschwindigkeitsinformationen können Stausituationen identifiziert und Stauhotspots prognostiziert werden. Dies kann zusätzlich in die Verfügbarkeitsberechnung der mobilen Speicher einfließen.

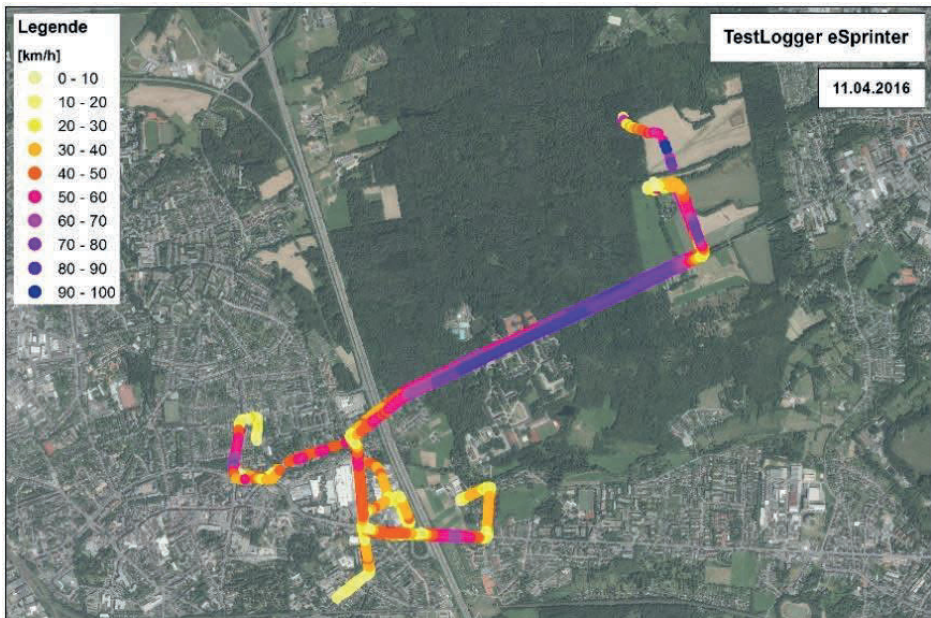


Abb. 3: Darstellung einer erfassten Route des GPS-Datenloggers [LO21]

Die Datenlogger wurden auch für die Erfassung von stationären Anwendungen konzipiert. Zur Veranschaulichung wurden die Stromzähler aus einer Projektanlage (kleines Bürogebäude mit Wohngebäude) [KO20] aus der Datenbank extrahiert und mit Microsoft Excel dargestellt. Diese Stromzähler bilden alle drei Phasen des Blockheizkraftwerks (BHKWs) und alle drei Phasen des Gesamtverbrauchszählers ab. Das BHKW ist wärmegeführt und erzeugt im Betrieb eine negative elektrische Verbrauchsleistung [W].

In der Abb. 4 ist ein Wochenverlauf der Energieflüsse dargestellt. Auf der Phase 1 (gelb) des Gesamtverbrauchs sind starke Schwankungen und eine geringe Grundlast zu erkennen. Phase 2 (blau) versorgt keine Grundlast, es sind jedoch starke Schwankungen und eine Lasterhöhung an den Arbeitstagen zu erkennen. Phase 3 (grün) unterliegt geringen Schwankungen und einer höheren Grundlast, auch die Lasterhöhung an den Arbeitstagen ist dort zu erkennen. Über das BHKW wird über alle drei Phasen die gleiche Leistung eingespeist, so kommt es zeitweise zu negativen Leistungen an den Gesamtverbrauchsstromzähler. Trotz des hohen Verbrauchs wird beim Betrieb des BHKWs Überschussenergie erzeugt und eingespeist. Bei einer unzureichenden Einspeisevergütung würde sich ein elektrischer Speicher anbieten, um den Eigenverbrauchsanteil zu erhöhen. Anhand des wärmegeführten Betriebs des BHKWs kann der Wärmebedarf abgeleitet werden. In der Nacht sinkt der Wärmebedarf - jedoch fällt der Bedarf wider Erwarten zum Wochenende nicht ab. Zu empfehlen wäre neben der Nachtabsenkung auch eine Temperaturabsenkung am Wochenende, wenn kein Bürobetrieb stattfindet.

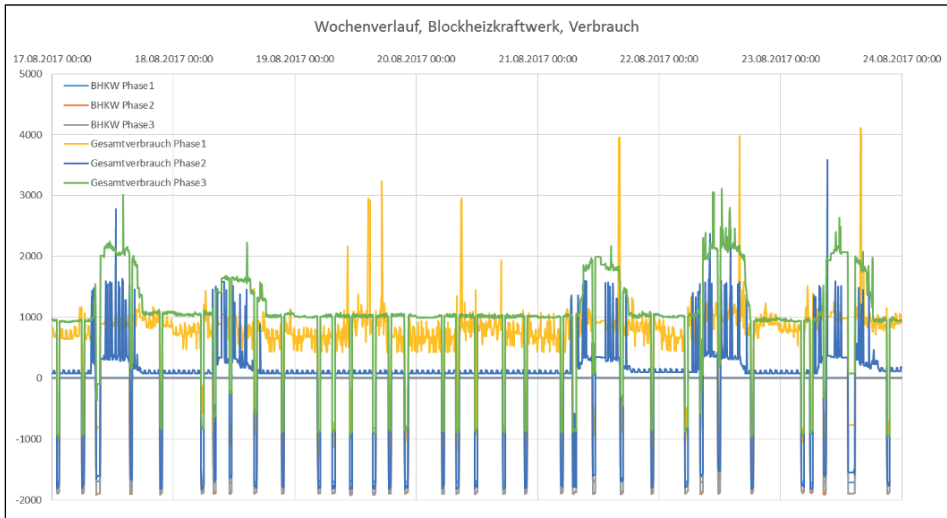


Abb. 4: Darstellung des Energieverbrauchs eines kleinen Bürogebäudes (Wochenverlauf)

Nach einer Jahresanalyse hat sich herausgestellt, dass im Winter durch den hohen Wärmebedarf das BHKW genügend Strom produziert, um einen Großteil des Strombedarfs zu decken. Im Sommer fehlt der Wärmebedarf und damit die hohe Stromerzeugung, weshalb sich eine Photovoltaikanlage zum Decken des Strombedarfs anbietet. Aus baurechtlichen Gründen ist es nicht möglich, auf der dargestellten Projektanlage eine Photovoltaikanlage zu installieren. Durch die Einbindung von elektrischen Fahrzeugen (mobilen Speichern) in das lokale Netz, ergibt sich die Möglichkeit den Strom des BHKWs in den Arbeitszeiten durch die Pendelfahrzeuge zwischen zu speichern. Weiterhin konnte ein Energietransfer von einem Wohngebäude mit Photovoltaikanlage zu dieser Projektanlage durch die Pendelfahrzeuge realisiert werden.

Bei drei weiteren Projektanlagen wurden die Daten zentral von einem kommerziell verfügbaren Energiemanagementsystem [EC21] erfasst. Es war jedoch nicht möglich, die gesammelten Sensordaten über eine Schnittstelle zur Verfügung zu stellen oder die Daten direkt auf die Datenbank hochladen zu lassen. Über einen entwickelten Service können die Daten über die verfügbare Cloudanbindung des kommerziellen Produktes abgerufen und auf der Datenbank gespeichert werden. Dieses Energiemanagementsystem bietet eine Vielzahl von Sensordaten, die in die flexible Datenbankstruktur integriert werden können. Jedoch empfiehlt es sich, vorab Auswertungsziele zu definieren und die Datenerfassung zu selektieren. Damit ist es möglich, unabhängig von weiteren Akteuren, alle benötigten Daten auf dem eigenen Server zu bündeln und für die weitere Verarbeitung bereitzustellen.

3 Fazit/Diskussion

Es hat sich in dem Forschungsprojekt gezeigt, dass die Datenlogger in großer Stückzahl eingesetzt und zuverlässig unterschiedlichste Daten an die Datenbank übertragen und gespeichert wurden. Bei der Auswertung der Fahrzeuge hat sich herausgestellt, dass die regelmäßigen Fahrten schnell identifiziert und klassifiziert werden. Die individuellen und ungeplanten Fahrten sind schwer vorherzusagen und in die Bewirtschaftung mit einzubeziehen. Trotz der drei unterschiedlichen Anwendungsfelder konnte in der kurzen Versuchszeit keine quantitative Aussage über die Nutzungspotenziale getroffen werden. Daher stellt sich für das weitere Vorgehen die folgende Forschungsfrage: „Wie hoch ist das Nutzungspotenzial der Lastverschiebung durch die Prognose von Energiesystemen und den eingebundenen mobilen Speichern zur verstärkten Nutzung von erneuerbaren Energien?“

Bei der Installation und bei der Nutzung der Datenlogger in den Fahrzeugen gab es teilweise Probleme in der Handhabung. Es kam vor, dass die Datenlogger nicht angesteckt, einfach in den Innenraum gelegt oder ganz entfernt wurden. Daraus ergibt sich die Forschungsfrage: „Wie hoch ist die Nutzerakzeptanz von solchen positionsbezogenen Datenerfassungssystemen bei der Nutzung von lokalen Netzen und welche möglichen Anreizprogramme gibt es zur Förderung der Akzeptanz?“

Bei einer großflächigen und kommerziellen Nutzung der Positionsdaten bietet es sich an, die Datenerfassung direkt in die Fahrzeugelektronik zu integrieren. Durch die Integration können weitere Fahrzeugparameter erfasst und die Gefahr von Bedienungsfehlern, wie auch die Ausfallwahrscheinlichkeit gesenkt werden. Eine weitere Möglichkeit zur Positionserfassung wäre die Nutzung von Mobiltelefonen mit entsprechenden Apps. Eine Herausforderung für diesen Ansatz ist die Identifikation des Fahrzeugs (mobilen Speicher) auf dem Mobiltelefon, sowie die notwendige Bereitschaft der Fahrer. Gerade im öffentlichen Bereich (Car-Sharing etc.) oder in Unternehmen (Flottenfahrzeuge, Lieferwagen etc.), bei denen kein vorkonfigurierbares Diensttelefon zur Verfügung steht, könnte die Bereitschaft des Fahrers, positionserfassende Software auf das private Mobiltelefon zu spielen, ein unüberwindbares Hindernis darstellen. Wichtig dabei wäre die Betrachtung des datenschutzrechtlichen Rahmens, wie auch schon in der zuvor aufgestellten Forschungsfrage, die Akzeptanz des Nutzers. Gerade bei der Akzeptanz zur Erfassung und Verarbeitung positionsbezogener Daten ist es von essenzieller Bedeutung, dass die Daten nur im Rahmen der Fahrzeuganalyse aufgenommen werden und die Anonymität der Nutzer gewahrt wird.

Bei den stationären Energiesystemen sind häufig zentrale Energiemanagementsysteme eingesetzt, die eine große Anzahl von Informationen für die Steuerung erfassen. Diese Energiemanagementsysteme stellen ihre gesammelten Daten nur selten über Schnittstellen zur Verfügung. Deshalb ist es mit hohem Aufwand verbunden, den Datenlogger in solche geschlossenen Systeme zu integrieren. Eine weitere Möglichkeit ist die direkte Einbindung der Energiemanagementsysteme in die Datenbank, was jedoch eine gewisse Flexibilität der Fremdsysteme voraussetzt. Viele dieser Energiemanagementsysteme

haben eigene Datenbanken und Cloud-Anbindungen, die den Kunden zur Verfügung gestellt werden. Mit angepassten Service-Anwendungen ist es möglich, auf diese Schnittstellen zurückzugreifen und die Daten zu erfassen. Eine Standardisierung solcher Datenschnittstellen würde die Datenerfassung erleichtern.

Bei der Auswertung der Daten können viele Energiequellen und Senken im Mittel gut prognostiziert werden. Gerade der wiederkehrende Jahreszyklus, bedingt durch das Wetter, spiegelt sich in vielen Energieflüssen wider. Aber auch prozessbedingte Abläufe können gut identifiziert und prognostiziert werden. Im privaten Bereich ist dies jedoch nur in Grenzen möglich.

Mit dem Nachweis der Funktionalität des Datenerfassungssystems und der daraus abzuleitenden Prognose von Verhaltensmuster, ist es möglich, die Bewirtschaftung von lokalen Netzen zu optimieren und mobile Speicher in diese Systeme mit einzubeziehen. Neben der Nutzung in lokalen Energiesystemen stellt sich die Frage, ob mit den Daten noch weitere Geschäftsmodelle entwickelt werden können, daher ergibt sich als weitere Forschungsfrage: „Welche neuen Geschäftsmodelle sind auf Basis der Datenerfassungssysteme möglich?“. Bei kommerziellen Ansätzen der Datenerfassung und der Datenhaltung im größeren Rahmen empfiehlt es sich, die physikalische Datenspeicherung auszulagern und Serviceleistungen einzukaufen. Es steigert die Flexibilität, wie auch die Skalierbarkeit und erhöht die Ausfallsicherheit. Für die Bewirtschaftung und Optimierung von lokalen Netzen entspricht eine zentrale Datenspeicherung und Verarbeitung jedoch nicht dem Autarkiegedanken vieler Kunden.

4 Literaturverzeichnis

- [AR16] Aretz, A.; Bost, M.; Gähns, S.; Hirsch, B.: Prosumer für die Energiewende. Institut für ökologische Wirtschaftsforschung, Berlin, 2016.
- [BU19] Bundesnetzagentur: Quartalsbericht Netz- und Systemsicherheit - Gesamtes Jahr 2019, 2019.
- [CH21] CHAdEMO, <https://chademo.com>, Stand: 03.05.2021.
- [DÜ13] Dütschke E.; Paetz A.G.; Wesche, J.: Integration Erneuerbarer Energien durch Elektromobilität—inwieweit sind Konsumenten bereit, einen Beitrag zu leisten? Springer, Berlin Heidelberg, 2013.
- [EC21] eco data, POWERDOG, www.eco-data.de, Stand: 03.05.2021.
- [HE10] Herb, F.: Alterungsmechanismen in Lithium-Ionen-Batterien und PEM-Brennstoffzellen und deren Einfluss auf die Eigenschaften von daraus bestehenden Hybrid-Systemen, Ulm, 2010.
- [HE14] Heinrichs H. U.: Analyse der langfristigen Auswirkungen von Elektromobilität auf das deutsche Energiesystem im europäischen Energieverbund, KIT, Karlsruhe, 2014.

- [HÜ19] Hüer, L.; Stadie, N.; Hagen, S.; Thomas, O.; Pfisterer H.-J.: Der CO₂-Kompass: Konzeption und Entwicklung eines Tools zur emissionsarmen Stromnutzung. Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn, 2019.
- [JI13] Jiang, Q; Xue, M; Geng, G.: Energy Management of Microgrid in Grid-Connected and Stand-Alone Modes. IEEE transactions on power systems, 2013.
- [KO20] Koenzen, U.; Rohrbach, R.; Broedersdorff, J.; Pfisterer, H.J.; Jahn, H.; Jungandreas, C.: Abschlussbericht (Gesamtvorhaben) LokSMART JETZT!, E-Mobilität im lokalen smart grid Bidirektionales Gleichstromladen in der gewerblichen Praxis – Kopplung von regenerativen Quellen und gewerblicher Elektromobilität im lokalen Smart Grid, DLR, Bonn, 2020.
- [KO21] Komarnicki, P; Lombardi, P; Styczynski Z.A.: Wirtschaftlichkeit elektrischer Energiespeichersysteme, Volume 1, S. 41-105 | 30, 2021.
- [LI12] Linsen, J.; Bickert, S.; Hennings, W.: Netzintegration von Fahrzeugen mit elektrifizierten Antriebssystemen in bestehende und zukünftige Energieversorgungsstrukturen, Jülich, 2012.
- [LO12] Logenthiran, T.; Srinivasan, D.; Shun, T. Z.: Demand Side Management in Smart Grid Using Heuristic Optimization, IEEE Transactions on Smart Grid 3/3, S. 1244 – 1252, 2012.
- [LO21] lokSMART, lokSMART Lokale smart grids www.loksmart.de, Stand: 03.05.2021.
- [MA21] Mathworks, <https://www.mathworks.com/products/matlab.html>, Stand 03.06.2021.
- [ME21] Meldung des BMWIs zum „Elektrischen-Energie-Gesetz 2021“ (EEG 2021), <https://www.bmwi-energiewende.de/EWD/Redaktion/Newsletter/2020/10/Meldung/topthema.html>, 2021.
- [NA14] Nardelli, P. H.J; Rubido, N; Wang, C.; Baptista, M.S.; Pomalaza-Raez; C.; Latva-aho, C.&M.: Models for the modern power grid. The European Physical Journal Special Topics 223, S. 2423–2437, 2014.
- [PA14] Parisio, A.; Rikos, E.; Glielmo L.: A Model Predictive Control Approach to Microgrid Operation Optimization. IEEE Transactions on Control Systems Technology, 22/5, S. 1813 - 1827, 2014.
- [RU20] Rumale, S.; Ashkar, H.A; Kerner, T.; Koya F.; Eitzenberger, M. Design and Implementation of an On-Board Vehicle CHAdeMO Interface for Vehicle-to-Grid Applications. IEEE, 2020.
- [SC14] Scholz, Y.; Gils, H. C.; Pregger, T.; Heide, D.; Cebulla, F.; Cao K.-K.; Hess, D.; Borggreffe, F.: Möglichkeiten und Grenzen des Lastausgleichs durch Energiespeicher, verschiebbare Lasten und stromgeführte KWK bei hohem Anteil fluktuierender erneuerbarer Stromerzeugung. DLR-Institut für Technische Thermodynamik, Stuttgart, 2014.
- [SC17] Schimanski, S.; Barta, M.; Schröder T.F.: Entwicklung eines autonomen Datenloggers zur Erfassung von Bewegungsdaten bei Elektro-Pkws für die Ableitung von nutzungszentrierten Dienstleistungsinnovationen. Springer Gabler, Wiesbaden, 2017.

- [SE14] Seddig, K.; Jochem, P.; Fichtner, W.: Nutzung von elektromobilen Flotten zur Lastverschiebung, 13. Symposium Energieinnovation, Graz, 2014.
- [WA18] Wawer, T.; Griese K.M.; Halstrup, D.; Ortman, M.: Stromspeicher im Quartier. Zeitschrift für Energiewirtschaft 42, S. 225–234, 2018.
- [WE02] Webster, J.; Watson R.T.: Analyzing the past to prepare for the future: Writing a literature review. MIS quarterly 26/2, 2002.
- [WI14] Wiest, P.; Eilenberger, S.; Tenbohlen, S.: Ausfallarbeit bei einem Einspeisemanagement für Photovoltaikanlagen im Niederspannungsnetz. IEEE Power and Energy, 2014.

Towards a transparency-oriented and integrating Service Registry for the Smart Living Ecosystem

Jonas Rebstadt,¹ Henrik Kortum,² Simon Hagen,³ Oliver Thomas⁴

Abstract: Many domains are increasingly dominated by interdependent services and data exchange between different actors, leading to the emergence of data ecosystems. As a result, service engineers are increasingly tasked with integrating existing service components and data sources into service systems and orchestrating them. In complex areas such as smart living, these tasks are even more difficult by the particular relevance of individual data protection requirements and the low fault tolerance of security-related systems. To address these issues, a central service registry for the domain smart living has been prototypically developed and evaluated, focusing especially on the transparency of data flows and the technical exchangeability of service components. In this way, added value is achieved for data providers and for data users by providing information on the forwarding of their own data as well as on the origin of the data and possible data quality.

Keywords: Transparency; Interoperability; Cyber-Physical Systems; Smart Living; Ecosystems

1 Introduction

Many offerings are providing their value by combining tangible products or product systems with intangible (technical) services [BT13]. This holds especially true for the domain of smart living. The smart living ecosystem links a variety of IoT devices and systems to provide smart solutions in a wide range of areas related to living and represents a multi-billion-dollar market [SD20]. However, even though the examination of the 'value added' from a business perspective is promising, especially for end-users in a private context many potentials remain unexploited, due to the complexity of integrating all components. A frequently used example of such a complex offering in the smart living domain [Ba19] is the intelligent gatekeeper - an automated, intelligent building door control system [Ko20]. Such a system enables the realization of various use cases, like automated door opening for workmen in the event of a water leakage, without the need for an occupant to be at home. An intelligent sensor system registers the defect at an early stage, reports it and later

¹ German Research Center for Artificial Intelligence, Smart Enterprise Engineering, Parkstraße 40, 49080 Osnabrück, Germany jonas.rebstadt@dfki.de

² German Research Center for Artificial Intelligence, Smart Enterprise Engineering, Parkstraße 40, 49080 Osnabrück, Germany henrik.kortum@dfki.de

³ German Research Center for Artificial Intelligence, Smart Enterprise Engineering, Parkstraße 40, 49080 Osnabrück, Germany simon.hagen@dfki.de

⁴ German Research Center for Artificial Intelligence, Smart Enterprise Engineering, Parkstraße 40, 49080 Osnabrück, Germany oliver.thomas@dfki.de

allows the workman to enter. It shows the multiple advantages from a business point of view, e.g. reduced time for ordered craftsmen due to an automated, biometrics based, entry, as well as reduced claims for delay. On the other hand, it also demonstrates several social and economic sustainability related obstacles that need to be tackled, to accomplish this comprehensive approach.

That said, this scenario will only become possible if a large number of previously unlinked products and services can be integrated into one (eco-)system. For example, in the context of housing, special security and data protection requirements must be met in order to achieve acceptance among all consumers involved. This results in the need for transparency, which is mainly in control of the service engineer (i.e. the developer) of the overall offer, who needs to take data flows and storage into account at an early stage of development. In addition to regulatory issues, there are also technical questions regarding the integration and interaction of systems from different manufacturers which need to be considered. These issues are related to both, social and economic sustainability. On one hand the link between products and services in data driven systems needs to be simplified [Ka19] and the exchange of data in ecosystems needs to be fostered [Ko21]. On the other hand, several roles required for the development of these systems need to be brought together and the access for potential customers needs to be enabled [SK20].

Science and practice have already tackled these questions from different directions and proposed approaches allowing combination of variously designed components. For example, the modularization of products and (technical) services [AS11] [JVD07] has been proposed, which allows the components to be orchestrated in regard of customer needs [Ba11]. [Ge20] extended this approach by tackling the relevance and challenges of solution space modeling out of the product system view. Furthermore, the standardization of interfaces plays an important role to ensure interoperability, as pursued by consortiums like the W3C⁵. However, the question arises how these different approaches can be brought together and used in a multi-layered domain like smart living to support the developer at an early stage. In the context of the ForeSight research project, therefore, two research questions have arisen, which are answered in the following paper:

1. *How to technically design a service registry, considering the integration of interoperable and domain-specific standards?*
2. *What added value is created for actors in the smart living ecosystem by disclosing service relations in regard to the derived requirements?*

These questions are addressed in a Design Science Research oriented way, focusing on the development and the evaluation of a service registry for smart living. For this purpose, central relevant concepts and the used methodical approach are explained in chapter 2 and chapter 3. In chapter 4 the derivation of requirements and the development of the prototype are presented, followed by an interview-based evaluation in chapter 5 and a concluding discussion section in chapter 6. Our work provides real-world value by providing service

⁵ The World Wide Web Consortium has for example published a standard for Web of Things, which is of high interest in the domain smart living: <https://www.w3.org/WoT/>

engineers with a tool that increases transparency in the data ecosystem smart living and helps them develop privacy-friendly, explainable solutions. Additionally, the evaluated impact of exemplary transparency increasing tools generated theoretical value for the research community.

2 Related Work

The smart living domain is a highly interconnected data ecosystem. In general, data ecosystems describe the interplay of various actors, such as organizations and individuals, sharing data in networks and using it as resource [OL18]. Each actor in a data ecosystem occupies one or more of three central roles characterized by different functions. The three central roles are data consumer (utilization and analysis of data as well as provision of feedback), data producer (production and publication of datasets), and intermediary (provision of services and orchestration) [Ca19] [OL18]. A cross-enterprise, data-centric approach enables the emergence of new platform-based business models and provides added value for all actors in a data ecosystem [Ko21]. Since the simultaneous interaction of various actors in a cooperative and competitive relationship [GGO21] using external data is fraught with uncertainty, establishing transparency is central to the success of a data ecosystem. It must be ensured that all data processing, including the legal, technical, organizational and procedural conditions, can be traced at any time [Ja19]. This holds especially true for the data ecosystem of smart living.

Unlike industrial data ecosystems, smart living concerns a very private area of human life and requires the collection and processing of personal data [Ko20]. Within the domain, there are a large number of different entities and sensitive assets that need to be properly protected. Smart living in general expands the smart home concept beyond the aspect of simple home automation to include other areas, such as intelligent energy management and health [Ba14] [ND15]. Smart living solutions therefore operate in an environment where a heterogeneous mix of service systems must be consolidated [Ko20]. A key characteristic of the smart living domain is thus the connection of different subdomains and stand-alone solutions with the aim to improve the quality of daily life [MH18]. The connection of these individual solutions leads to the emergence of a complex data ecosystem in which sensor manufacturers record raw data from homes, make it available to service providers (intermediaries), who in turn aggregate it, process it intelligently and make it available to data consumers within the ecosystem.

To allow an easy interaction and identification of the manifold number of components offered for smart homes, a semantic representation incorporating data as well as rules for logical reasoning is highly beneficial [BHL01]. This representation should allow comprehensive manipulation and retrieval of data, accessible by humans as well as by computer systems [Ku04]. In case of cyber-physical or smart service systems, incorporating products, services and information technology [BLW18], considerations of different languages and domains need to be taken into account. In the context of smart living, a combination of WoT and existing deliberations from GAIA-X self descriptions for systems seems to be promising

to be compliant to possible future regulations in the European Union. WOT is extending the scope of classical web services by integrating cyber-world as well as physical-world services. It intends to create an ecosystem by orchestrating services gracefully in a human centered and intelligent way [SW16] [ZGC11]. GAIA-X self descriptions tackle a very similar issue, trying to be machine-readable, machine-evaluable and including expressive semantics [Ot21]. It is an important component of the GAIA-X Ecosystem, simplifying the exchange of information about services or other involved entities and trying to increase the trust between different actors in the ecosystem.

3 Methodical Approach

Following the Design Science Research (DSR) methodology, this paper aims at generating a) prescriptive knowledge about how artifacts can be developed and applied to solve existing problems and b) descriptive knowledge, extracted from the applied evaluation [SV12]. Setting up on the build-evaluate pattern, the development and evaluation of artifacts are the core aspects of most DSR Projects. As shown in fig. 1, extended the build-evaluate pattern, including four evaluation phases (ex ante and ex post the construction). Even though the solution design has been part of the development process, the focus of the paper is the construction and its ex post evaluation.

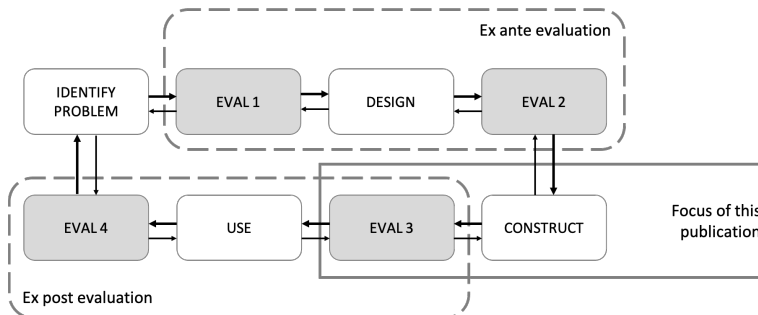


Fig. 1: DSR evaluation activities based on [SV12]

Due to the development of the described artifacts as domain specific implementations of GAIA-X, this paper sets up on defined and evaluated requirements, as further described in chapter 4.1. Using these requirements, three main components have been conceptually designed, a service self-description, a software architecture for the service registry itself and an improved type of visualization for the description of inter-service relations and data flows. These components have been developed and evaluated in a cooperative and iterative manner with weekly feedback meetings. The results of these steps are described in chapters 4.2, 4.3 and 4.4. The ex post evaluation (EVAL 3) happened in an artificial setting. It was carried out with five technical and domain experts who were surveyed with semi-structured interviews according to [MN07]. All interviewed company representatives are experts in the sense of [BLM09] interviewing and have several years of professional experience. An interview

guide consisting of 13 open-ended questions was prepared for conducting the interviews. For the evaluation, a multistage procedure for qualitative content analysis was applied, following the approach of [GL09]. The evaluation concept consists of the phases: (1) determining the research questions and analysis perspective, (2) developing a category system, (3) searching the texts for relevant information and assigning them to appropriate categories, (4) aggregating the information across different interviews, (5) analyzing the results to answer RQ2. The category system is based on the elicited requirements and distinguishes between positive and negative feedback. The categories as well as the assignment of the key statements to these categories are presented in chapter 5. Finally, we finish the evaluation by answering RQ2 in chapter 6 (5).

4 Prototype Development

4.1 Requirement Elicitation

Today, innovations and digital business models rely heavily on data and data streams, making them an important asset for companies requiring protection. Therefore, it is crucial for companies to retain data sovereignty and control over their data flows. In line with the European data strategy⁶, GAIA-X pursues the goal of establishing such a networked data structure for a sovereign, European data ecosystem by breaking down data silos and helping users avoid data lock-ins. The service registry presented in this paper has been developed as a prototype of the smart living domain originating from GAIA-X. Its requirements are therefore derived from the GAIA-X essential characteristics and requirements formalized in [Ho20], which are also and especially relevant for the smart living data ecosystem with its many heterogeneous actors, interconnected service systems, and sensitive data. To derive the requirements, which are summarized in tab. 1, we analyzed the goals of GAIA-X (see above) and firstly selected the ones suitable for the present case. To validate them they were discussed with a group of partners from the ForeSight project, which are on one hand familiar with GAIA-X and on the other hand familiar with the domain smart living. Since our artifact is a prototype, most of the requirements are in line with the evaluation criteria proposed by [SV12] for instantiations, namely effectiveness and efficiency. However, since the service self description (c.f. chapter 4.3) can be understood as a model (generic description of a service, which can be instantiated), too, the evaluation criteria completeness, internal consistency and level of detail apply as well.

4.2 Proposed Architecture for the Service Registry

To develop a GAIA-X compliant service registry for the domain smart living, the mentioned requirements need to be considered in architectural and design specific decisions of the

⁶ https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en, accessed 19th of April 2021

ID	Requirement	Description
R1	European data protection	The developed artifact must enable implementation of the European General Data Protection Regulation (GDPR), as well as case-specific adaptability of data protection scenarios for diverse protection classes.
R2	Openness and transparency	Developed artifacts must create transparency to reduce transaction costs and increase data availability and exchange.
R3	Authenticity and trust	Artifacts should provide a self-description, for example on certified data protection as well as compliance with regulatory criteria, in order to strengthen the mutual trust of ecosystem actors.
R4	Digital sovereignty and self-determination	Developed artifacts should enable data sovereignty and control of data sovereignty: Each user decides for himself or herself, on the basis of his or her own data classification, where his or her own data is stored and by whom it may be processed and for what purpose.
R5	Free market access and European value creation	The use of verifiably secure, open technologies in an open ecosystem is intended to promote competitiveness, especially in international comparison. Services should be linkable to enable secure use of data and algorithms as well as the movement of data along the value chain to applications.
R6	Modularity and interoperability	The artifacts should enable interoperability both in terms of technical and semantic standards and in terms of interconnectivity at network, data and service level between edge or cloud instances. This interoperability is intended to avoid lock-in effects and prevent the emergence of data silos.
R7	User-friendliness	GAIA-X services should be clear and intuitive for all stakeholders.

Tab. 1: Requirements for the service registry derived from GAIA-X

service registry. In the current and the following two chapters these decisions and the resulting artifact are presented in more detail. Its core task is the collection and presentation of available services. The information required for this is largely determined by the requirements R1 to R5 and smart living specific requirements with regard to cyber-physical service compositions. In addition to the design of the content, R6 also stipulates a high degree of standardization of the format, which is reflected in the definition of a uniform self-description that need to be stored in the service registry in a form that is both machine-readable and intuitively comprehensible to humans. Due to this, the self-description was designed in a format based on existing semantic standards, as further described in chapter 4.3 and an intuitive user interface needed to be developed. To achieve this combination in a time efficient manner, Django has been chosen as web-framework to develop a prototypical application. Django takes over basic settings for web development and also provides functions such as full user authentication or a pre-installed admin area [Fo]. By default, Django uses an object relational mapper on top of a relational database to process the contained information using e.g. the model view template (MVT) pattern [HK09]. MVT is a slightly varied version of the model view controller pattern [LD85] which divides the view into two separate components called view and template, while most controller functionality is handled by the Django

framework itself [Fo]. To include the developed self description directly into the Django framework the ORM and the relational database have been exchanged by mongoDB [Gy15] as document-oriented database and a corresponding object document mapper⁷. To offer the mentioned REST-API, an extension of Django called Django REST⁸ has been used. The resulting structure of the service registry is shown in fig. 2.

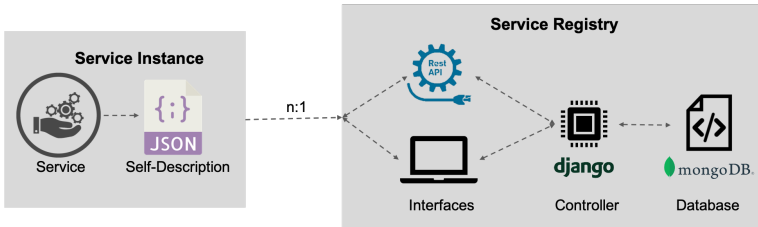


Fig. 2: General structure of the service registry

To integrate the prototype into the smart living application domain, the service registry has been dockerized and deployed in an open shift environment from an IT service provider of the housing industry. This facilitated further utilization and the possibility of a domain-specific evaluation, as described further in chapter 5.

4.3 Model Structure and Service Self Description

A key component of the model and thus a central feature of the registry concept is the self-description of each service. Each service registered in the registry is described by meta information using a Web of Things (WoT)-based description language. By describing services in a standardized format, the registry helps service engineers understand and orchestrate the components, thus promoting the creation of interoperable and transparent solutions within the data ecosystem. The registry supports the WoT standard classes (Properties, Actions and Events), as well as an extension of the Thing class with additional features of the respective service relevant for the data ecosystem. The extended attributes are divided into two subgroups: Attributes derived from GAIA-X requirements (GX) and smart living specific attributes (SL). Tab. 2 shows a detailed description of the extension made to the basic scheme.

4.4 View and Visualization of Services and Service Interdependencies

As described in the previous chapters, the core task of the service registry is the collection and presentation of services in a machine- and human-readable format. While the self descriptions and the REST-Interface are addressing the machine readability, in this chapter

⁷ <https://github.com/django-nonrel/mongodb-engine>.

⁸ <https://www.django-rest-framework.org/>.

Attribute	Description	GX	SL
Provider	Provider name of the respective service	x	x
Dependencies	Information about dependencies in the form of references to services and data sources used by the respective services	x	
Service ID	ID of the data supplying service		x
Data Type	Data type of the input data used by the respective service	x	x
Type of data use	Describes what the data is used for (e.g. training of an AI)		x
Smart readiness Indicator	Domain-specific assessment of the service with regard to its adaptability to the needs and preferences of individual users		x
Costs/billing unit	Contains information on possible costs arising from the use of the service and on the billing unit (e.g. per request)	x	x
Life Cycle stage	Contains information on the lifecycle status of the service in question (e.g. active, inactive, test)	x	
Location (physical)	Specifies the physical location (country) where the respective service is hosted	x	
Execution Node	References a particular GAIA-X node on which the service is hosted	x	
Security classification	Contains information about the security classification of the data provided and processed by the service in the form of a metric (0: none 6: very high)	x	
Processing of person-related data	Indicates whether personal data are processed that are subject to the GDPR	x	
Certification	List of certifications that the specific service has	x	

Tab. 2: Extension of the WoT standard by use case specific attributes

the human oriented visualization of information-based on the described requirements is presented. Even though the stored information is independent from the visualization and determined in the self-description, an intuitive visualization according to R7 can improve the usage of the service registry. According to this, the service registry is offering an overview of existing services and a detail view for each service. Due to the explicit relevance of execution location as well as provider and security information, these are already shown in the overview and part of the sort and search functionalities.

One aspect overarchingly relevant for requirements R1 to R5 is the possibility to understand necessary data flows and service dependencies of complex cyber-physical service compositions. Even though the required information regarding direct service relations in the presented self-description might be sufficient to understand service relations step by step, the representation of complex service relations keeps still quite hard. As one major contribution of this paper a dataflow representation has been developed to increase this transparency of service compositions for data consumer, data producer as well as intermediates. As shown in fig. 3 this representation can show the incorporated components in a cyber-physical system and help users to understand how their data will be used (R1 & R4) and help other actors to exchange services by for example more secure alternatives (R6).

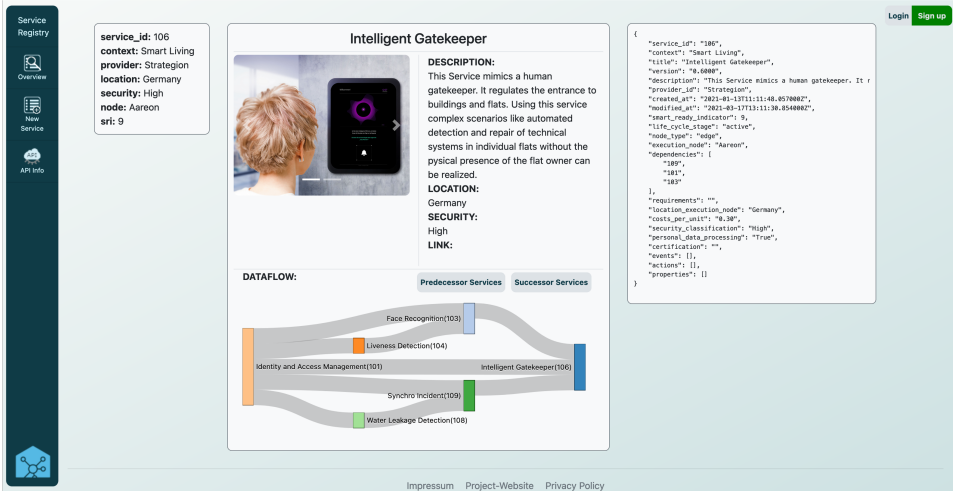


Fig. 3: Service registry GUI: Service details including a dataflow representation of complex service dependencies

Fig. 3 shows the use case of the intelligent gatekeeper outlined in chapter 1 for emergency access in the event of a water leak. Using this representation, possible users of the gatekeeper are able to identify all involved components, such as face recognition or liveness detection. Additionally, it can help service engineers to understand dependencies, if a service should be exchanged by another one or a complementary service should be developed, like in the case of the face recognition and its integration of a liveness detection service. Summarizing the implications of the presented architecture, the self-description as well as the view, we find that our service registry design, which focuses on the integration of interoperable and domain-specific standards, can answer RQ 1.

5 Evaluation

The evaluation of our prototype is part of the ex post evaluation according to [SV12] (c.f. chapter 3). It aims at verifying, that the previously described requirements (c.f. chapter 4.1) were met to judge the overall success of the implementation and to identify weaknesses and unexploited potentials for the next development iterations. We therefore transferred the requirements into a set of questions for the expert interviews to a) obtain an overall estimate per requirement, if the prototype is suitable, and b) to gain detailed feedback. Due to current travel restrictions the interviews were conducted via video conferences. If feasible the experts were granted access to our prototype and used it after the introduction of the interview. In case direct access was not possible the prototype was presented via screensharing and the interviewees could ask questions or request an in-depth presentation of certain features. To provide a structured overview, of the results we summarized them in

tab. 3. In addition to the requirement abbreviation (column one; the abbreviation referenced in chapter 4.1), the aggregated rating is listed on a scale from “fully met” (+ +) to “not met at all” (–) (column two) and specific feedback is mentioned in a condensed form (column three). The following section briefly describes the aspects rated by the experts.

Evaluation related to previously defined requirements		
Req.	Eval.	Requirement specific suggestions for extensions
R1	+	<ul style="list-style-type: none"> • More detailed description of the security level • Information on anonymization or pseudonymization of the data • Extension of the visualization function
R2	++	<ul style="list-style-type: none"> • More granular description of individual data sources of a service
R3	+	<ul style="list-style-type: none"> • Certification of the correctness and quality of services and data sources • Enabling the replacement of unwanted service components
R4	0	<ul style="list-style-type: none"> • Provide information about service calls • Clearer definition of the target group
R5	0	<ul style="list-style-type: none"> • Extension by detailed technical description of the services
R6	+	<ul style="list-style-type: none"> • Specification of rules, e.g. to avoid too high granularity of services
R7	+	<ul style="list-style-type: none"> • Integration of a development component for the orchestration of services
Further suggestions for expansion		
<ul style="list-style-type: none"> • Different domain registries should be interoperable • Consistent implementation of registries • Standardization of interfaces • Promote the interplay of different data ecosystems • Enabling clustering of services • Add information about services rollout • Extend service description and service classes (e.g. start service, processing service) • Increase data and information coverage • Support of established standards • Consideration of cross-domain requirements 		

Tab. 3: Results of the user evaluation

First of all, the requirement for digital sovereignty and self-determination (R4), which was rated by interviewees as neither “fully met” nor “not met at all”, should be named. This is primarily due to the different views on the extent to which the user is given options to actively influence the services listed in the service registry. On the one hand, it has been criticized that the prototype does not allow active interference with the services or their data flows. On the other hand, it has been stressed that the provision of information provides the opportunity to make an informed decision and thus increase sovereignty. The requirement of free market access and European value creation (R5) has also been evaluated neutrally, due to the differences between the two aspects merged into one requirement. A public service registry allows new or small players to easily enter the market (e.g. as in app stores), which was considered very positive. However, the extent to which this strengthens European value creation has been strongly questioned. Finally, modularity and interoperability (R6),

which, despite its generally positive assessment, has been interpreted differently. This mainly concerned the question of whether excessive granularity, as promoted by the service registry, has a positive or negative influence. On one hand, modularity tackles the increased understandability of the functionality of specific services as well as the increased ability to include already implemented functions. On the other hand, the need to modularize services can also foster the deployments of various subservices, decreasing the overview in a ecosystem based on the overload of information.

6 Discussion & Conclusion

This article presents the concept, development, and evaluation of a service registry for the domain smart living in the context of GAIA-X. The latter was achieved by extending a common standard to include domain-specific aspects (c.f. RQ2). The evaluation was used to verify functionality regarding requirements and suitability in supporting a service engineer (RQ2). This is also the key distinction to be made from existing systems, which offer broadly similar functionality. With developers as a target group and the GAIA-X smart living domain as a framework, the tool offers added value in developing innovative services consisting of products and services that have not yet been linked.

This has been confirmed by largely positive feedback on usefulness and usability during the, mainly technically oriented, evaluation. Some possible methodical limitations regarding the assumptions of the user centric elicitation of GAIA-X requirements including an ex ante evaluation and a quite small number of experts can be pointed out. However, regardless of this, extensive qualitative feedback was collected, which can form the basis for further development and evaluation steps. The feedback of the interviewed experts includes the possibility of directly influencing data flows (as opposed to purely destructive presentation) and the integration with other registrations within the meaning of GAIA-X funding (e.g. from other domains). Furthermore, specific instances of services and concrete physical devices should be integrated as instantiation of abstract services offered from providers. Next to this, currently all dependencies visualized in the service registry are based on the aggregated data of the service providers. To increase the trust and in order to prevent abuse, a neutral instance can be included, extending and certifying the identified dependencies.

In the context of this work, the registry was developed according to the specific requirements of the smart living data ecosystem. However, there is a need for transparency in many domains, therefore we want to encourage the application of the registry concept to other domains as well. Even more, the unwillingness to share data and services is a major obstacle to the establishment of data ecosystems and for the implementation of data-based business models [Ko21]. The service registry offers a way to increase this willingness by reducing uncertainty and creating transparency for all actors of the data ecosystem. It thus contributes to economic sustainability by reducing transaction costs.

Finally, the focus of the presented service registry has been the presentation of inter-service relations, neglecting the design and data processing of the system itself. To increase the transparency of services containing machine learning-based components, the interpretability

of the involved systems need to be considered in more detail to disclose the decision making process and prevent security threads or possibly discriminatory behavior. The latter aspect is of crucial importance in terms of social sustainability and is currently receiving a lot of attention in the scientific discourse on AI.

7 Acknowledgement

This research was conducted in the scope of the research project ForeSight (<https://foresight-plattform.de>) which is funded by the German Federal Ministry for Economic Affairs and Energy (BMWi). We would like to thank them for their support.

References

- [AS11] Araujo, L.; Spring, M.: Complex performance, process modularity and the spatial configuration of production. In (Caldwell, N.; Howard, M., eds.): *Procuring Complex Performance: Studies in Innovation in Product-Service Management*. Routledge, pp. 78–96, 2011, ISBN: 9780415800051.
- [Ba11] Bask, A.; Lipponen, M.; Rajahonka, M.; Tinnilä, M.: *Framework for modularity and customization: Service perspective*, 2011.
- [Ba14] Bauer, J.; Ketschau, A.; Michl, M.; Bürner, J.; Franke, J.: Die intelligente Wohnung als Baustein im Internet der Dinge. In: *Erste Transdisziplinäre Konferenz zum Thema Technische Unterstützungssysteme, die die Menschen wirklich wollen.*, Hamburg. Pp. 298–307, 2014.
- [Ba19] Bauer, J.; Hoffmann, H.; Feld, T.; Runge, M.; Hinz, O.; Mayr, A.; Förster, K.; Teske, F.; Schäfer, F.; Konrad, C.; Franke, J.: *ForeSight - Platform Approach for Enabling AI-based Services for Smart Living*. In (Pagán, J.; Mokhtari, M.; Aloulou, H.; Abdulrazak, B.; Cabrera, M. F., eds.): *How AI Impacts Urban Living and Public Health*. Springer International Publishing, Cham, pp. 204–211, 2019, ISBN: 978-3-030-32785-9.
- [BHL01] Berners-Lee, T.; Hender, J.; Lassila, O. The semantic web *Scientific american*, 284/5. pp. 34–43, 2001
- [BLM09] Bogner, A.; Littig, B.; Menz, W.: *Interviewing experts*. Springer, 2009.
- [BLW18] Beverungen, D.; Lüttenberg, H.; Wolf, V. Recombinant service systems engineering *Business & Information Systems Engineering*, 60/5. pp. 377–391, 2018
- [BT13] Boehm, M.; Thomas, O. Looking beyond the rim of one’s teacup: a multidisciplinary literature review of Product-Service Systems in Information Systems, Business Management, and Engineering & Design *Journal of Cleaner Production*, 51pp. 245–260, 2013

- [Ca19] Cappiello, C.; Gal, A.; Jarke, M.; Rehof, J. Data ecosystems—sovereign data exchange among organizations Dagstuhl Rep, 9/9. pp. 66–134, 2019
- [Fo] Foundation, D. S.: Django, <https://djangoproject.com>, accessed: 2021-04-23.
- [Ge20] Gembarski, P. C.: The Meaning of Solution Space Modelling and Knowledge-Based Product Configurators for Smart Service Systems. In (Borzemski, L.; Świątek, J.; Wilimowska, Z., eds.): Information Systems Architecture and Technology: Proceedings of 40th Anniversary International Conference on Information Systems Architecture and Technology – ISAT 2019. Springer International Publishing, Cham, pp. 28–37, 2020, ISBN: 978-3-030-30440-9.
- [GGO21] Gelhaar, J.; Groß, T.; Otto, B.: A Taxonomy for Data Ecosystems. In: Proceedings of the 54th Hawaii International Conference on System Sciences. P. 6113, 2021.
- [GL09] Gläser, J.; Laudel, G.: Experteninterviews und qualitative Inhaltsanalyse: als Instrumente rekonstruierender Untersuchungen. Springer-Verlag, 2009.
- [Gy15] Győrödi, C.; Győrödi, R.; Pecherle, G.; Olah, A.: A comparative study: MongoDB vs. MySQL. In: 2015 13th International Conference on Engineering of Modern Electric Systems. IEEE, pp. 1–6, 2015.
- [HK09] Holovaty, A.; Kaplan-Moss, J.: The definitive guide to Django: Web development done right. Apress, 2009.
- [Ho20] Hoppe, G.; Kraemer, P.; Biegel, F.; Boll, S.: GAIA-X: A Pitch Towards Europe - Status Report on User Ecosystems and Requirements, tech. rep. DOE-SLC-6903-1, 11019 Berlin: Federal Ministry for Economic Affairs and Energy (BMWi), 2020.
- [Ja19] Janßen, C.: Towards a system for data transparency to support data subjects. In: International Conference on Business Information Systems. Springer, pp. 613–624, 2019.
- [JVD07] Jacobs, M.; Vickery, S. K.; Droge, C. The effects of product modularity on competitive performance: Do integration strategies mediate the relationship? International Journal of Operations & Production Management, 27/10. pp. 1046–1068, 2007
- [Ka19] Kammler, F.; Hagen, S.; Brinker, J.; Thomas, O.: Leveraging the value of data-driven service systems in manufacturing: a graph-based approach. In: Proceedings of the 27th European Conference on Information Systems, Stockholm & Uppsala, Sweden. Vol. 27, 2019.
- [Ko20] Kortum, H.; Gravemeier, L. S.; Zarvic, N.; Feld, T.; Thomas, O.: Engineering of Data-Driven Service Systems for Smart Living: Application and Challenges. In: IFIP International Conference on Advances in Production Management Systems. Springer, pp. 291–298, 2020.

- [Ko21] Kortum, H.; Rebstadt, J.; Gravemeier, L. S.; Thomas, O. Data-based Customer-Retention-as-a-Service: Induktive Entwicklung eines datenbasierten Geschäftsmodells auf Basis einer Fallstudie der Automobilbranche HMD Praxis der Wirtschaftsinformatik, 58/3. pp. 1–15, 2021
- [Ku04] Kuck, G. Tim Berners-Lee's Semantic Web SA Journal of Information Management, 6/1., 2004
- [LD85] London, R.; Duisberg, R. Animating programs using Smalltalk IEEE Annals of the History of Computing, 18/08. pp. 61–71, 1985
- [MH18] Marrone, M.; Hammerle, M. Smart cities: A review and analysis of stakeholders' literature Business & Information Systems Engineering, 60/3. pp. 197–213, 2018
- [MN07] Myers, M. D.; Newman, M. The qualitative interview in IS research: Examining the craft Information and Organization, 17/1. pp. 2–26, 2007
- [ND15] Nikayin, F.; De Reuver, M. What motivates small businesses for collective action in smart living industry? Journal of Small Business and Enterprise Development, 22/2., 2015
- [OL18] Oliveira, M. I. S.; Lóscio, B. F.: What is a data ecosystem? In: Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age. Pp. 1–9, 2018.
- [Ot21] Otto, B.; Rubina, A. R.; Eitel, A.; Teuscher, A.; Schleimer, A. M.; Lange, C.; Stingl, D.; Loukipoudis, E.; Brost, G.; Böge, G.: GAIA-X and IDS, tech. rep., International Data Spaces Association, 2021.
- [SD20] Sovacool, B. K.; Del Rio, D. D. F. Smart home technologies in Europe: a critical review of concepts, benefits, risks and policies Renewable and sustainable energy reviews, 120p. 109663, 2020
- [SK20] Schoormann, T.; Kutzner, K.: Towards Understanding Social Sustainability: An Information Systems Research-Perspective. In: Proceedings of the International Conference on Information Systems, Munich, Germany. Vol. 51, 2020.
- [SV12] Sonnenberg, C.; Vom Brocke, J.: Evaluations in the science of the artificial—reconsidering the build-evaluate pattern in design science research. In: International Conference on Design Science Research in Information Systems. Springer, pp. 381–397, 2012.
- [SW16] Szilagy, I.; Wira, P.: Ontologies and Semantic Web for the Internet of Things - a survey. In: IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society. Pp. 6949–6954, 2016.
- [ZGC11] Zeng, D.; Guo, S.; Cheng, Z. The web of things: A survey JCM, 6/6. pp. 424–438, 2011

Workshop: IT-Governance und
Strategisches
Informationsmanagement
(ITG-SIM)

Auswahl einer IoT-Plattform für die Integration von operativen Anlagendaten an deutschen Flughäfen: Vorgehen und Lessons-Learned

Michael Freiberger¹, Sebastian Floerecke², Christoph Ertl³, Alexander Herzfeldt⁴

Abstract: Internet-of-Things (IoT)-Technologien eröffnen Flughäfen neue Möglichkeiten, Zustände verschiedener Anlagen am Standort zu erfassen, zu verarbeiten und damit zu steuern. Diese Technologien fungieren als Enabler neuartiger Geschäftsmodelle. Zur Realisierung solcher Geschäftsmodelle muss ein geeignetes IoT-Gesamtsystem – Sensorik, Übertragungstechnik und zentrale IoT-Plattform – geschaffen werden. Der Flughafen München, Analyseobjekt dieser Praxisstudie, beabsichtigt, eine IoT-Plattform zu beschaffen. Nachdem gegenwärtig über 600 Anbieter weltweit miteinander konkurrieren, deren Angebote jeweils unterschiedliche Funktionalitäten und Qualitätseigenschaften umfassen, handelt es sich dabei um ein komplexes Vorhaben. Die Forschung liefert bislang kein ausreichend dokumentiertes Vorgehensmodell für Unternehmen, um die für sie unter Berücksichtigung funktionaler und nicht-funktionaler Anforderungen geeignetste IoT-Plattform auszuwählen. Um diese Lücke zu adressieren, wurden in dieser Praxisstudie verschiedene Methoden miteinander kombiniert, um dadurch einen wissenschaftlich fundierten und praxisorientierten Ansatz zu entwickeln. Damit wurden vorselektierte IoT-Plattformen vergleichend bewertet. Unternehmensentscheider erhalten einen Leitfaden und einen initialen Katalog an Bewertungskriterien zur Auswahl einer IoT-Plattform für ihre unternehmensspezifischen Anwendungsfälle und daraus resultierenden Anforderungen.

Keywords: Internet-of-Things (IoT), IoT-Plattform, Tool-Vergleich, Requirements-Engineering, Bewertungsmatrix, Nutzwertanalyse, Experteninterviews, Fallstudie

1 Ausgangssituation und Zielsetzung

Die Flughafen München GmbH (FMG) betreibt mit ihren 22 Tochter- und Beteiligungsgesellschaften den Münchner Flughafen. Das bayerische Luftverkehrsdrehkreuz ist Deutschlands zweitgrößter Flughafen und Europas einziger Fünf-Sterne-Airport. Als sogenannter Full-Service-Operator bietet die FMG auf einer Fläche von 1.575 Hektar neben den Leistungen des Airport-Managements auch Geschäfte, Restaurants, Bürogebäude, Parkplätze und weitere Services an. Um auch zukünftig handlungsfähig und wirtschaftlich erfolgreich zu bleiben, ist die FMG gegenwärtig dabei, den Flughafencampus schrittweise als Smart-City zu realisieren. Dies erfordert unter anderem, technische Anlagen und Geräte derart zu vernetzen, dass dadurch ein positiver ökonomischer und ökologischer Effekt entsteht.

¹ Flughafen München, Servicebereich IT, Nordallee 25, 85356 Freising, michael.freiberger@munich-airport.de

² Universität Passau, Lehrstuhl für Wirtschaftsinformatik mit Schwerpunkt Informations- und IT-Service-Management, Innstraße 43, 94032 Passau, sebastian.floerecke@uni-passau.de

³ Flughafen München, Servicebereich TE, Nordallee 25, 85356 Freising, christoph.ertl@munich-airport.de

⁴ Technische Universität München, Fakultät für Informatik, Boltzmannstraße 3, 85748 Garching bei München, aherzfeldt@gmx.de

Während die Umsetzung für gewachsene Städte eine hochkomplexe Aufgabe darstellt [PF15], hat ein Flughafenbetreiber einen entscheidenden Vorteil: Als Eigentümer von Grundstücken, Gebäuden, Infrastruktur, Energieversorgung und vielem mehr kann er die Lösungen aus einer Hand planen und gestalten. Zudem ist die Fläche begrenzt. Voraussetzung der Strategie „Smart-City“ ist, Eigenschaften und Zustände verschiedener zentraler Anlagen und Geräte am Standort zu erfassen, zu verarbeiten, zu visualisieren und Anlagen so zu steuern. Dafür werden am Flughafen München bestehende Sensoren und Datenquellen, wie zum Beispiel die Gebäudeautomation, mit zusätzlichen Sensoren an noch nicht digital erfassten Objekten (Beispiel: Wasserzähler) verbunden. Ein spezifisches IoT-Funknetzwerk wurde bereits aufgebaut. Bislang umgesetzte Anwendungsfälle sind unter anderem die zustandsbasierte Optimierung des Lüftungsanlagenbetriebs, das Auslesen von Wasser- und Stromzählern und die Überwachung der Luftqualität in Besprechungsräumen.

Eine zentrale IoT-Plattform, die Daten sammelt, verarbeitet, anzeigt und auswertet und neben der Sensorik und der Datenübertragung den dritten integrale Bestandteil von IoT darstellt [PH14], wurde noch nicht realisiert. Die komplette IoT-Kette soll zukünftig sowohl FMG-internen (Fachabteilungen) als auch externen Kunden wie Bodenabfertiger, Reinigungsfirmen und Behörden bereitgestellt werden. Die FMG möchte insgesamt die Effizienz ihrer bestehenden Geschäftsprozesse erhöhen und zugleich ihren Kunden weitere Services anbieten. Die IoT-Services sollen den Service-Katalog im Bereich der digitalen Dienste erweitern und zusätzliche Umsätze generieren. Eine IoT-Plattform ist dafür unabdingbar. Ansonsten könnten Daten zwar erfasst werden, aber den Kunden nur als Rohdaten (Datenbank, Dateien oder Streams) zur Verfügung gestellt werden. Erfahrungsgemäß möchten Kunden allerdings meist keine zusätzliche Software erwerben, um auf Basis der generierten Daten einfache Graphen darzustellen oder Benachrichtigungen über Ereignisse zu erhalten. Deshalb soll hier eine IoT-Plattform zentral beschafft und mit den Daten der Kunden angereichert werden. Kunden sollen zukünftig mit einer seitens der FMG bereitgestellten Cloud-Anwendung ihre Daten einsehen und beliebig weiterverarbeiten können.

Die FMG steht somit vor der Herausforderung, eine geeignete IoT-Plattform auszuwählen. Da derzeit über 600 Anbieter miteinander konkurrieren, deren IoT-Plattformen jeweils unterschiedliche Funktionalitäten und Qualitätseigenschaften umfassen und damit für beliebige Anwendungsszenarien nicht gleichermaßen geeignet sind sowie aufwendig in die Unternehmenslandschaft eingebunden werden müssen [LP19], handelt es sich dabei um einen komplexen Entscheidungsprozess. In Ergänzung zum bereits in der Literatur vorhandenen Modell von Lempert und Pflaum [LP19] sind neben den funktionalen Anforderungen ebenfalls nicht-funktionale Anforderungen zu berücksichtigen, um eine IoT-Plattform auf strategischer Entscheidungsebene auszuwählen [SH20, Ti06].

Um eine Auswahlentscheidung zu treffen, ergeben sich aus Sicht der Autoren nachfolgende Fragen (Tabelle 1), die innerhalb dieses Praxisbeitrags beantwortet werden. Die Fragen lassen sich den beiden Phasen Analyse und Bewertung zuordnen.

Phase	Frage
Analyse	Wer sind die Stakeholder einer IoT-Plattform?
	Welche Anforderungen hat die FMG an eine IoT-Plattform?
	Welche Auswahlkriterien können von den Anforderungen abgeleitet werden?
	Was sind die gegenwärtig führenden IoT-Plattformen auf dem Markt?
Bewertung	Inwieweit werden die Kriterien seitens der einzelnen IoT-Plattformen erfüllt?
	Welche IoT-Plattform ist gemäß der Bewertung am besten geeignet?

Tab. 1: Leitende Fragestellungen an die Auswahl einer IoT-Plattform

Der Beitrag ist wie folgt aufgebaut: Kapitel 2 befasst sich mit den Grundlagen von IoT. Kapitel 3 beschreibt das Forschungsdesign. Danach wird für die Analyse ein dreistufiger Prozess, bestehend aus Stakeholderanalyse, Use-Case-Auswahl und Anforderungsermittlung, angewandt (Kapitel 4). Die Kapitel 5 und 6 beinhalten die Bewertung vorausgewählter IoT-Plattformen hinsichtlich ihrer Abdeckung der ermittelten Anforderungen beziehungsweise die Realisierung der Sieger-Plattformen via Proof-of-Concepts (PoCs). Am Ende werden die der Bewertung und der PoCs folgenden Schritte und abgeleitete Lessons-Learned vorgestellt sowie Limitationen und weiterer Forschungsbedarf thematisiert.

2 Grundlagen von Internet-of-Things

Internet-of-Things (IoT) beschreibt die Vernetzung physischer Gegenstände aller Art, die mit Sensoren, Aktuatoren und Rechenleistung ausgestattet sind [PH14]. Dabei werden Daten und Meta-Daten von Sensoren gesammelt, verknüpft und verarbeitet. Daraufhin werden diese Daten visualisiert, Alarme oder Benachrichtigungen versandt oder Aktionen von Aktoren ausgeführt [OR18]. Im privaten Umfeld (Smart-Home) liegt der Nutzen von IoT hauptsächlich in der Einsparung von Ressourcen, zum Beispiel im Bereich der Wärmeversorgung, und in der Erhöhung der Lebensqualität, etwa durch die Hausüberwachung [AZ17]. Im industriellen Kontext (Industrie 4.0) sind die Ziele anders gelagert. Hier können durch einen höheren Automatisierungsgrad menschliche Ressourcen von operativen zu strategischen Aufgaben verlagert und durch präzisere Vorhersagen Anlagenausfälle reduziert werden. Darüber hinaus lassen sich neue Geschäftsfelder erschließen und das Kundenerlebnis verbessern [LF14].

Ein IoT-System setzt sich aus den drei Bestandteilen Sensorik, Datenübertragungseinheit sowie Persistenz- und Auswertungseinheit (IoT-Plattform) zusammen [ATR16]. Die IoT-Plattform ist der zentrale Punkt, an dem sämtliche IoT-Geräte angebunden werden. Nach Guth et al. [GB16] lassen sich die von einer IoT-Plattform angebotenen Dienste in vier Kategorien einordnen: (1) Daten zentral speichern, (2) Daten grafisch aufbereiten, (3) Daten auswerten und (4) Geräte verwalten. Eine genauere Betrachtung möglicher Funktionen und Eigenschaften wird im Verlauf dieses Beitrags vorgenommen.

3 Forschungsdesign

In der wissenschaftlichen Literatur existiert kein etabliertes Vorgehensmodell für die Auswahl einer IoT-Plattform entsprechend den Anforderungen des strategischen IT-Managements [LP19]. Deshalb werden in vorliegender Praxisstudie verschiedene Methoden miteinander kombiniert, um hierfür einen wissenschaftlich fundierten, praxisorientierten Ansatz zu entwickeln. Die angewandte Forschungsmethodik entspricht einer Einzelfallstudie [Y118]. Hierbei wurde sich an den Richtlinien des gestaltungsorientierten Forschungsparadigmas der Wirtschaftsinformatik [HM04] orientiert. Das Vorgehen lässt sich in die beiden Phasen Analyse und Bewertung unterteilen:

Analyse: Zur Identifikation funktionaler und nicht-funktionaler Anforderungen wurde die Methodik des Requirements-Engineerings angewandt. Requirements-Engineering basiert auf einem iterativen, inkrementellen Prozess und bezeichnet ein systematisches Vorgehen zur Spezifikation und zum Management von Anforderungen [PR21]. Für die vorherige Ermittlung und Analyse von relevanten Stakeholdern (Stakeholderanalyse) im IoT-Auswahlprozess wurden die Fachabteilungen der FMG gemäß den von Bryson [Br04] beschriebenen Techniken untersucht. Dafür wurden drei Kriterien festgelegt: (1) mögliche Beteiligung per organisatorischer Verantwortungsordnung, (2) unternehmensinterne Relevanz bei der Ideenentwicklung für strategische Innovationen sowie (3) inhaltliche Unterstützung bei der Problemformulierung und der Lösungssuche.

Bewertung: Die anschließende Bewertung vorausgewählter IoT-Plattformen beruht auf einem nutzwertbasierten Scoring-Modell (Nutzwertanalyse) [Za14]: Um eine Bewertung durchführen zu können, wurde eine Bewertungsmatrix erstellt. Die dafür notwendigen Bewertungskriterien wurden aus den zuvor ermittelten Anforderungen abgeleitet. Dabei wurden Prioritäten festgelegt: Jede Anforderung erhielt eine Gewichtung zwischen 1 und 10. Schließlich wurden die Kriterien (deren entsprechende Gewichtung) mit den Erfüllungsgraden der IoT-Plattformen (siehe Kapitel 5.2) multipliziert und jeweils aufsummiert. Die Erfüllungsgrade wurden zum einen auf Basis von Experteninterviews mit Herstellervertretern beziehungsweise im Internet verfügbaren Herstellerinformationen und zum anderen aus den Erfahrungen FMG-interner Experten und Anwendern im Rahmen eines Tagesworkshops bestimmt. Die Rangfolge der IoT-Plattformen ergibt sich aus der jeweils erzielten Gesamtpunktzahl.

4 Analyse

Die Analyse bestand aus einem dreistufigen Prozess: Stakeholderanalyse, Use-Case-Auswahl und Anforderungsermittlung. Dieser Prozess wird nachfolgend beschrieben.

4.1 Stakeholderanalyse

Stakeholder eines Systems ist eine Person oder Organisation, die direkt oder indirekt Einfluss auf dessen Anforderungen hat [PR21]. Zur Identifikation und Analyse von Stakeholdern (Stakeholderanalyse) wurden die Fachabteilungen der FMG, wie in Kapitel 3 beschrieben, evaluiert. Tabelle 2 fasst die als relevant ermittelten Fachabteilungen (FA) und internen Gremien (IG) zusammen:

Stakeholder	Kurzbeschreibung
Betreiber der IT-Infrastruktur (FA1)	FA1 plant IT-Infrastruktur wie LAN, WLAN, Firewall und Netzwerkschränke und ist für das Thema IoT zuständig. Sie wird die IoT-Infrastruktur betreiben, neue Kunden anbinden und bei Bedarf Sensoren für Kunden zur Verfügung stellen.
Bereich für Konzernentwicklung / Digitalisierung (FA2)	FA2 ist daran interessiert, die Technologie IoT am Flughafen zu etablieren, um damit die Digitalisierung zu fördern.
Fachbereich Technik (FA3)	FA3 steht als möglicher Großkunde repräsentativ für alle Endkunden einer IoT-Lösung am Flughafen München.
Bereich Commercial (FA4)	FA4 betreibt die Parkflächen am Flughafen München und ist ein weiterer möglicher Kunde einer IoT-Lösung.
Demand-Management (FA5)	FA5 stellt als Verkaufsorgan der IT die Verbindung zwischen IT und Kunden dar.
Interner Wartungsdienstleister für die Sensoren (FA6)	FA6 ist für die Durchführung der Hardwarewartung an verbauten Sensoren zuständig.
IT-Sicherheit (FA7)	FA7 unterstützt bei der Klärung von Anforderungen an die Sicherheit der IT-Systeme im Allgemeinen.
Cloud-Board (IG1)	IG1 bewertet jedes Produkt mit Cloud-Anteilen und gibt eine Handlungsempfehlung für oder gegen dessen Verwendung ab.
Betriebsrat (IG2)	IG2 muss jede neu eingeführte Software im EDV-Ausschuss genehmigen, bevor sie eingesetzt werden darf. Dabei spielt die Wahrung der Mitarbeiterrechte eine zentrale Rolle.

Tab 2: Am Flughafen München als relevant ermittelte Stakeholder

4.2 Use-Case-Auswahl

Die Anforderungen an eine IoT-Plattform hängen im hohen Maße von den umzusetzenden, unternehmensspezifischen Use-Cases ab [GK05]. Deshalb wurden zu Beginn dieser Praxisstudie mehrere Use-Cases festgelegt, die exemplarisch für die Bewertung, Auswahl und prototypische Realisierung dienen sollen. Deren Auswahl orientierte sich zum einen am Reifegrad der vorhandenen Sensoren (Verfügbarkeit am Markt), zum anderen am wirtschaftlichen Potenzial (Amortisationsdauer < fünf Jahre). Die Use-Cases umfassen Pre-

Conditioned-Air (PCA)-Anlagen¹, Parkhauskassenautomaten, Klimaanlage im Terminal und Temperaturüberwachung in IT-Netzwerkschränken im Außenbereich.

4.3 Anforderungsermittlung

Um einen Überblick über die Anforderungen der beteiligten Fachabteilungen zu erhalten, wurde ein Tagesworkshop abgehalten. Neben FMG-Mitarbeitern wurden externe, wissenschaftliche Projektpartner hinzugezogen. Im Brainstorming-Verfahren wurden Anforderungen an verschiedene Teile einer IoT-Plattform zusammengetragen und dokumentiert. Die erarbeiteten Anforderungen wurden in die Kategorien Schnittstellen und Datenquellen, Anwendungen und Services, Analyse, Integration und Sensorik sowie Kommunikation strukturiert. Im Anschluss wurde jeder Aspekt ausführlich diskutiert und bei Bedarf verfeinert.

Insgesamt konnten dadurch 17 funktionale und 6 nicht-funktionale Anforderungen an eine IoT-Plattform ermittelt werden. Spezifiziert wurden jeweils die anfordernden Stakeholder, eine Beschreibung sowie eine Begründung. Zusätzlich wurden die Anforderungen nach ihrer Priorität in notwendig respektive optional unterschieden (siehe Muster in Tabelle 3).

Anforderer	Beschreibung	Begründung	Priorität
Stakeholdername und -rolle	Kurze Beschreibung und Erläuterung der Anforderung	Nachweis des Erfordernisses der Anforderung	Priorisierung nach „notwendig“ oder „optional“

Tab. 3: Matrix zur Erfassung und Beschreibung der Anforderungen an eine IoT-Plattform

5 Plattformauswahl

Die Plattformauswahl erfolgte in drei iterativen Schritten: Zunächst wurde eine Bewertungsmatrix entwickelt. Anschließend wurden grundsätzlich geeignete IoT-Plattformen identifiziert. Diese wurden dann anhand der Bewertungsmatrix bewertet. Der Prozess der Plattformauswahl wird nachfolgend beschrieben.

5.1 Entwicklung einer Bewertungsmatrix

Für die Bewertung wurde eine Matrix erstellt (Kapitel 5.3, Abbildung 4). Die Bewertungskriterien wurden aus den ermittelten Anforderungen abgeleitet. Zudem wurden Prioritäten definiert: Jede Anforderung wurde mit einer Gewichtung zwischen 1 und 10 versehen. Für optionale Anforderungen wurden Gewichte zwischen 1 und 5, für notwendige Anforder-

¹ PCA-Anlagen versorgen die Flugzeugkabine auf der Parkposition mit vorklimatisierter Frischluft. Dadurch müssen parkende Flugzeuge nicht mehr ihre Hilfsturbinen für die Klimatisierung einsetzen. Die PCA-Anlagen sind im Betrieb deutlich leiser und senken den Kerosinverbrauch sowie den damit verbundenen CO₂-Ausstoß.

rungen zwischen 6 und 10 gewählt. In die Gewichtung floss auch ein, welche Folgen eine Nichterfüllung nach sich ziehen würde. Die Gewichte wurden in Abstimmung mit der für die IoT-Plattform zuständigen Abteilung (FA1) festgelegt.

5.2 Ermittlung grundsätzlich geeigneter IoT-Plattformen

Zur Identifizierung grundsätzlich geeigneter IoT-Plattformen wurden die kommerziellen Plattformen von IoT-Analytics – einem Anbieter für Marktanalysen im IoT-Bereich – auf Basis von Kundenbewertungen analysiert. Open-Source-Varianten wurden unter anderem auf Geckflare, einem Herausgeber von Technologie-Artikeln, verglichen. Insgesamt gibt es nach IoT-Analytics [Io18] auf dem Weltmarkt etwa 620 IoT-Plattformen (Stand: 2018). Angesichts dieser hohen Anzahl ist es nicht möglich, alle im Detail zu betrachten. Im renommierten Gartner-Magic-Quadrant für IoT-Plattformen [Ga20] werden nur IoT-Plattformen einbezogen, die originär als solche entwickelt wurden. Die Angebote der Hyper-scaler wie Amazon Web Services, Google und Microsoft sind deshalb nicht darin enthalten. Sie bieten IoT-Connecting-Services im Rahmen ihrer Cloud-Plattform an, auf deren Basis Nutzer eine IoT-Plattform nach Maß entwickeln können.

Aufgrund der Marktgröße und der einhergehenden Intransparenz wurde sich bei der Vorauswahl an der Studie von Lempert und Pflaum [LP19] orientiert. Die Autoren trafen in ihrem Beitrag eine Selektion nach den Kriterien Jahresumsatz, Platzierung in Rankings von Beratungs- und Marktforschungsunternehmen sowie Zitierungshäufigkeit in wissenschaftlichen und nicht-wissenschaftlichen Publikationen. Deren Liste wurde im Fall der FMG um Siemens Mindsphere (von der FMG bereits getestet) und ThingsBoard (führende Open-Source-Lösung) erweitert. Somit wurden folgende sieben IoT-Plattformen in die Bewertung einbezogen: (1) PTC ThingWorx, (2) Microsoft Azure IoT, (3) IBM Watson IoT, (4) Amazon AWS IoT, (5) Google Cloud IoT, (6) ThingsBoard sowie (7) Siemens Mindsphere. Im weiteren Verlauf dieses Beitrags werden die betrachteten IoT-Plattformen mittels Buchstaben anonymisiert.

5.3 Bewertung der IoT-Plattformen

Um einen fundierten Einblick in die vorausgewählten IoT-Plattformen zu erhalten, wurden die in Fragen überführten Anforderungen aus Kapitel 3.3 (siehe Anhang) fünf Anbietern zur schriftlichen Beantwortung per E-Mail zugesandt. Die Unternehmen bevorzugten es, die Fragen im Rahmen eines etwa einstündigen Telefonats durch einen verantwortlichen Vertriebsmitarbeiter zu beantworten. Bei den Plattformen ThingsBoard und MindSphere wurde auf eine Kontaktaufnahme verzichtet, da eine umfassende Beurteilung basierend auf Prototypen aus anderen Projekten der FMG selbst vorgenommen werden konnte. Mit den Rückmeldungen der Unternehmen beziehungsweise der Selbsteinschätzung konnte die Bewertungsmatrix befüllt werden (Abbildung 4). Da nicht alle Fragen von jedem Unternehmen beantwortet wurden oder eine objektive Bewertung nicht möglich war, konnten nicht sämtliche Felder ausgefüllt werden. Der Gesamtpunktwert einer IoT-Plattform ist der Quotient aus der Summe der Produkte von Erfüllungsgrad und Gewicht aller bewer-

teten Kriterien sowie der Summe der Gewichte. Die IoT-Plattform mit dem höchsten Gesamtpunktwert eignet sich am besten für die FMG.

		Gewichte	Plattformen							
			A	B	C	D	E	F	G	H
Funktionale Bewertungskriterien	Anlegen einer Rechtestruktur	9	9	9	10	9	9	4	4	8
	User-Authentifizierung	10	9	9	9	9	9	9	9	9
	Mandantenfähigkeit	8	5	5	9	8	9	8	9	9
	Informationssicherheit	9	8	8	8	8	8	9	7	7
	Gerätekonnektivität	9	10	9	9	9	9	8	9	9
	Node-RED-Anbindung	4	9	9	9	9	9	10	9	9
	Geräteverwaltung	7	8	7	8	5	8	5	4	7
	Automatisierte Konfiguration der Plattform	7		8	8	9	8	9	9	10
	Datentransformation	7	8	9	5	9	7	6	3	6
	Ereignisverwaltung	7	8		8		7	8	7	8
	Daten- und Speicher- verwaltung	8	9	8	6	8	8	8	10	10
	einfache Datenanalyse	9	8	8	8	8	8	8	7	8
	erweiterte Datenanalyse	4	9	9	9	9	9	9	7	7
	Datenvisualisierung	8	8	7	7	6	7	5	9	9
	Datenmodellierung	7	8	6	6	6	6	8	7	8
	Anwendungsentwicklung	6	7	9	2	9	9	8	3	3
Unternehmensintegration	7	8	6	8	6	8	5	3	3	
Nicht-funktionale Bewertungskriterien	Erwartete Lebensdauer	8	9	10	8	10	9	7	5	5
	Support	8	8	8		8		8	4	9
	Lizenzkosten	7		5	5	5	7	5	10	9
	Benutzerfreundlichkeit	9	8					4	8	9
	Accounting	8		8	9	9	9			7
	Betrieb, Verwaltung und Wartung	8						5	7	7
Summenprodukt			1183	1177	1137	1194	1215	1161	1140	1346
Summe Gewichtungen von bewerteten Kriterien		174	144	150	149	150	149	166	166	174
Normierte Bewertung			8,2	7,8	7,6	8,0	8,2	7,0	6,9	7,7

Abb. 4: Final befüllte Bewertungsmatrix

Aus der finalen Bewertungsmatrix (Abbildung 4) geht hervor, dass zwei IoT-Plattformen (A und E) denselben Gesamtwert (8,2) aufweisen. Die Vorteile der Plattform A liegen vor allem in der einfachen Erstellung von Dashboards und Analysen via Drag-and-Drop und der möglichen Anbindung von Industrieanlagen. Plattform E bietet hingegen den Funktionsumfang von Cloud-Services verbunden mit einer umfassenden IoT-Plattform. Beide Plattformen erhielten nicht für alle Kriterien eine Bewertung, da eine objektive Beurteilung mit den Informationen aus den Gesprächen und den öffentlich verfügbaren Materialien nicht gewährleistet werden konnte. Dies betraf vornehmlich nicht-funktionale Bewertungskriterien.

Grundsätzlich ist erkennbar, dass alle in dieser Praxisstudie betrachteten IoT-Plattformen insgesamt relativ gut abgeschnitten haben (Gesamtpunktwert $\geq 6,9$). Dieser Umstand lässt sich unter anderem darauf zurückführen, dass gezielt nur führende Systeme einbezogen wurden. Manche Kriterien, wie User-Authentifizierung und Gerätekonnektivität, wurden durchweg sehr gut abgedeckt. Bestimmte Kriterien scheinen inzwischen zu den Basisanforderungen von IoT-Plattformen zu zählen. Bei anderen Kriterien, wie Unternehmensintegration und Geräteverwaltung, ergab sich dagegen ein heterogenes Ergebnis. Letzteres unterstreicht die Notwendigkeit eines fundierten und systematischen Tool-Vergleichs.

6 Realisierung

Die Bewertung zeigt, dass die beiden am höchsten bewerteten IoT-Plattformen dieselbe Gesamtpunktzahl erreichen. Jedoch konnten bei beiden einzelne nicht-funktionale Anforderungen noch nicht abschließend geklärt werden. Dazu zählen unter anderem „Betrieb, Verwaltung, Wartung“ und „Benutzerfreundlichkeit“. Deshalb wurde mit den Anbietern der Plattformen A und E sowie der Plattform G, die bereits als Prototyp in einem anderen Projekt eingesetzt ist, ein PoC durchgeführt. Hierbei wurden jeweils drei Use-Cases (PCA-Anlagen, Parkhauskassenautomaten und Klimaanlage) umgesetzt.

Aber auch nach Durchführung der PoCs konnte keine eindeutige Entscheidung herbeigeführt werden. Für Use-Cases mit einer geringen Anzahl von Sensoren, wie bei den PCA-Anlagen, eignet sich Plattform G besser, da hier Kriterien wie „Anwendungsentwicklung“ eine untergeordnete Rolle spielen und diese mit geringen Lizenzkosten punkten kann. Bei komplexeren Use-Cases hingegen hat Plattform E ihre Vorteile. So ist beispielsweise von einem Kunden angedacht, mithilfe einer IoT-Plattform ein Dispositionssystem für die Wartung von Kassenautomaten aufzubauen. Dafür ist die Entwicklungsmöglichkeit einer Anwendung in der IoT-Plattform Grundvoraussetzung.

Während der Durchführung der PoCs wurde deutlich, dass nicht alle Anforderungen von Beginn an ersichtlich gewesen waren. Beispielsweise wurde die Notwendigkeit einer Faktura-Funktion zur vereinfachten Verrechnung von variablen Kosten an den Kunden erst später erkannt und nachträglich hinzugefügt.

7 Lessons-Learned, Ausblick, Limitation und Forschungsbedarf

IoT-Plattformen sind hochkomplexe Lösungen und unterscheiden sich erheblich in ihrem Funktionsumfang, ihren Qualitätseigenschaften und ihren Lizenzmodellen. Zudem sind die Anforderungen von Use-Case zu Use-Case recht heterogen. Diese Erkenntnisse nehmen Einfluss auf die Methodik der Bewertung und die finale Plattformauswahl. Gegenwärtig erfüllt keine der hier betrachteten IoT-Plattformen alle Anforderungen und Anwendungsszenarien vollumfänglich. Die wirtschaftliche Bedeutung von IoT wird in der Zukunft weiter zunehmen: Für 2023 erwartet IoT-Analytics [Io18] im Bereich der IoT-Plattformen einen globalen Umsatz von 23 Milliarden US-Dollar. Dieser Wert entspricht einem Anstieg um 93 Prozent im Vergleich zu 2021.

Folgende, wesentliche Lessons-Learned, lassen sich aus dieser Praxisstudie ableiten:

1. Der Abgleich mit der Unternehmensarchitektur in einer frühen Phase des Auswahlprozesses kann zur Reduzierung potenzieller Plattformen beitragen. Denn durch die Definition obligatorischer Systeme und deren erforderlicher Schnittstellen können ungeeignete IoT-Plattformen frühzeitig ausgeschlossen und dadurch der Auswahlprozess ressourceneffizienter und zielgerichteter gestaltet werden.
2. Ein wesentliches Kriterium bei der Entscheidung auf Managementebene stellt die Wirtschaftlichkeit und das Potenzial der Wirtschaftlichkeit im Zuge der Einführung einer IoT-Plattform dar. Deshalb müssen Unternehmen ein besonderes Augenmerk auf die Skalierungseigenschaften der IoT-Plattformen legen. Diese können maßgeblich dazu beitragen, die Stückkosten und dadurch die Gesamtwirtschaftlichkeit positiv zu beeinflussen.
3. Ein oder mehrere konkrete, unternehmensspezifische Use-Cases sind elementar für die Erstellung eines adäquaten Anforderungskatalogs. Ansonsten besteht die Gefahr, wichtige Anforderungen zu übersehen oder Anforderungen zu betrachten, die schlussendlich nicht von Relevanz sind. Oft müssen auch Anforderungen, wie etwa der Funktionsumfang und der Preis, gegeneinander abgewogen werden. Naturgemäß existieren zwischen Anforderungen diverse Zielkonflikte, die es zu lösen gilt.
4. Bei der vergleichenden Bewertung einer IoT-Plattform ist es sinnvoll, zentrale Kriterien als sogenannte K.O.-Kriterien (Mindestbedingungen) zu definieren: Sobald eine IoT-Plattform ein solches Kriterium nicht im ausreichenden Maße erfüllt, braucht sie nicht weiter betrachtet werden.
5. Mehrere Use-Cases erhöhen die Wirtschaftlichkeit einer IoT-Plattform: Viele Use-Cases im IoT-Bereich sind Optimierungen, die nur vergleichsweise geringe monetäre Vorteile erzielen. Werden diese geringen Beträge aus verschiedenen Use-Cases aufsummiert, ist es einfacher, eine IoT-Plattform ökonomisch zu betreiben, da die Fixkosten verteilt werden können.
6. Die Prüfung der Anforderungen innerhalb eines PoCs ist präziser als innerhalb von Telefoninterviews mit Herstellervertretern und Herstellerdatenblättern. Erst bei ei-

nem PoC lässt sich der exakte Funktionsumfang ermitteln. Es ist davon auszugehen, dass Anbieter versuchen, sich in einem positiveren Licht darzustellen. Neben der tatsächlichen Funktionstauglichkeit der IoT-Plattform können im Rahmen eines PoC zudem eine Vielzahl nicht direkt messbarer Eigenschaften, wie zum Beispiel Anwenderfreundlichkeit der Bedienoberfläche, evaluiert werden.

7. Die großen Cloud-Anbieter wie Amazon Web Services, Google und Microsoft vertreiben keine separate IoT-Plattform, sondern stellen mit einem Service IoT-Konnektivität her und verwenden für die Auswertung und Darstellung ihre sonstigen Services. Somit entsteht zusätzlicher Aufwand für die Erstellung einer IoT-Plattform aus den einzelnen Services. Jedoch ist es dadurch möglich, die IoT-Plattform den eigenen Anforderungen entsprechend zu gestalten. Unternehmen müssen sich darüber bewusst sein, dass dafür weitreichendes Know-how erforderlich ist, über das viele, insbesondere KMUs, nicht verfügen [FL19].
8. Für das gesamte Auswahlverfahren gilt grundsätzlich, sich nicht an einem traditionellen Wasserfallmodell, sondern an einem agilen, iterativen Vorgehen zu orientieren. Denn bei einer gänzlich neuen Technologie werden die Möglichkeiten und eigenen Ziele erst dann klar, wenn man testet. Dadurch erkennt man, was möglich ist und was nicht. Dieser Lernprozess lässt sich nur über ein agiles Vorgehen erreichen. In dieser Studie wurde etwa der Anforderungskatalog nach und nach ergänzt.
9. Gerade bei den nicht-funktionalen Bewertungskriterien stellt die objektive Bewertung aufgrund mangelnder Messbarkeit, divergierender Erwartungen der Stakeholder sowie unterschiedlicher IT-Kenntnisse und -Fähigkeiten der beteiligten Anwender eine Herausforderung dar. Dennoch ist es für Unternehmen unabdingbar, diese weitestmöglich in die Bewertung einzubeziehen.

Die Ergebnisse dieser Praxisstudie unterliegen verschiedenen Limitationen, woraus sich Anknüpfungspunkte für weitere Forschungsarbeiten ergeben. Einerseits wurden lediglich sieben IoT-Plattformen in die Bewertung einbezogen. Andererseits handelt es sich um eine Einzelfallstudie an einem Flughafen, was die Generalisierbarkeit einschränkt. Forscher sollten den IoT-Plattform-Markt umfassender und für unterschiedliche Branchen mit deren spezifischen Use-Cases untersuchen.

Am Flughafen München werden weiterhin in PoCs die gewählten IoT-Plattformen getestet und auf die Umsetzung als finale Lösung vorbereitet. Parallel dazu erweitert das Unternehmen sein Funkübertragungsnetzwerk und installiert zusätzliche Sensoren. So soll mittelfristig das Ziel, den Flughafencampus als Smart-City zu realisieren, erreicht werden.

Mit diesem Beitrag erhalten Unternehmensentscheider im Digitalisierungsumfeld einen Leitfaden zur Auswahl einer IoT-Plattform für ihre unternehmensspezifischen Anwendungsfälle und daraus resultierenden Anforderungen. Zudem wird ein Katalog an Bewertungskriterien und abgeleiteten Fragen bereitgestellt, der einen geeigneten Startpunkt bietet und unternehmensspezifisch angepasst und erweitert werden kann. Einen besonderen Beitrag für Praktiker stellen die abgeleiteten Lessons-Learned dar. Für Wissenschaftler zeigt dieser Beitrag die Notwendigkeit des Ausbaus der Forschung im Bereich iterativer Vorgehensmodelle bei der strukturierten Durchführung von Entscheidungsprozessen.

Die Methodik des Bewertungsverfahrens kann auch für andere Technologien und Systeme, wie zum Beispiel Augmented-Reality oder Cloud-Computing, eingesetzt werden. Allerdings muss angemerkt werden, dass es sich im Fall von IoT-Plattformen um einen äußerst umfangreichen Anforderungskatalog handelt, der nicht nur funktionale und operative Software-Anforderungen, sondern auch Anforderungen bezogen auf die Hardware (Sensoren) beinhaltet. Bei anderen Themen sind es oftmals kleinere und spezifischere Projekte, die in der Regel weniger Stakeholder betreffen.

Literatur

- [ATR16] Abdmeziem, M. R.; Tandjaoui, D.; Romdhani, I.: Architecting the Internet of Things: State of the Art. In (Abdmeziem, M. R.; Romdhani, D. T. I.): Robots and Sensor Clouds. Springer, Cham. S. 55-75, 2011.
- [AZ17] Alaa, M.; Zaidan, A. A.; Zaidan, B. B.; Talal, M.; Kiah, M. L. M.: A Review of Smart Home Applications Based on Internet of Things. *Journal of Network and Computer Applications*, 97, S. 48-65, 2017.
- [Br04] Bryson, J.: What to Do When Stakeholders Matter – Stakeholder Identification and Analysis Techniques. *Public Management Review*, 6(1), S. 21-53, 2004.
- [FL19] Floerecke, S.; Lehner, F.: Der ungleiche Kampf um Marktanteile im Infrastructure-as-a-Service-Segment: Erfolgswirksame Geschäftsmodellcharakteristika regionaler Anbieter im Wettbewerb mit den Hyperscalern. *HMD – Praxis der Wirtschaftsinformatik*, 56(6), 1334-1354, 2019.
- [Ga20] Gartner: Magic Quadrant 2020 für industrielle IoT-Plattformen, 2020. <https://azure.microsoft.com/de-de/blog/microsoft-named-a-leader-in-gartner-s-2020-magic-quadrant-for-industrial-iot-platforms/>
- [GB16] Guth, J.; Breitenbucher, U.; Flkenthal, M.; Leymann, F.; Reinfurt, L.: Comparison of IoT Platform Architectures: A Field Study based on a Reference Architecture. 2. Cloudification of the Internet of Things, Paris, 2016.
- [GK05] Großmann, M.; Koschek, H.: Unternehmensportale: Grundlagen, Architekturen, Technologien. Springer, Berlin, Heidelberg, 2005.
- [HM04] Hevner, A. R.; March, S. T.; Park, J.; Ram, S.: Design Science in Information Systems Research. *Management Information Systems Quarterly*, 28(1), S. 75-105, 2004.
- [Io18] IoT Analytics: IoT Platforms Market Report 2018-2023, 2018. <https://iot-analytics.com/product/iot-platforms-market-report-2018-23/>
- [LF14] Lasi, H.; Fettke, P.; Kemper, H. G.; Feld, T.; Hoffmann, M: Industrie 4.0. *Wirtschaftsinformatik*, 56(4), S. 261-264, 2014.
- [LP19] Lempert, S.; Pflaum, A.: Vergleichbarkeit der Funktionalität von IoT-Software-Plattformen durch deren einheitliche Beschreibung in Form einer Taxonomie und Referenzarchitektur. *HMD – Praxis der Wirtschaftsinformatik*, 56(6), S. 1178-1203, 2019.

- [OR18] Oberländer, A. M.; Röglinger, M.; Rosemann, M.; Kees, A.: Conceptualizing Business-to-Thing Interactions – A Sociomaterial Perspective on the Internet of Things. *European Journal of Information Systems*, 27(4), S. 486-502, 2018.
- [PF15] Portmann, E.; Finger, M.: Smart Cities – Ein Überblick!. *HMD – Praxis der Wirtschaftsinformatik*, 52(4), S. 470-481, 2015.
- [PH14] Porter, M. E.; Heppelmann, J. E.: How Smart, Connected Products are Transforming Competition. *Harvard Business Review*, 92(11), S. 64-88, 2014.
- [PR21] Pohl, K.; Rupp, C.: *Basiswissen Requirements Engineering. Aus- und Weiterbildung nach IREB-Standard zum Certified Professional for Requirements Engineering Foundation Level*. dpunkt.verlag, Heidelberg, 2021.
- [SH20] Strey, L.; Hein, C.; Ritter, T.; Knauer, C.; Ganz, A.: *Modellbasierte Methode zur Ableitung nicht-funktionaler Anforderungen im Kontext der Softwaremodernisierung. Modellierung 2020*, Wien, 2020.
- [Ti06] Timm, I.: *IuK-Evaluations-Instrumente aus der Perspektive der Informatik*. In (Hülsmann, M.; Timm, I.): *Strategische IuK-Evaluation*. Gabler, Wiesbaden, S. 61-81, 2006.
- [Yi18] Yin, R. K.: *Case Study Research: Design and Methods*. SAGE, Thousand Oaks, 2018.
- [Za14] Zangemeister, C.: *Nutzwertanalyse in der Systemtechnik: eine Methodik zur multidimensionalen Bewertung und Auswahl von Projektalternativen*. BoD – Books on Demand, Norderstedt, 2014.

Anhang

Die für die Experteninterviews in Fragen überführten Anforderungen:

- Gibt es eine garantierte Dauer, wie lange die Plattform betrieben und supportet wird?
- Welche Supportverträge sind möglich?
- Welche Lizenzkosten fallen für den reinen Betrieb an?
- Wie hoch sind die Kosten pro Nutzer?
- Welche Gebühren sind pro Gerät zu entrichten (einmalig oder jährlich)?
- Welche Kosten fallen pro Nachricht an?
- Inwieweit können für unterschiedliche Nutzer abgekapselte Datenbereiche erstellt werden?
- Welche Protokolle werden für die Anbindung von Geräten unterstützt?
- Inwiefern ist die Verbindung zwischen Geräten und Plattform gesichert?
- Inwieweit ist es möglich, die Geräte zu verwalten (Beispiel: Updates einspielen)?
- Wie können Alarme gesendet werden?
- Wo und wie lange können die Daten gespeichert werden?
- Mit welchen Methoden können Daten in der Standardversion ausgewertet werden?
- Welche Arten der Datenvisualisierung stehen zur Verfügung?
- Welche Möglichkeiten zur Datenmodellierung sind gegeben?
- In welchen Programmiersprachen lassen sich eingebundene Anwendungen entwickeln?
- Welche Schnittstellen zur Anbindung von Anwendungen (Beispiel: SAP ERP) bietet die Plattform?
- Ist es möglich, Daten mit einer Drittsoftware aus der Plattform abzugreifen?
- Können historische Daten, zum Beispiel via CSV-Datei, importiert werden?

Outsourcing im Kontext agiler Produktentwicklung – Ein Erfahrungsbericht

Christoph Albers¹, Ralf Kneuper²

Abstract: Unterschiedliche Formen des Outsourcings können Unternehmen dabei unterstützen, einerseits Kosteneinsparungen zu generieren und andererseits auf Fachkräfte in- und ausländischer Märkte zuzugreifen. Am konkreten Beispiel der Flugdistribution der TUI AG wurde ein Vorgehen entwickelt, welches einerseits die Prozessintegration betrachtet, zusätzlich aber auch auf technische Voraussetzungen und das kulturelle Verständnis eingeht.

Keywords: Agile Entwicklung; Outsourcing; Offshoring

1 Einleitung und Zielsetzung

1.1 Einleitung

Outsourcing gilt als Ansatz, um den auf Unternehmen und insbesondere ihrer IT lastenden Kostendruck sowie den verbreiteten Fachkräftemangel zu adressieren. Dabei ist nicht nur die Evaluation und Beauftragung eines geeigneten Partners zu bewältigen, sondern auch die Zusammenführung der internen Prozesse mit denen des Partners, die gerade bei Nutzung agiler Vorgehensweise eine besondere Herausforderung darstellt. Davon ist auch die TUI AG betroffen, und der vorliegende Beitrag berichtet die dabei gemachten Erfahrungen.

1.2 Outsourcing und Offshoring

In der Literatur existiert keine einheitlich akzeptierte Definition des Begriffs „Outsourcing“. Im Folgenden verstehen wir darunter nach [Be18] die Verlagerung betrieblicher Aktivitäten in eine andere Organisation, den Leistungserbringer. Es gibt viele verschiedene Varianten des Outsourcings, abhängig beispielsweise von der Anzahl der Leistungserbringer, ihrem geografischen Standort, oder dem Umfang der Verlagerung.

Eng verwandt mit Outsourcing ist das „Offshoring“, bei dem es um die geografische Verlagerung betrieblicher Aktivitäten ins Ausland geht. Dies wird unterteilt in Nearshoring,

¹ TUI Group, Markets – Central Region, Karl-Wiechert-Allee 4, 30625 Hannover, Germany christoph.albers@tui.com

² IU Internationale Hochschule – Fernstudium, Kaiserplatz 1, 83435 Bad Reichenhall, Germany ralf.kneuper@iu.org

also die Verlagerung in das nahe Ausland (von Deutschland aus z.B. Ost- oder Südeuropa), und Farshoring, also die Verlagerung in das ferne Ausland, von Deutschland aus z.B. China oder Indien [Be18]. Als Gegensatz zum Offshoring existiert das Onshoring, bei dem betriebliche Aktivitäten an Leistungserbringer im Inland verlagert werden.

Im Allgemeinen wachsen mit dem geografischen Abstand die erzielbaren Kostenvorteile, gleichzeitig aber auch die durch die größeren sprachlichen, kulturellen und rechtlichen Barrieren verursachten Risiken, was zu höheren Transaktionskosten führt.

Um den möglichen Nutzen zu erzielen und die Risiken zu reduzieren, ist eine systematische Vorgehensweise beim Einsatz von (IT-)Outsourcing erforderlich, wie beispielsweise in [UW12] beschrieben. Im Kern besteht dieses Vorgehen aus den Steuerungsmechanismen Anforderungsmanagement, Vertragsmanagement, Kommunikationsmanagement, Risikomanagement, Performance-Management und Serviceverbesserung. Auch wenn das im Folgenden beschriebene Unternehmen sich beim Outsourcing nicht bewusst an diesem Modell orientiert hat, ist das verwendete Vorgehen sehr ähnlich, wobei in diesem Kontext die Mechanismen Anforderungsmanagement, Kommunikationsmanagement und Risikomanagement als besonders relevant betrachtet wurden und daher hier beschrieben werden.

2 Funktions- und Prozessverlagerung im Kontext agiler Teams

2.1 Outsourcing im Rahmen agiler Arbeitsprozesse

Da es sich bei allen betrachteten Formen des Outsourcings um den Fremdbezug von Leistungen handelt, steht ein Unternehmen unabhängig vom gewählten Ansatz vor der Aufgabe, die internen Prozesse mit denen des Outsourcing-Partners zu harmonisieren.

Eine besondere Herausforderung dabei ist, dass agiles Arbeiten besonderen Wert auf enge Zusammenarbeit zwischen Kunde und Softwareentwicklung legt. Ein zentrales Konzept des agilen Vorgehens besagt, dass die Zusammenarbeit mit dem Kunden einen höheren Stellenwert als Vertragsvereinbarungen hat, formuliert beispielsweise im agilen Manifest als Wert und konkretisiert durch definierte agile Prinzipien [Be01]. Im Vergleich zu plangetriebenem Vorgehen basieren agile Methoden viel mehr darauf, Konzepte auszuprobieren und schnelle Rückmeldungen einzuholen, als zu Beginn viel zu planen. Diese Ideen sind aber beim Outsourcing und insbesondere beim Offshoring deutlich schwieriger umzusetzen. Wie beispielsweise in [HSR20] analysiert betreffen die wichtigsten Problembereiche verteilter agiler Teams durchgängig „weiche“ Themen, insbesondere Kommunikation, Team-Konsens, Kultur.

2.2 Kulturelle Spezifika internationaler Kollaboration

Je größer die räumliche Distanz kollaborierender Menschen zueinander ist, desto größer ist meist auch die kulturelle Distanz. Darunter wird die Gesamtheit der Unterschiede zwischen zwei Gruppen oder Individuen anhand unterschiedlicher Dimensionen verstanden. Dabei ist nicht die kulturelle Distanz an sich das entscheidende Maß, sondern deren Auswirkungen auf die Zusammenarbeit. Beispielsweise kann diese kulturelle Distanz auf operativer Ebene dazu führen, dass durch eine hohe Machtdistanz Anforderungen bei Unstimmigkeiten nicht hinterfragt werden und in ein striktes Auftraggeber-/Auftragnehmer-Modell verfallen wird.

Eines der wichtigsten Problemfelder beim Outsourcing ist es daher, die Unternehmenskulturen der beteiligten Partner zusammenzuführen. Ungelöste kulturelle Barrieren und kulturell inkompatible Wertesysteme können das Scheitern des gesamten Vorhabens zur Folge haben und sind daher bereits in der Anbahnungsphase zu berücksichtigen. Quatram formuliert dies sogar als „Kernthese 2: Ohne nachhaltige Integration der internationalen Kollegen ist der Erfolg von Near- und Offshore gefährdet“ [Qu15].

3 Evaluation und Implementierung ausgewählter Outsourcing-Ansätze am Beispiel der Flugdistribution der TUI AG

3.1 Unternehmerischer Kontext

Die TUI Aktiengesellschaft (AG) „TUI“ ist mit über 60.000 Mitarbeitern das international führende Touristikunternehmen mit einem Konzernumsatz von rund 19 Milliarden Euro im Geschäftsjahr 2019 [TU]. In diesem Bericht betrachten wir die IT-Abteilung *Digital Flight Distribution*, deren Aufgabe es ist, den Vertrieb von Flugeinzelplätzen oder „Seat-Only“ als Ergänzung des Produktangebotes der TUI zu unterstützen.

Die Abteilung *Digital Flight Distribution* verantwortet alle technischen Bereiche, um dieses Geschäftsmodell umsetzen zu können. Dieses umfasst u.a. die Beschaffung und Verbindung zu den Airlines, das Bereitstellen eines Werkzeugs zur Umsetzung einer Preislogik und die Verwaltung von Distributionskanälen im Eigen- und Fremdvertrieb.

Dazu ist die Abteilung in zwei Teams *Flight Sourcing & Distribution* und *Ancillaries* gegliedert. Jedes der beiden Teams besteht aus Business-Analysten, ggf. Support-Analysten sowie Softwareentwicklern und wird von einem Product Owner fachlich geführt.

3.2 Arbeitsmethodik der Abteilung *Digital Flight Distribution*

Im Team Flight Sourcing & Distribution wird Scrum weitestgehend im Standard mit einigen Ergänzungen eingesetzt. Während eines zweiwöchigen Sprints finden diverse Prozesse, in

der Regel in Form von Besprechungen (im Scrum-Guide als Ereignisse bezeichnet), zur Steuerung der Produktentwicklung statt, die um zusätzliche Besprechungen wie in Tab. 1 dargestellt erweitert wurden.

<i>Name</i>	<i>Frequenz</i>	<i>Zielsetzung</i>	<i>Teilnehmer</i>
PO-Runde	wöchentlich	Abstimmung zwischen Product Owner und Abteilungsleiter	PO, Abteilungsleitung
Dev-Call	wöchentlich	Abstimmung zwischen Entwicklerteam und Leiter Softwareentwicklung über Technologien und IT-Strategie	Entwickler, Leiter Softwareentwicklung

Tab. 1: Scrum-Ergänzungen im Team *Flight Sourcing & Distribution*

Das Ancillaries-Team verwendet als Vorgehensmodell Kanban. Der Product Owner in diesem Team fungiert primär als Ansprechpartner für Anforderungen.

Entsprechend reduzierter gestalten sich die Prozesse der Produktentwicklung. Tab. 2 gibt eine Aufstellung über die regelmäßigen Prozesse im Team.

<i>Name</i>	<i>Frequenz</i>	<i>Zielsetzung</i>	<i>Teilnehmer</i>
Daily	täglich	Kurztreffen über den aktuellen Arbeitsfortschritt, die geplanten Aktivitäten, ggf. mit Aufzeigen von Problemen.	Entwickler, PO, BA, SA
Retro-spektive	4-wöchentlich	Reflexion des Entwicklungsteams über die Zusammenarbeit seit der letzten Retrospektive.	Entwickler, PO, BA, SA
PO-Runde	wöchentlich	Abstimmung zwischen Product Owner und Abteilungsleiter	PO, Abteilungsleitung
SA-Runde	2-wöchentlich	Bericht der SAs an die Abteilungsleitung v.a. über Auffälligkeit im Tagesgeschäft	SAs, Abteilungsleitung

Tab. 2: Prozessübersicht des Teams *Ancillaries*

3.3 Evaluation der Outsourcing-Ansätze

3.3.1 Vereinbarte Maßnahmen

Da in beiden Teams der Bedarf besteht, die Kosten zu reduzieren und dem Fachkräftemangel entgegen zu wirken, wurden die folgenden Outsourcing-Maßnahmen beschlossen.

Maßnahme 1: Farshoring des Airline-Sourcings. Unter dem Airline-Sourcing wird das Herstellen der technischen Verbindung zwischen den Buchungssystemen der Fluggesellschaften und den Systemen der TUI verstanden. Dies umfasst die Suche nach Flugangeboten,

die Abfrage von Preisen und Kontingenten sowie die Durchführung der eigentlichen Flugbuchung. Daher wurde entschieden, dieses Themenfeld an einen bereits im Unternehmen bekannten Farshoring-Partner in Indien zu übergeben. Dort soll ein Team aus drei Entwicklern in Vollzeit sowie einem Projektleiter in Teilzeit aufgebaut werden.

Maßnahme 2: Onshoring für die Entwicklung des Distributionssystems. Auch diese Maßnahme betrifft das Team Flight Sourcing & Distribution, da das Team keine ausreichenden Ressourcen hat, um neue und geschäftsrelevante Funktionen in der angemessenen Zeit und Umfang umsetzen zu können und offene Stellen nicht besetzt werden konnten. Dabei handelt es sich um komplexe Anforderungen, die ein umfassendes Verständnis touristischer und technischer Prozesse sowie einen hohen Abstimmungsaufwand erfordern. Zusätzlich werden hier auch streng vertrauliche Funktionalitäten wie die Preisbildung oder das Steuern der Produkte in die einzelnen Vertriebskanäle auf technischer Seite implementiert.

Daher wurde entschieden, auf einen ebenfalls bereits bekannten Onshoring-Partner mit Sitz in Deutschland zurückzugreifen. Von diesem werden zwei Entwickler bezogen, die teilweise an ihren jeweiligen Standorten und im Home-Office, oftmals aber auch vor Ort in den Räumlichkeiten der TUI tätig werden.

Maßnahme 3: Nearshoring für Betrieb und Entwicklung der Ancillary-Komponente. Die dritte hier betrachtete Maßnahme betrifft das Team Ancillaries, dessen Software bislang ausschließlich durch drei freiberufliche Softwareentwickler weiterentwickelt wird.

Hier soll die Entwicklung und der Betrieb der Anwendung im Ganzen übergeben werden. Der Product Owner, der Business-Analyst sowie die Support-Analysten bleiben als internes Team und fachliche Führung der Anwendung bestehen. Durch diese besondere Situation wurde entschieden, ein Entwicklungsteam als so genanntes *Virtual Captive Team*, das also der operativen Kontrolle des Auftraggebers unterliegt, bei einem Nearshoring-Partner in Portugal aufzusetzen. Das Team soll sich aus einem Projektleiter, drei Anwendungsentwicklern und einem Cloud-Spezialisten für DevOps zusammensetzen.

3.3.2 Kostenbetrachtung

Die direkten Kosten der Maßnahmen umfassen die tatsächlich anfallenden Ausgaben für die Arbeitsleistung, wie Stundenlöhne und Reisekosten, sowie Kosten für benötigtes Material, z.B. Laptops oder Lizenzen. Dazu kommen die indirekten Kosten, u.a. höhere Aufwände für die Kommunikation oder für detailliertere Beschreibungen von Anforderungen.

Die direkten Kosten pro Person und die damit verbundenen Bedingungen lassen sich aus den jeweiligen Verträgen mit den Outsourcing-Partnern erfassen. Wie zu erwarten stellt das Farshoring in Bezug auf die direkten Kosten die preisgünstigste Option dar, gefolgt

vom Nearshoring. Das Onshoring übersteigt dabei sogar leicht die direkten Kosten von festangestellten Mitarbeitern, was vor allem durch den Fachkräftemangel und das dadurch knappe Angebot an externen Experten zu erklären ist.

Wichtig ist, auch die indirekten Kosten (Transaktionskosten) für die Anbahnung, Vereinbarung, Kontrolle und Anpassung zu berücksichtigen, auch wenn diese nur schwer beziffert werden können. Je nach Rahmenbedingungen können sie die durch niedrigere Lohnkosten generierten Kostenvorteile in der Gesamtrechnung übersteigen.

Outsourcing-Maßnahmen ausschließlich aus Kostensicht zu betrachten wäre jedoch zu eindimensional. Die einzelnen Maßnahmen bergen zudem spezifischen Risiken in sich, welche im Nachfolgenden anhand ausgewählter Beispiele ausgewertet werden sollen.

3.3.3 Risikobetrachtung

Outsourcingmaßnahmen können Risiken beinhalten, welche es zu analysieren und bewerten gilt. Eines der Hauptrisiken für den vorgestellten Kontext ist, dass das Ziel der Kostenreduktion nicht erreicht wird kann, weil die indirekten Kosten den durch die direkten Kosten generierten Vorteil zu Mehrkosten umkehren. Noch gravierendere Auswirkungen enthält das Risiko, dass die gewünschten Ergebnisse nicht oder nicht in der erforderlichen Qualität geliefert werden, unabhängig von den damit verbundenen Kosten.

Um diese Risiken abzuschätzen, wurden Kolleginnen und Kollegen im Konzern, welche bereits Erfahrungen mit IT-Outsourcing aufweisen und die Maßnahmen kennen, um eine Bewertung gebeten. Aus den Wahrscheinlichkeiten in Kombination mit der geschätzten Schadenhöhe wurden dann die Risiken quantitativ abgeschätzt und mit dem erwarteten Nutzen verglichen.

Die Abschätzung bestätigte die Erwartung, dass das Farshoring die größten Potentiale hinsichtlich einer Kostenreduktion birgt, zugleich aber auch mit dem größten Risiko verbunden ist. Durch das Nearshoring werden moderate Kostenvorteile erwartet, dafür ist das Risiko, diese Kostenvorteile durch überhöhte indirekte Kosten aufzuheben, deutlich geringer als bei der Farshoring-Maßnahme. Für das Onshoring werden keine Kostenvorteile erwartet, sondern im Gegenteil Kostennachteile, da bekannt ist, dass die Lohnkosten bzw. die Stundensätze über denen von internen Mitarbeitern liegen. Hier liegt der erwartete Nutzen alleine in der Gewinnung der benötigten und intern nicht vorhandenen Ressourcen.

3.4 Implementierung der Outsourcing-Maßnahmen im Kontext agiler Teams

Nach der bisherigen Analyse bleibt die Frage offen, wie die Maßnahmen in den zuvor dargestellten Kontext integriert werden können. Hierzu wird im Nachfolgenden ein Lö-

sungsszenario entwickelt, in welchem die Prozesse des jeweiligen Outsourcing-Partners mit den internen Prozessen der Abteilung harmonisieren werden.

3.4.1 Zusammenführung interner und externer Prozesse des Teams *Flight Sourcing & Distribution*

Die erste Herausforderung bestand in der Einbindung des in Indien ansässigen Farshoring-Partners. Aufgrund der großen kulturellen und räumlichen Distanz wurde vereinbart, dass die Entwickler weitestgehend unabhängig vom Team Flight Sourcing & Distribution arbeiten und sich ausschließlich der zugewiesenen Aufgaben annehmen. Das indische Team nutzt ebenfalls einen agilen Entwicklungsprozess, allerdings in einer von Scrum abweichenden Form. Hier dauert eine Iteration zwei Wochen und besteht aus einer Planungsrunde, einer Entwicklungsphase mit täglichen Statustreffen und wöchentlichem schriftlichen Bericht, und einem anschließenden internen Review mit Präsentation der Ergebnisse.

Die Kommunikation wird ausschließlich über den Projektleiter des indischen Unternehmens umgesetzt. Aufgaben oder Prioritätsänderungen werden an den Projektleiter berichtet und von diesem in die Planungsrunde beim Farshoring-Partner eingebracht.

Eine explizite Teilnahme am Review der TUI ist nicht vorgesehen, da die gelieferten Ergebnisse in das Endprodukt, d.h. die Software, einfließen und durch die Softwareentwickler auf Seiten der TUI vorgestellt werden. Ebenso gibt es keine gemeinsame Retrospektive, auch wenn die Zusammenarbeit mit den indischen Kollegen als Bestandteil der Retrospektive des Teams Flight Sourcing & Distribution berücksichtigt wird.

Nach anfänglichen intensiven Einarbeitungsaufwänden hat sich diese Integration als sehr zielführend herausgestellt, was insbesondere auf die Abgrenzung des Airline-Sourcings zum übrigen Teil der Entwicklung zurückgeführt werden kann.

Darüber hinaus erhält das Team Flight Sourcing & Distribution aus der Maßnahme 2 Unterstützung durch Mitarbeiter eines Onshoring-Partners. Da die Mitarbeiter des Onshoring-Unternehmens direkt in die TUI-Teams und deren Prozesse eingebunden werden, ist hier keine explizite Prozessintegration erforderlich. Einschränkungen gibt es nur insofern, dass die Mitarbeiter des Onshoring-Partners nicht in alle internen Prozesse, insbesondere nicht in die PO-Runde und den Dev-Call eingebunden sind. Auf Grund der komplexen Anforderungen in diesem Bereich der Entwicklung ist "Vollintegration" des Onshoring-Partners rückwirkend als richtige Entscheidung zu bewerten. Insgesamt ergibt sich damit der in Abb. 1 dargestellte Prozess.

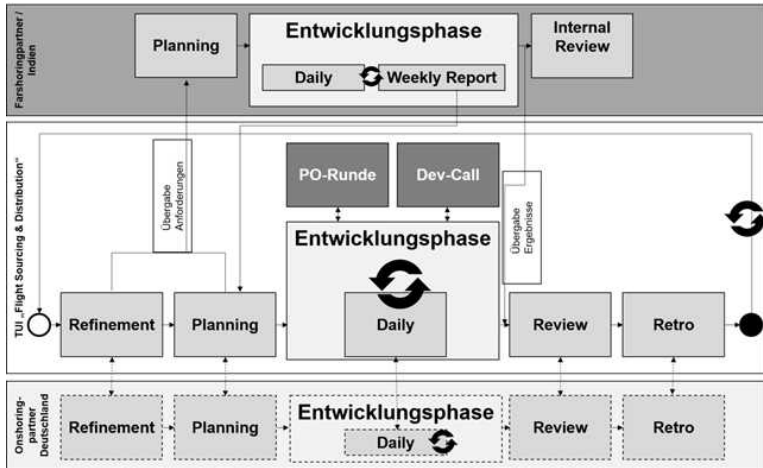


Abb. 1: Gesamtprozess des Teams *Flight Sourcing & Distribution* (eigene Darstellung)

3.4.2 Zusammenführung interner und externer Prozesse des Teams *Ancillaries*

Die dritte und damit letzte Maßnahme, deren Prozessintegration bestimmt werden soll, betrifft das Team *Ancillaries* und den Nearshoring-Partner in Portugal.

Im Gegensatz zum Nachbarsteam *Flight Sourcing & Distribution* wird im *Ancillaries*-Team Kanban eingesetzt. In einer jährlich erstellten Roadmap werden die wesentlichen Themen festgelegt und in Form einzelner Stories zu Entwicklungsaufgaben, Fehlerkorrekturen, operativen Themen etc. umgesetzt. Die einzelnen Stories durchlaufen einen vierstufigen Prozess mit den Schritten „In Planung“, „In Umsetzung“, „Test“ und „Erledigt“.

Dieser hochgradig dynamische Entwicklungsprozess mit Verzicht auf regelmäßige Planungsrunden erfordert eine ebenso hochgradig enge Verzahnung zwischen dem Team auf Seiten der TUI und dem Team in Porto.

Abb. 2 stellt diesen gemeinsamen Entwicklungsprozess dar, angelegt an den bereits etablierten Kanban-Prozessfluss. Die Abstimmung erfolgt über eine gemeinsame tägliche Besprechung, das „Daily“. Zur kontinuierlichen Prozessverbesserung erfolgt alle zwei Wochen eine Retrospektive. Um zu erreichen, dass das Team sich wie ein internes, verteiltes Team empfindet und die Unternehmensgrenzen weitestgehend ungeachtet bleiben, wurde vereinbart, dass jeweils alle Teammitglieder an allen Gruppenbesprechungen teilnehmen.

Diese Maßnahme ist zweifelsohne die tiefgreifendste von den dreien, da es um das Outsourcing einer gesamten Applikation inklusive dessen Betrieb geht. Entsprechend langwierig war der Einarbeitungsprozess. Der Gesamtprozess konnte letztendlich wie geplant umgesetzt werden.

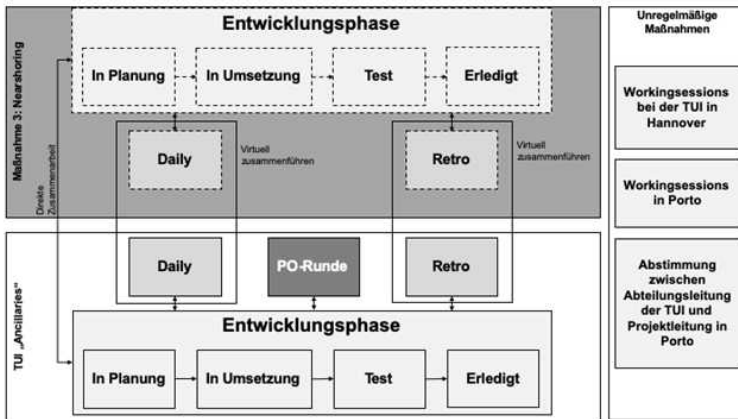


Abb. 2: Prozessintegration des Ancillaries-Teams (eigene Darstellung)

3.4.3 Technische Voraussetzungen

Für eine erfolgreiche Integration müssen neben der Harmonisierung der Prozesse die technischen Voraussetzungen für eine Kollaboration der Partner geschaffen werden. Dafür sind drei Themenbereiche von besonderer Bedeutung: Kommunikation, Dokumentation und Zugriff auf Entwicklungsumgebungen.

Für eine standortübergreifende Kommunikation wird eine Softwareunterstützung benötigt, welche Telefonie, Videotelefonie, Konferenzen, Chats und den Austausch von Dateien in Echtzeit auf mobilen und stationären Endgeräten ermöglicht. Daneben ist auch die Kommunikation per E-Mail möglich, jedoch hat sich gezeigt, dass insbesondere in agilen Produktentwicklungsteams der direkte, synchrone Austausch untereinander von großer Bedeutung ist. Da schon vor Beginn des Projekts ein Großteil der Kommunikation über MS Teams abgewickelt wurde, wird dies auch auf die drei Outsourcing-Partner ausgeweitet.

Neben der Kommunikation muss die Dokumentation von Ergebnissen, Arbeitsschritten etc. ermöglicht werden, damit parallel an Ergebnissen gearbeitet werden kann und gleichzeitig ein differenziertes Berechtigungsmanagement ermöglicht wird. Grundsätzlich wäre es denkbar, dass die TUI und der jeweilige Outsourcing-Partner ein jeweils eigenes Werkzeug für die Dokumentation einsetzen, wie in anderen Abteilungen der TUI praktiziert. Die Erfahrung hat jedoch gezeigt, dass dieses Modell zwar funktioniert, die Kommunikation aber erheblich erschwert und immer wieder zu Missverständnissen führt. Daher wurde entschieden, als gemeinsames Dokumentationssystem die bei TUI bereits etablierten Werkzeuge Jira und Confluence einzusetzen. Die Zugangsberechtigung erfolgt über eine zentrale Benutzerverwaltung innerhalb des Konzerns.

Der dritte und letzte Punkt betrifft den Zugriff auf die zu entwickelnden Produkte durch

die Outsourcing-Partner. Dazu sind mehrere Szenarien denkbar: Einerseits können die Systeme über eine Firewall-Konfiguration nach außen freigegeben werden, wobei allerdings ein großes Sicherheitsrisiko verbleibt. Deutlich sicherer, allerdings auch teurer, ist es, die betroffenen Personen mit einem VPN-Zugang auszustatten. Die dritte, hier gewählte Option besteht darin, das Softwareprodukt inklusive der Testsysteme entsprechend der bestehenden IT-Strategie der TUI in die Cloud auszulagern und den Partnern mittels Berechtigungskonzept Zugriff darauf zu gewähren.

3.4.4 Teamkultur

Wichtig für eine gute und erfolgreiche Teamkultur ist neben den beschriebenen strukturellen Maßnahmen das gegenseitige Vertrauen und das Gruppengefühl. Dazu gehört, dass jeder Mitarbeiter des Teams, unabhängig davon ob Mitarbeiter der TUI oder eines Outsourcing-Partners, das Vertrauen der anderen genießt, seine Aufgabe angemessen zu lösen. Um das zu erreichen wurde vereinbart, dass alle Teams eine *Definition of Done* festlegen, also eine im Team vereinbarte Definition, wann eine Aufgabe als abgeschlossen gilt (Mindestanforderungen an Code-Qualität, Test, Dokumentation etc.).

Zur Herstellung und Förderung eines Gruppengefühls ist neben den gemeinsamen Besprechungen der persönliche Austausch ein sehr wirksames Instrument. Auf Grund der unterschiedlichen Distanzen und der damit verbundenen Reisekosten wurde hier ein abgestuftes Vorgehen vereinbart (das dann auf Grund der Corona-Pandemie ausgesetzt werden musste). Da die Mitarbeiter des Onshoring-Partners einen Großteil der Arbeitsleistung in den Büros der TUI verrichten, ergibt sich hier eine regelmäßige, persönliche Zusammenarbeit. Zum persönlichen Austausch mit dem Nearshoring-Partner wurden quartalsweise Treffen abwechselnd bei der TUI in Hannover als auch beim Partnerunternehmen in Portugal vereinbart. Bedingt durch die Corona-Pandemie konnte der Austausch einmalig stattfinden und kann für die Teambildung als sehr gewinnbringend betrachtet werden. Bei der Zusammenarbeit mit dem Farshoring-Partner wird dieser Aufbau eines Gruppengefühls auf den Key-Account-Manager des Partnerunternehmens beschränkt, wie auch schon in der Beschreibung der gemeinsamen Besprechungen deutlich wurde. Dazu wurde das Vorgehen aus anderen TUI-Teams übernommen, ein bis zwei eintägige Treffen pro Jahr zwischen dem Key-Account-Manager und dem entsprechenden Team der TUI durchzuführen.

Eine weitere Maßnahme, die sich in anderen Teams der TUI als besonders positiv in Bezug auf das Gruppengefühl herausgestellt hat, ist die Vereinbarung, Telefonate, Konferenzen usw. stets mit eingeschalteter Kamera durchzuführen.

4 Fazit und Ausblick




Dieser Beitrag beschreibt einen Lösungsansatz zum Umgang mit den Herausforderungen des Outsourcings im Kontext agiler Produktentwicklung am Beispiel einer Fallstudie. Dabei

zeigt sich, dass nicht nur der potentielle Nutzen, insbesondere die Kostenersparnis, und die damit verbundenen Risiken bei den verschiedenen Ebenen des Outsourcings unterschiedlich sind, sondern auch unterschiedliche Vorgehensweisen zur Zusammenarbeit benötigt werden, um die unterschiedliche kulturelle und geografische Distanz bei Onshoring, Nearshoring und Farshoring zu überbrücken. Auf der rein technischen Ebene gibt es dagegen nur geringe Unterschiede zwischen diesen drei Ebenen des Outsourcings.

Literatur

- [Be01] Beck, K.; Beedle, M.; Bennekum, A. v.; Cockburn, A.; Cunningham, W.; Fowler, M.; Grenning, J.; Highsmith, J.; Hunt, A.; Jeffries, R.; Kern, J.; Marick, B.; Martin, R. C.; Mellor, S.; Schwaber, K.; Sutherland, J.; Thomas, D.: Principles behind the Agile Manifesto, <http://agilemanifesto.org/principles.html>, 2001.
- [Be18] Bendel, O.: Offshoring, <https://wirtschaftslexikon.gabler.de/definition/offshoring-54123/version-277177>, 2018.
- [HSR20] Haslinger, L.; Sikora, H.; Riedl, R.: Problembereiche verteilter agiler Teams: Literaturanalyse und Praxisimplikationen. HMD Praxis der Wirtschaftsinformatik 57/, S. 1014–1032, 2020.
- [Qu15] Quatram, J.: Nachhaltige Near-/Offshore-Integration in IT-Teams. Wirtschaftsinformatik & Management 7/, S. 40–48, 2015.
- [TU] TUI Group: Über die TUI Group, <https://www.tuigroup.com/de-de/ueber-uns/ueber-die-tui-group>.
- [UW12] Urbach, N.; Würz, T.: Ein integrierter Ansatz zur Steuerung von IT-Outsourcing-Vorhaben. HMD Praxis der Wirtschaftsinformatik 49/, S. 34–42, 2012.

Towards Augmenting Metadata Management by Machine Learning

Christopher Julian Kern ¹, Thomas Schäffer ², Dirk Stelzer ³

Abstract: Managing metadata is an important section of master data management. It is a complex, comprehensive and labor-intensive task. This paper explores whether and how metadata management can be augmented by machine learning. We deduce requirements for managing metadata from the literature and from expert interviews. We also identify features of machine learning algorithms. We assess 15 machine learning algorithms to determine their contribution to meeting the requirements and the extent to which they can support metadata management. Supervised and unsupervised learning algorithms as well as neural networks have the greatest potential to support metadata management effectively. Reinforcement learning, however, does not seem to be well suited to augment metadata management. Using Support Vector Machines and identification of metadata as an example, we show how machine learning algorithms can support metadata management.

Keywords: Artificial Intelligence, Machine Learning, Master Data, Metadata Management

1 Introduction

Digital transformation in companies, government agencies and our society as a whole requires appropriate data management. In a corporate context, managing master data is a key element of data management. The term master data refers to critical business objects of an organization. Master data management aims at planning, evaluating and controlling master data so that they can be used effectively and efficiently [Lo08]. Managing metadata is a central element of master data management [Lo08, HOÖ11]. Metadata is often referred to as data about content data [BNP99]. Hüner et al. define metadata [HOÖ11] as structured data referring to other data. Metadata focus either on technical characteristics [To99] or on features of applying content data [BNP99]. On the one hand, the challenges for metadata management are growing with increasing digitization, on the other hand, managing metadata is a complex, comprehensive and labor-intensive task [Lo08]. Many organizations will therefore need support for metadata management. One solution to this issue can be automation or at least automated support.

¹ EBS Universität für Wirtschaft und Recht, Management Group, Gustav-Stresemann-Ring 3, Wiesbaden,

65189, christopher.kern@ebs.edu,  <https://orcid.org/0000-0002-6864-2259>

² Hochschule Heilbronn, Institut für Wirtschaftsinformatik, Max-Planck-Str. 39

74081 Heilbronn, thomas.schaeffer@hs-heilbronn.de,  <https://orcid.org/0000-0001-8097-286X>

³ Technische Universität Ilmenau, Fachgebiet Informations- und Wissensmanagement, Postfach 100565,

Ilmenau, 98693, dirk.stelzer@tu-ilmenau.de,  <https://orcid.org/0000-0002-6757-9411>

Machine learning, a section of artificial intelligence, describes the ability of a machine or software to learn the execution of specific tasks. Machine learning is a powerful tool for analyzing large volumes of data [Mu12]. It is therefore reasonable to study to what extent machine learning is suitable to support metadata management. In this paper, we explore how machine learning algorithms may help to fully or partially automate selected tasks in metadata management. Using artificial intelligence for master data management has received considerable attention. The Gartner Group has labelled it Augmented MDM [JW20]. This paper focusses on assessing whether and how machine learning algorithms may be used to augment metadata management.

Prior research has only sparsely assessed the field of automating operations performed on metadata. Ganesan et al. have explored automatic identification of entities and their relationships which were then visualized based on metadata [Ga20]. Murthy et al. [Mu10] have considered automated extraction of metadata from a set of master data objects. This research was rather prototypical and the results are more than ten years old. Therefore, this topic deserves a fresh look and a more detailed analysis.

The objective of this paper is to identify to what extent selected machine learning algorithms may support metadata management tasks. We aim to answer two questions:

- RQ1: What are key requirements for augmenting metadata management?
- RQ2: Which machine learning algorithms support implementing these requirements?

2 Methodology

Figure 1 illustrates our research approach. The first step is based on the functional architecture for master data management proposed by Otto and Hüner [OH09]. This architecture refers to master data management as a whole. It consists of numerous sub-architectures, of which metadata management is one. We also conducted a literature review to elicit functional requirements for tools supporting metadata management. We followed the approach proposed by Fettke [Fe06] and analyzed the results according to Webster and Watson [WW02]. To identify relevant literature we used Google Scholar, Web of Science, the AIS eLibrary, IEEE Xplore and Springer Link. We used the search term [„metadata management“ OR „Metadatenmanagement“] and analyzed each article’s keywords and abstract. We considered 29 publications as depicted in Fig. 1. We included articles on metadata management in the context of master data management. From the findings of the review, we derived a set of requirements via a requirements elicitation process according to ISO 29148 [IS18]. We identified 37 requirements and summarized them in seven categories.

In the second step, we discussed the practical relevance of each requirement and prioritized the requirements with 15 participants of the master data management roundtable at the Heilbronn University of Applied Sciences. The participants are involved in strategic or operative management of master data in eight German companies. We also conducted four expert interviews. Each of the experts is responsible for master data management and works either in digital sales, process management or R&D. Our questionnaire consists of 25 closed-ended questions as well as six open-ended questions.

We evaluated answers to the closed-ended questions using a qualitative content analysis proposed by Mayring [Ma15]. We present our findings in form of a concept matrix based on Webster and Watson which is depicted in table 1 [WW02].

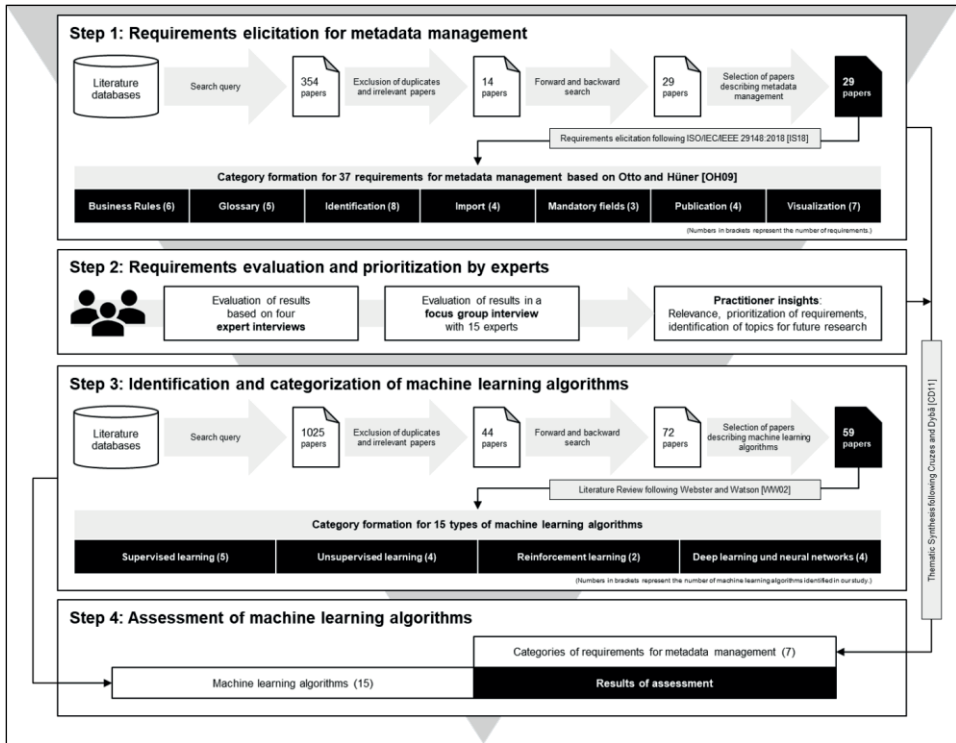


Fig. 1: Research approach

In the third step, we performed a second literature review to identify and categorize machine learning algorithms. We followed a similar approach as in the first review described in step 1. We used the search term („master data management“ OR „metadata management“ OR „Stammdatenmanagement“ OR „Metadatenmanagement“ OR „algorithms“ OR „Algorithmen“) AND („machine learning“ OR „maschinelles Lernen“). After analyzing keywords and abstracts, we considered 59 publications. We explored papers published between 2005 and 2020. We included articles on machine learning and machine learning algorithms, respectively. We studied 15 types of machine learning algorithms and grouped them into four categories, namely *supervised learning*, *unsupervised learning*, *reinforcement learning* and *neural networks*. We identified functions and features, potential fields of application as well as benefits and downsides when applying these algorithms.

The fourth step consists of assessing selected machine learning algorithms. We assessed the algorithms based on an argumentative-deductive approach using the 37 requirements identified in step 1. We used the approach presented by Cruzes and Dyba [CD11] and

extracted a model of higher-order themes from our results, merging sub-requirements into requirements. We used a four-eye-principle as proposed by Peffers et al. [Pe12]. We present our findings in section 3.4.

3 Findings

3.1 Requirements for Metadata Management

Subsequent to the first literature review, we performed a requirements elicitation process following ISO 29184 [IS18]. This resulted in 37 requirements, which we grouped in seven categories. Six of these categories were originally proposed by Otto and Hüner [OH09]. We added the category *identification* of metadata as this is regarded also as an important area of metadata management in the literature. In the following passage, we briefly explain the seven categories of metadata management tasks (see figure 1).

- Documentation of *business rules* seeks to document process related guidelines, partly formalized rules and individual knowledge of staff members. Documented business rules are essential for process automation [OÖ16]. These rules can be formed with the help of metadata definitions. They refer to individual master data elements or to a set of these data.
- Creation of a *glossary* or dictionary is a prerequisite for identifying, documenting and maintaining master data. A glossary can be created and updated using metadata. It supports documentation of and communication on business processes for which master data is required [Lo08].
- *Identification* means creating metadata for already existing master data. These metadata must be continuously updated afterwards. Collecting all definitions of master data in an organization creates a basis for metadata management [Lo08]. These definitions can be documented in a data catalogue or data dictionary [Hi21].
- *Metadata import* helps to consolidate metadata stored in different formats into a coherent system. Import of metadata from heterogeneous systems is an element of integration processes as described by Loshin [Lo08].
- *Mandatory fields* must contain valid values if master data is inserted into a database or updated [Hi18]. Managing mandatory fields is a prerequisite to ensure adequate quality of master data. Metadata is used to define mandatory or optional fields.
- *Metadata publication* denotes the process of making metadata available for operational information systems. This service supports a coherent use of master data. Glossaries and data dictionaries help to maintain a consistent use of definitions and formats when integrating information systems [Lo08].
- The aim of *visualization* is to display master data in a graphical form. Visualization uses metadata to represent data objects graphically, e. g., with Entity-Relationship- or UML-Diagrams.

3.2 Evaluated and Prioritized Requirements

As already mentioned, we conducted expert interviews to evaluate and prioritize the results of our requirement elicitation process. In the following focus group interviews, we summarized and discussed the results of the interviews with a larger group of experts. We asked them to prioritize the seven metadata management tasks described above. The experts assigned priority one to the following tasks: documentation of *business rules*, *identification* of metadata and *visualization* of master data. Priority two was assigned to definition of *mandatory fields* and metadata *import* and priority three to creation of a *glossary* and metadata *publication*. These priorities provide valuable guidance for the selection of practical projects to support metadata management by machine learning.

3.3 Selected Machine Learning Algorithms

In step 3, we identified and categorized four types of machine learning algorithms. In the following step, we assess how they can be used to augment metadata management.

Supervised learning denotes the ability of algorithms to allocate a given set of inputs to a set of outputs [GGA20]. We further evaluated four subtypes of supervised learning, presented by Sindhu et al. [SS20], namely ‘support vector machines’, ‘naïve bayes’, ‘linear regression’ and ‘decision trees’. We also subsume ‘semi-supervised learning’ as another sub-category of supervised learning.

Unsupervised learning seeks to find patterns in a given set of data [GAM20]. Nock and Nielsen [NN06] describe four types of algorithms as key examples: ‘K-means-clustering’, ‘fuzzy c-means’, ‘gaussian-expectation-maximization’ and ‘harmonic k-means-clustering’.

Reinforcement learning trains algorithms with a set of reinforcements, ‘rewards’ and/or ‘punishments’ [Ru16]. Weber describes ‘q-learning’, ‘temporal difference learning’ and ‘generative adversarial networks’ as instances of reinforcement learning [We20a].

Neural networks are based on the assumption that mental activity consists primarily of the electrochemical activity of networks of neurons [Ru16]. In a neural network, ‘artificial neurons’ - arranged in a network - perform the computation. The basic architecture consists of three layers of artificial neurons: an input layer, a hidden layer and an output layer [Ha09]. There are numerous variations of neural networks [Ru16]. We decided to address ‘auto-encoders’, ‘convolutional neural networks’ and ‘recurrent neural networks’.

3.4 Selected Findings of Assessments

In step 4 of our research process, we assessed the extent to which the algorithms meet the requirements. Table 1 shows the results of the assessment. For reasons of space, we cannot show all results in detail here, but only at the level of the requirement categories. The original presentation showing all 37 requirements is available from the authors upon request. We illustrate the results using Harvey-Balls. A full black ball represents a strong match between the requirements and the functionality of an algorithm, a white ball expresses that an algorithm does not match the requirements at all.

	Requirements						
	Business rules	Glossary	Identification	Import	Mandatory fields	Publication	Visualization
Machine learning algorithms							
Supervised learning							
Support vector machines	○	●	●	●	○	●	●
Naïve bayes	●	●	●	●	○	●	●
Decision trees	●	●	○	○	○	○	○
Semi-supervised learning	●	●	●	○	●	●	●
Linear regression	○	○	○	○	○	○	○
Unsupervised learning							
K-means-clustering	●	●	○	○	○	●	●
Fuzzy-c-means	●	●	○	○	○	●	●
Gaussian-EM	●	●	●	○	○	●	●
Harmonic-k-means-clustering	●	●	○	○	○	●	●
Reinforcement learning							
Q-learning	○	○	○	○	○	○	○
Temporal difference learning	○	○	○	○	○	○	○
Generative adversarial networks	○	○	○	○	○	○	○
Neural networks							
Auto encoders	●	●	○	●	●	●	●
Convolutional neural networks	○	○	●	○	○	○	○
Recurrent neural networks	●	●	●	○	○	●	○

Tab. 1: Assessment of machine learning algorithms with metadata management requirements

Four types of algorithms were assessed as not suitable for meeting any requirement. ‘Linear regression’ primarily aims at predicting future states. This does not correspond with any of the requirements. The same applies to *Reinforcement Learning Algorithms*, namely ‘q-learning’ and ‘temporal difference learning’. These algorithms are primarily suitable for problems that can be solved with the help of Markov chains. ‘Generative adversarial networks’ are particularly fit for generating sets of data, e.g., for manipulating or generating images.

The experts have ranked the identification of metadata as one of the most important functional requirements. We have assessed ‘support vector machines’ (SVM) as the algorithm with the highest potential to support identifying metadata. Therefore, we focus on this task to illustrate how machine learning algorithms can support metadata management.

SVM are among the most popular supervised learning algorithms. Algorithms are trained to draw boundaries through a given set of data and to group them into classes [Ru16].

Training SVM needs a large number of data of sufficient quality [We20b]. Bahmani et al. [BBV17] use SVM for data cleansing and for the detection of duplicates to be merged into a single data object. Singh et al. [STS16] name text classification as the main application of SVM. Advantages are high accuracy of the algorithms, avoidance of overfitting and a good suitability for generalizing facts. Disadvantages are the complexity of SVM, the extensive training effort needed and that performance depends on the choice of appropriate parameters. In the following paragraphs, we address the eight requirements that we have summarized in figure 1 and table 1 under the heading *identification* of metadata.

In the context of SVM, we primarily assessed classification-, entity detection- and entity resolution problems. Classification problems connect a given set of inputs with a discrete number of outputs [Mu12]. Text classification problems use certain words or word combinations to characterize an object [STS16]. Duplicates belong to certain groups of objects which can be characterized by their attributes [BBV17]. Conflicts resulting from duplicates can be resolved by either replacing them with a single object [BBV17] or by connecting the two objects – which also aids in representing semantic aspects [HOÖ11, Lo08].

The first requirement is ‘The algorithm shall list all definitions of master data objects that exist within an organization in the form of metadata’. This is feasible because SVM can use text classification to recognize data objects and definitions. However, this requires considerable training.

The second requirement, ‘The algorithm shall recognize and list similarities and differences of master data definitions in the form of metadata’, can easily be satisfied, as it is a duplicate detection problem in combination with text recognition. The amount of training required will probably also be high.

The third requirement is ‘The algorithm shall standardize and harmonize similar definitions of master data objects in the form of metadata’. The implementation is feasible since text recognition is the main task here as well. Training algorithms will probably be time-consuming. The results may need to be improved by experts.

The fourth requirement is ‘If similar definitions of master data objects cannot be standardized, the algorithm shall name them differently in the metadata according to a uniform scheme’. The implementation is also quite simple, since this again is a text recognition problem. However, the naming scheme must already exist for the results to be usable.

The fifth requirement is ‘If similar definitions cannot be standardized, but data objects are similar, the algorithm shall connect the metadata of data objects’. This is possible because linking definitions is a classification problem. However, the complexity of the problem substantially increases the effort required to train the algorithm. The results may also need to be improved by experts.

The sixth requirement, ‘If the existing definitions of master data objects do not match newly determined definitions, the algorithm shall replace them in the metadata’, can be implemented well with SVM as this is mainly a comparative text recognition problem.

The seventh requirement is ‘If two definitions of master data objects exist at the same hierarchical level and if they have the same fields, the algorithm shall indicate that they may be synonyms and establish a link in the referring metadata’. SVM are well suited to support this task, since it is a classification and text recognition issue. As with previous

requirements, the training effort is probably significant and the results may need to be improved by experts.

The eighth requirement ('The algorithm shall enter definitions of master data objects into a repository') does not require machine learning. Therefore, we have not considered this requirement in our study.

4 Conclusion

4.1 Summary

We have identified 37 requirements for augmenting metadata management in the context of master data management and grouped these requirements in seven categories. Experts evaluated our findings and prioritized the requirements. On this basis, we can now answer RQ 1: Key requirements for augmenting metadata management are documentation of business rules, identification of metadata and visualization of master data.

We have also identified 15 types of machine learning algorithms and grouped them in four categories. We explored which algorithms are fit for meeting which of the requirements identified in the first step of our research. We answer RQ 2 as follows: Supervised and unsupervised learning algorithms and, with certain restrictions, neural networks have the greatest potential to support metadata management effectively. Reinforcement learning, however, does not seem to be suited to augment metadata management.

A more detailed assessment showed that SVM might best support the identification of metadata. Unsupervised machine learning algorithms are appropriate to augment visualization of metadata. These algorithms help to identify and classify elements to be visualized and their relationships. Unfortunately, no algorithm is well suited to support documentation of business rules.

4.2 Limitations

We need to acknowledge some limitations to our research. First, our database for eliciting requirements for algorithms to support metadata management was rather small. While there is a large body of literature on metadata management, sources that address managing metadata referring to master data are rare. The number of experts we interviewed was also limited. A more extensive number of relevant publications and more interviews could lead to a more comprehensive list of requirements. Second, our research has only been able to provide an overview of the topic. We took a comprehensive look at metadata management and assessed 15 algorithms against 37 requirements. It is likely that a more detailed analysis of specific subtasks of metadata management would lead to requirements that are more specific. Focusing on one or a few algorithms would allow assessing the contribution of these algorithms more precisely. Third, we have worked and argued on a conceptual level only. However, we have not practically tested the extent to which concrete algorithms are suitable for meeting specific requirements.

4.3 Implications

Our results can help practitioners who want to support managing metadata through artificial intelligence to pre-select machine learning algorithms. Investigations that are more concrete must be carried out in specific application areas. We believe that our work offers a useful basis for future research. As mentioned above, an expansion of the database with more publications and more interviewees, a more detailed analysis of the contribution of selected algorithms to concrete subtasks of metadata management, and a real-world testing of algorithms in practical applications are promising research opportunities. It would also be interesting to explore in which cases the benefits of using machine learning algorithms exceed the costs involved. We have limited our analysis to assessing the contribution of machine learning algorithms for augmenting metadata management. Research that is more extensive could consider the entire area of master data management.

5 Bibliography

- [BBV17] Bahmani, Z.; Bertossi, L.; Vasiloglou, N.: ERBlox: Combining Matching Dependencies with Machine Learning for Entity Resolution. In *International Journal of Approximate Reasoning*, 2017, 83; pp. 118–141.
- [BNP99] Burnett, K.; Ng, K. B.; Park, S.: A comparison of the two traditions of metadata development. In *J. Am. Soc. Info. Sci.*, 1999, 50; pp. 1209–1217.
- [CD11] Cruzes, D. S.; Dybå, T.: Recommended Steps for Thematic Synthesis in Software Engineering. In (Ed. IEEE Computer Society): *ESEM 2011 Proceedings*, 2011; pp. 275–284.
- [Fe06] Fettke, P.: State-of-the-Art des State-of-the-Art. In *WIRTSCHAFTSINFORMATIK*, 2006, 48; pp. 257–266.
- [Ga20] Ganesan, B. et al.: Link Prediction using Graph Neural Networks for Master Data Management, 2020.
- [GAM20] Guérin, E.; Aydin, O.; Mahdavi-Amiri, A.: Artificial Intelligence. In (Eds. Guo, H.; Goodchild, M. F.; Annoni, A.): *Manual of Digital Earth*. Springer, Singapore, 2020; pp. 357–385.
- [Ha09] Haykin, S. S.: *Neural networks and learning machines*. Pearson, New York, 2009.
- [Hi21] Eds. Hildebrand, K. et al.: *Daten- und Informationsqualität. Auf dem Weg zur Information Excellence*. Springer Vieweg, Wiesbaden, 2021.
- [HOÖ11] Hüner, K. M.; Otto, B.; Österle, H.: Collaborative management of business metadata. In *International Journal of Information Management*, 2011, 31; pp. 366–373.

- [IS18] ISO/IEC/IEEE: Systems and Software Engineering: Life Cycle Processes Requirements Engineering (ISO/IEC/IEEE 29148), 2018.
- [JW20] Judah, S.; White, A.: Hype Cycle for Data and Analytics Governance and Master Data Management, 2020, 05.04.2021.
- [Lo08] Loshin, D.: Master Data Management. Elsevier, Amsterdam, 2008.
- [Ma15] Mayring, P.: Qualitative Inhaltsanalyse. Grundlagen und Techniken. Beltz Verlag, Basel, 2015.
- [Mu10] Murthy, K. et al.: Content-Aware Master Data Management: Proceedings of the 16th COMAD, Nagpur, India, 2010; pp. 206–210.
- [Mu12] Murphy, K. P.: Machine learning. A probabilistic perspective. MIT Press, Cambridge, 2012.
- [NN06] Nock, R.; Nielsen, F.: On weighting clustering. In IEEE transactions on pattern analysis and machine intelligence, 2006, 28; pp. 1223–1235.
- [OH09] Otto, B.; Hüner, K. M.: Funktionsarchitektur für unternehmensweites Stammdatenmanagement. Universität St. Gallen, St. Gallen, 2009.
- [OÖ16] Otto, B.; Österle, H.: Corporate Data Quality. Springer, Berlin, 2016.
- [Pe12] Peffers, K. et al.: Design Science Research Evaluation. In (Eds. Peffers, K.; Rothenberger, M.; Kuechler, B.): Design science research in information systems. Springer, Berlin, 2012; pp. 398–410.
- [Ru16] Russell, S. J. et al.: Artificial intelligence. A modern approach. Pearson, Boston, 2016.
- [SS20] Sindhu Meena, K.; Suriya, S.: A Survey on Supervised and Unsupervised Learning Techniques. In (Eds. Kumar, L. A.; Jayashree, L. S.; Manimegalai, R.): AISGSC. Springer, Cham, 2020; pp. 627–644.
- [STS16] Singh, A.; Thakur, N.; Sharma, A.: A review of supervised machine learning algorithms: 2016 3rd INDIACom, 2016; pp. 1310–1315.
- [To99] Tozer, G. V.: Metadata management for information control and business success. Artech House, Boston, 1999.
- [We20a] Ed. Weber, F.: Künstliche Intelligenz für Business Analytics. Springer Fachmedien, Wiesbaden, 2020.
- [We20b] Ed. Wennker, P.: Künstliche Intelligenz in der Praxis. Springer Fachmedien Wiesbaden, Wiesbaden, 2020.
- [WW02] Webster, J.; Watson, R. T.: Analyzing the Past to Prepare for the Future: Writing a Literature Review. In MIS Quarterly, 2002, 26; pp. xiii–xxiii.

State-of-the-Art der IT-Governance-Forschung

Eine Literaturanalyse über Themen und Forschungsbedarfe

Rodrigo Raab¹, David Treu², Sebastian Straßburg³ und Helmut Beckmann⁴

Abstract: State-of-the-Art Beiträge verfolgen das Ziel, den aktuellen Kenntnisstand innerhalb eines Forschungsgebiets systematisch zu untersuchen und darzustellen. Die vorliegende Arbeit repräsentiert einen State-of-the-Art Beitrag im Forschungsgebiet IT-Governance. IT-Governance ist ein integraler Bestandteil des strategischen Informationsmanagements und nimmt in Unternehmen eine eminent wichtige Schlüsselstellung für die Generierung eines hohen IT-Wertbeitrags ein. In Anbetracht der Wichtigkeit dieses Forschungsgebiets hat die vorliegende Arbeit zum Ziel, mithilfe einer systematischen Literaturanalyse ein Review existierender Beiträge in der IT-Governance-Forschung durchzuführen. So wird der aktuelle Stand der Forschung in Form zentraler Themen dargestellt, inklusive einer Konzeptmatrix und einem abgeleiteten Research Framework. Zudem sollen Forschungsbedarfe für die Zukunft präsentiert werden.

Keywords: IT Governance, Review, State of the Art, Forschungsbedarfe

1 Einleitung

Der Forschungsbereich IT-Governance umfasst Methoden, welche mithilfe der IT die Abdeckung der Geschäftsziele und die angemessene Überwachung von Risiken gewährleisten sollen [MZK03]. Besonders aufgrund der sich wandelnden Rolle der IT in Unternehmen von einer operativen hin zu einer strategischen Bedeutung nimmt eine effektive IT-Governance die Stellung eines kritischen Erfolgsfaktors ein (vgl. [DV09], [Lu17], [AG16]). IT-Governance hat einen großen Einfluss auf die Generierung eines hohen Wertbeitrags aus der IT und damit zusammenhängend auf die Unternehmensperformance (vgl. [WPG06], [VR17]). Die Wichtigkeit der IT-Governance induziert den Bedarf, im Sinne einer State-of-the-Art-Forschung zu untersuchen, durch welche Themen sich IT-Governance charakterisieren lässt und welche Forschungsbedarfe in der Literatur identifiziert werden können. Die Forschungsfrage lautet wie folgt:

Wie ist der State-of-the-Art im Forschungsbereich IT-Governance?

¹ Hochschule Heilbronn, Fakultät für WV, Max-Planck-Str. 39, 74081 Heilbronn, rraab@stud.hs-heilbronn.de

² Hochschule Heilbronn, Fakultät für WV, Max-Planck-Str. 39, 74081 Heilbronn, dtreu@stud.hs-heilbronn.de

³ Hochschule Heilbronn, Fakultät für WV, Max-Planck-Str. 39, 74081 Heilbronn, strassburg@stud.hs-heilbronn.de

⁴ Hochschule Heilbronn, Fakultät für WV, Max-Planck-Str. 39, 74081 Heilbronn, helmut.beckmann@hs-heilbronn.de

2 Forschungsmethodik

Das gewählte Verfahren für die vorliegende Literaturanalyse orientiert sich an Fettke [Fe06] sowie Webster und Watson [WW02]. Beide Vorgehensmodelle sind in der Wissenschaft weit verbreitet und eignen sich für die durchzuführende Literaturanalyse. Zunächst wurden gemäß dem Phasenmodell von Fettke [Fe06] im Anschluss an die Problemformulierung in ausgewählten Literaturdatenbanken und Konferenzen mithilfe von Suchtermen in deutscher und englischer Sprache erste Beiträge akkumuliert.

Es erfolgte keine Eingrenzung des Zeitraums der Suche, da sämtliche existierende relevante Beiträge im Zeitverlauf identifiziert und analysiert werden sollten. Die Suche mit den Suchtermen erfolgte innerhalb der Titel potenzieller Forschungsbeiträge, um die Suche auf solche Beiträge zu limitieren, die die Thematik als Hauptaspekt ihrer Untersuchung behandeln. Die nach einem ersten Suchlauf identifizierten 981 potenziellen Beiträge wurden in einem nächsten Schritt daraufhin untersucht, ob sie für die vorliegende Literaturanalyse relevant und wissenschaftlich geeignet sind. Zudem wurden Duplikate ausgefiltert. So wurden insgesamt 834 Beiträge (85 %) von der weiteren Untersuchung ausgeschlossen. Im nachfolgenden Analyseschritt wurden die Abstracts der 147 verbliebenen Beiträge untersucht. Dies führte zur näheren Auswahl von insgesamt 18 Forschungsbeiträgen. Schließlich wurden mit einer Vorwärts- und Rückwärtssuche gemäß Webster und Watson [WW02] 4 weitere Beiträge gesammelt. Anschließend erfolgte eine Kategorisierung der Literatur in die verschiedenen thematischen Schwerpunkte, um die Basis für eine strukturierte Literaturanalyse sowie die Anfertigung einer Konzeptmatrix zu schaffen. Schließlich wurden die 22 ausgewählten spezifischen Forschungsbeiträge im Detail analysiert. Das Ergebnis dieser Analyse kann in den nachfolgenden Kapiteln sowie in der Konzeptmatrix in Kapitel 3.4 begutachtet werden.

3 Zentrale Themen

3.1 Mechanismen der IT-Governance

IT-Governance-Mechanismen können durch drei Komponenten charakterisiert werden: *Prozesse, Strukturen und Beziehungsmechanismen*. Gemäß Van Grembergen et al. [VDG04] beziehen sich Strukturen dabei auf die Existenz von verantwortlichen Funktionen wie IT-Führungskräfte oder einer Vielzahl von IT-Ausschüssen. Prozesse hingegen beschäftigen sich mit der strategischen IT-Entscheidungsfindung und Überwachung. Zu den Beziehungsmechanismen gehören die Beteiligung von Unternehmen oder der IT und Partnerschaften sowie der strategische Dialog und das gemeinsame Lernen [VDG04]. Alle drei Bereiche sind relevant und müssen kombiniert werden zur Schaffung eines ganzheitlichen Ansatzes, der eine effektive und effiziente IT-Governance in der gesamten Organisation fördert [APM13].

Ein Problem, welches in der gegenwärtigen IT-Governance-Forschung festgestellt werden kann, ist die unvollständige und unübersichtliche Definition der zentralen Mechanismen. Viele Forschungsbeiträge listen neue Mechanismen auf, entfernen dafür andere, fassen mehrere zu einem zusammen oder nutzen andere Benennungskonventionen für den gleichen Aspekt. Aufgrund der fehlenden Einheitlichkeit besteht ein starker Bedarf nach Konsolidierung. Tab. 1 gibt einen ersten Überblick über die häufigsten adressierten Strukturen, Prozesse und Beziehungsmechanismen und leistet dahingehend einen Beitrag, wichtige IT-Governance-Mechanismen und ihre Benennungshäufigkeit in Forschungsbeiträgen übersichtlich darzustellen.

Mechanismus & Quelle		[DV04]	[VDG04]	[DV09]	[APM13]	[Lu17]	[LHP18]	[De20]	Σ
Strukturen	IT Steering Committee (at executive/senior management level)	x	x	x	x	x	x	x	7
	IT Strategy Committee (at level of board of directors)	x	x	x	x		x	x	6
	IT Project Steering Committee		x	x	x	x	x	x	6
	Integration of Governance / Alignment Tasks in Roles & Responsibilities	x	x	x	x		x	x	6
Prozesse	Service Level Agreements (SLAs)	x	x	x	x	x	x	x	7
	Portfolio Management	x	x	x	x	x	x	x	7
	IT Balanced Scorecard	x	x	x	x	x	x	x	7
	Strategic Information Systems Planning (SISP)	x	x	x	x	x	x	x	7
	Frameworks (COBIT, ITIL, ...)	x	x	x	x	x	x	x	7
Beziehungen	Cross-functional Business/IT Training	x	x	x	x	x	x	x	7
	Cross-functional Business/IT Job Rotation	x	x	x	x	x	x	x	7
	Business IT Collocation	x	x	x	x		x	x	6

Tab. 1: Übersicht über IT-Governance-Mechanismen in der Literatur

Die sieben in der Tab. 1 abgebildeten Beiträge beschäftigen sich im Detail mit IT-Governance-Mechanismen und die Forscher nehmen darin eine um Vollständigkeit bemühte Auflistung dieser Mechanismen vor. Dargestellt sind sämtliche Mechanismen, die in mindestens 6 der 7 Beiträge genannt wurden. Eine wichtige Aufgabe für künftige

Forschungsarbeiten besteht darin, diese Tabelle zu erweitern, um einen konsolidierten und gesamtheitlichen Überblick über die in der Literatur referenzierten Mechanismen zu erhalten und so schließlich die zentralen Mechanismen ableiten zu können. Genau wie es keine einheitliche Definition für IT-Governance gibt, gibt es auch keine einheitliche Lösung bei der Umsetzung der vorgestellten Strukturen, Prozesse und Beziehungsmechanismen. Eine „universell beste IT-Governance-Struktur existiert nicht“ [BG05, 703]. Dementsprechend liegt es an jedem Unternehmen, individuell für dessen spezifischen Kontext die richtige Kombination aus der Großzahl der vorgestellten Mechanismen auszuwählen und umzusetzen. Dieser Kontext kann dabei etwa durch die Größe, Branche oder die Kultur des Unternehmens charakterisiert sein [De20]. Dabei ist zu berücksichtigen, dass die Mechanismen auf koordinierte Weise zusammenarbeiten. Bestimmte Strukturen lassen sich etwa ohne entsprechende Prozesse nicht implementieren. Außerdem ist darauf zu achten, dass die Beziehungsmechanismen nicht nur in den Anfangsstadien der IT-Governance-Implementierung beachtet werden, sondern auch bei dessen Einbettung in den Geschäftsbetrieb hinreichende Berücksichtigung finden [LHP18]. In der zukünftigen Forschung sollte daher untersucht werden, welche Mechanismen unter welchen Umständen besonders gut miteinander kombinierbar sind, oder bei welchen das nicht der Fall ist.

3.2 Kontingenzfaktoren

Nachfolgend werden neun Kontingenzfaktoren beschrieben, die gemäß Pereira und Mira da Silva [PM12] hauptsächlich unterschieden werden. Kontingenzfaktoren lassen sich als organisationsspezifische Faktoren definieren, die Auswirkungen auf die IT-Governance haben können [PM12] und sowohl von interner als auch von externer Natur sein können [XLB08]. Ein relevanter Faktor ist die *Unternehmenskultur* (I). Diese wirkt sich meist stark auf das Veränderungspotenzial eines Unternehmens aus und hat Auswirkungen auf die Strukturen, Prozesse und Beziehungsmechanismen eines Unternehmens [ARH14]. Auch die *Organisationsstruktur* (II) spielt eine bedeutende Rolle. Eine effektive IT-Governance wird unter anderem dadurch bestimmt, wie die IT organisiert ist und wo die IT-Entscheidungskompetenz innerhalb der Organisation angesiedelt ist. Es werden allen voran zentralisierte, dezentralisierte und föderale IT-Organisationen unterschieden [DV04]. Die *Organisationsgröße* (III) ist ein weiterer Kontingenzfaktor. Diese hat dahingehend Auswirkungen, dass beispielsweise kleinere Organisationen aufgrund von fehlenden Ressourcen oft nicht in der Lage sind, typische IT-Best-Practices zu erfüllen. Dies ist in größeren Unternehmen mit mehr Ressourcen einfacher möglich [Co10]. Ein weiterer Einflussfaktor ist die *Industrie* (IV). Dabei hängt die Bedeutung der IT-Governance stark mit der Branche zusammen und wird durch diese beeinflusst [WPG06]. Der fünfte Kontingenzfaktor sind die *regionalen Unterschiede* (V). So spielen beispielsweise Aspekte wie die lokalen Gesetze, Sprache oder nationale Informationsinfrastrukturen eine Rolle (vgl. [WT03], [PM12]). Ein weiterer Kontingenzfaktor ist der *Reifegrad* (VI). Die Messung des Reifegrads der IT-Governance ist ein Mittel, um deren Erfolg zu bewerten [DL07]. Durch die Nutzung solcher

Reifegradmodelle kann die eigene Implementierung mit Standards und Best Practices als Benchmark verglichen werden. So können Optimierungspotenziale und Handlungsempfehlungen abgeleitet werden [De20]. Zwischen dem Reifegrad und der Performance der IT-Governance existiert eine Korrelation [Si08]. Auch die *Strategie* ist ein relevanter Kontingenzfaktor (VII). Aspekte, die durch IT-Governance ermöglicht werden sollen, sind in diesem Zusammenhang Operational Excellence [LHP18] sowie Business-IT-Alignment [De20]. Das Business-IT-Alignment hat zum Ziel, die IT an die Bedürfnisse und Anforderungen des Geschäfts auszurichten und dadurch einen hohen IT-Geschäftswert zu generieren. Ein Beispiel, das die beiden Kontingenzfaktoren Reifegrad und Strategie miteinander kombiniert, sind die „Strategic Alignment Maturity Levels“ [Lu00]. Bei der *Ethik* (VIII) müssen Elemente wie ethische Grundsätze und Richtlinien sowie Sanktionen und Belohnungen berücksichtigt werden [LHP18]. Ein Ergebnis aus einer Interview-Studie zeigt, dass die Ethik ein wichtiger Aspekt bei der IT-Governance ist [MA10]. Zudem gibt es eine Wechselbeziehung zwischen der Governance und der ethischen Entscheidungsfindung in Projekten [Mü13]. Der Kontingenzfaktor *Vertrauen* (IX) ist zu berücksichtigen auf Individual-, Gruppen- und Systemebene [LHP18]. Vertrauen und Zusammenarbeit sind essenzielle Faktoren für erfolgreiche Governance [MP10].

3.3 Frameworks und Standards

Zur Unterstützung der Umsetzung von IT-Governance in Unternehmen existieren zahlreiche Frameworks und Standards, die Orientierungshilfen und Werkzeuge bereitstellen. Levstek et al. [LHP18] unterteilen existierende IT-Governance-Rahmenwerke in sieben Kategorien. Die erste Kategorie ist die *IT-Service-Delivery* (I), unter die die weit verbreiteten Frameworks COBIT, ITIL und CMMI fallen. Die zweite Framework-Kategorie ist die *IT Value Delivery* (II). Als Repräsentant dieser Kategorie kann Val IT genannt werden. Eine weitere Framework-Kategorie ist die *Informationssicherheit* (III). Levstek et al. [LHP18] führen hier als Beispiel den ISO-Standard 27001 auf. Dieser beschreibt, wie Organisationen Risiken der Informationssicherheit begegnen und ein Informationssicherheits-Managementsystem aufsetzen, pflegen und verbessern können. Als vierte Framework-Kategorie werden *Business Standards* (IV) aufgeführt. Diese Kategorie kann unterteilt werden in COSO und die Standards SSAE 16 und ISAE 3402. Die fünfte Framework-Kategorie ist das *Projektmanagement* (V). Ein Beispiel für ein solches Framework ist PMBOK. Ein Framework für einen prozessbasierten Ansatz für das Management beliebiger Projekte ist PRINCE2. Eine weitere Kategorie ist die *Leistungsmessung* (VI). Ein bekanntes Framework ist hier die IT-Balanced-Scorecard [LHP18]. Die siebte und letzte Kategorie wird unter *Allgemeines* (VII) aufgeführt. Hierzu zählt etwa das Framework Six Sigma für Prozessverbesserung und Qualitätsmanagement [LHP18].

3.4 Zusammenhang zwischen zentralen Themen der IT-Governance-Literatur

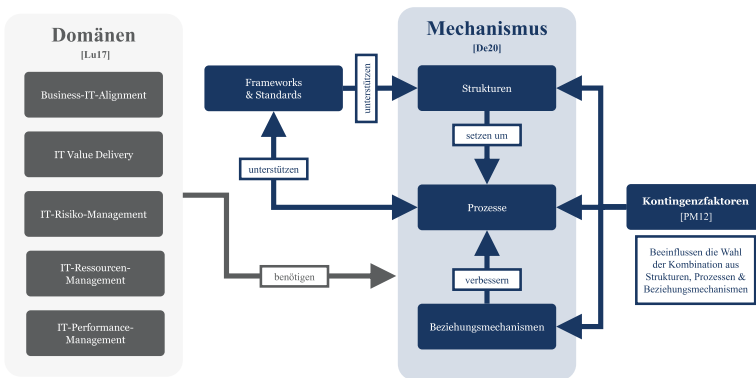


Abb. 1: IT-Governance Research Framework (eigene Darstellung)

Die eigene Darstellung (Abb. 1) zeigt ein aus den Ergebnissen der Literaturanalyse abgeleitetes „IT Governance Research Framework“ zur Darstellung des Zusammenhangs zwischen den wichtigsten Themengebieten in der IT-Governance-Literatur. Die Kontingenzfaktoren von Pereira und Mira da Silva [PM12] beschreiben den Kontext des Unternehmens und beeinflussen so die Auswahl der zu implementierenden Kombination aus Strukturen, Prozessen und Beziehungsmechanismen gemäß (beispielsweise) De Haes et al. [De20]. Bei den Strukturen handelt es sich dabei um Instanzen, welche für die Durchführung und Umsetzung der Prozesse verantwortlich sind. Diese wiederum können durch den Einsatz geeigneter Beziehungsmechanismen optimiert werden. Die Prozesse stehen in einem engen Zusammenhang zu Frameworks und Standards. Durch diese können Prozesse etwa implementiert, bewertet und kontinuierlich angepasst werden. Frameworks und Standards unterstützen somit Strukturen der IT-Governance bei der Umsetzung der Prozesse. Auf einer Metaebene sind die fünf Domänen der IT-Governance angesiedelt. Das Business-IT-Alignment beschreibt wie schon zuvor dargelegt die Passfähigkeit zwischen der strategischen Ausrichtung von Geschäft und IT. Bei der IT Value Delivery liegt der Fokus darin, die Auswahl und Durchführung von IT-Investitionen durch den zugrundeliegenden Geschäftswert zu lenken. Auch das IT-Risikomanagement ist eine wichtige Domäne, da mit der zunehmenden Abhängigkeit von Organisationen zu der IT die Wahrscheinlichkeit, dass externe Einflüsse die Unternehmensziele gefährden, signifikant ansteigt. Das IT-Ressourcen-Management gewährleistet, dass IT-Investitionen mit Zielen und Bedürfnissen einer Organisation kongruent sind und Ressourcen optimal zugewiesen werden. Das IT-Performance-Management beinhaltet schließlich die Messung und Bewertung der Performance mithilfe von Werkzeugen wie beispielsweise dem COBIT-Framework. Die Performance der IT-Governance-Domänen korreliert dabei mit der Effektivität der IT-Governance einer Organisation [Lu17]. Diese Domänen benötigen und nutzen Methoden und Werkzeuge aus der Klasse der IT-Governance-Mechanismen. Die Konzeptmatrix in Tab. 2 mit den 22 Forschungsbeiträgen zeigt den

aktuellen Forschungsstand im Bereich IT-Governance. Die Beiträge wurden hinsichtlich ihrer zentralen Themen (Mechanismen: *M*, Kontingenzfaktoren: *K*, Frameworks: und Standards: *F*, Domänen: *D*) sowie dem Vorhandensein von Anregungen für zukünftige Forschungsarbeiten (*Z*) kategorisiert. An der hohen Anzahl adressierter Forschungsbedarfe kann erkannt werden, dass es innerhalb des Themengebiets IT-Governance noch zahlreiche Problemstellungen gibt, die künftig zu untersuchen sind.

Referenz	M	K	F	D	Z	Referenz	M	K	F	D	Z
[APM13]	X		X		X	[MA10]			X		
[ARH14]	X				X	[Mü13]	X				X
[BG05]		X	X		X	[MZK03]			X	X	X
[Co10]			X	X	X	[NM16]			X		X
[De13]			X		X	[PM12]	X	X	X		X
[De20]	X	X	X	X	X	[Si08]			X		
[DL07]	X	X			X	[TLB17]					X
[DV04]	X		X			[VDG04]	X		X	X	
[DV09]	X		X		X	[VR17]			X		X
[LHP18]	X	X	X		X	[WPG06]	X		X	X	X
[Lu17]	X		X	X	X	[XLB08]	X	X			X

Tab. 2: Konzeptmatrix

4 Forschungsbedarf

Nachfolgend sollen identifizierte Forschungsbedarfe aus der Literatur vorgestellt werden. Als problematisch kann angesehen werden, dass der IT-Governance-Bereich mitsamt seinen Mechanismen und Kontingenzfaktoren nicht vollständig ausgereift ist (vgl. [Lu17], [TLB17]). Dabei ist zu prüfen, welche Mechanismen für welche Unternehmen besonders relevant sind in Abhängigkeit vom zugrundeliegenden Unternehmenskontext. Außerdem ist zu untersuchen, welche der Mechanismen zusammenwirken können oder müssen und welche sich nicht kombinieren lassen. Auch die Identifikation neuer Mechanismen soll Gegenstand zukünftiger Forschung sein [APM13]. Ebenfalls ist die Eignung bestimmter Kontingenzfaktoren basierend auf dem Unternehmenskontext näher zu erforschen [PM12]. Daraus lässt sich ein Bedarf an anpassungsfähigen Modellen ableiten, die für einen spezifischen Unternehmenskontext flexibel konfigurierbar sind. Zudem existiert keine einheitliche Definition für IT-Governance, was die Implementierung in der Praxis erschwert (vgl. [Lu17], [WPG06]). Ferner besteht der Bedarf, den Zusammenhang von IT-Governance und Corporate Governance durch sämtliche Unternehmensebenen hinweg zu untersuchen (vgl. [Lu17], [DL07]). Der Fokus der IT-Governance-Forschung bewegt

sich zudem hauptsächlich auf taktischer und operativer Managementebene, sodass die strategische Ebene tiefer erforscht werden sollte [TLB17]. Auch Debreceny [De13] präsentiert in seiner Forschungsarbeit Anregungen für zukünftige Forschungsarbeiten. Diese umfassen die Untersuchung des Zusammenhangs zwischen IT-Governance und der Wertgenerierung aus IT-Investments, Kriterien für eine erfolgreiche IT-Governance-Umsetzung unter Berücksichtigung des Zusammenspiels zwischen Business und IT, Einflüsse von technischen Neuentwicklungen wie Security Governance und Cloud Computing sowie der Zusammenhang zwischen IT-Governance und dem Risikomanagement. Zusammenfassend hält er fest, dass die IT-Governance-Forschung nicht ausgereift ist und mehr qualitative Forschungsarbeiten benötigt. Hervorgehoben wird von ihm der Bedarf nach detaillierten Fallstudien zu Ausrichtung, Erfolgen und Misserfolgen der IT-Governance. Darüber hinaus sollen Herausforderungen und Einführungsbarrieren für IT-Governance künftig nach unterschiedlichen Implementierungsstufen differenziert werden [NM16]. Ein Governance-Modell zur Analyse des Einflusses neuer Trends kann zudem helfen, sich ergebende Herausforderungen frühzeitig zu identifizieren [AG16]. Zukünftige Forschungsthemen sind ferner die Erforschung der Schlüsselvariablen für die einzelnen Mechanismen und der Einfluss effektiver Governance-Modelle auf die Unternehmensperformance [VR17].

5 Diskussion und Ausblick

Die vorliegende Arbeit präsentiert den State-of-the-Art der IT-Governance Forschung. Zentrale Themen der IT-Governance sind die Mechanismen, Kontingenzfaktoren sowie Frameworks und Standards. Um eine erfolgreiche Implementierung von IT-Governance in Unternehmen zu gewährleisten, müssen alle drei Themen berücksichtigt werden. Verschiedene Herausforderungen bei der Implementierung von IT-Governance gilt es zu beachten, die allen voran auf der Ausrichtung zwischen Business und IT und dem hohen Komplexitätsgrad existierenden Framework basieren. Die IT-Governance ist zudem noch nicht vollständig ausgereift, was auch daran ersichtlich wird, dass noch zahlreiche wichtige Forschungsbedarfe für das Themengebiet vorhanden sind. Die Definition von Mechanismen und Kontingenzfaktoren in der Literatur ist nicht abgeschlossen. Als ein großes Problem kann zudem die Auflistung von unterschiedlich vielen Mechanismen und Faktoren mit verschiedenen Benennungen erwähnt werden, was die Generierung eines gesamtheitlichen Überblicks deutlich erschwert. Das Forschungsgebiet der IT-Governance muss sich zudem um einheitliche Definitionen und Benennungskonventionen bemühen, sodass Forscher basierend auf einem konsolidierten Stand neue Forschungsaufgaben und Erkenntnisse ableiten können. Allgemeine Erkenntnisse gilt es in Zukunft für spezifische Kontexte zu untersuchen: in welcher Situation eignet sich welche Kombination von Mechanismen am besten? Hierfür muss eine konsolidierte Darstellung der zentralen Mechanismen und Kontingenzfaktoren aus der Literatur abgeleitet werden. Zusammenfassend kann die IT-Governance als kritischer Erfolgsfaktor für Unternehmen aufgefasst werden, der noch weiter erforscht werden muss.

Literatur

- [AG16] Awais, M.; Gill, A. Q.: Enterprise IT Governance: Back to Basics. In: ISD, Katowice, 2016.
- [APM13] Almeida, R.; Pereira, R.; Mira da Silva, M.: IT Governance Mechanisms. In (Falcao e Cunha, J.; Snene, M.; N6voa, H., Hrsg.): Exploring Services Science. Springer, S. 186-199, 2013.
- [ARH14] Aasi, P.; Rusu, L.; Han, S.: The Influence of Culture on IT Governance. In: HICSS, Waikaloa, 2014.
- [BG05] Brown, A. E.; Grant, G. G.: Framing the Frameworks. Commun. Assoc. Inf. Syst. 15/05, S. 696-712, 2005.
- [Co10] Cochran, M.: Proposal of an Operations Department Model to provide IT Governance in Organizations That Don't Have IT C-level Executives. In: HICSS, Waikoloa, 2010.
- [De13] Debreceeny, R. S.: Research on IT Governance, Risk, and Value: Challenges and Opportunities. J. Inf. Syst. 1/13, S. 129-135, 2013.
- [De20] De Haes, S.; Van Grembergen, W.; Joshi, A.; Huygh, T.: Enterprise Governance of Information Technology, 3. Auflage, Springer, 2020.
- [DL07] Dahlberg, T.; Lahdelma, P.: IT Governance Maturity and IT Outsourcing Degree. In: HICSS, Waikoloa, 2007.
- [DV04] De Haes, S.; Van Grembergen, W.: IT Governance and its Mechanisms. Information Systems Control Journal 1/04, S. 27-33, 2004.
- [DV09] De Haes, S.; Van Grembergen, W.: An Exploratory Study into IT Governance Implementations and its Impact on Business/IT Alignment. Inf. Syst. Manag. 2/09, S. 123-137, 2009.
- [Fe06] Fettke, P.: State-of-the-Art des State-of-the-Art. Wirtschaftsinformatik 4/06, S. 257-266, 2006.
- [LHP18] Levstek, A.; Hovelja, T.; Pucihar, A.: IT Governance Mechanisms and Contingency Factors. Organizacija 4/18, S. 286-310, 2018.
- [Lu00] Luftman, J.: Assessing Business-IT Alignment Maturity. Commun. Assoc. Inf. Syst. 4/00, 2000.
- [Lu17] Lunardi, G. L.; Becker, J. L.; Maçada, A. C. G.; Dolci, P. C.: The Impact of Adopting IT Governance on Financial Performance: An Empirical Analysis among Brazilian Firms. Int. J. Account. Inf. Syst. 1/17, S. 66-81, 2017.
- [MA10] Maidin, S. S.; Arshad, N. H.: IT Governance Practices Model in IT Project Approval and Implementation in Malaysian Public Sector. In: ICEIE, Kyoto, 2010.

- [MP10] Memiyanti, A. R.; Putera, M. S.: Ethical Leadership and Employee Trust: Governance Perspective. In: IEEE ICIFE, Chongqing, 2010.
- [Mü13] Müller, R.; Andersen, E. S.; Kvalnes, Ø.; Shao, J.; Sankaran, S.; Rodney Turner, J.; Biesenthal, C.; Walker, D.; Gudergan, S.: The Interrelationship of Governance, Trust, and Ethics in Temporary Organizations. *Proj. Manag. J.* 4/13, S. 26-44, 2013.
- [MZK03] Meyer, M.; Zarnekow, R.; Kolbe, L. M.: IT Governance: Begriff, Status quo und Bedeutung. *Wirtschaftsinformatik* 4/03, S. 445-448, 2003.
- [NM16] Nicho, M.; Muamaar, S.: Towards a Taxonomy of Challenges in an Integrated IT Governance Framework Implementation. *JITIM* 2/16, S. 1-32, 2016.
- [PM12] Pereira, R.; Mira da Silva, M.: A Literature Review: Guidelines and Contingency Factors for IT Governance. In: EMCIS, München, 2012.
- [Si08] Simonsson, M.; Johnson, P.; Ekstedt, M.; Rocha Flores, W.: IT Governance Decision Support using the IT Organization Modeling and Assessment Tool. *IJITM* 2/08, S. 167-189, 2008.
- [TLB17] Turel, O.; Liu, P.; Bart, C.: Board-Level Information Technology Governance Effects on Organizational Performance. *Inf. Syst. Manag.* 2/17, S. 117-136, 2017.
- [VDG04] Van Grembergen, W.; De Haes, S.; Guldentops, E.: Integration Strategies and Tactics for Information Technology Governance. In (Van Grembergen, W., Hrsg.): *Strategies for Information Technology Governance*. IGI Publishing, S. 1-36, 2004.
- [VR17] Vejseli, S.; Rossmann, A.: The Impact of IT Governance on Firm Performance. In: PACIS, Langkawi, 2017.
- [WPG06] Webb, P.; Pollard, C.; Ridley, G.: Attempting to Define IT Governance: Wisdom or Folly? In: HICSS, Kauai, 2006.
- [WT03] Weisinger, J. Y.; Trauth, E. M.: The Importance of Situating Culture in Cross-Cultural IT Management. *IEEE Trans. Eng. Manag.* 1/03, S. 26-30, 2003.
- [WW02] Webster, J.; Watson, R. T.: Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly* 2/02, S. xiii-xxiii, 2002.
- [XLB08] Xue, Y.; Liang, H.; Boulton, W. R.: Information Technology Governance in Information Technology Investment Decision Processes. *MIS Quarterly* 1/08, S. 67-96, 2008.

IT-Compliance - Themenschwerpunkte und Frameworks

Eine systematische Literaturanalyse zum State-of-the-Art

Sebastian Straßburg,¹ Sebastian Kahlert,² Helmut Beckmann³

Abstract: Bekannte Compliance Skandale aus der Vergangenheit, wie zum Beispiel die Datenschutzskandale bei Sony und Lidl, sorgten dafür, dass Compliance in Unternehmen sich von einem Nischenthema der Rechtsabteilung zu einer Kernaufgabe des Top-Managements entwickelt hat. Durch die voranschreitende Digitalisierung und die dadurch steigende Bedeutung und Abhängigkeit von Informationssystemen zur Unterstützung von Geschäftsprozessen steigt auch die Notwendigkeit, diese Systeme an gesetzliche und regulatorische Vorgaben und Richtlinien auszurichten. Diese Vorgaben stellen den Gegenstand der IT-Compliance dar. In den letzten Jahren ist die Zahl von Gesetzen, Vorgaben und Richtlinien im Rahmen der IT-Compliance gestiegen, was die Gefahr einer Unübersichtlichkeit bei der Erfüllung von Compliance-Anforderungen an die IT und so auch das Risiko für Regelverstöße erhöht. Aus diesem Grund liefert der vorliegende Beitrag in einer systematischen Literaturanalyse einen Überblick über die aktuellen Themenschwerpunkte und Frameworks der IT-Compliance, mit denen sich die Wissenschaft beschäftigt.

Keywords: IT-Compliance; State-of-the-Art; Frameworks; Rahmenwerk

1 Einleitung

Durch die steigende Bedeutung von Informationssystemen und die damit einhergehende steigende Anzahl an Gesetzen und Regeln, die erfüllt werden müssen, ist das Thema IT-Compliance für Unternehmen zunehmend unübersichtlicher geworden [SW12a, Le17]. So haben Unternehmen bereits Schwierigkeiten herauszufinden, welche IT-Compliance Anforderungen sie überhaupt erfüllen müssen und welches Framework für sie am besten geeignet ist [SW12a].

Die zentrale Forschungsfrage dieser Arbeit lautet demnach: **Welche Themenschwerpunkte können in der wissenschaftlichen Literatur im Bereich IT-Compliance identifiziert werden?** Dabei sollen mithilfe der Themenschwerpunkte der State-of-the-art ermittelt werden. Für die Beantwortung der Forschungsfrage wird zusätzlich die Teilfrage „*Welche Frameworks werden in der IT-Compliance gegenwärtig eingesetzt?*“ aufgestellt.

¹ Hochschule Heilbronn, Fakultät für WV, Max-Planck-Str. 39, 74081 Heilbronn, strassburg@stud.hs-heilbronn.de

² Hochschule Heilbronn, Fakultät für WV, Max-Planck-Str. 39, 74081 Heilbronn, skahlert@stud.hs-heilbronn.de

³ Hochschule Heilbronn, Fakultät für WV, Max-Planck-Str. 39, 74081 Heilbronn, helmut.beckmann@hs-heilbronn.de

Das Ziel der Literaturanalyse ist es, dem Leser, mithilfe der Themenschwerpunkte, eine Übersicht über den aktuellen Stand der Forschung und die eingesetzten IT-Compliance Frameworks zu geben. Der vorliegende Beitrag ist wie folgt gegliedert: In Kapitel 2 werden relevante Begriffe zum Thema erläutert, um eine einheitliche Ausgangslage zu schaffen. Kapitel 3 befasst sich mit der methodischen Vorgehensweise der Literaturanalyse. In Kapitel 4 werden die identifizierten Themenschwerpunkte, sowie die Frameworks präsentiert. Abschließend werden die Ergebnisse zusammengefasst und Implikationen für die weitere Forschung abgeleitet.

2 Grundlagen und Begriffe

2.1 IT-Compliance

Der Begriff Compliance wird in verschiedenen wissenschaftlichen Disziplinen verwendet. Im Kontext der IT-Compliance wird folgende Definition betrachtet: „IT-Compliance bezeichnet einen Zustand, in dem alle die IT des Unternehmens betreffenden und verbindlich vorgegebenen bzw. als verbindlich akzeptierten Vorgaben nachweislich eingehalten werden.“ [K114a, S. 20]. Dabei kann zwischen Compliance von IT und Compliance durch IT unterschieden werden [SW12a]. Bei Compliance von IT ist das Ziel, dafür zu sorgen, dass die im Unternehmen eingesetzten IT-Systeme die gesetzlichen und anderen regulatorischen Vorgaben erfüllen, während bei der Compliance durch IT dagegen IT-Systeme als Hilfsmittel zur Erfüllung der Vorgaben eingesetzt werden [SW12a]. Hierbei gibt es zahlreiche verschiedene Vorgaben, die sowohl nationalen, supranationalen als auch internationalen Charakter haben [SW12a].

2.2 Framework

Laut [WB11] wird der Begriff „Framework“ weder in der Wissenschaft noch in der Praxis präzise verwendet. Um die verschiedenen Regelwerke für die jeweiligen internen Unternehmensprozesse anzuwenden, sodass die Anforderungen für die IT-Compliance erfüllt werden, werden sogenannte Frameworks verwendet. Frameworks sind hierbei Rahmenwerke, die die verschiedenen rechtlichen Anforderungen und Methoden zur Erfüllung dieser Anforderungen beinhalten. Des Weiteren beinhalten die Frameworks auch Angaben zu den Anwendungsbereichen [WB11].

3 Methodik zur Literaturanalyse

Zur Erstellung der Literaturanalyse werden die Methoden von [WW02, Fe06] verwendet. [WW02] empfehlen bei der Erstellung einer Literaturanalyse die Verwendung eines konzeptzentrierten Organisations-Frameworks, dass den Vergleich verschiedener Literaturinhalte

ermöglicht [WW02]. [Br09] verlangen eine striktere Dokumentation und Darstellung des gesamten Suchprozesses bei wissenschaftlichen Arbeiten, da diese von anderen Forschern aufgegriffen werden. Für die Literatursuche wurden die Datenbanken AIS Electronic Library, Association for Computing Machinery Digital Library, IEEE Xplore Digital Library, Google Scholar, Springer Link und EBSCO Information Services, sowie die Konferenzen Americas Conference on Information Systems (AMCIS), European Conference on Information Systems (ECIS) und International Conference on Information Systems (ICIS) durchsucht. Die Suche beschränkte sich dabei auf Publikationen ab 2010, um nur möglichst aktuelle Themenschwerpunkte zu identifizieren. Dabei wurde in den Titeln, Abstracts und Keywords mit den folgenden Suchbegriffen gesucht: [{"IT-Compliance"}] und [{"IT-Compliance"}] AND "Framework" OR ("Management System") OR "Rahmenwerk"]. So konnten insgesamt 441 Publikationen identifiziert werden. Nachdem die Duplikate entfernt wurden, konnten die Abstracts von 223 Publikationen auf thematische Relevanz geprüft werden. Nach der Analyse des Abstracts, der Schlüsselwörter oder, wenn nötig, des vollständigen Artikels, blieben für unseren Forschungsschwerpunkt 41 relevante Publikationen übrig. Neben der Datenbanksuche empfehlen [WW02] die Durchführung einer Rückwärts- und Vorwärtsuche, wodurch die Anzahl der für diese Arbeit relevanten Publikationen auf 46 erhöht werden konnte.

4 Ergebnisse der Literaturanalyse

Tabelle 1 gibt einen Überblick über die identifizierten Themenschwerpunkte. Der Schwerpunkt IT-Outsourcing wird in 15 Arbeiten erwähnt, während IT-GRC in 12 und Frameworks in 11 Arbeiten vorkommen. Danach folgen Arbeiten zu den Anforderungen mit 11 und zum Mittelstand mit 5 Erwähnungen.

4.1 IT-GRC

Aus der Konzeptmatrix geht hervor, dass die IT-Compliance häufig im Kontext von IT – Governance, Risk und Compliance (IT-GRC) erwähnt wird [Da10, RWS10a, Ma15, KS17, VMM17, NKR17, JK20b]. Im Unternehmenskontext ist einen monolithischer Ansatz der drei Teilbereiche aufgrund von zunehmender Anzahl gesetzlicher Regelungen und Vorschriften, steigender Komplexität und Kosten nicht mehr möglich, sodass ein integrierter IT-GRC-Ansatz entwickelt wurde [GPS18]. In einem wechselseitigen Dreieck stellt die IT-Governance eine nachhaltige strategische Steuerung und Entwicklung der IT sicher, während das IT-Risikomanagement die systematische, zyklische Identifikation, Analyse, Bewertung und Behandlung der IT-Risiken übernimmt. Als drittes Element sorgt die IT-Compliance für die Einhaltung aller internen und externen Vorgaben im Kontext der IT [Kn13, Ig16].

Literatur	IT-GRC	IT-Out-sourcing	Anforderungen	Mittelstand	Framework
[ABT15]			X		
[Ad19]		X			
[APT17]		X			
[BN14]		X			
[Bi10]			X		
[BHP10]					X
[Da10]	X				
[DR20]		X		X	
[DRS10]		X		X	
[EG14]		X			
[Fa12]		X			X
[GPS18]	X				X
[Gu10]			X		
[HH14]		X			
[Ha17]			X		
[Ig16]	X				
[JK20a]	X		X		
[JK20b]	X			X	
[KEU17]		X			
[K112]			X		
[K113]			X		
[K114b]					X
[K119]					X
[KB15]			X		
[Kn13]	X				
[KS17]	X				X
[Ko18]				X	
[Le20]			X		
[MT11]		X			
[Ma15]	X				
[NKR17]	X				X
[OR13]		X			
[Pa11]		X			
[PTT15]		X		X	
[PD13]					X
[PL12]		X			
[RWS10b]	X				
[RWS10a]	X				
[RSB13]					X
[RW11]		X			
[SW12a]			X		
[SW12b]			X		
[SW13]		X			
[VMM17]	X				
[WB11]					X

Tab. 1: Identifizierte Themenschwerpunkte

4.2 IT-Outsourcing

In der betrachteten Literatur wird IT-Outsourcing allgemein, sowie auch Cloud Computing, als besondere Form des IT-Outsourcings, behandelt. Beim IT-Outsourcing handelt es sich um eine alternative Möglichkeit, IT-Compliance im Unternehmen sicher zu stellen, indem IT-Prozesse und somit auch IT-Compliance-Aufgaben auf externe Dienstleister übertragen werden [PL12, DR20]. In der Literatur werden dabei vor allem die damit verbundenen Risiken betrachtet [Fa12, SW13]. So fehlt es zum Beispiel beim IT-Dienstleister an Loyalität und an Kenntnisse über die Unternehmenskultur, IT-Systeme und internen Prozesse [Fa12, SW13]. Vor allem muss die Geschäftsleitung daran denken, dass auch beim Outsourcing von IT-Compliance-Aufgaben die Risiken sowie die Verantwortung weiterhin beim eigenen Unternehmen bleiben [Pa11, DR20].

Beim Cloud Computing beschäftigen sich die Autoren ebenfalls mit der Vielzahl von Risiken für die IT-Sicherheit und Compliance, welche die Reputation von Unternehmen verletzen und wirtschaftliche Schäden nach sich ziehen können [BN14, HH14, APT17, KEU17, Ad19]. Trotzdem sind sich die Autoren einig, dass Cloud Computing kein Hype-Thema ist und dauerhaft Geschäftsmodelle, IT-Strukturen und Unternehmensprozesse nachhaltig verändern wird [OR13, BN14]. Vor allem können hier mittelständische Unternehmen viel Geld durch Personaleinsparungen sparen [DRS10, DR20].

4.3 Anforderungen

Eine grundlegende Aufgabe der IT-Compliance ist es zunächst festzulegen, welche Vorgaben als relevant angesehen werden, also welche Regelwerke mit IT-Bezug zu berücksichtigen sind [K113]. Diese lassen sich in die folgenden drei Gruppen identifizieren. Die erste Gruppe umfasst rechtliche Vorgaben, wie zum Beispiel aus Gesetzen anderer Staaten und Rechtsakten der Europäischen Union (EU) [SW12a, K113]. Außerdem gibt es unternehmensexterne Regelwerke, Normen, wo auch Frameworks und Standards, wie die „Information Technology Infrastructure Library“ (ITIL) und die „Control Objectives for Information and Related Technology“ (COBIT) dazugehören [SW12a, K113]. Neben Vorgaben, die von außerhalb des Unternehmens kommen, gibt es zudem auch unternehmensinterne Regelwerke, wie z. B. IT-Richtlinien und -Hausstandards und innerbetriebliche Verhaltenskodizes [K113]. Darüber hinaus gibt es auch den Datenschutz der oftmals als größeres Thema in der Literatur behandelt wird. Durch die Datenschutzgrundverordnung (DSGVO) ist jedes Unternehmen, das in der EU tätig ist, zum Datenschutz verpflichtet [Ha17]. Dabei liegt es im Interesse des Unternehmens, Verstöße gegen solche Richtlinien und Gesetze zu verhindern, um drohende Bußgelder und Strafen zu entgehen [ABT15]. Das Zusammenspiel von IT-Compliance und Datenschutz kann in Unternehmen laut [KB15] Probleme aufwerfen, da die konkreten systemtechnischen und organisatorischen Anforderungen nicht klar bestimmt sind. So gibt es zum Beispiel keine klare Aussage dazu, wie viele Mitarbeiter in einem Unternehmen mit der IT-Compliance betraut sein sollten oder ob es einen IT-Compliance-Beauftragten

oder gar eine IT-Compliance Abteilung geben muss [Bi10, KB15]. Die Problematik, die in Verbindung zu den einzuhaltenden Anforderungen in der Literatur behandelt wird ist, dass ihre steigende Anzahl zu höheren Gefahren von Intransparenzen bei ihrer Erfüllung führen kann [SW12a]. Aus diesem Grund ist, aus Sicht der Autoren, die Kenntnis aller geltenden Anforderungen essenziell, um mögliche Regelverstöße und damit verbundenen Strafen zu vermeiden [Gu10, K112, SW12b].

4.4 Mittelstand

Im Hinblick auf die Umsetzung von IT-Compliance wird in der betrachteten Literatur überwiegend auf den Mittelstand eingegangen. Und zwar wird hier vor allem die Vernachlässigung von IT-Compliance Maßnahmen erwähnt [DR20]. Während große Unternehmen die benötigten Ressourcen besitzen, um die betriebsinternen Prozesse konform an die Vorgaben der IT-Compliance anzupassen, fehlen diese meistens bei den KMUs [DR20]. Zusätzlich ist laut [Ko18], die sogenannte IT-Compliance-Kultur hier eher schwach ausgeprägt, was vor allem an einer anderen „Gesellschafterstruktur“ liegt [Pa11]. Darüber hinaus wird die Realisierung von Compliance-Maßnahmen oftmals nicht mit allen Beteiligten des Unternehmens gemeinsam, sondern separat in einzelnen Abteilungen, durchgeführt. Ebenso sorgt eine mangelhafte Kommunikation zu einer erschwerten Integrierung und Akzeptanz neuer IT-Compliance Vorgaben und verhindert eine erfolgreiche Umsetzung von IT-Compliance in mittelständischen Unternehmen [Pa11].

4.5 Frameworks

Literatur	Identifizierte Frameworks						
	COBIT	ITIL	IT-Grundschutz	ISO/IEC 27002	COSO	CMMI	Sonstige
[BHP10]							X
[DR20]	X	X		X	X		
[GPS18]							X
[K114a]	X						
[K119]	X						
[KS17]	X					X	
[NKR17]	X	X		X			X
[PTT15]							X
[PD13]	X	X					
[RSB13]	X						
[WB11]	X	X	X	X			
Σ	8	4	1	3	1	1	4

Tab. 2: Identifizierte IT-Compliance Frameworks

Die folgende Konzeptmatrix in Tabelle 2 zeigt die, in der betrachteten Literatur, identifizierten IT-Compliance Frameworks. Wie aus der Konzeptmatrix zu entnehmen ist, wird vor allem COBIT als mögliches IT-Compliance Framework erwähnt. COBIT beinhaltet ein Prozessmodell mit international akzeptierten IT prozessbezogenen Kontrollzielen, die umgesetzt und erfüllt werden sollten, um eine verlässliche Anwendung der Informationstechnologie zu gewährleisten [WB11]. Dabei werden verschiedene Versionen, wie COBIT 4.0 und COBIT 5.0 erwähnt, wobei in COBIT 5.0 die Compliance-Thematik besonders erweitert wurde [K114a, K119]. Außerdem sind auch Adaptionen von COBIT möglich [RSB13]. Eine weitere etablierte Lösung ist ITIL, eine Sammlung von Best-Practices im Bereich IT-Service-Management, die sich mittlerweile als De-facto-Standard im IT-Service Management etabliert hat [WB11]. Auch der von der International Organization for Standardization entworfene Standard ISO/IEC 27002 wird von einigen Autoren erwähnt [PD13, NKR17, DR20]. Neben diesen drei Lösungen gibt es auch den IT-Grundschutz als Lösung, die vom Bundesamt für Sicherheit in der Informationstechnik (BSI) veröffentlicht wurde, COSO, ein Rahmenwerk zur Entwicklung des eigenen Risikomanagements oder auch das CMMI-Framework, welches ebenfalls im Kontext von IT-Compliance angewandt werden kann [WB11, KS17, DR20]. Darüber hinaus werden in der Literatur auch vollständig neue Lösungen entwickelt und präsentiert [BHP10, PTT15, GPS18].

5 Diskussion und Ausblick

Die vorliegende Arbeit zeigt, dass die IT-Compliance häufig im Rahmen von IT – GRC und IT-Outsourcing behandelt wird. Darüber hinaus beschäftigt sich vor allem die jüngere Literatur mit den Herausforderungen und Schwierigkeiten, mit denen mittelständische Unternehmen bei der Umsetzung von IT-Compliance konfrontiert werden. Ein großer Teil der Literatur befasst sich ebenfalls mit verschiedenen IT-Compliance Frameworks, in denen, neben Eigenlösungen, vor allem bereits anerkannte und etablierte Frameworks wie COBIT und ITIL erwähnt werden. Die Aussagekraft dieses Papiers ist jedoch begrenzt. Aufgrund der begrenzten Zeit und des großen Umfangs des Themas konnte nur eine begrenzte Anzahl von Datenbanken und Konferenzen im Rahmen der Literaturrecherche durchsucht werden. Dies bedeutet, dass die Möglichkeit besteht, dass weitere empirische Forschung, die eine größere Anzahl von Quellen betrachtet, zusätzliche Themenschwerpunkte und Frameworks identifizieren könnte. Darüber hinaus könnte die Nutzung weiterer Forschungsmethoden, wie z. B. Experteninterviews oder Umfragen, ebenfalls neue Themen und Frameworks identifizieren, die möglicherweise noch nicht in der Literatur erwähnt werden.

Da sich ein Großteil der Literatur zum Thema IT-Outsourcing mit den drohenden Risiken auseinandersetzt, ist für die zukünftige Forschung im Bereich der IT-Compliance der Fokus auf die Risikominderung beim IT-Outsourcing zu empfehlen. Außerdem stellt die Implementierung von IT-Compliance Frameworks bei mittelständischen Unternehmen ein Problem dar. Hier fehlt es auch an empirischen Untersuchungen, die sich mit den Gründen und Motiven befasst. Aus diesem Grund sind auch hier weitere Forschungen nötig.

Literaturverzeichnis

- [ABT15] Abawi, D. F.; Bausen, K.; Terrasi, D.: Compliance-konforme Ermittlungsverfahren in IuK-Systemen. *Wirtschaftsinformatik & Management*, (4):48–55, 2015.
- [Ad19] Adam, C. et al.: Cognitive Compliance: Analyze, Monitor and Enforce Compliance in the Cloud. In: *Proceedings of CLOUD 2019*. S. 234–242, 2019.
- [APT17] Adelmeyer, M.; Petrick, C.; Teuteberg, F.: IT-Risikomanagement von Cloud-Dienstleistungen im Kontext des IT-Sicherheitsgesetzes. *HMD Praxis der Wirtschaftsinformatik*, (1):111–123, 2017.
- [BHP10] Bonazzi, R.; Hussami, L.; Pigneur, Y.: Compliance Management is Becoming a Major Issue in IS Design. In: *Information Systems: People, Organizations, Institutions, and Technologies*, S. 391–398. 2010.
- [Bi10] Bierekoven, C.: IT-Compliance für nationale Unternehmen - die wachsende Herausforderung. In: *Steuerung der IT im Klinikmanagement: Methoden und Verfahren*. 2010.
- [BN14] Bagban, K.; Nebot, R.: Governance und Compliance im Cloud Computing. *HMD Praxis der Wirtschaftsinformatik*, (3):267–283, 2014.
- [Br09] vom Brocke, J. et al.: Reconstructing the giant: On the Importance of Rigour in Documenting the Literature Search Process. In: *Proceedings of ECIS 2009*. 2009.
- [Da10] Dameri, R. P.: From IT Compliance Cost to IT Governance Benefits: An Italian Business Case. In: *Management of the Interconnected World*, S. 317–324. 2010.
- [DR20] Deistler, N.; Rentrop, C.: IT-Compliance in KMU – State of the art. *HMD Praxis der Wirtschaftsinformatik*, (5):1047–1057, 2020.
- [DRS10] Doelitzscher, F.; Reich, C.; Sulistio, A.: Designing Cloud Services Adhering to Government Privacy Laws. In: *Proceedings of CIT 2010*. S. 930–935, 2010.
- [EG14] El-Gazzar, R. F.: A Literature Review on Cloud Computing Adoption Issues in Enterprises. In: *Proceedings of TDIT 2014*. S. 214–242, 2014.
- [Fa12] Falk, M.: *IT-Compliance in der Corporate Governance*. 2012.
- [Fe06] Fettke, P.: State-of-the-Art des State-of-the-Art: Eine Untersuchung der Forschungsmethode "Reviewinnerhalb der Wirtschaftsinformatik". *Wirtschaftsinformatik*, (4):257–266, 2006.
- [GPS18] Govindji, S.; Peko, G.; Sundaram, D.: A Context Adaptive Framework for IT Governance, Risk, Compliance and Security. S. 14–24, 2018.
- [Gu10] Guggenberger, J. M.: *Aufbau und Ablauf einer IT-Integration*. 2010.
- [Ha17] Hamann, U.: In der Schokoladenfabrik: IT-Sicherheit spielerisch verbessern. *Wirtschaftsinformatik & Management*, (4):40–47, 2017.
- [HH14] Haas, A.; Hofmann, A.: Risiken aus der Nutzung von Cloud-Computing-Diensten: Fragen des Risikomanagements und Aspekte der Versicherbarkeit. *ZVersWiss*, (4):377–407, 2014.

- [Ig16] Iguer, H. et al.: Including EAS-SGR IT Risk Framework in an IT GRC Global Framework. In: Proceedings of UNet 2015. S. 525–533, 2016.
- [JK20a] Johannsen, A.; Kant, D.: IT-Governance, Risiko- und Compliance-Management (IT-GRC) – Ein Kompetenz-orientierter Ansatz für KMU. HMD Praxis der Wirtschaftsinformatik, (5):1058–1074, 2020.
- [JK20b] Johannsen, A.; Kant, D.: IT-Governance, Risiko- und Compliance-Management (IT-GRC) für KMU – Literaturanalyse und Ansatzbildung. Ergebnisse einer Studie durch das Institut für Betriebliche Integration und Digitalisierung (IBID). 2020.
- [KB15] Kniese, R.; Bühlmann, O.: IT-Compliance als Teil der Corporate Governance. Wirtschaftsinformatik & Management, (4):34–47, 2015.
- [KEU17] Kalaiprasath, R.; Elankavi, R.; Udayakumar, R.: Cloud Security and Compliance - A Semantic Approach in End to End Security. International Journal on Smart Sensing and Intelligent Systems, (4):482–494, 2017.
- [K112] Klotz, M.: Regelwerke der IT-Compliance - Klassifikation und Übersicht, Teil 1: Rechtliche Regelwerke. SIMAT Arbeitspapiere, (04-12-020):1–80, 2012.
- [K113] Klotz, M.: Regelwerke der IT-Compliance - Klassifikation und Übersicht, Teil 2: Normen. SIMAT Arbeitspapiere, (05-13-024):1–49, 2013.
- [K114a] Klotz, M.: IT-Compliance: Begrifflichkeit und Grundlagen. SIMAT Arbeitspapiere, (06-14-028):1–75, 2014.
- [K114b] Klotz, M.: IT-Compliance nach COBIT: Gegenüberstellung zwischen COBIT 4.0 und COBIT 5. SIMAT Arbeitspapiere, (06-14-025):1–50, 2014.
- [K119] Klotz, M.: IT-Compliance nach COBIT 2019. SIMAT Arbeitspapiere, (11-19-034):1–62, 2019.
- [Kn13] Knoll, M.: Sicherstellung einer ordnungsgemäßen IT. HMD Praxis der Wirtschaftsinformatik, (1):6–19, 2013.
- [Ko18] Kolek, E.: IT-Sicherheit der Digitalisierung in Kleinen und Mittleren Unternehmen: Eine literaturbasierte und empirische Studie von Effekten und Barrieren. In: Proceedings of MKWI 2018. S. 1706–1717, 2018.
- [KS17] Knoll, M.; Strahringer, S.: IT-GRC-Management im Zeitalter der Digitalisierung. 2017.
- [Le17] Legner, C. et al.: Digitalization: Opportunity and Challenge for the Business and Information Systems Engineering Community. Business & Information Systems Engineering, (4):301–308, 2017.
- [Le20] Leeser, D. C.: Digitalisierung in KMU kompakt. 2020.
- [Ma15] Mayer, N. et al.: An ISO Compliant and Integrated Model for IT GRC (Governance, Risk Management and Compliance). In: Proceedings of EuroSPI 2015. S. 87–99, 2015.
- [MT11] Martens, B.; Teuteberg, F.: Towards a Reference Model for Risk and Compliance Management of IT Services in a Cloud Computing Environment. In: Application Management, S. 135–163. 2011.

- [NKR17] Nicho, M.; Khan, S.; Rahman, M.S.M.K.: Managing Information Security Risk Using Integrated Governance Risk and Compliance. In: Proceedings of ICCA 2017. S. 56–66, 2017.
- [OR13] Opala, O. J.; Rahman, S. M.: An Exploratory Analysis of the Influence of Information Security on the Adoption of Cloud Computing. In: Proceedings of SOSE 2013. S. 165–170, 2013.
- [Pa11] Paulus, S.: Outsourcing von GRC-Aufgaben im Mittelstand. Risk, Fraud & Compliance, (1):8, 2011.
- [PD13] Pereira, R.; Da Silva, M. M.: IT Compliance Management Process Modeling based on Best Practices Reference Models and Qualitative Data. In: Proceedings of EDOC 2013. S. 178–187, 2013.
- [PL12] Puksta, K.; Laurins, P.: Outsourcing Document Management System/Application as Key Factor for Compliant Process Implementation and Increasing Management Efficiency. Regional Formation and Development Studies, (8):167–176, 2012.
- [PTT15] Pekker, W.; Tannhäuser, C.; Turowski, K.: Best Practices als Treiber der Prozessharmonisierung am Beispiel des ISO-19011-Standards. HMD Praxis der Wirtschaftsinformatik, (1):150–160, 2015.
- [RSB13] Ramos, E. C.; Santoro, F. M.; Baião, F.: Watch Out and Improve IT: Adapting COBIT 5.0 Framework Based on External Context Discovery. In: Proceedings of IC3K 2012. S. 426–439, 2013.
- [RW11] Reimer, H.; Wegener, C.: IT Compliance und Governance: Cloud Computing erweitert den Horizont. Datenschutz und Datensicherheit, (5):363–363, 2011.
- [RWS10a] Racz, N.; Weippl, E.; Seufert, A.: A Frame of Reference for Research of Integrated Governance, Risk and Compliance (GRC). In: Proceedings of IFIP CMS 2010. S. 106–117, 2010.
- [RWS10b] Racz, N.; Weippl, E.; Seufert, A.: A Process Model for Integrated IT Governance, Risk, and Compliance Management. In: Proceedings of DB&IS 2010. S. 1–15, 2010.
- [SW12a] Strasser, A.; Wittek, M.: IT-Compliance. Informatik-Spektrum, 35(1):39–44, 2012.
- [SW12b] Strasser, A.; Wittek, M.: IT-Compliance-Officer: Verantwortung, Aufgaben, Befugnisse und Qualifikationen. Wirtschaftsinformatik & Management, (4):66–71, 2012.
- [SW13] Strasser, A.; Wittek, M.: Dienstleistungsspektrum der IT-Compliance. HMD Praxis der Wirtschaftsinformatik, 50(4):33–41, 2013.
- [VMM17] Vunk, M.; Mayer, N.; Matulevičius, R.: A Framework for Assessing Organisational IT Governance, Risk and Compliance. In: Proceedings of SPICE 2017. S. 337–350, 2017.
- [WB11] Weber, S.; Bartmann, D.: Standards/IT-Frameworks zur Sicherstellung der IT-Compliance Konformität in hochflexiblen Geschäftsprozessen. In: Dienstorientierte IT-Systeme für hochflexible Geschäftsprozesse, S. 313–339. 2011.
- [WW02] Webster, J.; Watson, R. T.: Analyzing the Past To Prepare for the Future : Writing a Literature Review. MIS Quarterly, (2):13–23, 2002.

Workshop: Stand und den
Herausforderungen des
Geschäftsprozessmanagements
(ZuGPM)

7. Workshop zum Stand und den Herausforderungen des Geschäftsprozessmanagements

Ralf Laue¹, Jana-Rebecca Rehse² und Thorsten Schoormann³

Der Workshop zum Stand, den Herausforderungen und Impulsen des Geschäftsprozessmanagements (ZuGPM) findet 2021 zum siebenten Male im Rahmen der GI-Jahrestagung INFORMATIK statt. Für den Workshop konnten Beiträge eingereicht werden, die Forschungsergebnisse vorstellen, bereits veröffentlichte Arbeiten zusammenfassen oder neue Forschungsideen in einer frühen Phase diskutieren. Darüber hinaus gibt es auch in diesem Jahr wieder das schon in den vergangenen Jahren bewährte Freistil-Format – die Möglichkeit, Themen auch auf andere Weise als den üblichen Vortrag mit anschließender Frage-Antwort-Runde zu präsentieren.

Beim ZuGPM-Workshop 2021 können wir ein Jubiläum begehen, das vermutlich den wenigsten Teilnehmern bewusst ist: Vor genau 100 Jahren wurde durch das US-amerikanische Ehepaar Gilbreth die erste graphische Prozessmodellierungssprache vorgeschlagen. In einem eingeladenen Auftaktvortrag stellt Ralf Laue diese historische Modellierungssprache vor und vergleicht die Modellierung des Jahres 1921 mit der des Jahres 2021.

Das Fachprogramm zeichnet sich, wie das auch in den Vorjahren der Fall war, durch ein breites Spektrum an behandelten Themen aus.

Zunächst folgt ein Diskussionsbeitrag von Joachim Stocker, Nils Herda und Jan Jürjens zu Widerstandsfähigkeit von Prozessen gegenüber dem Ausfall von Ressourcen.

Anschließend werden Beiträge vorgestellt, die sich im vorliegenden Konferenzband finden. Als Ergebnis des Begutachtungsprozesses wurden fünf Langbeiträge und zwei Kurzbeiträge zur Veröffentlichung im Tagungsband angenommen.

Thomas Bauer entwickelte seine bereits in den Workshops der vergangenen Jahre thematisierten Gedanken zur Modellierung menschlicher Ressourcen in Geschäftsprozessmodellen weiter. Er thematisiert insbesondere die Modellierung von Eskalationen und Stellvertretungen. Ebenfalls mit der Zuordnung menschlicher Ressourcen zu Aufgaben in Geschäftsprozesse befasst sich der Beitrag von Felix Holz, Christoph Gibcke, Sven Erdmann, Michael Fellmann und Birger Lantow. Ihre Literaturstudie untersucht, inwiefern menschliche Faktoren wie Wohlbefinden und Gesundheit der Mitarbeiter bei der Aufgabenzuweisung in der vorhandenen Literatur berücksichtigt werden.

¹ Westsächsische Hochschule Zwickau, ralf.laue@fh-zwickau.de

² Universität Mannheim, rehse@uni-mannheim.de

³ Universität Hildesheim, schoor@uni-hildesheim.de

Lars Drewes und Volker Nissen stellen die Ergebnisse einer online durchgeführten experimentellen Studie zum Einfluss bürokratischer Hürden auf die Prozessakzeptanz vor.

Eine Literaturlauswertung von Fabian Sauer, Manuel Wittig und Helmut Beckmann behandelt Prozesse im Bereich des IT-Service-Managements. Eines der Ergebnisse der Literaturlauswertung ist, dass Cloud-Technologien im IT-Service-Management zunehmend an Bedeutung gewinnen. Mit diesen Technologien befassen sich auch Klaus Lehner und Sebastian Floercke. Sie identifizierten 16 Faktoren für den Erfolg von cloudbaiserten Geschäftsprozessmanagement-Lösungen.

Der Beitrag von Clemens Schreiber diskutiert die Herausforderung, die Variabilität von Prozessen zu messen, wobei auch andere Aspekte als Unterschiede im Kontrollfluss einbezogen werden.

Stephan Leible, Marc Frerichs und Markus Nüttgens entwickeln die Vision einer Plattform, die das Arbeiten mit Modellen unabhängig von Werkzeugen und Modellierungsmethoden erlaubt.

Der Workshop wäre nicht möglich gewesen ohne die Unterstützung durch das Organisationsteam der Jahrestagung INFORMATIK 2021.

Ein besonderer Dank geht darüber hinaus auch an das ZuGPM-Programmkomitee, das den Workshop mit sorgfältig erstellten Gutachten unterstützt hat:

Michael Fellmann, Universität Rostock

Matthias Goeken, Hochschule der Deutschen Bundesbank

Constantin Houy, Deutsches Forschungszentrum für Künstliche Intelligenz

Kathrin Kirchner, Dänemarks Technische Universität

Agnes Koschmider, Universität Kiel

Birger Lantow, Universität Rostock

Luise Pufahl, Technische Universität Berlin

Kurt Sandkuhl, Universität Rostock

Rainer Schmidt, HAW München

Alle im Workshop vorgestellten Themen sind bestens geeignet, um eine rege Diskussion anzuregen. Wir wünschen daher allen Teilnehmern des digitalen Workshops viele neuen Erkenntnisse, spannende Diskussionen und interessante Kontakte – auf dass sich inspirierende und fruchtbare Gespräche auch nach dem Workshop fortsetzen mögen.

September 2021

Ralf Laue, Jana-Rebecca Rehse und Thorsten Schoormann

100 Jahre graphische Geschäftsprozessmodellierung

Eingeladener Vortrag

Ralf Laue¹

Abstract: 1921 hielten Frank B. and Lillian Moller Gilbreth bei der jährlichen Tagung der American Society of Mechanical Engineers eine Präsentation unter dem Titel „Process Charts“ [GG21]. Sie stellten eine graphische Notation vor, mit der Arbeitsabläufe erfasst und vom Management diskutiert werden können. Das Ehepaar Gilbreth gehörte damit zu den Pionieren einer Bewegung, die unter dem Namen „Scientific Management“ bekannt wurde.

In Deutschland wurden die von den Gilbreths vorgeschlagenen Darstellungen von Fritz Nordsieck unter den Bezeichnungen „Arbeitsablaufplan“ und „Zergliederungsplan“ bekanntgemacht [N32].

Neben einer graphischen Notation stellt der Beitrag von Frank B. and Lillian Moller Gilbreth auch eine Methode der Anwendung der als Process Charts bezeichneten Prozessvisualisierungen vor. Im Vortrag wird gezeigt, wie aktuell einige der Aussagen in dieser Methodenbeschreibung auch einhundert Jahre später noch sind. Andererseits wird auch aufgezeigt, wie sich die Prozessmodellierung weiterentwickeln musste, insbesondere durch die Ersetzung von Hand- durch Maschinenarbeit, die Unterstützung von Prozessen durch Rechentechnik sowie durch die Zunahme von Prozessen, deren Ausführung Wissen und Kreativität erfordert.

Keywords: Geschäftsprozessmodellierung, Process Charts

Literaturverzeichnis

- GG21 Gilbreth, F. B.; Gilbreth, L.M.: Process Charts, The American Society of Mechanical Engineers, presented at the Annual Meeting, December 5 to 9, 1921.
- N32 Nordsieck, F.: Die schaubildliche Erfassung und Untersuchung der Betriebsorganisation, C. E. Poeschel Verlag, Stuttgart, 1932.

¹ Westsächsische Hochschule Zwickau, ralf.laue@fh-zwickau.de

Effiziente Modellierung von Eskalationen und Stellvertretungen in Geschäftsprozessen

Thomas Bauer¹

Abstract: Eskalationen und Stellvertretungen sind wichtig, um bei der Ausführung von Geschäftsprozessen (GP) auf Ausnahmesituationen reagieren zu können. Hierbei erfordert die Festlegung der jeweils gewünschten Zielperson(en) ähnlich komplexe Regeln, wie die Zuordnung von Aktivitäten zu regulären Bearbeitern. Zusätzlich sind unterschiedliche Regeln für verschiedene Aktivitäten, Originalbearbeiter, etc. erforderlich. Sehr viele solche Regeln zu definieren, würde aber einen großen Aufwand für die GP-Designer bedeuten. Deswegen sollte eine Regel für mehrere Aktivitäten verwendbar sein. Dies wird möglich, indem sie sich auf die reguläre Bearbeiterzuordnung der betroffenen Aktivität beziehen kann, z.B. auf die dort festgelegte Rolle. Da in einer Bearbeiterzuordnung mehrere Rollen gefordert sein können, muss zusätzlich festgelegt werden, welche von diesen referenziert werden soll. Deshalb führen wir für Objekte des Organisationsmodells einen benutzerdefinierbaren SubType ein. Dieser kann bei der Spezifikation der Zielpersonen einer Eskalation oder Stellvertretung verwendet werden, um die gewünschte Rolle auszuwählen. Dadurch wird eine solche Regel für viele unterschiedliche Aktivitäten verwendbar.

Keywords: Geschäftsprozess, Organisatorische Perspektive, Eskalation, Stellvertretung.

1 Motivation

Die organisatorische Perspektive von Prozess Management Systemen (PMS) definiert, welche Personen eine bestimmte Aktivität in einem Geschäftsprozess (GP) bearbeiten sollen (mittels einer sog. Bearbeiterzuordnung) und auch, wie sich das PMS in Ausnahmesituationen verhalten soll. [BL20] stellt hierzu Ansätze aus der Wissenschaft vor, und zeigt, dass heutzutage einige Aspekte in kommerziellen PMS nur unzureichend umgesetzt sind. So ist es i.d.R. nicht möglich, mächtige Bearbeiterzuordnungen auf einfache (z.B. graphische) Art und Weise zu definieren.

Startet oder beendet ein Bearbeiter eine Aktivität nicht innerhalb der vorgegebenen Frist, so wird eine *Eskalation*² ausgelöst. Dann wird z.B. ein Vorgesetzter informiert oder die Aktivität wird einem anderen Benutzer zugeordnet. Ist ein Bearbeiter längere Zeit abwesend (z.B. in Urlaub), so legt eine *Stellvertreterregelung* fest, wer diese Aktivität stattdessen bearbeiten soll. Eskalationen und Stellvertretungen haben gemeinsam, dass zusätzliche Personen involviert werden sollen. Diese Zielpersonen müssen in beiden Fällen abhängig von dem Typ der betroffenen Aktivität, dem Originalbearbeiter und dem

¹ Hochschule Neu-Ulm, Wileystr. 1, 89231 Neu-Ulm, thomas.bauer@hnu.de

² Im Folgenden ist mit Eskalation ausschließlich ein solches Verhalten gemeint, d.h. Aktionen, die ausgelöst werden, wenn der Start oder die Beendigung einer Aktivität des GP nicht rechtzeitig erfolgt.

Prozesskontext definiert werden können. Deshalb bestehen für beide Themen ähnliche Anforderungen. In [BL20] wurde aufgezeigt, dass für die Modellierung solcher Zielpersonen, aktuell keine geeigneten wissenschaftlichen Konzepte existieren, und dass Eskalationen und Stellvertretungen in kommerziellen PMS heutzutage oft nur rudimentär umgesetzt sind. Deshalb beschäftigt sich dieser Beitrag mit dieser Problemstellung.

Das Problem hierbei ist, dass die Festlegung der Regeln zur Spezifikation der Zielpersonen (Target Person Rule: TPR) ohne einen zu großen Aufwand für die GP-Designer erfolgen soll (d.h. effizient). Das ist nur erreichbar, wenn möglichst wenige solche Regeln erforderlich sind. Hierzu muss eine einzelne TPR für viele unterschiedliche Originalbearbeiter und Aktivitäten gültig sein. Dies ist z.B. sehr einfach möglich, wenn die Zielpersonen einer Eskalation oder Stellvertretung mittels einer festen Rolle definiert werden können (z.B. „*Role = Arzt*“). Es sollen aber auch komplexe TPR verwendbar sein, wie z.B. „*Role = HW-Entwickler \wedge selbe Projektzugehörigkeit wie der Originalbearbeiter*“ (hierbei soll das Organisationsmodellobjekt Projekt eine Gruppe sein, z.B. *Entwicklung C-Klasse, Entwicklung S-Klasse*). Damit diese TPR für viele Aktivitäten und Bearbeiter unterschiedlicher Projekte (in identischer Form) verwendet werden kann, muss sie sich auf die „normale“ Bearbeiterzuordnung der betroffenen Aktivität beziehen können, weil ausschließlich diese auf das (jeweils unterschiedliche) Projekt verweist.

Eine solche Referenzierung soll natürlich nicht nur für Projekte (also Gruppen) möglich sein, sondern für alle Objekttypen des Organisationsmodells: Im Folgenden wird von den Objekttypen Rolle, Gruppe, Kompetenz und Organisationseinheit (letztere inkl. einer Hierarchie) ausgegangen, wobei eine Erweiterung leicht möglich ist (s. Abschnitt 3.1). Verallgemeinert man die oben dargestellte TPR zu „*selbe Rolle und selbe Gruppe wie die Originalaktivität*“, dann kann diese zusätzlich Regel z.B. auch für Aktivitäten verwendet werden, in deren Bearbeiterzuordnung eine andere Rolle gefordert ist, wie z.B. „*Role = Testfahrer \wedge selbe Projektzugehörigkeit wie der Originalbearbeiter*“. Die Notwendigkeit zur Definition einer weiteren TPR entfällt also. Schließlich können in einer Bearbeiterzuordnung auch mehrere Rollenzugehörigkeiten gefordert sein, von denen nur manche für die TPR relevant sind, z.B. „*Role = HW-Entwickler \wedge Role = Senior Entwickler \wedge selbe Projektzugehörigkeit wie Vorgängeraktivität X*“. Für die TPR soll in diesem Beispiel die zweite Rollenzuordnung nicht berücksichtigt werden. Deshalb muss irgendwie ausgedrückt werden, welche der beiden Rollenzugehörigkeiten relevant ist. Die direkte Verwendung des Rollennamens (hier: *HW-Entwickler*) verbietet sich, weil eine so eng gefasste Regel dann nicht mehr für andere Aktivitäten verwendet werden könnte, z.B. mit Bearbeiterzuordnung „*Role = Testfahrer*“. Ebenso wäre eine solche Regel nicht bei Bearbeiterzuordnungen verwendbar, bei denen sich der Rollename dynamisch ergibt, z.B. aus einer Prozessvariablen (d.h. Anwendungsdaten).

Um dieses Problem zu lösen, müssen unterschiedliche Arten von Organisationsmodellobjekten (z.B. unterschiedliche Arten von Rollen) unterscheidbar gemacht werden. Hierzu führen wir sog. SubTypes ein. Jedem Organisationsmodellobjekt (d.h. jeder Rolle, jeder Gruppe, etc.) kann ein SubType zugeordnet werden. So wird z.B. der Rolle *HW-Entwickler* der SubType *Tätigkeit* und der Rolle *Senior Entwickler* der SubType

Skill-Level zugeordnet. Die Gruppen *Entwicklung C-Klasse* und *Entwicklung S-Klasse* sind dem SubType *Projekt* zugeordnet. Dadurch wird es möglich, die Zielpersonen von Eskalationen bzw. Stellvertretungen folgendermaßen zu definieren: die Rolle mit SubType *Tätigkeit* und die Gruppe vom SubType *Projekt* müssen so sein, wie bei der (normalen) Bearbeiterzuordnung dieser Aktivität (d.h. aber auch: eine dort ggf. vorkommende Rolle mit SubType *Skill-Level* ist für diese TPR nicht relevant). Eine einzige solche TPR wird dadurch für sehr viele Originalbearbeiter und für Bearbeiterzuordnung vieler Aktivitäten verwendbar. Wie in den Beispielen dargestellt, dürfen diese Bearbeiterzuordnungen sogar mittels booleschen Operationen aus mehreren Teilen zusammengesetzt sein und sich auf Prozessvariablen und Bearbeiter von Vorgängeraktivitäten beziehen (d.h. abhängige Bearbeiterzuordnungen).

Im Abschnitt 2 werden der Stand von Wissenschaft und Technik und die Herausforderungen dargestellt. Das entwickelte Konzept wird Abschnitt 3 erläutert. Der Beitrag schließt mit einer Zusammenfassung und einem Ausblick.

2 Aktuell bestehende Herausforderungen

Im Folgenden werden relevante Problemstellungen anhand eines (vereinfachten, aber realitätsnahen) Anwendungsbeispiels detailliert erläutert. Dann wird der aktuelle Stand von Forschung und kommerziellen PMS dargestellt, die Anforderungen an die Festlegung der Zielpersonen beschrieben und daraus die Forschungsfrage abgeleitet.

2.1 Problemstellung und Anwendungsbeispiel

Einem Benutzer eines PMS kann mehr als eine Rolle zugeordnet sein. Deshalb kann die Definition der Zielpersonen einer Eskalation oder Stellvertretung (TPR) nicht einfach mittels Regeln wie z.B. „*selbe Rolle wie der Originalbearbeiter*“ erfolgen. Dann wäre nämlich unklar, welche seiner Rollen gemeint ist. Dasselbe gilt für Gruppen, Kompetenzen, etc. Außerdem bestehen Bearbeiterzuordnungen häufig aus mehreren Teilen, die mit booleschen Operationen verknüpft sind. Wie bereits beschrieben, sollen auch solche Bearbeiterzuordnungen Grundlage für eine TPR sein können. Hierzu müssen bestimmte Teile einer solchen Bearbeiterzuordnung referenziert werden können. Es ist also ein Mechanismus notwendig, um genau diese benötigten Teile von Bearbeiterzuordnungen in einer TPR auswählen zu können. Die daraus resultierenden Problemstellungen werden im Folgenden an Beispielen erläutert, die später bei der Darstellung des Lösungsansatzes aufgegriffen werden.

Der Mitarbeiter (User) u sei aufgrund seiner aktuellen Aufgaben Mitglied in den beiden Gruppen (Group) *Entwicklung C-Klasse* und *Entwicklung S-Klasse*. Diese Gruppenmitgliedschaften können in einer TPR nicht direkt übernommen werden, weil dann die Mitglieder beider Gruppen enthalten wären, die konkrete Eskalation oder Stellvertretung aber beispielsweise eine Aufgabe speziell in der C-Klassen-Entwicklung betrifft.

Deshalb wird nicht die Gruppenzugehörigkeit des ursprünglichen Bearbeiters verwendet, sondern die Bearbeiterzuordnung der betroffenen Aktivität. Diese ist für die Akt. *Bauteil entwerfen* in Abb. 1a dargestellt: Außer zwei Rollen (Role, siehe ①, ②) wird hier auch eine Gruppenzugehörigkeit gefordert, wobei sich der Gruppenname aus der Prozessvariablen *Projektname* dynamisch zur Runtime der Prozessinstanz ergibt ③. Abb. 1b zeigt einige der im Organisationsmodell des PMS gespeicherten Objekte einschließlich ihres SubType. Diese werden in der TPR (Abb. 1c) verwendet. Die Regel legt fest, dass die Zielpersonen dieselbe Rolle mit SubType *Tätigkeit* haben müssen ①, und derselben Gruppe mit dem SubType *Projekt* zugeordnet sein müssen ②, wie in der Bearbeiterzuordnung (Abb. 1a) festgelegt. Das Projekt hat sich dynamisch aus der Prozessvariablen *Projektname* ergeben, was in der TPR aus Abb. 1c aber nicht erkennbar ist (d.h. diese wird dadurch nicht komplizierter).

a) ... → **Bauteil entwerfen** → ...

Role = "HW-Entwickler" ^ ①
 Role = "Senior-Entwickler" ^ ②
 Group = Value(Projektname) ③

b)

Nr	Name	Type	SubType
1	HW-Entwickler	Role	Tätigkeit
2	Testfahrer	Role	Tätigkeit
3	SW-Entwickler	Role	Tätigkeit
4	SW-Architekt	Role	Tätigkeit
5	Senior-Entwickler	Role	Skill-Level
6	Entwicklung C-Klasse	Group	Projekt
7	Entwicklung E-Klasse	Group	Projekt
8	Entwicklung S-Klasse	Group	Projekt
9	Elektronik-Entwicklung	Group	Projekt
10	Betriebsrat	Group	Gremium
11	Englisch A1	Competence	Sprachkenntnis
12	Englisch A2	Competence	Sprachkenntnis

c) **Spezifikation der Zielpersonen (TPR):**
 sameRole(Tätigkeit) ^ ①
 sameGroup(Projekt) ②

d) ... → **Steuergeräte-SW entwickeln** → ...

Role = "SW-Entwickler" ^ ①
 Group = "Elektronik-Entwicklung" ②

e) ... → **SW-Review** → ...

Role = "SW-Architekt" ^ ①
 Group = "Entwicklung S-Klasse" ^ ②
 Competence = "Englisch A2" ③

Abb. 1: a), d) und e) Bearbeiterzuordnungen, b) Objekte des Organisationsmodells, c) eine TPR

Die in Abb. 1d dargestellte Aktivität hat eine ähnliche, jedoch abweichende Bearbeiterzuordnung: So muss der Bearbeiter hier die Rolle *SW-Entwickler* haben ① und er muss kein *Senior-Entwickler* sein. Außerdem ist das zugehörige Projekt fest durch den Wert *Elektronik-Entwicklung* vorgegeben ②. Abb. 1e zeigt die Bearbeiterzuordnung der Akt. *SW-Review* eines Prozessstyps von S-Klasse-Entwicklungsprojekten. Deshalb ist außer der Rolle ① auch die Gruppe (das Projekt) ② fest vorgegeben. Zusätzlich ist noch eine bestimmte Kompetenz ③ erforderlich. Obwohl die in Abb. 1a, d und e dargestellten Bearbeiterzuordnungen recht unterschiedlich sind, kann für sie alle die in Abb. 1c dargestellte TPR verwendet werden: Diese bezieht sich auf die in den Bearbeiterzuordnungen vorkommenden Tätigkeiten (Type *Role* mit SubType *Tätigkeit*) und Projekte (Type *Group* mit SubType *Projekt*), unabhängig davon, auf welche Art und Weise diese in den Original-Bearbeiterzuordnungen festgelegt werden. Deshalb ergeben sich bei diesen Aktivitäten z.B. die unterschiedlichen Rollen *HW-Entwickler*, *SW-Entwickler* bzw. *SW-Architekt* (durch dieselbe TPR aus Abb. 1c).

In Abb. 1 wurden einige „fachliche Aktivitäten“ diskutiert. Zusätzlich zur in Abb. 1c dargestellten TPR, können für andere Aktivitäten abweichende Regeln erforderlich sein. Auch kann es unterschiedliche Regeln für Eskalationen und für Stellvertretungen geben oder mehrere Regeln für jeweils unterschiedliche Originalbearbeiter.³ Insbesondere für „organisatorische Aktivitäten“, z.B. innerhalb von Teams oder Abteilungen (d.h. der Linienorganisation), sind typischerweise andere TPR erforderlich, weil die Zielpersonen dann z.B. Vorgesetzte sind (anstatt Projektmitglieder). So wird die in Abb. 2a dargestellte Genehmigung vom Abteilungsleiter ① desjenigen Mitarbeiters durchgeführt, der die Akt. X (Antragstellung) ausgeführt hat ②. Auch die in Abb. 2c dargestellte Bestellgenehmigung erfolgt durch einen Abteilungsleiter ①. Allerdings ergibt sich dessen Abteilung dadurch, dass aus der betroffenen Kostenstelle (wurde vom Antragsteller eingegeben) die Abteilungsbezeichnung ermittelt und in der Prozessvariablen *AbtDerKostenstelle* gespeichert wurde. Dieser Wert wird nun verwendet, um die betroffene Organisationseinheit zu ermitteln ② (z.B. Zeile 13 oder 14 in Abb. 2b). Für beide Aktivitäten gilt die in Abb. 2d dargestellte TPR: Zielpersonen müssen derselben *OrgUnit* mit SubType *Abteilung* angehören ①. Die Regel zur Festlegung ihrer Rolle hängt vom *Geschäftsbereich* des Originalbearbeiters ab: Gehört dieser zum Geschäftsbereich *PKW*, dann muss die Zielperson ein stellvertretender Abteilungsleiter ② oder ein Teamleiter ③ sein. Bei anderen Geschäftsbereichen handelt es sich um den Leiter des der Abteilung übergeordneten Centers ④. Diese Regel ist zwar recht komplex (und deshalb erfordert ihre Erstellung einen gewissen Aufwand), sie kann aber für mehrere unterschiedliche Aktivitäten verwendet werden.

a) ... → **Urlaubsantrag genehmigen** → ...

Role = "Abteilungsleiter" ^ ①
OrgUnit = Abteilung(Actor(X)) ②

c) ... → **Bestellung genehmigen** → ...

Role = "Abteilungsleiter" ^ ①
OrgUnit = Value(AbtDerKostenstelle) ②

d) **Spezifikation der Zielpersonen:**
 sameOrgUnit(Abteilung) ^ ①
 (if hasOrgUnit(Geschäftsbereich) = "PKW" then
 Role = "AbtLeiterStellvertreter" v ②
 Role = "Teamleiter" ③
 else
 Role = "Centerleiter") ④

Nr	Name	Type	SubType
13	Sitzentwicklung	OrgUnit	Abteilung
14	Blinkerentwicklung	OrgUnit	Abteilung
15	Crash-Tests	OrgUnit	Center
16	Teamleiter	Role	Leitungsfunktion
17	Abteilungsleiter	Role	Leitungsfunktion
18	AbtLeiterStellvertreter	Role	Leitungsfunktion
19	Centerleiter	Role	Leitungsfunktion
20	PKW	OrgUnit	Geschäftsbereich
21	LKW	OrgUnit	Geschäftsbereich

Abb. 2: a) und c) Bearbeiterzuordnungen, b) Organisationsmodellobjekte und d) eine TPR

Die dargestellten Beispiele enthalten bereits abhängige Bearbeiterzuordnungen, die auf den Bearbeiter einer Vorgängeraktivität bzw. Prozessvariablen verweisen (Abb. 2a und c). Das zu entwickelnde Konzept muss zusätzlich weitere Arten abhängiger Bearbeiterzuordnungen unterstützen. Die Schwierigkeit hierbei ist, dass eine abhängige Bearbeiterzuordnung wie z.B. „selber Bearbeiter wie Akt. X“ keine Rollen, Gruppen, etc. spezifiziert, die durch eine TPR referenziert werden könnten. Eine weitere bereits dargestellte Anforderung ist, dass in den TPR Fallunterscheidungen vorkommen können (if/else in Abb. 2d). Schließlich sind TPR, ebenso wie Bearbeiterzuordnungen, oft mittels boole-

³ Auf die Darstellung solcher Fälle wurde hier zur Verbesserung der Übersichtlichkeit verzichtet.

scher Operationen aus mehreren Teilen zusammengesetzt. Durch die Einführung der benutzerdefinierten SubTypes werden einzelne Teile solcher Bearbeiterzuordnungen referenzierbar. Dadurch entfällt der Aufwand für GP-Designer, für jede Kombination von vielen unterschiedlichen Aktivitäten und zahlreichen unterschiedlichen Personengruppen (z.B. Geschäftsbereiche, Mitarbeiterarten), jeweils eine separate TPR zu erstellen.

2.2 Stand von Forschung und Technik

Im Folgenden wird wissenschaftliche Literatur zu Eskalationen und Stellvertretungen, sowie der aktuelle Stand einiger kommerzieller PMS darstellt. Dabei wurden Bizagi Studio V.11.2.3 [Bi21], IBM Business Automation Workflow V20.0.0.2 [IBM20], K2 Cloud V.4.0 [K21] und Signavio Workflow Accelerator V.13.6.0 [Si21] betrachtet.

Eskalationen

Eskalationen werden in der wissenschaftlichen Literatur zu PMS zwar erwähnt (z.B. in [Gr08, Ru05]), es gibt aber keine Arbeiten, die sich hiermit umfassend beschäftigen. Deshalb werden im Folgenden Arbeiten vorgestellt, die spezielle Arten bzw. Aspekte von Eskalationen betrachten: [ARD07] verwendet erwartete Zeiten für die Beendigung von Aktivitäten, um schon vor dem Verpassen einer Frist, geeignet reagieren zu können. Hierzu wird ein sehr mächtiger Eskalationsmechanismus, speziell für das betrachtete Szenario der Überlastung eines Call-Centers, entwickelt. [PR98] minimiert die Anzahl an Eskalationen, indem Deadlines automatisch angepasst werden, um Verzögerungen auszugleichen. Die von uns betrachtete Fragestellung (d.h. die effiziente Modellierung der Zielpersonen einer Eskalation) wird in all diesen Arbeiten nicht erwähnt.

Die betrachteten kommerziellen PMS bieten unterschiedliche Funktionalitäten. So ermöglicht Bizagi keine Eskalationen. Bei K2 kann ein einzelner Benutzer als Zielperson festgelegt werden, der abhängig von der betroffenen Aktivität ist. Bei Signavio sind auch mehrere Zielpersonen möglich. Auch IBM ermöglicht mehrere Zielpersonen, und es können zudem mehrere Eskalationen (parallel und nacheinander) definiert werden.

Stellvertretungen

[Ba09] definiert Anforderungen, wie z.B. kontext-abhängige und mehrstufige Stellvertretungen, oder eine Aktivität dem Stellvertreter bei Rückkehr des Originalbearbeiters wieder zu entziehen. Hierfür werden Algorithmen zur Berechnung der Stellvertreter vorgestellt. Die Arbeit enthält jedoch kein Konzept zur Modellierung solcher Zielpersonen. Auch andere Arbeiten zum Thema Stellvertretungen machen hierzu keine Vorschläge: In [Mu04] werden Stellvertreter mittels Rollen definiert. [Ru05] betrachtet Stellvertretungen als vom PMS durchgeführte Delegation. [HD05] erwähnt die Notwendigkeit von Stellvertretungen, ohne Lösungskonzepte vorzustellen. [RM98] fordert „kontextbezogene Stellvertretungen“, betrachtet jedoch die Modellierung solcher Stellvertreterregelungen nicht. [HS99] definiert Ersatzressourcen (d.h. auch Bearbeiter) mittels einer SQL-artigen Sprache. [DW15] ermittelt die optimale Ersatzressource mittels Process-Mining.

Bizagi unterstützt Stellvertretungen nicht. K2 ermöglicht unterschiedliche Stellvertreter für verschiedene Aktivitäten. Bei IBM kann für jeden Benutzer eine Liste von Stellvertretern definiert werden. Die Stellvertreter können aber nicht abhängig von der betroffenen Aktivität oder von Prozessdaten sein. Zudem wird nur der erste anwesende Stellvertreter dieser Liste berücksichtigt. Wie auch bei den Eskalationsmechanismen bietet keines der PMS effiziente Regeln zur Festlegung der Zielpersonen, die zudem sehr leistungsfähig sind (d.h. komplexe, vom Prozesskontext abhängige Regeln).

2.3 Anforderungen an die Festlegung von Zielpersonen und Forschungsfrage

Nachfolgend werden die im Abschnitt 2.1 erläuterten Anforderungen (Requirements) zusammengefasst und anschließend die Forschungsfrage vorgestellt.

- **Req1:** Es soll für den GP-Designer möglich sein, die Zielpersonen mit geringem Aufwand (d.h. effizient) zu modellieren. Deshalb muss eine einzelne solche Regel (TPR) für mehrere Aktivitäten und Originalbearbeiter verwendbar sein.
- **Req2:** Es sollen beliebig viele unterschiedliche Regeln definierbar sein, die jeweils für einen unterschiedlichen Kontext gültig sind. Dieser muss beliebig definierbar sein, d.h. sich z.B. auf die betroffenen Originalbearbeiter und Aktivitäten, sowie den Prozesskontext (die aktuellen Anwendungsdaten der Prozessinstanz) beziehen können.
- **Req3:** Die Bearbeiterzuordnung einer Aktivität darf (weiterhin) beliebig komplex sein, d.h. sie kann mittels boolescher Operationen aus mehreren Teilen zusammengesetzt sein, und Fallunterscheidungen enthalten.
- **Req4:** Teile einer Bearbeiterzuordnung können abhängig z.B. (vom Bearbeiter) von Vorgängeraktivitäten oder von Prozessdaten sein.
- **Req5:** Die Definition der Zielpersonen (TPR) kann sich auf die Bearbeiterzuordnung der Aktivität beziehen und muss dort beliebige organisatorische Objekte (z.B. eine Rolle) referenzieren können. Dies muss auch möglich sein, wenn die Bearbeiterzuordnung komplex ist (Req3) oder abhängige Teile enthält (Req4).

Diese Anforderungen stellen ein komplexes Problem dar, das in der wissenschaftlichen Literatur und in kommerziellen PMS (vgl. Abschnitt 2.2) noch ungelöst ist. Deshalb beschäftigt sich die vorliegende Arbeit mit folgender Forschungsfrage: Wie kann die Festlegung der Zielpersonen von Eskalationen und Stellvertretungen so gestaltet werden, dass eine große Mächtigkeit erreicht wird, für Aktivitäten weiterhin die benötigten Bearbeiterzuordnungen verwendbar sind, und dennoch ein möglichst geringer Aufwand bei der Erstellung entsprechender Regeln entsteht?

3 Konzept zur effizienten Modellierung der Zielpersonen

Im Folgenden wird erläutert, wie Organisationsmodell-Objekte inkl. SubType definiert und später in einer TPR verwendet werden. Außerdem wird auf die Aspekte komplexer (d.h. zusammengesetzter) TPR und abhängiger Bearbeiterzuordnungen eingegangen.

3.1 Erweiterung und Verwendung des Organisationsmodells

Wie bereits erwähnt, basiert der hier vorgestellte Ansatz darauf, dass die Objekte des Organisationsmodells einem benutzerdefinierten SubType zugeordnet werden. Es ist also die Festlegung zusätzlicher Information erforderlich. Diese Information wird bewusst nicht als Teil von Bearbeiterzuordnungen definiert, weil dies sonst bei jeder Bearbeiterzuordnung jeder einzelnen Aktivität erfolgen müsste, was zu einem großen Aufwand für die GP-Designer führen würde. Das bedeutet, die Bearbeiterzuordnungen bleiben gegenüber einer „klassischen GP-Modellierung“ unverändert. Die Festlegung der Zusatzinformation SubType erfolgt im Organisationsmodell bei der Erstellung der Organisationsmodell-Objekte (z.B. einer Rolle). Solche Objekte werden selten neu erstellt und sie werden selten verändert (ihnen werden lediglich öfters neue oder andere Benutzer zugeordnet). Deshalb erfordert die Festlegung der SubTypes nur einen geringen Aufwand.

Das diesem Beitrag zugrundeliegende Organisationsmodell unterscheidet die Objekttypen Role, Group, Competence und OrgUnit (vgl. [Ru05]). Diese Objekttypen werden jedoch weitgehend auf identische Art und Weise verwendet, so dass der Ansatz leicht um beliebige weitere Typen erweitert werden kann. Jedes Objekt des Organisationsmodells kann einem SubType zugeordnet werden (z.B. die Rolle *HW-Entwickler* dem SubType *Tätigkeit*, vgl. Abb. 1b). Bei der Spezifikation einer TPR kann darauf mittels vom PMS bereitgestellten Funktionen Bezug genommen werden. So legt *sameRole(S)* fest (vgl. Abb. 1c), dass die Zielpersonen diejenige Rolle des SubType *S* besitzen müssen, die in der Bearbeiterzuordnung dieser Aktivität festgelegt wurde (z.B. bei SubType = *Tätigkeit* die Rolle *HW-Entwickler*, vgl. Abb. 1a). Ähnlich funktioniert die Funktion *hasRole(R)*, mit der abgefragt werden kann, ob der Originalbearbeiter der Aktivität der Rolle *R* zugeordnet ist oder nicht. Dies wird z.B. für Fallunterscheidungen (if) genutzt. Dem GP-Designer stehen analoge Funktionen für Organisationsmodell-Objekte des Typs Group (d.h. *sameGroup(S)* und *hasGroup(G)*), Competence und OrgUnit zur Verfügung.

Bei der Auswertung einer TPR (folgt in Abschnitt 3.2) werden intern nicht die oben dargestellten Funktionen verwendet. Stattdessen werden die TPR nach Abschluss der Modellierung (d.h. zur Buildtime) automatisch in generic Target Person Rules (*gTPR*) umgewandelt, die auf Funktionen basieren, in deren Namen der Typ des jeweils betroffenen Organisationsobjekts nicht enthalten ist. Diese Funktionen werden bei Anfragen an die Organisationsmodell-Datenbank zur Ermittlung der entsprechenden Benutzer verwendet. Eine dieser generischen Funktionen ist *sameOrgObject(Type, SubType)*. Diese wird z.B. verwendet, um den vom GP-Designer modellierten Ausdruck *sameRole(Tätigkeit)* in *sameOrgObject(Role, Tätigkeit)* umzuwandeln. Die andere generische Funktion ist *hasOrgObject(Type, SubType)*. Sie ist erforderlich, um z.B. *hasGroup(Betriebsrat)* in *hasOrgObject(Group, Betriebsrat)* umzuwandeln. Vorteil dieser generischen Funktionen ist, dass die Schnittstelle der Organisationsdatenbank nur diese beiden Funktionen unterstützen muss, und dennoch beliebige Typen von Organisationsmodell-Objekten verwendet werden können. Die im Weiteren dargestellten Inhalte bleiben ebenfalls für beliebige zusätzliche Typen unverändert. Die Umformung der vom GP-Designer modellierten TPR in die *gTPR* mit generischen Funktionen kann

automatisch erfolgen, weil beide Arten von Ausdrücken inhaltlich dasselbe beschreiben. Dem GP-Designer sollten aber (weiterhin) „typisierte“ Funktionen (z.B. *sameRole(R)*) zur Verfügung gestellt werden, weil diese Personen solche Funktionen gewöhnt sind.

Zwischen Mitarbeitern bzw. den von ihnen besetzten Stellen existiert eine Hierarchie. Diese wird von normalen Bearbeiterzuordnungen verwendet (z.B. „*Vorgesetzter vom Bearbeiter der Akt. X*“). Solche Hierarchien sind, abhängig vom konkreten Organisationsmetamodell des verwendeten PMS, evtl. bereits vorhanden. Sie können im hier vorgestellten Ansatz zur Ermittlung von Zielpersonen von Eskalationen und Stellvertretungen in einer TPR ebenfalls verwendet werden.

Es kann der Eindruck entstehen, dass die vorgeschlagene Lösung unnötig komplex ist, weil ein Objekt des Organisationsmodells nun zwei „Kategorien“ zugeordnet ist (z.B. dem Type *Role* und SubType *Tätigkeit*). Die Alternative hierzu wäre, auf die Einführung der SubTypes zu verzichten und stattdessen zu ermöglichen, dass der Benutzer zusätzliche Types selbst definiert. Dies ist allerdings kaum umsetzbar, weil die Metamodelle von PMS üblicherweise von einer fest vorgegebenen Menge an Typen von Organisationsmodell-Objekten ausgehen (vgl. [BL20]). Auch die wissenschaftliche Literatur betrachtet üblicherweise keine frei durch den Benutzer definierbaren Typen von Organisationsobjekten [Ru05]. Dies wäre auch aufwendig und kaum nützlich, weil in einer Bearbeiterzuordnung dann anstatt „*Role = HW-Entwickler*“ der selbst definierte Type verwendet werden müsste, also z.B. „*Role_Tätigkeit = HW-Entwickler*“. Da der SubType bei „normalen“ Bearbeiterzuordnung jedoch gar nicht relevant ist, würde dies deren Definition sogar verkomplizieren. Der Vorteil bei der Festlegung einer TPR fällt im Vergleich hierzu kaum ins Gewicht, weil vergleichsweise wenige TPR definiert werden müssen. Dort ergäbe sich in diesem Beispiel zudem ein nicht wirklich einfacherer Ausdruck, nämlich *same(Role_Tätigkeit)* anstatt *sameRole(Tätigkeit)*. Generell ist es bei beiden Alternativen möglich, dass der GP-Designer sowohl Funktionen wie *sameRole*, als auch die benötigten Werte von Type und SubType graphisch in einer GUI aus einer Liste auswählt. Für den Benutzer ist der hier vorgestellte Ansatz also einfach handhabbar

3.2 Spezifikation und Ermittlung der Zielpersonen

Im Folgenden werden die Regeln zur Festlegung von Zielpersonen vollständig eingeführt und es wird beschrieben, wie diese PMS-intern verwendet werden.

Beim GP-Design wird eine Menge *ESR* (Escalation and Substitution Rules) zur Definition von Zielpersonen festgelegt: $ESR = \{R_1, R_2, \dots\}$. Jede dieser Regeln besteht aus einer Bedingung *Cond* (Condition), die ihren Gültigkeitsbereich definiert, und einer Regel TPR zur Festlegung der Zielpersonen: $R_i = (Cond, TPR)$. Hierbei ist die Bedingung *Cond* nicht weiter problematisch. Dies ist eine boolesche Regel, die beliebige Attribute verwenden kann (z.B. *Cond = „ActivityClass = Entwicklungstask \wedge hasOrgUnit(Geschäftsbereich) = PKW“*). Diese können auch benutzerdefiniert sein, wie z.B. *ActivityClass* (ein dieser Aktivität zugeordneter Klassenname). Außerdem können

z.B. organisatorische Objekte oder Werte von Prozessvariablen verwendet werden, und es kann definiert werden, ob die Regel für Eskalationen, Stellvertretungen oder für beides gültig ist. Ergibt die Bedingung *Cond* den Wert *True*, ist die TPR im aktuell vorliegenden Fall anzuwenden. Im Folgenden liegt der Fokus auf dieser Regel TPR.

Zur Buildtime definiert der GP-Designer Bearbeiterzuordnungen *AA* (Actor Assignment) für Aktivitäten *a*, z.B. $AA(a) = „Role = HW-Entwickler“$. In diesem Fall hat die Rolle *HW-Entwickler* den SubType *Tätigkeit* (vgl. Abb. 1b). Außerdem definiert er TPR, z.B. für diese Aktivität *a* die Regel $TPR = „sameRole(Tätigkeit)“$. Selbstverständlich sind hier auch komplexere Regeln möglich (vgl. Abschnitt 3.3 und 3.4). Wie bereits erläutert, werden die modellierten TPR in ein nicht-typisiertes (d.h. generisches) Format umgewandelt, in diesem Beispiel $gTPR = „sameOrgObject(Role, Tätigkeit)“$.

Ist nun zur Runtime eine Eskalation oder Stellvertretung für den Originalbearbeiter *OriginalActor* dieser Aktivität *a* erforderlich, so ermittelt das PMS anhand deren Bedingung *Cond* die passende Regel R_i . Die daraus resultierende *gTPR* wird verwendet, um diejenigen Organisationsmodell-Objekte zu ermitteln, denen die Zielpersonen angehören müssen. Im obigen Beispiel ergibt sich also aus $gTPR = „sameOrgObject(Role, Tätigkeit)“$ die Rolle *HW-Entwickler*, weil dies diejenige Rolle mit SubType *Tätigkeit* ist, die in $AA(a)$ enthalten ist. Dieses konkrete Organisationsmodell-Objekt $O = HW-Entwickler$ wird verwendet, um aus der *gTPR* eine *rTPR* (resolved TPR) abzuleiten. Diese hat die Form $OrgObjectType = O$, im dargestellten Beispiel also $rTPR = „Role=HW-Entwickler“$. Diese *rTPR* wird verwendet, um die entsprechenden Zielpersonen mit der Funktion $getUsers(rTPR)$ von der Organisationsmodell-Datenbank zu erfragen, hier also mit $getUsers(„Role=HW-Entwickler“)$. Funktionen wie $hasRole()$, $hasGroup()$ werden analog behandelt, wobei diese aber *True* oder *False* zurückliefern, abhängig davon, ob der Originalbearbeiter *OriginalActor* dem ermittelten Organisationsobjekt *O* angehört (z.B. der Rolle *HW-Entwickler*).

Selbstverständlich können bei der Festlegung einer TPR auch „klassische“ Regeln verwendet werden (d.h. so wie bei normalen Bearbeiterzuordnungen ohne SubType). Dann entfällt die eben beschriebene Umformung. Beispielsweise ergeben sich für die TPR „*Role = HW-Entwickler*“ die identischen *gTPR* und *rTPR* und damit direkt die Anfrage $getUsers(„Role=HW-Entwickler“)$.

3.3 Komplexe Regeln für Bearbeiterzuordnungen

Die Anforderung Req3 besagt, dass eine TPR und damit eine *gTPR* eine komplexe (d.h. zusammengesetzte) Regel sein kann. Dies wird folgendermaßen behandelt: Wenn eine *gTPR* eine Bedingung (if) (vgl. Abb. 2d) enthält, so wird diese zuerst aufgelöst, indem der für den betroffenen Bearbeiter *OriginalActor* relevante Fall ausgewählt wird (z.B. anhand der aktuellen Werte von Prozessvariablen oder vorheriger Bearbeiter). Dadurch entsteht eine Formel, die ausschließlich boolesche Operationen enthält. Beispielsweise entsteht aus der in Abb. 2d dargestellten TPR im *else*-Fall $gTPR = „sameOrg-$

Object(OrgUnit, Abteilung) ∧ Role = Centerleiter“. Die mit booleschen Operationen verknüpften Teile werden nun, so wie bereits erläutert, einzeln umgeformt, um die *rTPR* zu berechnen. Falls sich z.B. bei der Bearbeiterzuordnung aus Abb. 2a aus dem Bearbeiter dieser Aktivität die Abteilung *Sitzentwicklung* ergibt, entsteht die *rTPR* = „*OrgUnit = Sitzentwicklung ∧ Role = Centerleiter*“, wobei der zweite Teil unverändert aus der *gTPR* übernommen wurde. Diese *rTPR* wird dann von der Funktion *getUsers(rTPR)* zur Abfrage der gesuchten Benutzer verwendet.

Es wäre einfach, zu fordern, dass für alle Regelteile einer TPR, die sich auf eine Bearbeiterzuordnung beziehen (z.B. mittels *sameRole()*), eine entsprechende Teilbedingung auch tatsächlich in jeder referenzierten Bearbeiterzuordnung enthalten sein muss. Dann wäre diese TPR jedoch nicht für Aktivitäten verwendbar, bei denen einer dieser Teile fehlt. Die Folge wäre, dass hierfür eine weitere TPR nötig wird, also mehr Regeln und mehr Aufwand für die GP-Designer entstehen. Das folgende Beispiel zeigt, dass solche zusätzlichen Regeln nicht unbedingt erforderlich sind: Angenommen, für eine Aktivität sind keine besonderen Sprachkenntnisse (wie z.B. *Englisch A2*, vgl. Abb. 1b) erforderlich. Außerdem existiert bereits eine (ansonsten passende) TPR, welche dieselben Sprachkenntnisse (*Competence* mit SubType *Sprachkenntnis*) fordert, wie bei der Bearbeiterzuordnung. Da in dieser jedoch gar keine Sprachkenntnisse gefordert sind, sind diese auch für etwaige Stellvertreter nicht notwendig. Die bereits vorhandene TPR kann also auch für solche Aktivitäten verwendet werden, indem diese Bedingung geeignet ignoriert wird. Im Folgenden wird erläutert, wie ein PMS das umsetzen kann.

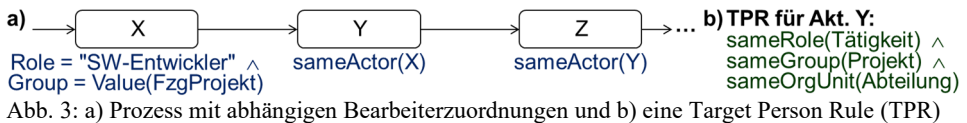
Da jede boolesche Formel in die Disjunktive Normalform (DNF) umgeformt werden kann (automatisch vom PMS), wird im Folgenden angenommen, dass die *gTPR* in diesem Format vorliegt, z.B.: $gTPR = (A1 \wedge A2 \wedge A3) \vee (B1 \wedge B2) \vee (C1)$

Ein in der *gTPR* enthaltener Teilausdruck (z.B. *sameOrgObject(Type, SubType)*) kann sich auf einen Objekttyp *Type* (z.B. eine Rolle) mit einem SubType (z.B. *Tätigkeit*) beziehen, der in der Bearbeiterzuordnung dieser Aktivität nicht enthalten ist. Es wird also etwas „nicht Vorhandenes“ referenziert. Dies wird vom PMS folgendermaßen behandelt:

1. Trifft dies für einen Teil einer Konjunktion (d.h. einer AND-Verknüpfung) zu, aber nicht für alle ihre Teilausdrücke (z.B. *A2* der in DNF dargestellten Regel), so wird dieser Teilausdruck ignoriert (d.h. aus der Regel gestrichen). Damit ist nur noch der Rest der Konjunktion relevant, im Beispiel also $A1 \wedge A3$, womit sich die Gesamtregel $gTPR' = (A1 \wedge A3) \vee (B1 \wedge B2) \vee (C1)$ ergibt. Dies ist der Fall, der im obigen Beispiel der nicht relevanten Sprachkenntnisse greift.
2. Trifft dies für alle Teile einer Konjunktion zu (z.B. für *B1* und *B2* oder für *C1*), dann werden diese Teile ignoriert, d.h. die gesamte Konjunktion entfällt, weil eine nicht verwendbare Bedingung nicht zu zusätzlichen Zielpersonen führen soll. Bei nicht relevantem *B1* und *B2* verbleibt von obiger Formel beispielsweise $gTPR' = (A1 \wedge A2 \wedge A3) \vee (C1)$.
3. Trifft dies für alle Teile einer Regel zu, so ist sie für diese Aktivität nicht verwendbar. Es liegt also ein Modellierungsfehler vor. Dieser kann bereits zur Buildtime automatisch erkannt und damit verhindert werden.

3.4 Berücksichtigung abhängiger Bearbeiterzuordnungen

Entsprechend der Anforderung Req4 sollen auch abhängige Bearbeiterzuordnungen durch eine TPR referenziert werden können. So soll die Akt. Y in Abb. 3a vom selben Bearbeiter ausgeführt werden, wie die Vorgängerakt. X. Allerdings enthält der Ausdruck (Actor Assignment) $AA(Y) = sameActor(X)$ keinen Rollennamen, der z.B. von dem in der TPR enthaltenen Teil $sameRole(Tätigkeit)$ referenziert werden könnte (vgl. Abb. 3b). Der tatsächliche Bearbeiter der Akt. X und damit auch der Akt. Y steht jedoch vor dem Start von Y bereits fest. Es ist aber nicht möglich, einfach die ihm zugeordnete Rolle zu verwenden, da er mehrere Rollen mit SubTyp *Tätigkeit* innehaben kann (z.B. *SW-Entwickler* und *SW-Architekt*). Dasselbe gilt für andere Typen organisatorischer Objekte, wie Kompetenzen (z.B. *Englisch A1* und *Französisch B2*).



Um unabhängige Bearbeiterzuordnungen zu erhalten, wird aus der (ggf. abhängigen) Bearbeiterzuordnung $AA(a)$ eine unabhängige (independent) Bearbeiterzuordnung $iAA(a)$ abgeleitet, sobald die Akt. a ausführbar wird. Diese neue Regel $iAA(a)$ wird dann von Folgeaktivitäten verwendet, welche die Bearbeiterzuordnung der Akt. a referenzieren. i) Im einfachsten Fall enthält $AA(a)$ keine Abhängigkeiten und wird unmittelbar übernommen, d.h. $iAA(a) = AA(a)$. ii) Enthält eine Bearbeiterzuordnung Prozessvariablen, so werden diese durch ihren aktuellen Wert ersetzt. Der in Abb. 3a dargestellte Teil von $AA(X)$ $Group = Value(FzgProjekt)$ wird also umgeformt in z.B. $Group = Entwicklung\ C-Klasse$. iii) Im Falle einer Bearbeiterzuordnungen des Typs $AA(Y) = sameActor(X)$ wird $iAA(Y)$ gebildet, indem die unabhängige Bearbeiterzuordnung (iAA) der referenzierten Aktivität X übernommen wird: $iAA(Y) = iAA(X)$. Im Beispiel aus Abb. 3a ergibt sich also z.B. $iAA(Y) = „Role = SW-Entwickler \wedge Group = Entwicklung\ C-Klasse“$. Die hierdurch entstandene unabhängige Bearbeiterzuordnung $iAA(Y)$ kann dann von der in Abb. 3b dargestellten TPR, so wie in Abschnitt 3.3 beschrieben, problemlos verwendet werden. Diese Vorgehensweise kann über beliebig viele Stufen hinweg angewandt werden. So ergibt sich im Beispiel aus Abb. 3a $iAA(Z)$ analog als $iAA(Y)$.

Organisatorische Objekte des Typs *OrgUnit* stellen eine Besonderheit dar, weil jeder Benutzer üblicherweise maximal einer *OrgUnit* jeder Ebene angehören kann, d.h. maximal einem Team, einer Abteilung, einem Center, etc. Damit entfällt das bereits erläuterte Problem, dass (für den vorgegebenen SubType) z.B. die Rolle des tatsächlichen Bearbeiters der referenzierten Akt. X nicht eindeutig ist. Deshalb ist für den Fall *OrgUnit* ein verbessertes Verhalten realisierbar: Beim Start der Akt. Y aus Abb. 3a ist der Bearbeiter (User) u der Akt. X bereits bekannt. Dieser ist auch der regulär vorgesehene Bearbeiter für die Akt. Y. Der in der TPR der Akt. Y enthaltene Teil $sameOrgUnit(Abteilung)$ (Abb. 3b) kann sich also konkret auf diesen Bearbeiter u beziehen. Gehört er z.B. zur Abteilung *Sitzentwicklung*, so muss auch die Zielperson dieser Abteilung angehören. Die

rTPR erhält deshalb die Teilbedingung $OrgUnit = Sitzentwicklung$. Dies ist möglich, obwohl weder die reguläre Bearbeiterzuordnung $AA(Y)$ noch $iAA(Y)$ eine Vorgabe für die $OrgUnit$ enthält. So wird eine genauere Anwendung der TPR möglich, als wenn dieser Teil durch die in Abschnitt 3.3 vorgestellte „Standard-Vorgehensweise“ ignoriert wird. Ob dies überhaupt gewünscht ist, sollte der GP-Designer jedoch auswählen können.

Eine ähnliche Vorgehensweise ist möglich, wenn die Akt. Y einem Vorgesetzten des Bearbeiters u der Akt. X zugeordnet ist. Dies ist ein relevanter Fall, weil solche Bearbeiterzuordnungen bei Prüf- oder Genehmigungsaufgaben häufig vorkommen. Der Unterschied zum eben vorgestellten Fall ist, dass mittels des Organisationsmodells, ausgehend vom Bearbeiter u, der Vorgesetzte u' der geforderten Ebene ermittelt wird (z.B. der Teamleiter). Dessen Abteilung wird dann von der rTPR der Akt. Y referenziert. Analog kann auch bei anderen Bearbeiterzuordnungen vorgegangen werden, die einen von u abhängigen, aber eindeutigen Benutzer ergeben, wie z.B. „Frauenbeauftragter von ...“.

Eine Bearbeiterzuordnung kann sich auch auf die Organisationseinheit einer Vorgängerakt. X beziehen. So kann $AA(Y)$ einen Teil $sameOrgUnit(X, Abteilung)$ enthalten. Dieser spezifiziert diejenigen Personen, die derselben *Abteilung* (dem SubType) wie der Bearbeiter der Akt. X angehören. Auch dies ist ein relevanter Fall, insb. im Kombination mit dem 4-Augen-Prinzip, wenn z.B. eine Überprüfung durch einen anderen Mitarbeiter derselben Abteilung erfolgen soll. In solchen Fällen wird $iAA(Y)$ erzeugt, indem die $OrgUnit$ (dieser Ebene) des tatsächlichen Bearbeiters der Akt. X aus dem Organisationsmodell ermittelt wird (z.B. die Abteilung *Blinkerentwicklung*). Diese ersetzt dann den entsprechenden Teil aus der Regel $AA(Y)$, so dass sich z.B. $iAA(Y) = „Abteilung = Blinkerentwicklung \wedge nicht User43“$ ergibt (*User43* hat die Akt. X bearbeitet). Die nun in $iAA(Y)$ spezifizierte Abteilung *Blinkerentwicklung* kann durch die in Abb. 3b dargestellte TPR problemlos referenziert werden (mittels $sameOrgUnit(Abteilung)$).

4 Zusammenfassung und Ausblick

Eskalationen und Stellvertretungen sind in der Praxis wichtig, um geeignet auf die verzögerte Bearbeitung von Aktivitäten und Abwesenheit regulärer Bearbeitern zu reagieren. Deshalb wird dies von einigen kommerzielle PMS unterstützt. Allerdings werden meist nur sehr einfache Regeln (entspricht einer TPR) zur Festlegung der entsprechenden Zielpersonen angeboten. Da zudem keine ausreichend mächtigen und dennoch leicht handhabbaren wissenschaftlichen Ansätze existieren, wird in diesem Beitrag ein entsprechendes Konzept vorgestellt. Es basiert auf der Einführung eines SubType für organisatorische Objekte. Dieser ermöglicht die Referenzierung bestimmter Teile einer regulären Bearbeiterzuordnung durch eine TPR auf eine Art und Weise, welche die Verwendung der TPR für viele unterschiedliche Aktivitäten erlaubt. Dies ist sogar bei komplexen (d.h. zusammengesetzten) und abhängigen Bearbeiterzuordnungen bzw. TPR möglich.

Da die Ermittlung der Zielpersonen aus einer TPR algorithmisch recht einfach ist, wird eine prototypische Umsetzung kaum zu neuen Erkenntnissen führen. Interessanter wäre

eine Evaluation, wie leicht die GP-Designer die TPR definieren können, und ob mit dem Ansatz alle Anforderungen aus der Praxis erfüllt werden können.

Literaturverzeichnis

- [ARD07] Aalst, W. M. van der; Rosemann, M.; Dumas, M.: Deadline-based Escalation in Process-Aware Information Systems. In *Decision Support Systems*, 2007; S. 492–511.
- [Ba09] Bauer, T.: Stellvertreterregelungen für Task-Bearbeiter in prozessorientierten Applikationen. In *Datenbank-Spektrum*, 2009, 9; S. 40–51.
- [Bi21] Bizagi 11.2.3 BPM Suite User Guide. <https://help.bizagi.com/bpm-suite/en/11.2.3/>, 10.2.2021.
- [BL20] Bauer, T.; Laue, R.: Stand der anwendungsnahen Forschung und Technik für die organisatorische Perspektive von Geschäftsprozessen. In *Proc. Informatik 2020*, 6. Workshop zum Stand, den Herausforderungen und Impulsen des Geschäftsprozessmanagements, 2020; 605–619.
- [DW15] Dulai, T.; Werner-Stark, A.: A Database-Oriented Workflow Scheduler with Historical Data and Resource Substitution Possibilities. In *Proc. 4th Int. Conf. on Operations Research and Enterprise Systems*, 2015; S. 325–330.
- [Gr08] Großkopf, A.: An Extended Resource Information Layer for BPMN, Hasso-Plattner-Institute for IT Systems Engineering, Potsdam, 2008.
- [HD05] Hochmüller, E.; Dobrovnik, M.: Flexibility Issues in Workflow Management Systems. In *Proc. Business Process Modeling, Development and Support*, 2005, 5.
- [HS99] Huang, Y. N.; Shan, M. C.: Policies in a Resource Manager of Workflow Systems: Modeling, Enforcement and Management. In *Proc. 15th Int. Conf. on Data Engineering*, 1999.
- [IBM20] IBM Business Automation Workflow V20.0.0.2 Documentation. https://www.ibm.com/support/knowledgecenter/en/SS8JB4_20.x, 10.2.2021.
- [K21] K2 Cloud: Low Code Digital Process Automation. <https://www.nintex.com/process-automation/k2-software/>, 10.2.2021.
- [Mu04] Mühlen, M. Zur: Organizational Management in Workflow Applications – Issues and Perspectives. In *Information Technology and Management Journal*, 2004, 5.
- [PR98] Panagos, E.; Rabinovich, M.: Reducing Escalation-Related Costs in WFMSs. In *Workflow Management Systems and Interoperability*, 1998; S. 107–128.
- [RM98] Rosemann, M.; Mühlen, M. Zur: Modellierung der Aufbauorganisation in Workflow-Management-Systemen. In *EMISA-Forum*, 1998; S. 78–86.
- [Ru05] Russell, N. et al.: Workflow Resource Patterns: Identification, Representation and Tool Support. In *Proc. Int. Conf. on Advanced Information Systems Engineering*, 2005.
- [Si21] Signavio Workflow Accelerator. <https://www.signavio.com/products/workflow-accelerator>, 10.2.2021.

Der Einfluss bürokratischer Hürden auf die Prozessakzeptanz

Eine experimentelle Untersuchung am Beispiel eines generischen Einkaufsprozesses

Lars Drewes¹ und Volker Nissen²

Abstract: Gemäß der Prozessakzeptanztheorie hat die Akzeptanz von Prozessen einen Einfluss auf die korrekte Ausführung eben dieser. Sollen Abweichungen in der Prozessausführung verhindert werden, ist es notwendig zu verstehen, welche Faktoren einen Einfluss auf die Akzeptanz aufweisen. In der vorliegenden Untersuchung wird in einer online durchgeführten, experimentellen Studie der Einfluss von Bürokratie auf die Prozessakzeptanz untersucht. Die Basis für das Experiment bildet ein generischer Einkaufsprozess, der online über Amazon's Mechanical Turk bereitgestellt wird. Die Prozessakzeptanz wird mit Hilfe eines Fragebogens aufgeteilt in die drei Dimensionen der Einstellung (kognitiv, affektiv, konativ) gemessen. Die Ergebnisse zeigen, dass eine Freigabeerfordernis der eigenen Handlung mehr akzeptiert wird, als ein Prozess, bei dem diese nicht erforderlich ist. Dieses kontraintuitive Ergebnis kann dahingehend interpretiert werden, dass die Teilnehmer nichts falsch machen wollen und eine zusätzliche Bestätigung erwünschen. Da im Experiment sehr große Bestellsummen genutzt worden sind, soll zukünftig der Einfluss dieser auf die Akzeptanz untersucht werden.

Keywords: GPM; Prozessakzeptanz; Bürokratie; Experiment; Amazon's Mechanical Turk;

1 Einleitung

Die jahrzehntelang vorherrschende Funktionsorientierung in Unternehmen ist mittlerweile durch eine Prozessorientierung zur Verbesserung der Unternehmensleistung und Wettbewerbsfähigkeit abgelöst worden. Gutes Prozessdesign und aktives Prozessmanagement sind jedoch wesentliche Voraussetzungen, um Prozessinnovation, Kostensenkung und Kundenzufriedenheit zu erreichen [BK12, HC06, MN14, Mü15]. Außerdem müssen Prozesse tatsächlich so ausgeführt werden, wie es im Design vorgesehen ist. Die Entscheidung, einen Prozess wie gewünscht zu initiieren und auszuführen, basiert hierbei unter anderem auf der Akzeptanz, die der Prozessausführende (Akzeptanzsubjekt) dem Prozess (Akzeptanzobjekt) innerhalb eines bestimmten (Akzeptanz-)Kontextes zuschreibt. Ist diese Akzeptanz nicht vorhanden oder zu gering, kann es zu Modifikationen des Prozesses oder sogar einer Verweigerung der Ausführung kommen. Es ist auch möglich, dass

¹ Technische Universität Ilmenau, Wirtschaftsinformatik für Dienstleistungen, 98693 Ilmenau, lars.drewes@tu-ilmenau.de

² Technische Universität Ilmenau, Wirtschaftsinformatik für Dienstleistungen, 98693 Ilmenau, volker.nissen@tu-ilmenau.de

dadurch inoffizielle Prozesse, sogenannte Schattenprozesse, entstehen oder dass ähnliche, bereits existierende Ersatzprozesse eine wesentlich höhere Akzeptanz erfahren und aus diesem Grund bevorzugt werden [MN14, Mü15].

Solche Prozessabweichungen können nicht nur zu betriebswirtschaftlichen Nachteilen, wie z. B. dem Verlust von Kundenzufriedenheit und damit Umsatzeinbußen führen, sondern auch ein potenzielles Risiko für Mensch und die Umwelt darstellen. Ein Beispiel ist der anhaltende Ausbruch von MRSA-Bakterien in Krankenhäusern, der auf die mangelnde Beachtung von Hygieneprozessen zurückzuführen ist. Ein Ableiten in die Illegalität ist bei Nichtakzeptanz von Prozessen auch möglich, wenn z. B. Embargoprozesse umgangen werden, was zu empfindlichen Strafen führen kann [Mi08, MN14, Mü15].

Aufgrund der obigen Ausführungen ist es von großer Relevanz, die Einflussfaktoren auf die Prozessakzeptanz zu untersuchen und zu verstehen. Hierzu wird die Theorie zur Prozessakzeptanz von Müllerleile [Mü19] genutzt. Diese definiert und konkretisiert das Konzept der Prozessakzeptanz und schlägt dafür auch ein Messkonstrukt vor. Weiterhin werden dort eine Reihe von Einflussfaktoren auf die Akzeptanz anhand einer qualitativen Studie identifiziert, unter anderem der Faktor „Bürokratie“.

Um die Einflüsse und Zusammenhänge dieser Faktoren genauer zu verstehen und Optimierungspotentiale sowie Maßnahmen ableiten zu können, sind weitere Untersuchungen notwendig. Diese sollen dazu dienen die Effektstärken sowie mögliche Wechselwirkungen zu identifizieren. Im Rahmen der hier präsentierten Studie wird der Einflussfaktor „Bürokratie“ näher untersucht. Gewählt wird dieser Faktor, da sich bürokratische Merkmale in sämtlichen Organisationen und damit auch in betrieblichen Prozessen auffinden lassen [Ba17]. Kann der Einfluss von Bürokratie auf die Akzeptanz geklärt werden, so ergeben sich konkrete Verbesserungspotentiale betrieblicher Prozesse.

Die Untersuchung soll folgende Forschungsfragen beantworten:

FF1: Wie kann der Faktor „Bürokratie“ operationalisiert werden?

FF2: Weist ein niedriger Bürokratiegrad positive Effekte auf die Prozessakzeptanz auf?

2 Methodik

Zur Untersuchung der Effekte des Faktors Bürokratie auf die Prozessakzeptanz der am Prozess beteiligte Personen, wird im Rahmen dieser Studie ein Online-Experiment durchgeführt. Hierzu wird ein fiktives Beschaffungssystem mit dem dazugehörigen Einkaufsprozess implementiert (Kap. 3.3) und auf einem Server zur Verfügung gestellt. Dies ermöglicht den Teilnehmern³ des Experimentes den Prozess aktiv und direkt zu erleben. Zur Akquirierung der Teilnehmer wird die Crowdsourcing-Plattform Amazon's Mechanical

³ Aus Gründen der Lesbarkeit wird innerhalb dieser Arbeit das generische Maskulinum verwendet.

Turk (MTurk) genutzt. Begründet wird diese Wahl mit der kosteneffizienten Durchführbarkeit einer Studie mit großer Stichprobengröße [HRZ11]. Die Vergleichbarkeit von so durchgeführten Studien mit konventionellen Studien, die vor Ort in der Praxis stattfinden, ist mehrfach in verschiedenen Wissenschaftsdisziplinen belegt worden [BHL12, HRZ11, PCP10].

Potenzielle Störvariablen, wie Alter, Geschlecht, Herkunft, Bildungsgrad, etc., werden mittels einer Blockrandomisierung der Teilnehmer kontrolliert. Hierzu wird die notwendige Stichprobengröße in gleich große Blöcke geteilt, denen jede Ausprägung der unabhängigen Variablen jeweils einmal zugeteilt wird. Dies erlaubt eine zufällige Zuteilung der Ausprägungen und erhöht die Chance einer Gleichverteilung [SR18, Mu02]. Die Zufälligkeit wird im Fall dieser Studie durch den Pseudozufallszahlengenerator der Klasse `Random` der Sprache Java umgesetzt [OroJ]. Zur Eliminierung von Erwartungseffekten und anderen psychologischen Störfaktoren wird eine Doppelblindstudie durchgeführt. Die Teilnehmer wissen nicht in welcher Gruppe sie sich befinden bzw. wie viele und welche Gruppen existieren. Auf der anderen Seite wird der Versuchsleiter durch einen Server und damit einer künstlichen Entität dargestellt, die keine Erwartungen und Emotionen übertragen kann [DB16, SR18].

Im Rahmen online-basierter Experimente müssen neben solchen Mechanismen weitere Kontrollen implementiert werden, die ungewolltes Verhalten innerhalb der Studiendurchführung verhindern. Dazu zählen unter anderem bewusste Manipulation als auch unbewusstes Verhalten, die in verzerrten Resultaten münden. Beispielsweise können Crowdsourcing-Plattformen von Personen ausgenutzt werden, um sich schnell und ohne viel Aufwand monetär zu bereichern, indem die zur Verfügung stehenden Aufgaben nicht persönlich, sondern durch implementierte Bots durchgeführt werden. Diese führen die Aufgaben automatisiert aus und erzielen somit die mit der Aufgabe verbundene Entlohnung. Aus der Sicht des Forschers generieren Bots aber unnütze Daten, da die Aufgabenbewältigung lediglich aus dem Aspekt der Nutzenmaximierung des Entwicklers stattfindet. Fragebögen werden dementsprechend beispielsweise mit zufälligen Antworten gefüllt. Zur Unterbindung von solchem Versuchen der persönlichen Bereicherung wird, wie Mason und Suri [MS12] empfehlen, ein CAPTCHA⁴ zur Differenzierung von Mensch und Maschine genutzt [ABL04]. Innerhalb dieser Studie gestaltet sich das CAPTCHA in Form einer Auswahl eines von drei möglichen EPK-Prozessmodellen, welches mit einer textuellen Beschreibung übereinstimmt. Jeder Fehlversuch wird in den Meta-Daten dokumentiert.

Weiterhin gibt es immer wieder die Tendenz, dass Teilnehmer die gestellte Aufgabe nicht gewissenhaft erfüllen. Dies führt zu unüberlegten Antworten, die wiederum als Ausreißer in der statistischen Auswertung erscheinen. Hinzu kommt, dass einige Teilnehmer die Instruktionen nur überfliegen oder willkürlich antworten. Insgesamt führt dies zu einer Minderung der Teststärke der experimentellen Studie. Als Gegenmaßnahme wird ein „in-

⁴ Completely Automated Public Turing Test To Tell Computers and Humans Apart [CaoJ].

structional manipulation check“ (IMC) verwendet [OMD09]. Mit Hilfe dieses methodischen Tools kann identifiziert werden, ob die Teilnehmer einer Studie die gegebenen Instruktionen, Fragen und Antwortmöglichkeiten vollständig lesen und verstehen. Der IMC besteht aus einem längeren Einführungstext sowie einer Fragestellung mit vorgegebenen Antwortmöglichkeiten. Die Frage selbst ist einfach gestaltet und basiert auf dem vorher durchgeführten Prozess, so dass jeder Teilnehmer diese ohne weitere Probleme lösen kann. Der eigentliche Test und damit die Prüfung auf Aufmerksamkeit wird durch eine im Text eingebettete Anweisung realisiert. Erst das Befolgen der Anweisung, hier das Klicken auf den Titel des Tests, führt zum erfolgreichen Abschluss des IMC. Die Teilnehmer, die dieser Anweisung nicht Folge leisten und stattdessen die Frage beantworten, fallen durch den Test und können diesen wiederholen. Dies soll die Teilnehmer dazu zwingen die angeführten Informationen vollständig zu erfassen [OMD09]. Auch hier wird jeder Fehlversuch in den Meta-Daten des Versuchs dokumentiert.

3 Aufbau des Online-Experimentes

3.1 Operationalisierung der abhängigen Variablen „Prozessakzeptanz“

Innerhalb der von Müllerleile erarbeiteten Theorie der Prozessakzeptanz wird der Begriff unter Einbezug der allgemeinen Akzeptanzdefinition nach Lucke [Lu95] wie folgt definiert: „Prozessakzeptanz ist die affirmativ positive Grundeinstellung der Prozessbeteiligten gegenüber einem Prozess. Diese Einstellung resultiert aus einer inneren, rational oder motivational-emotional geprägten Überzeugung, durch die der Prozess gebilligt und inhaltlich in seiner Gesamtheit anerkannt wird. Diese Überzeugung mündet in einem Verhalten, welches einer unveränderten Ausführung des vorgegebenen Prozesses entspricht [Mü19]“. Das zur Theorie dazugehörige Messkonstrukt [Mü19] wurde unter Anwendung der C-OAR-SE Methodik von Rossiter [Ro11] entwickelt und in die drei Dimensionen der Einstellungsakzeptanz unterteilt: Kognitiv, affektiv und konativ [MM86]. Erstere umfasst die Einstellung gegenüber einem Objekt, die auf dem verfügbaren Wissen und dem Verstand basiert. Die affektive Dimension erfasst motivational-emotionale Empfindungen und die konative bezieht sich auf die innere Bereitschaft eine Handlung durchzuführen, auch wenn diese letztendlich nicht durchgeführt werden muss [Mü19]. Jede der drei Dimensionen umfasst eine Menge von Frage-Antwort-Paaren, die mit dem DLF IIST Binary Ansatz [Ro11] operationalisiert worden ist. Dahingehend kann der Teilnehmer lediglich zwischen den Antwortoptionen „Yes“ und „No“ wählen. Er wird also gezwungen eine Antwort abzugeben. Tab. 1 zeigt beispielhaft die kognitive Dimension des hier verwendeten Fragebogens. Die Antworten werden entsprechend ihrer Ladung nach effektkodiert [BS16]. Dabei gibt die Ladung an, ob die Frage dem Prozess gegenüber positiv gerichtet ist oder nicht. Bei einer positiven Ladung wird „Yes“ mit 1 und „No“ mit -1 kodiert. Bei einer negativen Ladung entsprechend umgekehrt. Zusätzlich zu diesen Fragen enthält jede Dimension eine fünfstufige bipolare Likert-Skala, die im Intervall [-2; 2] kodiert wird. Für die kognitive Dimension entsprechend:

[Overall I reject this way of purchasing material] -2 -1 0 1 2 [Overall I approve this way of purchasing material]

Zur Bildung der drei Akzeptanzkennzahlen werden die kodierten Antworten pro Dimension aggregiert. Diese fließen in die statistischen Auswertungen ein.

Frage	Skala	Ladung
I think this way of managing purchases is useful.	Yes/No	+
I think this way of managing purchases is important.	Yes/No	+
I think this way of managing purchases is complicated.	Yes/No	-
I think this way of managing purchases is laborious.	Yes/No	-
I think this way of managing purchases takes a long time.	Yes/No	-
I think I'm happy with the result.	Yes/No	+
I think this way of managing purchases is decent.	Yes/No	+

Tab. 1: Kognitive Dimension des Messkonstruktes nach Müllerleile [Mü19]

3.2 Definition und Operationalisierung der unabhängigen Variablen

Als Ergebnis seiner qualitativen Forschung stellte Müllerleile [Mü19] die These auf, dass eine zu überbordende Bürokratie innerhalb eines Prozesskontextes dessen Ausführung behindert und damit die Akzeptanz negativ beeinflusst. Gemäß seinen theoretischen Überlegungen sollte ein ausgewogenes Verhältnis zwischen bürokratischen und damit starren Elementen und Flexibilität angestrebt werden. Zur Untersuchung dieser Ausführungen sollen nun folgende theoretisch-inhaltliche Hypothesen⁵ aufgestellt werden, deren Richtung auch durch die Annahme, dass Bürokratie die Amtsinhaber und damit die Prozessbeteiligten in Ihrer Handlungsfreiheit einschränkt, begründet wird [Oe05]: Sinkt der Grad der Bürokratie in einem Prozess, so steigt die Prozessakzeptanz, die eine am Prozess beteiligte Person diesem beimisst, in der kognitiven/affektiven/konativen Dimension.

Bevor nun eine konkrete Operationalisierung der Variablen Bürokratie durchgeführt werden kann, muss dieses theoretische Konstrukt zunächst definiert werden. Ebenso muss festgelegt werden, was unter einer überbordenden Bürokratie verstanden werden soll. Müllerleile selbst gibt zur Beschreibung des Begriffes lediglich die Beispiele „Anzahl an Formularen im Prozess“ bzw. „Wiederholungshäufigkeit von Formularen“ an [Mü19]. Dies erscheint hier viel zu ungenau, da lediglich ein kleiner Teil bürokratischer Möglichkeiten erfasst wird.

Wesentlich geprägt wurde der heutige Begriff der Bürokratie von Max Weber im frühen 20. Jh., dessen Schriften, im Hinblick auf Einfluss und entstandenen Diskussionen große Bedeutung in der Wissenschaft aufweisen [AR72, Ba17, We22]. Auch heute noch gilt das von ihm entwickelte Modell als wesentliches Element innerhalb theoretischer Analysen

⁵ Da der Unterschied zwischen den Hypothesen lediglich die Dimensionen der Einstellung betrifft, werden diese hier zur Vereinfachung zusammengefasst.

und dem praktischen Entwurf formaler Organisationsstrukturen [Ba17, BJ09, Oe05]. Weber, der die Bürokratie als Herrschaftsform ansieht, hat den Begriff selbst zwar nie abschließend definiert, aber sechs konstitutive Merkmale erarbeitet. Diese sollen den Idealtyp einer bürokratischen Ordnung widerspiegeln und werden nachfolgend erläutert [AR72, Ba17, WW90].

Arbeitsverteilung und Befehlsgewalt: Kompetenzen werden prinzipiell durch zuvor fest definierte Regeln vergeben. Die entsprechende Aufgabenverteilung inklusive der inhärenten Handlungs- und Entscheidungsrechte steht mit dieser Vergabe fest und ist an die Stelle gebunden. Damit werden diese von der zum Zeitpunkt der Betrachtung ausübenden Person, dem Amtsinhaber, entkoppelt [Ba17, WW90].

Amtshierarchie: Innerhalb einer bürokratischen Ordnung sollte es ein festgelegtes System aus hierarchischer Über- und Unterordnung geben. Dabei besitzen die übergeordneten Positionen Weisungs- und Kontrollrechte gegenüber den entsprechenden untergeordneten Stellen [Ba17, WW90].

Aktenmäßigkeit der Verwaltung: Die Erfüllung der einem zugewiesenen Aufgaben muss in schriftlicher Form mittels Aktennotizen, Formularen oder anderen Dokumenten erfolgen. Zweck dieser strikten Schriftlichkeit ist die nachhaltige Dokumentation und Kontrolle der Aufgabenbewältigung sowie der Kontinuität. Relevant wird dies unter anderem beim Wechsel des Amtsinhabers [Ba17, WW90].

Amtsspezifische Schulung: Zur Ausübung, der an die Stelle gebundenen Aufgaben und der damit verbundenen Wahrnehmung von Kompetenzen, ist es erforderlich entsprechende fachliche Schulungen zu erhalten. Beförderungen sollen durch diese determiniert werden [Ba17, Oe05, WW90].

Amtszeit: Die Bearbeitung der zugeteilten Aufgaben hat die gesamte Arbeitszeit zu beanspruchen. Privates muss aus dem Arbeitsalltag entfernt werden [BTC14, WW90].

Regelgebundenheit der Amtsführung: „Die Amtsführung [...] erfolgt nach generellen, mehr oder minder festen und mehr oder minder erschöpfenden, erlernbaren Regeln [WW90]“. Diese beziehen sich auf die zu erbringenden Leistungen und Kommunikationswege und stellen sicher, dass eine Gleichbehandlung sichergestellt wird [Ba17].

Anhand der Aufschlüsselung der bürokratischen Merkmale kann festgehalten werden, dass der Faktor Bürokratie ein durchaus komplexes Konstrukt darstellt. Da die Variation all dieser Merkmale zur selben Zeit ein nahezu unmögliches Unterfangen darstellen würde, gilt es zunächst ein für den ausgewählten Prozess sinnvolles Merkmal zu identifizieren und anschließend dessen Variationen festzulegen.

Unter Bewertung der implementierbaren Möglichkeiten wird der Faktor Bürokratie in dieser Untersuchung durch das Merkmal Arbeitsteilung und Befehlsgewalt operationalisiert. In unternehmerischen Einkaufsprozessen, wie dem Standard P2P-Prozess in einem SAP

ERP-System, sind oft Freigabeaktivitäten für Bestellanforderungen oder Bestellungen enthalten. Die Kompetenz solche Freigaben durchzuführen liegt entsprechend oft bei den Leitern der Struktureinheiten, wodurch sie die Entscheidungsrechte besitzen⁶, sowohl über eine grundsätzliche Freigabebedürftigkeit zu entscheiden, als auch die Freigabe selbst zu treffen. Es besteht aber weiterhin die Möglichkeit, dass die Kompetenz, die Entscheidung über den Kauf zu treffen, den Sachbearbeitern selbst zugeordnet wird. Daraus kann die Variation abgeleitet werden, dass in einer Gruppe der Leiter entscheidet, und damit eine Freigabe erfolgen muss (Gruppe 1), und in der anderen Gruppe der Sachbearbeiter und damit Teilnehmer selbst (Gruppe 2). In der Praxis ist dies ebenso denkbar und in ERP-Systemen beispielsweise auch so umsetzbar, da eine Freigabebedürftigkeit nicht zwingend notwendig ist bzw. vorsieht, dass die Kompetenz bis zu einem gewissen Rahmen beim Sachbearbeiter liegt und erst bei Erreichen bestimmter Grenzen beim Leiter. Beispielsweise könnte diese Grenze durch Warengruppen oder der Höhe des Nettobestellwertes festgelegt werden. Da die zweite Variante hier die Hürden, die der Sachbearbeiter und damit der Teilnehmer auf sich nehmen muss, verringert, soll dies dementsprechend als Verringerung der Bürokratie gesehen werden. Die Freigabe durch den Leiter wiederum wird als überbordende Bürokratie definiert, da diese in der ersten Variante immer erforderlich ist.

	Gruppe 1 (mit Freigabe)	Gruppe 2 (ohne Freigabe)
Wartezeit auf Laden der Bestellanforderungen	1,5 Sekunden	1,5 Sekunden
Wartezeit auf Freigabe	7,5 Sekunden	entfällt
Wartezeit auf Abschluss des Sendeprozesses	1 Sekunden	8,5 Sekunden
Gesamte Wartezeit im Prozess	10 Sekunden	10 Sekunden

Tab. 2: Wartezeit der einzelnen Funktionen

Die Freigabe im Prozess wird durch den Server durchgeführt und wird mit einer festen Zeit in Höhe von 7,5 Sekunden implementiert. Damit nun die gemessene Akzeptanz nicht dadurch verfälscht wird, dass die durch die Freigabe determinierte Wartezeit entfällt, wird in der Variante ohne Freigabe die Wartezeit auf das erfolgreiche Absenden entsprechend erhöht. Zusammen mit der Wartezeit auf das Laden der Bestellanforderungen, welche ebenso fest implementiert ist, beträgt die Gesamtwarezeit in beiden Varianten insgesamt 10 Sekunden (siehe Tab. 2). Zehn Sekunden erscheinen zwar als sehr gering, allerdings zeigt eine bereits durchgeführte Studie, dass auch geringe Wartezeiten einen signifikanten Einfluss auf die Prozessakzeptanz aufweisen [DN21]. Demnach sollte diese in allen Gruppen gleich gehalten werden.

⁶ Innerhalb des Prozesses wird postuliert, dass die Freigaberegeln vom Einkaufsleiter und nicht von höheren Instanzen festgelegt werden.

Aus der Operationalisierung werden folgende empirisch-inhaltliche Hypothesen⁷ abgeleitet: Sinkt der Grad der Bürokratie in einem Prozess gemessen durch die Reduktion des Merkmals Arbeitsteilung und Befehlsgewalt in Form der Entfernung einer Freigabeaktivität in einem Beschaffungssystem, so steigt die Prozessakzeptanz, die eine am Prozess beteiligte Person diesem beimisst, in der kognitiven/affektiven/konativen Dimension.

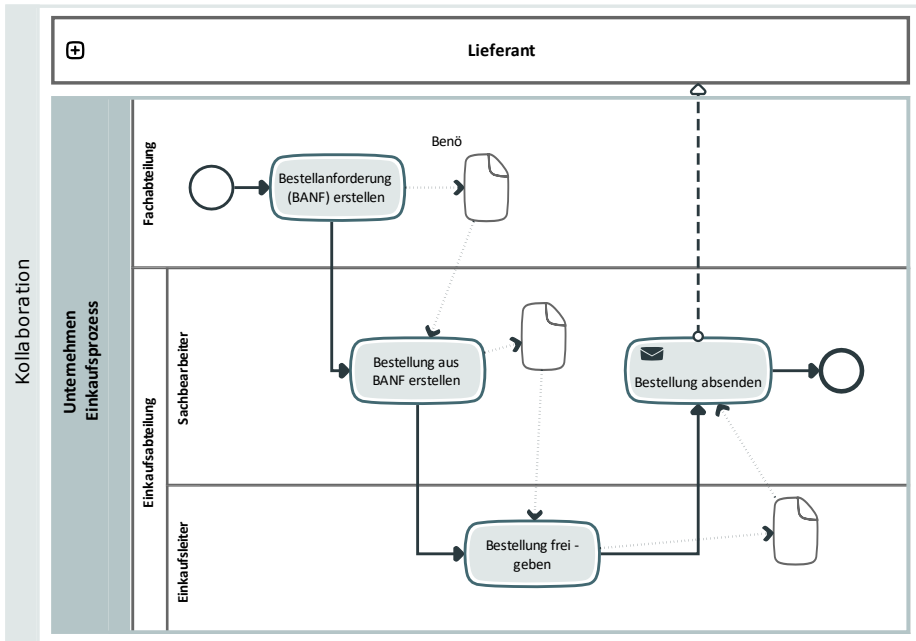


Abb. 1: Verwendeter generischer Einkaufsprozess

3.3 Prozessdesign

Die Basis für das durchzuführende Experiment bildet ein generischer Einkaufsprozess (Abb. 1). Ziel dahinter ist, es eine möglichst große Nähe zu operativen Unternehmensprozessen und zeitgleich zu Prozessen des alltäglichen Privatgebrauchs zu schaffen. Dadurch soll die externe Validität und damit die Generalisierbarkeit der Untersuchungsergebnisse erhöht werden [DB16]. Der Prozess, einschließlich aller Variationen der unabhängigen Variablen wird mit der Skriptsprache JavaScript und HTML umgesetzt. Publiziert wird

⁷ Da der Unterschied zwischen den Hypothesen lediglich die Dimensionen der Einstellung betrifft, werden diese hier zur Vereinfachung zusammengefasst.

dieser anschließend über einen JATOS⁸-Server, der in Verbindung mit MTurk betrieben wird [LKF15].

Der Prozess selbst läuft wie folgt ab: Nachdem der Teilnehmer in das Experiment mittels textueller Beschreibung eingeführt worden ist, sowie das CAPTCHA gelöst hat, startet der eigentliche Prozess mit der Erstellung einer Bestellanforderung (BANF) durch einen Mitarbeiter einer Fachabteilung. Im Rahmen der Untersuchung wird dieser durch den Experimentalserver repräsentiert. Jeder Teilnehmer erhält somit zunächst eine Bestellanforderung. Diese wird in einem speziellen Cockpit angezeigt. Der Teilnehmer selbst repräsentiert den Sachbearbeiter der Einkaufsabteilung. Jede Bestellanforderung muss in eine entsprechende Bestellung umgesetzt werden. Diese muss in einer der beiden Gruppen im Anschluss vom Einkaufsleiter, also hier vom Server, freigegeben werden. Nach Freigabe kann der Teilnehmer die Bestellung absenden. Sollte der Teilnehmer falsches Material bestellen wollen, wird die Freigabe verweigert. Innerhalb der anderen Gruppe ist die Freigabe nicht notwendig und der Teilnehmer kann die Bestellung direkt absenden. Im gesamten Experiment erhält der Teilnehmer zwei Bestellanforderungen, so dass er zwei Prozessinstanzen seriell durchlaufen muss und damit das Vorgehen wiederholt. Im Anschluss an den Prozess folgt der IMC gefolgt vom Fragebogen zur Erfassung der Prozessakzeptanz in den genannten drei Dimensionen.

3.4 Versuchsplan und Stichprobenberechnung

Da zwei Ausprägungen bezüglich der unabhängigen Variablen gewählt werden, wird für die Untersuchung ein Zweigruppenplan verwendet. Zur Berechnung der optimalen Stichprobengröße wird die Verwendung eines zweiseitigen t-Tests postuliert. Im Einklang mit Empfehlungen in der Literatur wird α auf 0,05 und β auf 0,80 gesetzt [DB16]. Damit auch eine kleine Effektstärke nachgewiesen werden kann, wird $d=0,2$ (kleiner Effekt nach Cohen [Co88]) gesetzt. Daraus errechnet sich mit Hilfe des Tools G*Power (Version 3.1.9.7) [Fa07] eine optimale Stichprobengröße in Höhe von 788 Teilnehmern. Diese wird gleichmäßig auf beide Gruppen verteilt.

4 Datenanalyse

Das Experiment wurde mit einer optimalen Stichprobengröße in Höhe von 788 Teilnehmern im Zeitraum 15.01.2021 – 26.01.2021 über MTurk durchgeführt. 18 Teilnehmer mussten aufgrund fehlender Daten aus dem Datenset entfernt werden. Eine Imputation war an dieser Stelle nicht möglich, da diese 18 Teilnehmer die Studie vor Erreichen des Fragebogens komplett abgebrochen haben. Dementsprechend basieren die weiteren Auswertungen auf den Daten der restlichen 770 Teilnehmer. Gruppe 1 (mit Freigabe) sind 370 Teilnehmer und Gruppe 2 (ohne Freigabe) 400 Teilnehmer randomisiert zugeordnet worden. Das Ungleichgewicht zwischen den Gruppen wird darin begründet, dass Teilnehmer

⁸ Just another tool for online surveys (JATOS)

die Studie gestartet und damit vom Randomisierungsalgorithmus zugeordnet worden sind, die Studie allerdings direkt wieder abgebrochen haben. Somit wurde trotz Blockrandomisierung kein exaktes Gleichgewicht erzielt. Der Großteil der Teilnehmer (70,6 %) stammt aus den USA, gefolgt von Indien (9 %) und Großbritannien/Nord-Irland (4,8 %). Der Rest verteilt sich auf 18 weitere Länder. Das mittlere Alter der Teilnehmer liegt bei 34,61 Jahren, wobei der jüngste Teilnehmer 16 und der älteste 79 Jahre alt ist. Die Geschlechterverteilung liegt bei 54,7 % männlich, 44,4 % weiblich und 0,9 % divers. 50,3 % der Teilnehmer geben an bereits solche Einkaufssysteme bedient zu haben.

Zur Prüfung der Hypothesen soll nun für jede Dimension der Einstellung ein statistischer Signifikanztest⁹ durchgeführt werden. Da die Hypothesen allerdings gerichtet sind, muss zunächst über einen Mittelwert-Vergleich geprüft werden, ob die Daten entsprechend der in den Hypothesen unterstellten Richtung verlaufen. In Tab. 3 sind die Mittelwerte pro Dimension und Gruppe aufgeschlüsselt. Ein höherer Mittelwert entspricht dabei einer höheren Akzeptanz. Wie zu sehen ist, verlaufen damit die Mittelwerte entgegen der Hypothesenrichtung. Die Hypothesen können damit direkt verworfen werden [DB16]. Da allerdings die Ergebnisse dennoch zur Verbesserung des Prozessentwurfs bzw. der -optimierung genutzt werden können, soll weiter auf statistische Signifikanz der entgegengerichteten Hypothesen geprüft werden.

Gruppe	Kognitiv	Affektiv	Konativ
Mit Freigabe	M = 5,79 SD = 3,97	M = 7,66 SD = 4,76	M = 4,53 SD = 4,64
Ohne Freigabe	M = 5,64 SD = 3,81	M = 6,56 SD = 5,41	M = 3,86 SD = 4,72

Tab. 3: Mittelwert und Standardabweichung der Akzeptanzwerte in den Dimensionen pro Gruppe

Da ein Zweigruppen-Experiment durchgeführt worden ist, fällt die Wahl hier auf einen unabhängigen t-Test. Zunächst werden die dem Test entsprechenden Voraussetzungen geprüft. In der affektiven Dimension liegen zwei Datensätze vor, deren Akzeptanzwert über dem 3-fachen des Interquartilsabstandes zum 1. Quartil liegt und damit als extreme Ausreißer klassifiziert werden. Da jedoch kein Hinweis vorliegt, dass diese Datenpunkte auf Mess- oder Dateneingabefehlern basieren, werden sie nicht entfernt. Da die Stichprobengröße > 30 ist, kann auf die Prüfung der Normalverteilungsvoraussetzung verzichtet werden [Bo60, BS16, SB92]. Dennoch wird zur weiteren Verifikation ein nicht-parametrischer Test (Mann-Whitney-U-Test) angeschlossen. Die Voraussetzung der Varianzhomogenität wird mittels Levene Test geprüft (Tab. 4).

⁹ Alle statistischen Tests dieser Studie sind mit IBM SPSS v27 berechnet worden.

Dimension	F	Sig
Kognitiv	,116	,733
Affektiv	13,137	,000
Konativ	,557	,456

Tab. 4: Levene Test

Im Fall der affektiven Dimension kann, aufgrund $p < 0,001$, keine Varianzhomogenität festgestellt werden. Aus diesem Grund wird für diese Dimension anstatt des t-Tests ein Welch-Test durchgeführt [Ru06].

Dimension	Test	T	Df	Sig.	Cohen's d
Kognitiv	t-Test	,541	768	,294	
Affektiv	Welch-Test	3,000	766,009	,002	,215
Konativ	t-Test	1,989	768	,024	,143

Tab. 5: Signifikanztests der Dimensionen

In Tab. 5 sind die Ergebnisse der Signifikanztests abzulesen. Da SPSS immer zweiseitig prüft, wurden die p-Werte für eine einseitige Prüfung entsprechend halbiert [BS16]. In der affektiven als auch in der konativen Dimension kann ein statistisch signifikanter Unterschied zwischen den Akzeptanzmesswerten der Gruppe mit Freigabe und der Gruppe ohne Freigabe festgestellt werden ($p < ,05$). Diese Feststellung kann durch einen Mann-Whitney-U-Test bestätigt werden. Die Signifikanzwerte liegen dort bei Sig. (Affektiv) = ,006 und Sig. (Konativ) = ,034. Zur abschließenden Bewertung der entgegengesetzten Hypothese (\overline{ETH}) „Sinkt der Grad der Bürokratie in einem Prozess gemessen durch die Entfernung einer Freigabeaktivität, so *sinkt* die Prozessakzeptanz, die eine am Prozess beteiligte Person diesem beimisst, [...]“ müssen die erreichten Effektstärken miteinbezogen werden. Da keine vergleichbaren Studien in diesem Feld zur Bewertung hinzugezogen werden können, wird ein mittlerer Effekt nach Cohen [Co88] in Höhe von $d = .5$ zur Bewertung genutzt, da dieser am häufigsten in einem breiten Querschnitt psychologischer Studien vorkommt [SR18]. Unter Anbetracht der vorliegenden Informationen kann \overline{ETH} nun, gemäß dem Paradigma nach Westermann und Hager [WH82] angenommen und als bedingt bewährt klassifiziert werden.

5 Diskussion der Ergebnisse und Ausblick

Die Ergebnisse des durchgeführten Experimentes bestätigen nicht, die aufgestellte Vermutung, dass eine Reduktion einer Freigabeerfordernis die Akzeptanz bezüglich eines Prozesses verbessert. Allerdings bezieht sich diese Erkenntnis lediglich auf die empirisch-inhaltliche Hypothese, die nach Wahl des Merkmals und der Operationalisierung aufgestellt worden ist. Die theoretisch-inhaltliche Hypothese, die sich allgemein auf den Faktor Bürokratie bezieht, kann hier nicht abschließend bewertet werden. Wie in der Definition

des Faktors zu erkennen ist, bildet diese ein komplexes Konstrukt aus konstitutiven Merkmalen. Diese müssen alle für sich empirisch untersucht werden. Zudem beschränkt sich dieses Experiment mit dem Freigabeerfordernis auf eine mögliche Ausprägung des Merkmals Arbeitsteilung und Befehlsgewalt.

Wie ist nun das hier erhaltene Ergebnis im Hinblick auf die Praxis bzw. auf Prozessentwurf und -optimierung zu interpretieren. Es entsteht der Eindruck, dass die Teilnehmer bei Ihrer Bestellung eine Kontrolle der eigenen Handlungen wünschen. Der durch die Bestellanforderungen angegebene Bestellwert lag im Experiment bei 14.950 € und 25.000 €. Die Vermutung liegt nahe, dass die Teilnehmer solche Summen lieber vom Verantwortlichen bestätigt bekommen, als dass sie etwas falsch machen. Der stärkste Effekt auf die Akzeptanz ist in der affektiven und damit motivational-emotionalen Dimension vorzufinden. Dies würde die Vermutung unterstützen. Der erste Ansatz für eine Empfehlung für die Praxis, der daraus abgeleitet werden kann, ist es, in solchen Fällen mit erheblicher finanzieller Tragweite, ein Vier-Augen-Prinzip, bspw. in Form einer Freigabe, einzuführen. Falls solche Maßnahmen bereits etabliert sind, zeigt die durchgeführte Untersuchung, dass diese erhalten bleiben sollten.

Weiterhin wäre es von Interesse zu verstehen, ob sich die emotionale Einstellung der Teilnehmer hinsichtlich der Freigabe ändert, wenn der Bestellwert deutlich kleiner ist. Lediglich 13,4% der Teilnehmer haben laut Ihren Angaben bereits in einer Einkaufsabteilung und demnach vermutlich mit großen Bestellwerten gearbeitet. Demnach soll zukünftig in einem weiteren Experiment überprüft werden, ob niedrigere Bestellwerte zu anderen Resultaten führen und ausgefeiltere Freigabestrategien demnach empfehlenswert erscheinen. Die Studie hat insgesamt gezeigt, dass ein Experiment mit 788 Teilnehmer in kurzer Zeit weltweit durchgeführt werden kann und signifikante Ergebnisse liefert. Somit stellt diese Art der Untersuchung eine effiziente Alternative zu konservativen Feldstudien dar. Aus diesem Grund sollen zukünftig weitere Faktoren mittels dieser Vorgehensweise experimentell auf Ihren Einfluss auf die Prozessakzeptanz untersucht werden.

Literaturverzeichnis

- ABL04 Ahn, L. von; Blum, M.; Langford, J.: Telling humans and computers apart automatically. *Communications of the ACM* 2/04, S. 57–60, 2004.
- AR72 Albrow, M.; Roeckl, U.: *Bürokratie*. Dt. Erstausg. List, München, 1972.
- Ba17 Bach, N.; Brehm, C.; Buchholz, W.; Petry, T.: *Organisation. Gestaltung wertschöpfungsorientierter Architekturen, Prozesse und Strukturen*. 2., vollständig überarbeitete und erweiterte Auflage. Springer Gabler, Wiesbaden, 2017.
- BHL12 Berinsky, A. J.; Huber, G. A.; Lenz, G. S.: Evaluating Online Labor Markets for Experimental Research. *Amazon.com's Mechanical Turk. Political Analysis* 03/12, S. 351–368, 2012.
- BJ09 Bogumil, J.; Jann, W.: *Verwaltung und Verwaltungswissenschaft in Deutschland. Einführung in die Verwaltungswissenschaft*. 2., völlig überarb. Aufl. VS Verl. für Sozialwiss., Wiesbaden, 2009.

- BK12 Becker, J.; Kahn, D.: Der Prozess im Fokus. In (Becker, J.; Kugeler, M.; Rosemann, M. Hrsg.): Prozessmanagement. Ein Leitfaden zur prozessorientierten Organisationsgestaltung. Springer Gabler, Berlin, Heidelberg, S. 3–16, 2012.
- Bo60 Boneau, C. A.: The effects of violations of assumptions underlying the test. *Psychological Bulletin* 60, S. 49–64, 1960.
- BS16 Bortz, J.; Schuster, C.: Statistik für Human- und Sozialwissenschaftler. Extras online. Limitierte Sonderausgabe, 7., vollständig überarbeitete und erweiterte Auflage. Springer, Berlin, Heidelberg, 2016.
- BTC14 Bonazzi, G.; Tacke, V.; Corti, A. Hrsg.: Geschichte des organisatorischen Denkens. Springer VS, Wiesbaden, 2014.
- CaoJ Carnegie Mellon University: The Official CAPTCHA Site. <http://www.captcha.net/>, 05.06.2019.
- Co88 Cohen, J.: Statistical power analysis for the behavioral sciences. 2. ed. Erlbaum, Hillsdale, NJ, 1988.
- DB16 Döring, N.; Bortz, J.: Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften. 5. vollständig überarbeitete, aktualisierte und erweiterte Auflage. Springer, Berlin, Heidelberg, 2016.
- DN21 Drewes, L.; Nissen, V.: The Effect of Process Length on Process Acceptance. In (Reussner, R. H.; Koziolok, A.; Heinrich, R. Hrsg.): 50. Jahrestagung der Gesellschaft für Informatik INFORMATIK 2020. Back to the Future, S. 635–648, 2021.
- Fa07 Faul, F.; Erdfelder, E.; Lang, A.-G.; Buchner, A.: G*Power 3. A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods* 2/07, S. 175–191, 2007.
- HC06 Hammer, M.; Champy, J.: Reengineering the corporation. A manifesto for business revolution. Paperback ed., rev. and updated. Collins Business Essentials, New York, 2006.
- HRZ11 Horton, J. J.; Rand, D. G.; Zeckhauser, R. J.: The online laboratory: conducting experiments in a real labor market. *Experimental Economics* 3/11, S. 399–425, 2011.
- LKF15 Lange, K.; Kühn, S.; Filevich, E.: "Just Another Tool for Online Studies" (JATOS): An Easy Solution for Setup and Management of Web Servers Supporting Online Studies. *PloS one* 6/15, e0130834, 2015.
- Lu95 Lucke, D.: Akzeptanz. Legitimität in der „Abstimmungsgesellschaft“. VS Verlag für Sozialwissenschaften, Wiesbaden, s.l., 1995.
- Mi08 Militz, M.: Strategien zur Eingrenzung multiresistenter Erreger im Krankenhaus. *Trauma und Berufskrankheit* S1/08, S. 140–145, 2008.
- MM86 Müller-Böling, D.; Müller, M.: Akzeptanzfaktoren der Bürokommunikation. Oldenbourg, München, 1986.
- MN14 Müllerleile, T.; Nissen, V.: When Processes Alienate Customers. Towards a Theory of Process Acceptance. In (van der Aalst, W. et al. Hrsg.): S-BPM ONE - Scientific Research. Springer International Publishing, Cham, S. 171–180, 2014.

- MS12 Mason, W.; Suri, S.: Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods* 1/12, S. 1–23, 2012.
- Mu02 Muche, R.; Rohlmann, F.; Büchele, G.; Gaus, W.: Randomisierung in klinischen Studien in der Rehabilitationsforschung: Grundlagen und praktische Aspekte. *Die Rehabilitation* 5/02, S. 311–319, 2002.
- Mü15 Müllerleile, T.; Ritter, S.; Englisch, L.; Nissen, V.; Joenssen, D. W.: The Influence of Process Acceptance on BPM. An Empirical Investigation. In (Aveiro, D.; Caetano, A. Hrsg.): 2015 IEEE 17th Conference on Business Informatics (CBI). IEEE, Piscataway, S. 125–132, 2015.
- Mü19 Müllerleile, T.: *Prozessakzeptanz. Theoretische und empirische Untersuchung der Akzeptanz und Ablehnung betrieblicher Prozesse*. 1. Auflage 2019. Springer Gabler, Wiesbaden, 2019.
- Oe05 Oelsnitz, D. von der: Max Webers Bürokratiemodell noch gültig in der globalen Wissensgesellschaft? *Das Wirtschaftsstudium* 11/05, S. 1384–1389, 2005.
- OMD09 Oppenheimer, D. M.; Meyvis, T.; Davidenko, N.: Instructional manipulation checks. Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology* 4/09, S. 867–872, 2009.
- OroJ Oracle: Random (Java Platform SE 8). <https://docs.oracle.com/javase/8/docs/api/java/util/Random.html#nextInt-->, 24.03.2021.
- PCP10 Paolacci, G.; Chandler, J.; Panagiotis, I. G.: Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making* 5/10, S. 411–419, 2010.
- Ro11 Rossiter, J. R.: *Measurement for the Social Sciences. The C-OAR-SE Method and Why It Must Replace Psychometrics*. Springer Science+Business Media LLC, New York, 2011.
- Ru06 Ruxton, G. D.: The unequal variance t-test is an underused alternative to Student's t-test and the Mann–Whitney U test. *Behavioral Ecology* 4/06, S. 688–690, 2006.
- SB92 Sawilowsky, S. S.; Blair, R. C.: A more realistic look at the robustness and Type II error properties of the t test to departures from population normality. *Psychological Bulletin* 2/92, S. 352–360, 1992.
- SR18 Sedlmeier, P.; Renkewitz, F.: *Forschungsmethoden und Statistik für Psychologen und Sozialwissenschaftler. Für Psychologen und Sozialwissenschaftler*. 3., aktualisierte und erweiterte Auflage. Pearson, Hallbergmoos, 2018.
- We22 Weber, M.: *Wirtschaft und Gesellschaft. Grundriss der Sozialökonomik*. J. C. B. Mohr (Paul Siebeck), Tübingen, 1922.
- WH82 Westermann, R.; Hager, W.: Entscheidung über statistische und wissenschaftliche Hypothesen: Zur Differenzierung und Systematisierung der Beziehungen. *Zeitschrift für Sozialpsychologie* 13/82, S. 13–21, 1982.
- WW90 Weber, M.; Winckelmann, J.: *Wirtschaft und Gesellschaft. Grundriss der verstehenden Soziologie*. 5., rev. Aufl., Nachdr., Studienausg. Mohr, Tübingen, 1990.

Ein systematisches Literatur-Review des IT- Prozessmanagements

State-of-the-Art des Teilbereichs IT-Service-Management

Fabian Sauer¹, Manuel Wittig² und Helmut Beckmann³

Abstract: Das IT-Service-Management (ITSM) als Teilbereich des IT-Prozessmanagements gewinnt in Unternehmen durch den immer stärkeren Einsatz von IT-Services, welche zur Leistungserbringung wichtig sind, sowohl in der Praxis als auch in der Literatur an Bedeutung. Um die Veränderungen dieses Bereichs sowie die darin fokussierten Themen übersichtlich darzustellen, bedarf es eines systematischen State-of-the-Art Literaturreviews. Im vorliegenden Beitrag haben wir das ITSM im Rahmen der letzten fünf Jahre (2015-2020) betrachtet, relevante Thematiken identifiziert, sowie Veränderungen aufgezeigt. Neben der inhaltlichen Betrachtung von 96 relevanten Arbeiten wurde zur Risikoreduktion des subjektiven Einflusses der Autoren eine Keyword-Betrachtung vorgenommen. Im Rahmen dieser Arbeit wurde ersichtlich, dass sich das ITSM-Framework „Information Technology Infrastructure Library“ (ITIL) als führendes Rahmenwerk dieses Forschungsgebiets auch weiterhin bewährt. Außerdem konnte aufgezeigt werden, dass sich vermehrt mit Technologien wie der Cloud oder dem Machine Learning befasst wird, um diese im Rahmen des ITSM zu fokussieren.

Keywords: IT-Prozessmanagement, IT Process Management, IT-Service-Management, State-of-the-Art, Wirtschaftsinformatik, Information Systems

1 Einleitung

In der heutigen wettbewerbsintensiven Geschäftsumgebung spielt die Informationstechnologie (IT) eine entscheidende Rolle bei der Unterstützung von Geschäftsfunktionen und der Erfüllung von Geschäftsanforderungen. Ein effektives und effizientes IT-Management hat dabei das Potenzial, das Geschäft zu transformieren und dessen Leistung zu optimieren [Je16, To10]. Im Gegensatz zu den traditionellen technologieorientierten Ansätzen konzentriert sich das IT-Service-Management (ITSM)

¹ Hochschule Heilbronn, Institut für Wirtschaftsinformatik, Max-Planck-Straße 39, 74081 Heilbronn, fsauer@stud.hs-heilbronn.de

² Hochschule Heilbronn, Institut für Wirtschaftsinformatik, Max-Planck-Straße 39, 74081 Heilbronn, mwittig@stud.hs-heilbronn.de

³ Hochschule Heilbronn, Institut für Wirtschaftsinformatik, Max-Planck-Straße 39, 74081 Heilbronn, helmut.beckmann@hs-heilbronn.de

auf die Verwaltung des IT-Betriebs als Service. IT-Serviceanbieter müssen dabei nicht nur mehr Technologien und ihre interne Organisation betrachten, sondern müssen sich zudem auf die Qualität ihrer Dienstleistungen und Kundenbeziehung konzentrieren [Ga09].

1.1 Problemdefinition

Die Anzahl an wissenschaftlichen Beiträgen, welche das ITSM behandeln, steigt kontinuierlich an. Es mangelt allerdings an detaillierten systematischen Literatur-Reviews, welche die Thematik seit 2015 zusammenfassen [Pr13, YMK15], um Wissenschaftlern und Praktikern einen Überblick über den Forschungsstand zu ermöglichen. Die Relevanz dieses Problems ergibt sich aus dem Umstand, dass aufgrund der sich ständig weiterentwickelnden Natur des Fachgebiets eine aktuelle Literaturübersicht wichtig und wertbringend ist. Ausgehend von der dargelegten Problemstellung einer fehlenden systematischen Literaturübersicht, ergeben sich die folgenden zentralen Forschungsfragen dieses Beitrags:

- Welche zentralen Themen des ITSM wurden seit 2015 in der Literatur behandelt?
- Welche Trends lassen sich in der Literatur identifizieren?

Zielsetzung des Forschungsvorhabens ist es, relevante Themen und Schwerpunkte der Forschung im Bereich ITSM seit dem Jahr 2015 zu dokumentieren und zu analysieren, um Wissenschaftlern einen kompakten Überblick über relevante Themen, Entwicklungen und potenzielle Forschungslücken zu ermöglichen.

1.2 Forschungsstand

Bisherige State-of-the-Art Beiträge befassten sich im Schwerpunkt mit der Implementierung [Ma14], kritischen Erfolgsfaktoren [TCT09], messbaren Vorteilen [GCT10], Prozessverbesserungen [Me12] sowie mit Technologien [CT10, He12] des ITSM. State-of-the-Art Beiträge, welche sich speziell mit relevanten Themen des ITSM befassen, lassen sich in [Pr13] und [Fo17] identifizieren. [Pr13] untersuchen hierbei relevante Themen der wissenschaftlichen Literatur, ausgehend von der Vergangenheit bis 2013. Im Kontrast dazu untersucht [Fo17] vor allem Themen, welche für die Anwender in der Praxis von Relevanz sind.

1.3 Aufbau der Arbeit

Nach Darlegung der zentralen Problem- und Fragestellung sowie des Forschungsstands beinhaltet der folgende Abschnitt 2 für ein einheitliches Verständnis die Erörterung und Definition von Kernbegriffen. Abschnitt 3 befasst und definiert die verwendete Methodik des Literatur-Reviews. Abschnitt 4 umfasst die Darstellung der Ergebnisse und die

Beantwortung der Forschungsfrage. Abschließend erfolgt im Abschnitt 5 eine Zusammenfassung, kritische Würdigung sowie Darlegung weiteren Forschungsbedarfs.

2 Grundlagen

2.1 Geschäftsprozessmanagement

Nach Aguilar-Savén [Ag04] kann ein Geschäftsprozess als eine Kombination von unterschiedlichen Aktivitäten eines Unternehmens verstanden werden, wobei logische und zeitliche Abhängigkeiten zwischen den Aktivitäten dargestellt werden. Ein Prozess verfolgt hierbei ein festgelegtes Ziel, welches durch die Aktivitäten erreicht wird. Als eine Sammlung von Werkzeugen, Techniken und Methoden ist das Geschäftsprozessmanagement ein zentraler Forschungsbereich. Durch die immer stärker wachsenden Anforderungen der Globalisierung, Standardisierung, Agilität und Digitalisierung des Marktes ist die Relevanz der Prozessmodellierung und -verbesserung in den letzten Jahren stark angestiegen. Verbesserungen der Prozesse sind dabei meist mit einer Senkung der Kosten, der Zeit oder möglicher Fehler und der damit einhergehenden Steigerung der Qualität und Kundenzufriedenheit verbunden [Du13].

2.2 IT-Management

Das IT-Management als eine zentrale Funktion innerhalb eines Unternehmens in Zeiten der Digitalen Transformation wird in der Literatur häufig auf unterschiedliche Weisen definiert, wodurch keine allgemeingültige Definition des Begriffs IT-Management existiert [MB15]. In dieser Arbeit wird entsprechend der Arbeit von Krcmar [Kr15] das IT-Management als eine Funktion definiert, welche das Ziel verfolgt, die IT anhand von Analysen und Gegenmaßnahmen zu steuern, die Geschäftsprozesse zu unterstützen sowie als Enabler innerhalb des Unternehmens zu fungieren. Dabei steht die Wirtschaftlichkeit der IT, deren Strategien sowie die Service- und Kundenorientierung im Fokus des IT-Managements.

2.3 IT-Prozesse

Wie auch das bereits beschriebene Geschäftsprozessmanagement sind die IT-Prozesse für die Leistungserbringung eines Unternehmens von hoher Relevanz. Tiemeyer [Ti20] hat hierfür in seiner Arbeit eine mögliche IT-Prozesslandschaft eines Unternehmens dargestellt, welche die grundlegenden IT-Prozesse beinhaltet. Beispiele hierfür sind die IT-Beschaffung, die Applikationsentwicklung oder auch das IT-Service-Management. Aufgrund der Vielfalt möglicher IT-Prozesse, wird sich diese Arbeit nachfolgend mit dem IT-Service-Management befassen.

2.4 IT-Service-Management

Nach Krcmar [Kr15] kann die Leistungserbringung der IT, welche nach Tiemeyer [Ti20] als kritische Ressource zu verstehen ist, in Bezug auf die Unterstützung der Geschäftsprozesse inklusive der dort enthaltenen Komponenten als IT-Services verstanden werden. Ziel hierbei ist es, zu jeder Zeit die Serviceempfänger (meist in Form von Geschäftsprozessen) bestmöglich zu unterstützen und so den Unternehmensbetrieb aufrecht zu erhalten. Der Input für einen solchen IT-Service setzt sich in der Regel aus einer Kombination unterschiedlicher Faktoren zusammen. Dazu zählen die IT-Mitarbeiter mit deren Know-How, die bereits betrachteten IT-Prozesse, aber auch die dazugehörigen Informationstechnologien [Kr15].

Die Leistung der IT, diese Ressourcen entsprechend einzusetzen, um Services zu entwickeln, zu betreiben, aber auch zu warten, wird als IT-Service-Management (ITSM) bezeichnet. Das ITSM stellt dabei einen prozessorientierten und serviceorientierten Ansatz dar [Ga09, Kr15]. Um den Kunden zu jeder Zeit den Service zuverlässig erbringen zu können und so die Servicevereinbarungen einhalten zu können, ist es notwendig, die Prozesse zur Serviceerbringung in möglichst hoher Qualität zu erbringen, diese sicherzustellen und damit einhergehend die betroffenen Prozesse zu standardisieren [Kr15]. Hierfür werden Werkzeuge und Frameworks eingesetzt, welche sowohl aus dem Geschäftsprozessmanagement bekannt sind [Ga09] als auch ITSM spezifisch entwickelt wurden.

3 Methodik der Forschung

Um die Herausforderungen eines Literatur-Reviews kontrollierbarer zu machen, verwendet dieses Forschungsvorhaben die Aufbau- und Handlungsempfehlungen nach Fettke [Fe06] in Verbindung mit dem konzeptzentrierten Organisations-Framework von Webster und Watson [WW02]. Diese Erstellung der Konzeptmatrix wird zudem durch das Vier-Augenprinzip angelehnt an das Vorgehen von Peffers et al. [Pe12] ergänzt, um Inkonsistenzen bei der Interpretation der Ergebnisse vorzubeugen. Die Konzeptmatrix wurde dabei in zwei Iterationen erstellt. In der ersten Iteration wurden von beiden Autoren unabhängig potenzielle Dimensionen aus den Arbeiten extrahiert, generalisiert und anschließend zusammengetragen. In der zweiten Iteration fand im Anschluss eine ebenfalls unabhängige Einordnung der Beiträge in die zuvor definierten Dimensionen statt. Eine Abstimmung der Ergebnisse, sowie gegebenenfalls gemeinsame Sichtung der Arbeiten bei Abweichungen bildet das Ende der zweiten Iteration und hat die finale Konzeptmatrix als Ergebnis.

Die Literatursuche wurde am 01.11.2020 unter Verwendung des zentralen Suchterms "IT-Service-Management" durchgeführt, um das gesamte Spektrum dieser Thematik sowohl im deutsch- als auch englischsprachigen Raum erfassen zu können, wodurch keine

Einschränkung potenziell relevanter Themen vorgenommen wurde. Des Weiteren wurden folgende Ausschlusskriterien für die Literatursuche definiert:

- Angabe der Suchfelder: Abhängig von den verwendeten Literaturdatenbanken werden die Suchterme nur auf Titel, Abstract und Schlüsselwörter angewendet, um potenziell relevante Quellen zu identifizieren.
- Zeitraum der Suche: Der Zeitraum der Suche wird auf die Jahre 2015 bis 2020 eingeschränkt. Begründet wird diese Entscheidung damit, den State-of-the-Art erfassen zu wollen.
- Art der Publikationen: Zur Wahrung der Qualität und Wissenschaftlichkeit wird die Suche eingeschränkt auf Journale und Konferenzbände.

4 Ergebnisse

Der folgende Abschnitt umfasst die Ergebnisse der Literaturanalyse. Der Prozess der Literatursuche und -auswahl, kompakt dargestellt in Tabelle 1, resultierte in 96 relevanten Beiträgen, welche das IT-Service-Management als Forschungsschwerpunkt untersuchten.

Name der Datenbank	Suchfelder	Treffer	Relevant
AISel ⁴	Titel, Abstract	39	14
ACM ⁵	Titel, Abstract, Keywords	15	10
Google Scholar ⁶	Titel	241	32
IEEE ⁷	Titel, Keywords	304	29
Science Direct ⁸	Titel, Abstract, Keywords	207	11
Summe		806	96

Tab. 1: Darstellung der Suchergebnisse

Abbildung 1 stellt eine aggregierte Form der dort angefertigten Konzeptmatrix dar, welche alle identifizierten Dimensionen miteinander in Beziehung setzt. Dadurch wird das paarweise gemeinsame Auftreten dargestellt. Aus dieser Abbildung wird ebenso ersichtlich, dass besonders die Dimensionen „ITIL“, „COBIT“ und „ISO/IEC 20000“ in

⁴ <http://aisel.aisnet.org/>

⁵ <https://dl.acm.org/>

⁶ <https://scholar.google.com/>

⁷ <http://ieeexplore.ieee.org/>

⁸ <http://www.sciencedirect.com/>

den Arbeiten, beispielsweise von Hermita et al. [HSG20], thematisiert werden. Hierbei handelt es sich um zwei ITSM-Frameworks, welche in der Literatur und der Praxis etabliert sind, sowie um den ISO-Standard, welcher als Grundlage für diese Frameworks dient. Auch die Dimension „CMMI“ wird häufig betrachtet und stellt ein Reifegradmodell dar, welches zur Bewertung der Prozessqualität dient. Arbeiten, welche sich mit diesem Thema befassen haben, sind die von Wulf et al. [WWB15], Zajac [Za16] und Reinehr und Malucelli [RM19].

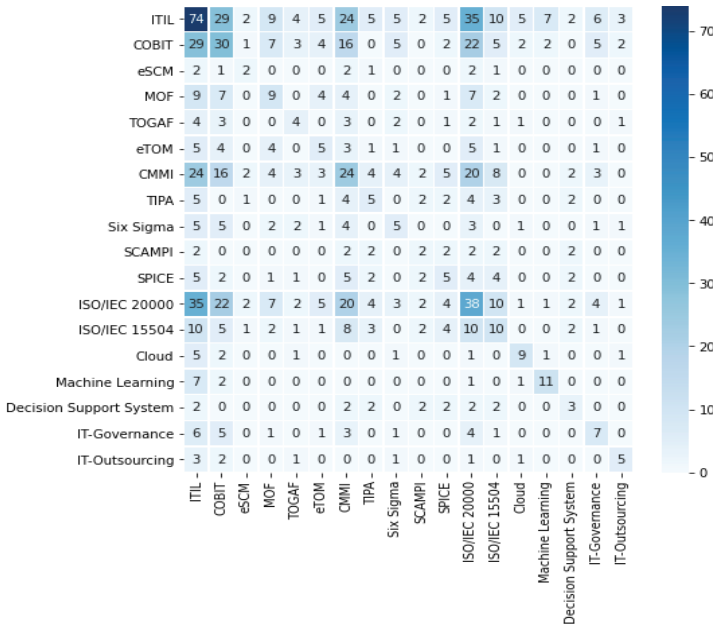


Abb. 1: Aggregierte Konzeptmatrix

Die Dimensionen wurden, entgegen der dargestellten Aggregation, zudem in Kategorien eingeteilt, so dass ähnliche Dimensionen einer Gruppe zugeordnet werden können und so eine direkte Beziehung zueinander aufweisen. Die relevantesten und am häufigsten eingesetzten Dimensionen sind hierbei die „ITSM-Frameworks“, „Process Improvement Model“ und die „Standards“. Zur Betrachtung der aktuellen Trends und der Einsatzszenarien wurden zudem Technologien sowie IT-Governance und IT-Outsourcing als strategische Dimensionen ergänzt. Da nicht alle Arbeiten in diese Selektion an Dimensionen eingeteilt werden können, wurde die Dimension „Fokus“ als Freitextfeld dazu verwendet, in wenigen Worten die Arbeit zu beschreiben.

In Bezug auf die Dimensionen und die Dimensionskategorien zeigt Abbildung 2 eine quantitative Übersicht über die fünf häufigsten Dimensionen und kennzeichnet diese entsprechend ihrer Kategorie farblich.

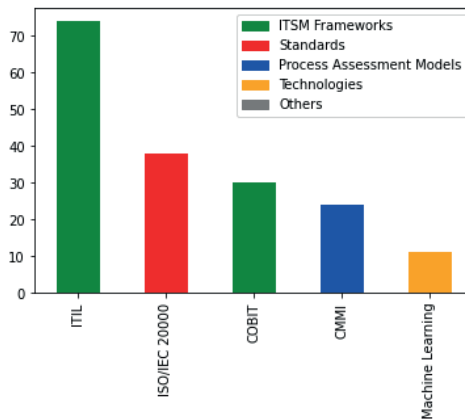


Abb. 2: Fünf häufigste Dimensionen nach Kategorie

Auch in dieser Abbildung ist gut ersichtlich, dass die bereits angesprochenen Dimensionen und damit verbundenen Kategorien besonders häufig in den Arbeiten auftreten. Von den 96 relevanten Arbeiten thematisieren demnach 74 Arbeiten das Thema ITIL aus den ITSM-Frameworks, was 77% aller Arbeiten bedeutet. Ebenso kann aus der Grafik entnommen werden, dass die Dimension auf dem fünften Platz, das Machine Learning, die am häufigsten identifizierte Technologie innerhalb der Konzeptmatrix ist. Eine Hypothese für diese Beobachtung könnte die zunehmende Relevanz von Machine Learning und der damit verbundenen Optimierungsmöglichkeiten in Unternehmen [Ga20] sein. Durch Prozessautomatisierungen oder -unterstützungen wie zum Beispiel durch Robotic Process Automation (RPA) oder Machine Learning Modelle können Arbeitsabläufe mit hoher Qualität binnen kürzester Zeit durchgeführt und unterstützt werden. Ein Beispiel für eine derartigen Automatisierung stellt die Arbeit von Zhou et al. [Zh17] dar. Aber auch die Relevanz des Machine Learnings für andere Fachbereiche eines Unternehmens kann möglicherweise eine Ursache dafür sein, dass sich das ITSM vermehrt mit diesen Themen befassen muss. Dies ist auch in der nachfolgenden Gegenüberstellung der Abbildungen 3 und 4 erkennbar. Dabei wurden die Jahre 2015 und 2019 gegenübergestellt sowie das Auftreten der zehn häufigsten Dimensionen in Relation zur Anzahl der Arbeiten betrachtet. Das Jahr 2020 wurde hierbei aufgrund der geringeren Anzahl an Arbeiten und der damit verbundenen geringeren Vergleichbarkeit bewusst nicht als Referenz genutzt. Konträr zu den bisher getroffenen Aussagen zu den häufigsten Dimensionen ist außerdem erkennbar, dass die Technologie der Cloud ebenfalls im Jahr 2019 in mehr Arbeiten thematisiert wird als im Jahr 2015. Im Jahr 2019 übersteigt diese Dimension zudem das Machine Learning. Arbeiten aus dem Bereich der Cloud sind unter anderem die von Schneider et al. [SPA19] oder auch Landis [La19].

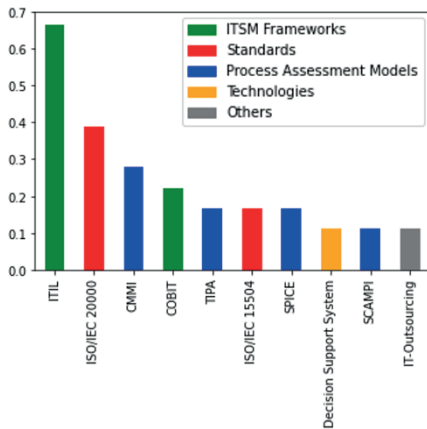


Abb. 3: Relative Verteilung 2015

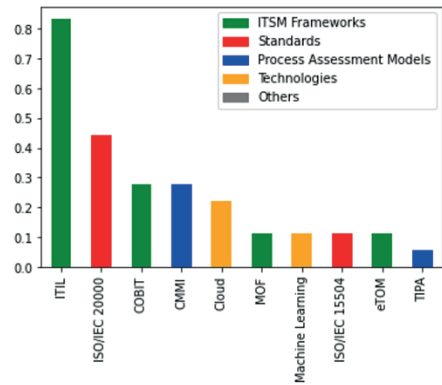


Abb. 4: Relative Verteilung 2019

Da bei der Erstellung einer Konzeptmatrix die subjektive Bewertung des Betrachters in der Identifikation von Dimensionen einen Einfluss auf die Ergebnisse hat, werden nachfolgend die vom jeweiligen Autor selbst festgelegten Keywords quantitativ betrachtet. Das Vorgehen folgt demnach der Arbeit von Marrone und Hammerle [MH17], so dass hier die grundlegenden Thematiken objektiv ohne Einfluss des Lesers ermittelt werden können. Abbildung 5 stellt eine Übersicht über die 25 häufigsten Keywords dar. Die Summe aller Keywords, welche im Rahmen dieser Arbeit betrachtet wurden, beläuft sich auf 261 Stück (2,7 individuelle Keywords pro Arbeit). Da sich alle Arbeiten entsprechend des möglichst allgemeingültigen Suchterms mit dem IT-Service-Management befassen, ist es nicht hervorzuheben, dass Keywords diesbezüglich in dieser Abbildung berücksichtigt werden. Auch die aufgestellte Konzeptmatrix, welche in aggregierter Form betrachtet wurde, kann in der Aussage bestätigt werden, dass die Dimension „ITIL“ in den meisten Arbeiten berücksichtigt wird. Auch andere Dimensionen wie das „IT-Governance“ oder die „Cloud“ finden sich hier wieder. Ein möglicher Grund, weshalb in den Keywords Dimensionen wie „COBIT“ oder „ISO/IEC 20000“ nicht an erwarteter Stelle wiedergefunden werden, könnte der Fokus der jeweiligen Arbeit sein. Häufig werden diese Frameworks und Standards als Vergleich genutzt und stehen nicht im Hauptfokus der jeweiligen Arbeit.

Hingegen gibt es auch in Abbildung 5 Keywords, welche in der Konzeptmatrix nicht berücksichtigt wurden. Beispiele wie „continuous improvement“ oder „quality management“ sind hier jedoch den Improvement Dimensionen wie beispielsweise „CMMI“ in Teilen zuzuordnen. Auch die Keywords wie „sla management“, „service integration“ oder „incident management“ sind vielmehr der Dimension von ITIL zuzuordnen, weshalb diese Keywords in der Konzeptmatrix nicht wiedergefunden werden können.

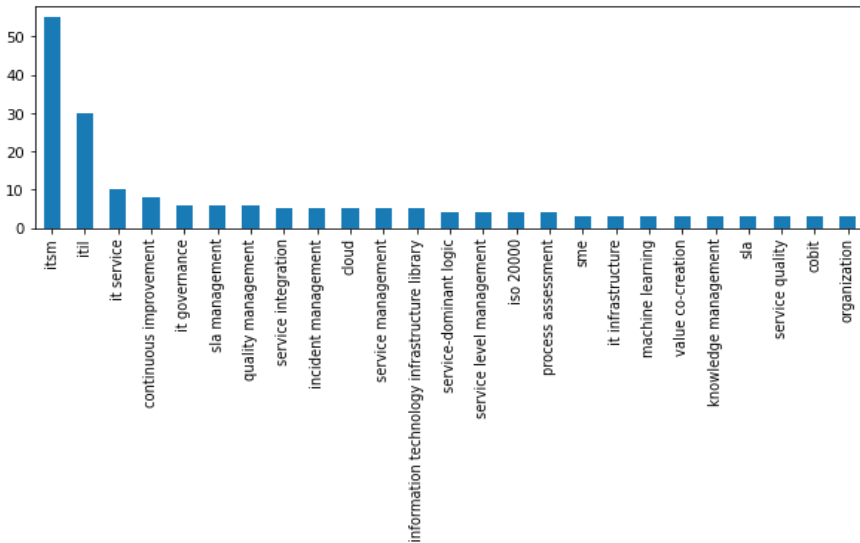


Abb. 5: Quantitative Betrachtung der Keywords

Abschließend lässt sich die Betrachtung der Keywords verwenden, um die erstellte Konzeptmatrix zu bestätigen. Auch hier ist zu erkennen, dass die Cloud als eine mögliche Technologie bereits in den am häufigsten eingesetzten Keywords vertreten ist und somit bereits zwischen 2015 und 2020 ein relevantes Thema im Bereich des IT-Service-Managements darstellt. Ebenfalls kann deutlich bestätigt werden, dass ITIL als führendes Framework auch in den betrachteten Arbeiten die zentrale Rolle spielt und so als etabliertes Framework auch weiterhin eingesetzt wird.

Vergleicht man die gewonnen Erkenntnisse mit Unternehmensumfragen von Forbes im Jahr 2017 [Fo17], wird ersichtlich, dass es neben Gemeinsamkeiten auch grundlegende Unterscheidungen gibt. Zunächst lässt sich auch hierbei erkennen, dass ITIL mit Abstand am häufigsten in Unternehmen eine Rolle spielt, was auch den Ergebnissen dieses Beitrags entspricht. Ebenso ist COBIT und ISO/IEC 20000 unter den häufigsten fünf Antworten aufgelistet. Die Unterscheidung findet sich jedoch darin wieder, dass der zweite (eTOM) und vierte Platz (MOF) dieser Umfrage in der Literaturanalyse nicht oder nicht in diesem quantitativen Rahmen repräsentiert waren. Dies könnte darauf zurückgeführt werden, dass es sich bei der Umfrage nicht um einen Stand der Wissenschaft handelt, sondern um einen Stand der Technik innerhalb der angewandten Praxis.

5 Diskussion und Ausblick

Aufgrund der zunehmenden Relevanz des IT-Service-Managements als zentrale Funktion innerhalb eines Unternehmens und Teilbereichs des Prozessmanagements innerhalb der

IT sowie des zugleich fehlenden Vorhandenseins von Literaturreviews innerhalb dieses Bereichs, hat sich diese Arbeit mit der Frage befasst, welche Themen das ITSM in den letzten fünf Jahren seit 2015 dominiert haben und welche Veränderungen zu erkennen sind. Da das ITSM für die Qualität der angebotenen IT Services und den damit verbundenen Kosten von großer Wichtigkeit ist, befassen sich in den letzten Jahren immer mehr akademische Arbeiten mit dieser Thematik, was zu einer natürlichen Entwicklung dieses Forschungsgebiets führt und in dieser Arbeit betrachtet wird.

Mit Hilfe des Suchterms „IT-Service-Management“ wurden infolge der Relevanz dieser Thematik unterschiedliche Literaturdatenbanken durchsucht und ein möglichst objektives Bild des aktuellen Stands innerhalb des ITSM erzielt. Dabei wurden 96 relevante Arbeiten betrachtet und innerhalb einer Konzeptmatrix eingeordnet. Durch ein Vier-Augen-Prinzip wurde hierbei die Qualität sichergestellt. Die Auswertung dieser Arbeiten in den Jahren 2015 bis 2020 (jeweils inklusive) hat ergeben, dass die bereits etablierten Themengebiete ITIL, ISO 20000 und COBIT von hoher Relevanz sind. Aber auch die Qualitätssicherung und -verbesserung der Services in Form von Reifegradmodellen wie CMMI sind hierbei vertreten. Außerdem werden vereinzelte Technologien wie das Machine Learning oder die Cloud immer stärker im ITSM thematisiert. Diese Veränderung wurde ebenfalls in der Gegenüberstellung der Jahre 2015 und 2019 ersichtlich. Bei der Betrachtung der beiden Jahre konnte ebenso dargestellt werden, dass sich die führenden Thematiken innerhalb des Forschungsgebiets nur geringfügig verändern und so nahezu unverändert bleiben. Stattdessen verlieren Frameworks wie TIPA und SPICE sowie Technologien wie Decision Support Systems an Bedeutung und werden durch ITSM-Frameworks wie eTOM und MOF oder auch durch die bereits angesprochenen innovativeren Technologien wie die Cloud oder Machine Learning ersetzt.

Eine weitere Betrachtung, welche die Subjektivität der Einordnung der Arbeiten in die Konzeptmatrix mindert, war die Auswertung der verwendeten Keywords innerhalb der Literatur. Dabei wurden 261 individuelle Keywords identifiziert. Die hierbei führenden Keywords unterstrichen die bereits identifizierten Erkenntnisse des aktuellen Stands der letzten fünf Jahre, indem die Thematiken ITIL sowie die Servicequalität und -verbesserung besonders häufig vertreten waren. Auch die Cloud sowie Managementbereiche wie die IT Governance wurden ebenso in den Keywords identifiziert.

Trotz der Verwendung der Keywords muss jedoch angemerkt werden, dass der subjektive Einfluss der Autoren auf die Literatursuche und -auswahl nicht vollständig eliminiert werden konnte. Ebenso ist der Einsatz eines Vier-Augen-Prinzips kein Garant dafür, dass die Konzeptmatrix nahezu objektiv aufgestellt wurde und keine Voreingenommenheit oder Vorkenntnisse Einfluss nehmen konnten. Durch die transparente Darstellung der Literatúrauswahl sowie der simultanen Verwendung und Analyse der Konzeptmatrix sowie der Keywords wurde eine möglichst objektive Ansicht angestrebt. Ebenso ist die Qualität der eingesetzten Literaturdatenbanken sowie des Suchterms zu hinterfragen, da diese zu einer fundamentalen Eingrenzung der Literatur führen können. So ist es denkbar,

dass die Ergebnisse dieser Arbeit in weiteren Betrachtungen validiert werden, welche den Fokus auf weitere Datenbanken oder Suchterme legen, welche beispielsweise Synonyme des ITSM verwenden. In dieser Arbeit wurde bewusst dagegen entschieden, weitere Suchterme einzusetzen, um ein allgemeingültiges und uneingeschränktes Bild des aktuellen Stands zu erzielen.

Ein weiterer Forschungsbedarf, welchen es zu fokussieren gilt, ist die detailliertere Betrachtung der Literatur in Bezug auf das ITSM-Framework ITIL. Da dieses Framework auch in diesem Literaturreview als besonders relevant identifiziert wurde, wäre eine Einordnung der Literatur in die darin enthaltenen Phasen vorstellbar. Aufgrund der bereits angesprochenen Allgemeingültigkeit dieser Arbeit wurde auch hierauf bewusst verzichtet, um keine Einschränkung vorzunehmen.

6 Literaturverzeichnis

- [Ag04] Aguilar-Savén, R.: Business process modelling: Review and framework. In: International Journal of Production Economics Bd. 90, Elsevier, Nr. 2, S. 129–149, 2004.
- [CT10] Cater-Steel, A.; Tan, W.: The role of IT service management in green IT. In: Australasian Journal of Information Systems, Band 17.1, S. 107–125, 2010.
- [Du13] Dumas, M.; La Rosa, M.; Mendling, J.; Reijers, H. A.: Fundamentals of business process management (Vol. 1, p. 2). Heidelberg: Springer, 2013.
- [Fe06] Fettke, P.: State-of-the-Art des State-of-the-Art. In: Wirtschaftsinformatik Bd. 48 2006.
- [Fo17] Forbes Insights: Delivering Value to Today's Digital Enterprise. In: Forbes Insights. Forbes, Inc., Jersey City, 2017.
- [GCT10] Gacenga, F.; Cater-Steel, A.; Toleman, M.: An international analysis of IT service management benefits and performance measurement. In: Journal of Global Information Technology Management Bd. 13, Routledge, Nr. 4, S. 28–63, 2010.
- [Ga20] Gartner, Inc., <https://www.gartner.com/smarterwithgartner/5-trends-drive-the-gartner-hype-cycle-for-emerging-technologies-2020/>, Stand: 27.04.2021
- [Ga09] Galup, S.D.; Dattero, R.; Quan, J.; Conger, S.: An overview of IT service management. In: Communications of the ACM Bd. 52, Nr. 5, S. 124–127, 2009.

- [He12] Heininger, R.: IT Service Management in a Cloud Environment: A Literature Review. In: SSRN Electronic Journal, Elsevier BV, 2012.
- [HSG20] Hermita, E.S.; Suchayo, Y.G.; Gandhi, A.: Designing IT Service Management at Indonesia Internet Domain Names Registry Association's Helpdesk Function. In: ACM International Conference Proceeding Series. New York, NY, USA: Association for Computing Machinery, S. 10–15, 2020.
- [Je16] Jeong, B.K.; Yoon, T.; Keun, B.; Tom, J.; Yoon, E.: Improving IT Process Management Through Value Stream Mapping Approach: A Case Study. In: JISTEM Bd. 13, Universidade de São Paulo - USP, Nr. 3, S. 1807–1775, 2016.
- [Kr15] Krcmar, H.: Informationsmanagement. In: Informationsmanagement. Berlin, Heidelberg: Springer Berlin Heidelberg, S. 85–111, 2015.
- [La19] Landis, M.: IT Service Management aus der Cloud – ein Technologievergleich IT Service Management from the Cloud—a Technology Comparison. In: HMD Praxis der Wirtschaftsinformatik Bd. 56, Springer Fachmedien Wiesbaden GmbH, Nr. 2, S. 385–390, 2019.
- [MB15] Mangiapane, M.; Büchler, R.: Modernes IT-Management: Methodische Kombination von IT-Strategie und IT-Reifegradmodell, 2015.
- [MH17] Marrone, M.; Hammerle, M.: Relevant Research Areas in IT Service Management: An Examination of Academic and Practitioner Literatures. In: Communications of the Association for Information Systems Bd. 41, Association for Information Systems, Nr.1, S. 517-543, 2017.
- [Ma14] Marrone, M.; Gacenga, F.; Cater-Steel, A.; Kolbe, L.: IT service management: A cross-national study of ITIL adoption. In: Communications of the Association for Information Systems Bd. 34, Association for Information Systems, Nr. 1, S. 865–892, 2014.
- [Me12] Mesquida, A.; Mas, A.; Amengual, E.; Calvo-Manzano, J.: IT service management process improvement based on ISO/IEC 15504: A systematic review. In: Information and Software Technology Bd. 54, Elsevier, Nr. 3, 2012.
- [Pr13] Proehl, T.; Erek, K.; Limbach, F.; Zarnekow, R.: Topics and applied theories in IT Service Management. In: Proceedings of the Annual Hawaii International Conference on System Sciences, S. 1367–1375, 2013.

- [Pe12] Peffers, K.; Rothenberger, M.; Tuunanen, T.; Vaezi, R.: Design science research evaluation. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Bd. 7286 LNCS: Springer, Berlin, Heidelberg, S. 398–410, 2012.
- [RM19] Reinehr, S.; Malucelli, A.: Maturity in IT Service Management: A Longitudinal Study. In: Communications in Computer and Information Science. Bd. 1060: Springer Verlag, S. 99–110, 2019.
- [SPA19] Schneider, D.; Plate, F.; Auth, G.: Evaluation von ITSM-Tools für Integration und Management von Cloud-Diensten am Beispiel von ServiceNow Digital Transformation of Higher Education View project, 2019.
- [TCT09] Tan, W.; Cater-Steel, A.; Toleman, M.: Implementing it service management: A case study focussing on critical success factors. In: Journal of Computer Information Systems Bd. 50, Nr. 2, S. 1–12, 2009.
- [Ti20] Tiemeyer, E.: Handbuch IT-Management IT Management Qualifizierung View project Current issues of project magement View project, 2020.
- [To10] Tony Lee, T.: Optimizing IT process management. In: ACM SIGSOFT Software Engineering Notes Bd. 35, Association for Computing Machinery (ACM), Nr. 4, S. 1–10, 2010.
- [WW02] Webster, J.; Watson, R.: Analyzing the Past to Prepare for the Future: Writing a Literature Review. In: MIS Quarterly Bd. 26, Nr. 2, 2002.
- [WWB15] Wulf, J.; Winkler, T.; Brenner, W: Measuring IT Service Management Capability: Scale Development and Empirical Validation, Universität Osnabrück, 2015.
- [YMK15] Yazici, A.; Mishra, A.; Kontogiorgis, P.: IT Service Management (ITSM) Education and Research: Global View. In: International Journal of Engineering Education Bd. 31, Nr. 4, S. 1071–1080, 2015.
- [Za16] Zajac, A.: Maturity Models in IT Service Management: An Overview. In: Information Technology for Practice, S. 365, 2016.
- [Zh17] Zhou, W.; Xue, W.; Baral, R.; Wang, Q.; Zeng, C.; Li, T.; Xu, J.; Liu, Z.; U. A.: STAR: A System for Ticket Analysis and Resolution. In: KDD Bd. 17, Association for Computing Machinery, S. 2181–2190, 2017.

Identifikation erfolgsrelevanter Faktoren von Cloud-basierten Geschäftsprozessmanagement-Lösungen

Franz Lehner, Sebastian Floerecke¹

Abstract: Die Forschung befindet sich, was die Auswahl, die Einführung und die Verwendung von Cloud-basierten Geschäftsprozessmanagement (GPM)-Lösungen anbelangt, noch relativ am Anfang und bietet deshalb der betrieblichen Praxis diesbezüglich nur wenig Hilfestellung. Vor diesem Hintergrund besteht das Ziel dieser Untersuchung darin, mittels einer qualitativ-empirischen Querschnittanalyse derartige Lösungen zu identifizieren und hinsichtlich ihrer erfolgsrelevanten Faktoren zu analysieren. Die erfolgsrelevanten Faktoren werden zum einen aus der Wahrnehmung der Anbieter und zum anderen der Anwender betrachtet und im Nachgang miteinander verglichen. Dazu werden alle Faktoren gezählt, die bei der Nutzung einer spezifischen Cloud-basierten GPM-Lösung als wichtig, nützlich und wirtschaftlich vorteilhaft eingestuft werden. Als Ergebnis konnten aus der Kunden- und der Anbieterperspektive insgesamt 16 erfolgsrelevante Faktoren ermittelt werden. Die beiden Perspektiven weisen eine hohe Übereinstimmung auf. Durch die Ergebnisse erhalten Kunden Unterstützung bei der Anbieter- beziehungsweise Toolauswahl, indem sie die abgeleiteten erfolgsrelevanten Faktoren in ihren Auswahlprozess einbeziehen können. Anbieter können im Gegenzug ihre Cloud-basierten GPM-Lösungen entsprechend den Kundenwünschen gestalten und bewerben.

Keywords: Geschäftsprozessmanagement (GPM), Geschäftsprozessmanagementsystem (GPMS), Cloud-Computing, Erfolgsrelevante Faktoren, Internetrecherche, Querschnittanalyse

1 Ausgangssituation und Zielsetzung

Geschäftsprozessmanagement (GPM) bezeichnet einen Bereich, der Methoden, Techniken, Tools und Software verwendet, um betriebliche Prozesse zu entwerfen, durchzuführen, zu steuern und zu analysieren, an denen Menschen, Organisationen, Anwendungen, Dokumente und andere Informationsquellen beteiligt sind [WHW03]. Zur Unterstützung und teilweisen automatisierten Umsetzung von anfallenden GPM-Aufgaben existiert für Unternehmen ein großes Software-Tool-Angebot. Nach einer Studie des Marktforschungsinstituts Fortune Business Insights [Fo20] ist der weltweite Umsatz des GPM-Software-Markts in den vergangenen Jahren stetig angestiegen. Im Jahr 2018 betrug er rund 8,63 Milliarden US-Dollar. Schätzungen zufolge wird sich dieser bis 2026 mehr als verdoppeln [Fo20].

Wie in vielen Bereichen ist auch bei GPM-Software-Lösungen der verstärkte Einsatz von Cloud-Computing als Bereitstellungsmodell wahrnehmbar [Ga21]. Eine Cloud-Vari-

¹ Universität Passau, Lehrstuhl für Wirtschaftsinformatik mit Schwerpunkt Informations- und IT-Service-Management, Innstraße 43, 94032 Passau, {franz.lehner, sebastian.floerecke}@uni-passau.de

ante bietet Kunden im Vergleich zu traditionellen On-Premise-Systemen vielfältige Vorteile. So entfallen Anschaffungs- und Wartungskosten für die zugrundeliegenden Infrastrukturressourcen, zudem können die Lösungen mit vergleichsweise geringem Aufwand eingeführt werden. Zu den weiteren Vorzügen zählen unter anderem die nutzungsabhängige Abrechnung und die gebotene Elastizität und Skalierbarkeit [Ri15; SJ15].

Neben Großunternehmen, die vielfach bereits On-Premise-Varianten im Einsatz hatten, profitieren im besonderen Maße kleine und mittlere Unternehmen (KMUs) vom Cloud-Computing. Aufgrund der mit On-Premise-GPM-Lösungen verbundenen relativ hohen Anschaffungskosten war für dieses Kundensegment eine Nutzung ökonomisch oftmals nicht sinnvoll [GD20; Gz20; Ri15]. Angesichts der Potenziale von Cloud-Lösungen prognostiziert Fortune Business Insights [Fo20] für die kommenden Jahre insbesondere bei KMUs einen vermehrten Einsatz für das GPM. Cloud-Varianten bieten hier unmittelbar verfügbare, leistungsstarke Lösungen verbunden mit flexiblen Preismodellen. Daraus ergeben sich besonders für KMUs ein deutliches Kostensenkungspotenzial sowie weitere Chancen wie Produktivitätssteigerungen oder die Reduktion der Time-to-Market [Gz20]. Dem stehen allerdings Herausforderungen beim Business-IT-Alignment gegenüber. Neben den üblichen Nachteilen von KMUs, wie einem geringeren Grad an Standardisierung und Formalisierung sowie knappen Personalressourcen, fehlt durch den starken Bezug zum operativen Geschäft häufig die Zeit für langfristige Erwägungen im Umgang mit neuen Technologien. KMUs können zudem wegen ihrer oft geringeren IT-Kompetenzen die Vorteile nur begrenzt nutzen. Die Abstimmung zwischen den technischen Anforderungen eines Cloud-gestützten GPM mit den betrieblichen Anforderungen ist aber von zentraler Bedeutung [GD20].

Die gegenwärtig auf dem Markt erhältlichen Cloud-basierten GPM-Lösungen umfassen für gewöhnlich die Prozessmodellierung, das Prozessdesign sowie weitere für die Prozessausführung und -steuerung erforderlichen Anwendungen [Ga21]. Die Forschung befindet sich jedoch, was deren Auswahl, Einführung und Verwendung betrifft, noch relativ am Anfang und bietet deshalb der betrieblichen Praxis nur wenig Hilfestellung [Ri15; SJ15]. Die zunehmende Bedeutung ist aber inzwischen in der wachsenden Zahl an Veröffentlichungen sichtbar [AL20; CS17; GFA18; Gz20; LNB19]. Das Ziel dieses Beitrags ist, Cloud-basierte GPM-Lösungen zu identifizieren und hinsichtlich ihrer erfolgsrelevanten Faktoren zu analysieren. Die erfolgsrelevanten Faktoren sollen dabei zum einen aus der Wahrnehmung der Anbieter, zum anderen der Anwender betrachtet und anschließend gegenübergestellt werden. Die Untersuchungsergebnisse sollen Anbietern dabei helfen, ihre Cloud-basierten GPM-Lösungen entsprechend den Kundenwünschen zu gestalten und zu bewerben. Anwender erhalten im Gegenzug Unterstützung bei der Anbieter- beziehungsweise Toolauswahl, indem sie die abgeleiteten erfolgsrelevanten Faktoren in ihrem Auswahlprozess berücksichtigen können. Zur Erreichung dieses Forschungsziels wird eine systematische Internetrecherche und Auswertung in Form einer qualitativ-empirischen Querschnittanalyse [WH07] durchgeführt.

Der Beitrag ist wie folgt aufgebaut: Kapitel 2 beschreibt überblicksmäßig die Grundlagen von GPM, die damit verbundenen Kernaufgaben und die Entwicklung zu Cloud-

gestützten Lösungen. Kapitel 3 erklärt das forschungsmethodische Vorgehen. Die Ergebnisse werden in Kapitel 4 präsentiert. Am Ende geht der Beitrag auf den Nutzen für die Wissenschaft und die Praxis sowie Limitationen ein und zeigt weiteren Forschungsbedarf auf.

2 Vom klassischen GPM zu Cloud-gestützten Lösungen

Hammer und Champy [HC93] definieren den Begriff Geschäftsprozess als eine Sammlung von Aktivitäten, die bestimmte Ressourcen als Input nutzen und dadurch einen Output mit einem Mehrwert für den Kunden generieren. Geschäftsprozesse lassen sich anhand ihrer Funktion und ihrer Kernkompetenzen in die Arten Kern-, Management- und Supportprozesse unterteilen. Die bewusste Klassifizierung und vor allem ein systematisches Management sind für Unternehmen von strategischer Bedeutung [LWS08].

GPM wurde erstmals von Hammer und Champy [HC93] in den frühen 1990er Jahren vorgestellt – damals noch unter der Prämisse der Neuorganisation und Restrukturierung. Eine häufig zitierte Definition stammt von van der Alst et al. [WHW03]. Die Autoren definieren GPM als einen Bereich, der Methoden, Techniken, Tools und Software verwendet, um betriebliche Prozesse zu entwerfen, durchzuführen, zu steuern und zu analysieren, an denen Menschen, Organisationen, Anwendungen, Dokumente und andere Informationsquellen beteiligt sind [WHW03].

Es existieren verschiedene Möglichkeiten, die von Unternehmen bei der Umsetzung von GPM zu bewältigenden Aufgaben in einer strukturierten Form darzustellen. Am häufigsten ist in der Literatur die Abbildung in Form eines Kreislaufs vorzufinden [Ha10]. Der sogenannte GPM-Kreislauf unterteilt sich auf höchster Ebene in mehrere Phasen, innerhalb derer einzelne Aufgaben für ein erfolgreiches, vollumfängliches GPM zu realisieren sind. Die Phasen sind voneinander abhängig und beeinflussen sich gegenseitig [We12]. Tabelle 1 zeigt eine Übersicht der Phasen mit den jeweiligen Kernaufgaben. Ihre Umsetzung charakterisiert die sogenannte prozessorientierte Organisation, in der letztlich das optimale Zusammenwirken aller Funktionen (unabhängig von ihrer aufbauorganisatorischen Einordnung) gewährleistet werden soll, um ein gemeinsames, über die einzelne betriebliche Funktion hinausgehendes Ziel – die Erfüllung einer Kundenanforderung – zu verfolgen. Unterstützt wird die prozessorientierte Organisation dabei durch ein GPM, das durch eine Abstimmung aller betrieblichen Ressourcen und Abläufe die Realisierung strategischer Ziele einer Organisation vereinfachen beziehungsweise erst ermöglichen soll. Das GPM verfolgt die nachhaltige Verbesserung der Wirtschaftlichkeit einer Organisation. Daraus lassen sich im Wesentlichen drei Unterziele ableiten, nämlich Ausrichtung an der Wertschöpfungskette, Kundenorientierung sowie Integration der betrieblichen Funktionsbereiche. Davon leiten sich wiederum die Planung, die Überwachung und Steuerung sowie die kontinuierliche Verbesserung der Geschäftsprozesse als Einzelaufgaben ab.

Phase	Kernaufgaben
-------	--------------

(1) Planung, Analyse und Design	<ul style="list-style-type: none"> • Identifikation und Spezifikation von Prozessen • Planung der Anforderungen und deren Umsetzung • Festlegung von Key-Performance-Indikatoren (KPIs) • Spezifikation von Richtlinien und Geschäftsregeln • Modellierung der identifizierten Prozesse • Validierung, Simulation und Verifikation der Prozessmodelle • Planung von Kapazitäten
(2) Konfiguration	<ul style="list-style-type: none"> • Auswahl, Integration und Konfiguration einer GPM-Architektur • Überführung der Prozessmodelle in ausführbare Prozessmodelle • Einbindung der ausführbaren Prozessmodelle in eine geeignete Software • Testen des Softwaresystems • Deployment des Softwaresystems
(3) Ausführung	<ul style="list-style-type: none"> • Ausführung und Steuerung von Prozessinstanzen anhand der Prozessmodelle • Monitoring und Überwachung der Prozessinstanzen • Kontrolle und Anpassung der Prozessinstanzen • Aufgabenzuweisung • Echtzeitfehlererkennung und Handlung
(4) Evaluation	<ul style="list-style-type: none"> • Analyse, Bewertung und Diagnose ausgeführter Prozessinstanzen • Auswertung der festgelegten KPIs • Konformitätsanalyse • Anpassung der Prozessmodelle • Process-Mining / Process-Activity-Monitoring • Optimierung der Prozesse

Tab. 1: Phasen und Kernaufgaben des Geschäftsprozessmanagements [Ha10; We12]

Das GPM ist heute in der Praxis fest etabliert und galt aus Sicht der Forschung längere Zeit als weitgehend abgeschlossen. Aufgrund der aktuellen technologischen Entwicklungen wie Big-Data, Data-Analytics, Process-Mining, Robotic-Process-Automation (RPA), Künstliche Intelligenz, Blockchain und Internet-der-Dinge ist allerdings erneut Bewegung entstanden, weshalb sich die Forschung wieder intensiv mit dem GPM beschäftigt [AL20; HL21]. In Verbindung mit weiteren digitalen Innovationen geht man inzwischen von der Sinnhaftigkeit der Nutzung von Cloud-Technologien aus. Traditionelle On-Premise-Anwendungen, Hosting-Lösungen und Cloud-Services schließen sich dabei nicht aus, sondern werden teilweise auch gemischt eingesetzt. Sowohl für das übergeordnete GPM als auch für die erwähnten Einzelaufgaben existiert mittlerweile ein entsprechendes Softwareangebot [AL20; GFA18; Gz20; LNB19]. Der Einsatz von Cloud-Technologien für diese Anwendungen begann vor etwa zehn Jahren und hat sich seither kontinuierlich weiterentwickelt [Gz20; SK14].

Aus einer anwendungsbezogenen Perspektive sind für das Verständnis der Entwicklung jedoch weder der GPM-Kreislauf noch eine Strukturierung nach den bekannten Cloud-Service-Modellen geeignet. Daher wird zunächst von einer Differenzierung zwischen den übergeordneten Managementaufgaben und der operativen Prozessausführung und -steuerung ausgegangen. Die Unterstützung erfolgt im Wesentlichen auf diesen beiden Ebenen. Die operative Prozessausführung stellt Business-Process-as-a-Service (BPaaS) im engeren Sinne dar. Dabei wird ein definierter Geschäftsprozess vom Service-Provider unter der Kontrolle des Diensteanbieters über die Cloud durchgeführt. Eine Konfiguration

(vergleichbar dem Customizing bei SAP-Anwendungen) ist zumeist nur eingeschränkt möglich. Auch wenn diese begriffliche Unterscheidung nicht immer streng vorgenommen wird, müsste man die Unterstützung des GPM als Business-Process-Management-as-a-Service (BPMaaS) bezeichnen. Hier geht es um die Unterstützung aller weiteren Aufgaben, die mit der Prozesssteuerung und -weiterentwicklung verbunden sind [SK14]. Letzteres erfordert zumindest grundlegende Kenntnisse der Methoden des GPM. Nachdem die softwaretechnische Unterstützung oft nicht prozessübergreifend, sondern nur auf der Ebene von Einzelsystemen erfolgt, zeigt sich dies auch in der Cloud-Unterstützung, die sich dann ebenfalls auf einzelne Teilaufgaben des GPM (beziehungsweise die dafür eingesetzte Software) beschränkt. Beispiele sind die Unterstützung der Prozessmodellierung, der Prozesssimulation oder des RPA. Wenn das GPM allerdings über diese Einzelsysteme hinaus unterstützt werden soll, kann auf entsprechende Business-Process-Management-Suiten (BPMS) beziehungsweise Geschäftsprozessmanagementsysteme (GPMS) und Process- oder Workflow-Engines zurückgegriffen werden. Viele Anbieter bieten diese Systeme inzwischen auch als Cloud-Variante an, wodurch sich der Bogen zum BPM-Kreislauf wieder spannt. Zusätzliche Entwicklungen finden sich unter Begriffen wie Dynamic-BPM (DBPM) und Elastic-BPM (EBPM).

3 Forschungsmethodisches Vorgehen

Zur Ermittlung Cloud-basierter GPM-Lösungen und deren Analyse hinsichtlich sowohl aus Anbieter- als auch Anwenderperspektive wahrgenommenen, erfolgsrelevanten Faktoren wurde eine qualitativ-empirische Querschnittanalyse [WH07] durchgeführt. Querschnittanalysen beinhalten eine einmalige Erhebung über eine größere Anzahl von Objekten, die anschließend ausgewertet wird. Das Ergebnis ist ein Querschnittsbild über die Stichprobenteilnehmer, das Rückschlüsse auf die Grundgesamtheit zulässt [WH07]. Unter erfolgsrelevanten Faktoren werden im Rahmen dieser Untersuchung alle Faktoren verstanden, die bei der Nutzung einer spezifischen Cloud-basierten GPM-Lösung als wichtig, nützlich und wirtschaftlich vorteilhaft eingestuft werden. Aufgrund fehlender objektiver Daten wird der Erfolgsbegriff damit bewusst breit gefasst. Der durchlaufene Forschungsprozess lässt sich in die beiden Phasen **(I) Datenerhebung** und **(II) Datenauswertung** unterteilen:

Die Phase **(I) Datenerhebung** bestand aus einer manuellen Google-Suche nach relevanten Einträgen innerhalb von Anbieterinformationen einerseits und Anwenderberichten andererseits. Dafür wurde eine Vielzahl an Suchwortkombinationen wie etwa „Cloud Business Process Management“ und „Cloud BPM Suite“ (sowie ihre deutschsprachigen Äquivalente) verwendet. Die erhaltenen Internetseiten wurden gemäß Anbieterinformationen und Anwenderberichten separiert. Als Erstes wurde die Anbieterperspektive betrachtet. Hierbei konnten Cloud-basierte GPM-Lösungen unterschiedlichster Hersteller identifiziert werden. Die Internetauftritte der Tool-Hersteller wurden jeweils nach relevanten Formulierungen in Bezug auf erfolgsrelevante Faktoren zu ihrer Lösung durchsucht. Konnten derartige Formulierungen ermittelt werden, wurde die betreffende GPM-

Lösung, deren Anbieter und die konkreten Formulierungen in einer Tabelle dokumentiert. Andernfalls wurde die Software-Lösung nicht weiterverfolgt.

Für die Beschaffung relevanter Informationen aus der Anwenderperspektive wurden die durch die im Rahmen der Suche erhaltenen Internetseiten hinsichtlich Foren, Communities, Technologie-Review-Plattformen und ähnliches gefiltert, da angenommen wurde, dort diesbezüglich am ehesten fündig zu werden. Entsprechend der Anbieterperspektive wurden die ermittelten und als geeignet eingestuften Bewertungen auf Formulierungen zu erfolgsrelevanten Faktoren hin untersucht. Die Funde wurden in einer gesonderten Tabelle dokumentiert. Durch die erhobenen Anwenderberichte konnten weitere Cloud-basierte GPM-Lösungen identifiziert werden, die im Nachgang wiederum aus Anbieterperspektive beleuchtet wurden.

Nach der systematischen Datenerhebung wurden die Formulierungen, die entsprechende Erfolgshinweise beinhalten, analysiert (**II Datenauswertung**). Die unterschiedlichen Perspektiven (Anbieter und Anwender) wurden dabei bewusst differenziert betrachtet, um etwaige Unterschiede herausarbeiten zu können. Die Formulierungen wurden den in einer Studie von Riemann [Ri15] ermittelten Erfolgsfaktoren und Nutzenbeiträgen bei Cloud-basierten GPM-Lösungen zugeordnet. Dadurch konnten die Formulierungen in ein gemeinsames, initiales Kodierschema eingeordnet werden, was die Analyse erleichterte, aber bei Bedarf angepasst und erweitert wurde (auf Anfrage bei den Autoren erhältlich). Das Vorgehen orientierte sich an den Richtlinien der Qualitativen Inhaltsanalyse nach Mayring [Ma15]. Nach erfolgter Zuordnung wurden die Faktoren hinsichtlich ihrer Auftretshäufigkeit sortiert, um dadurch erste Aussagen zu deren jeweils wahrgenommenen Wichtigkeit ableiten zu können. Schließlich wurden die Ergebnisse der Anbieter- und der Anwenderperspektive gegenübergestellt.

4 Ergebnisse

Anbieterperspektive Im Rahmen der Untersuchung konnten 36 Cloud-basierte GPM-Lösungen mit entsprechenden Formulierungen zu erfolgsrelevanten Faktoren identifiziert werden. Bei dieser Auswahl dominieren Software-Lösungen, die sich sowohl den Cloud-Computing-Servicemodellen Platform-as-a-Service (PaaS) als auch Software-as-a-Service (SaaS) zuordnen lassen. Zwei reine PaaS- und drei reine SaaS-Varianten vervollständigen die Stichprobe. Die Anbieter Nintex, ProcessMaker, OpenText und Software AG sind mit mehreren Lösungen vertreten. Die GPM-Lösungen unterstützen Nutzer bei der Umsetzung des GPM, indem sie umfangreiche GPM-Suiten über die Cloud zur Verfügung stellen. Der Funktionsumfang umfasst insbesondere die Prozessmodellierung, das Prozessdesign, die Prozesssimulation, die Prozessautomatisierung, das Prozessmonitoring, die Prozessanalyse und die Anbindung von Drittsystemen und -services.

	bedarfsgerechte Bereitstellung / Skalierbarkeit / flexible Anpassung	bessere Kostenkontrolle / Kostenreduktion	Steigerung der Performance / Effizienz / Produktivität	kurze Markteinführungszeit	umfassende Sicherheit / Datenschutz / Standards / Compliance	geringer Implementierungsaufwand	kontinuierliche Updates / Bereitstellung neuester Technologien	Zuverlässigkeit / hohe Verfügbarkeit / Robustheit der Umgebung	orts- und zeitunabhängiger Zugriff / Mobility	Kollaboratives Arbeiten	Funktionalitäten in einer Umgebung eines Anbieters	Transparenz	zentralisiertes Echtzeit-Prozessmanagement	schnellerer ROI	Σ
Adonis		X			X	X	X				X				5
Agiloft	X				X	X	X	X							5
Appian		X		X		X	X	X	X	X			X	X	9
AuraQuantic	X	X	X		X			X							5
BIC	X	X				X	X				X			X	6
Bizagi	X	X			X		X	X							5
BP Logix	X	X						X	X						4
Bonitasoft	X	X	X		X		X	X		X					7
Camunda	X		X			X				X			X		5
Celonis	X	X	X	X	X		X		X	X	X	X			10
Cflow			X	X	X			X	X	X					6
Creatio	X		X												2
Comindware	X	X	X		X	X		X	X		X	X			9
Flokzu			X	X					X			X			4
Fujitsu	X	X	X	X											4
Genpact		X	X	X											3
IBM	X	X		X		X		X	X	X		X	X	X	10
iGfrax	X	X			X				X		X				5
Kissflow			X		X	X		X	X	X					6
Nintex	X	X		X		X	X								5
OpenText	X	X	X		X	X	X			X				X	8
Oracle			X	X					X	X					4
PegaSystems	X	X		X	X		X	X				X			7
ProcessMaker	X	X			X	X							X		5
Quixy	X				X				X				X		4
SAP	X		X	X					X			X			5
Signavio	X	X	X	X		X	X	X	X			X			9
Software AG	X	X		X					X	X	X	X			7
TIBCO	X			X					X			X		X	5
Zoho	X	X			X	X		X	X						6
Σ	23	20	15	14	15	13	11	13	16	10	6	9	5	5	

Tab 2: Erfolgsrelevante Faktoren aus Anbietersicht

Durch die Analyse der ermittelten Formulierungen konnten 14 erfolgsrelevante Faktoren bei Cloud-basierten GPM-Lösungen aus Sicht von Anbietern abgeleitet werden. Tabelle 2 fasst die Zuordnung der Software-Anbieter zu den Faktoren zusammen. Dabei wird die Summe über die Unternehmen sowie die Faktoren ausgewiesen. Wie die Häufigkeiten zeigen, werden folgende Faktoren seitens der Anbieter besonders oft als erfolgsrelevant herausgestellt: (1) bedarfsgerechte Bereitstellung / Skalierbarkeit / flexible Anpassung, (2) bessere Kostenkontrolle / Kostenreduktion, (3) orts- und zeitunabhängiger Zugriff / Mobility, (4) Steigerung der Performance / Effizienz / Produktivität, (5) Umfassende Sicherheit / Datenschutz / Standards / Compliance sowie (6) kurze Markteinführungszeit. Von den 14 erfolgsrelevanten Faktoren werden von einem Hersteller höchstens zehn abgedeckt (Celonis und IBM). Zum Zwecke der Nachvollziehbarkeit illustriert Tabelle 3 beispielhaft für zwei Cloud-basierte GPM-Lösungen, Adonis und AuraQuantic, die Zuordnung der Formulierungen zu erfolgsrelevanten Faktoren.

GPM-Lösung/ -Anbieter	Formulierung auf dem Internetauftritt	Erfolgsfaktorzuordnung
Adonis	„Sofort startklar“	geringer Implementierungsaufwand
	„Alles aus einer Hand“	Funktionalitäten in einer Umgebung eines Anbieters
	„Neuester Stand“	kontinuierliche Updates / Bereitstellung neuester Technologien
	„Kosteneffizientes Szenario“	bessere Kostenkontrolle / Kostenreduktion
	„Umfassende Sicherheit“	Umfassende Sicherheit / Datenschutz / Standards / Compliance
AuraQuantic	„Technicians specialized in AuraQuantic Cloud take care of security“	Umfassende Sicherheit / Datenschutz / Standards / Compliance
	„AuraQuantic Cloud is monitored to ensure optimal performance“	Steigerung der Performance / Effizienz / Produktivität
	„Only pay for what you use without hefty initial investments“	bessere Kostenkontrolle / Kostenreduktion
	„Achieve maximum efficiency and reduce IT costs, as it helps organizations to automate and control their business processes through applications in the Microsoft Azure Cloud“	Steigerung der Performance / Effizienz / Produktivität
	„Stable storage in the Cloud with high availability and scalability“	Zuverlässigkeit / hohe Verfügbarkeit / Robustheit der Umgebung bedarfsgerechte Bereitstellung / Skalierbarkeit / flexible Anpassung
	„One of the most important advantages of AuraQuantic Cloud is its scalability, rendering it unnecessary to purchase oversized hardware for a possible future need“	bedarfsgerechte Bereitstellung / Skalierbarkeit / flexible Anpassung
	„Simple and reliable security backups of the servers in the Cloud.“	Umfassende Sicherheit / Datenschutz / Standards / Compliance

Tab 3: Zuordnung der Anbieterformulierungen zu erfolgsrelevanten Faktoren (Auszug)

Anwenderperspektive Relevante Formulierungen zu Cloud-basierten GPM-Lösungen wurden in insgesamt 56 Anwenderberichten gefunden. Abbildung 1 stellt die 16 resultierenden erfolgsrelevanten Faktoren mit Angabe der jeweiligen Häufigkeiten dar. Aufgrund der oftmals generellen, anbieterneutralen Aussagen von Anwendern wurde auf die explizite Zuordnung zu Anbietern verzichtet.

Wie Abbildung 1 zu entnehmen ist, finden Anwender Cloud-basierter GPM-Lösungen nachfolgende Faktoren für besonders erfolgsrelevant: (1) orts- und zeitunabhängiger Zugriff / Mobility, (2) bedarfsgerechte Bereitstellung / Skalierbarkeit / flexible Anpassung, (3) geringer Implementierungsaufwand, (4) kontinuierliche Updates / Bereitstellung neuester Technologien, (5) bessere Kostenkontrolle / Kostenreduktion, (6) Steigerung der Performance / Effizienz / Produktivität, (7) Kollaboratives Arbeiten sowie (8) kurze Markteinführungszeit.

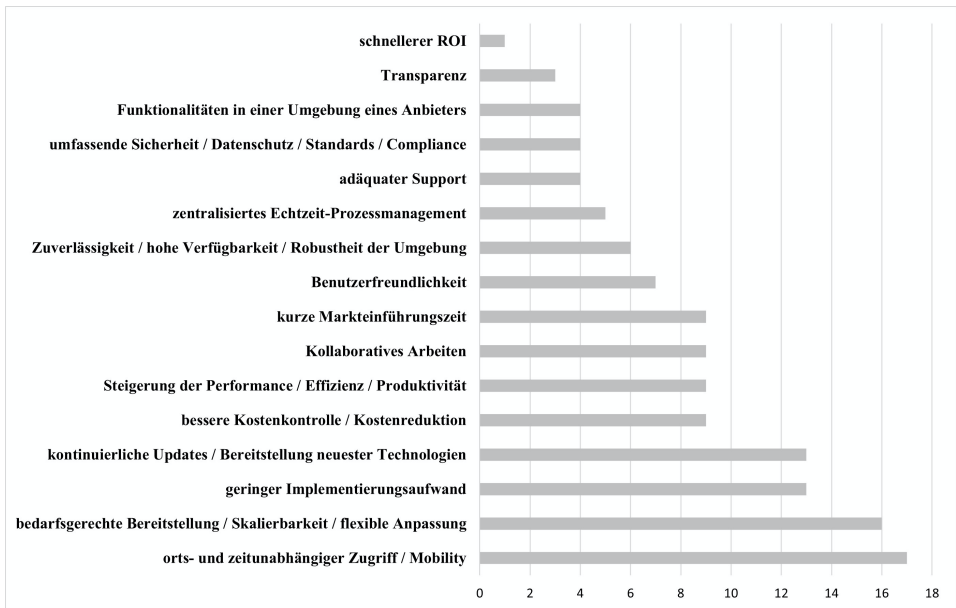


Abb. 1: Erfolgsrelevante Faktoren aus Anwendersicht

Vergleich der Anbieter- und der Anwenderperspektive Perspektivenübergreifend ergaben sich in Summe 16 erfolgsrelevante Faktoren. Bei genauerer Betrachtung der Faktoren fällt auf, dass sich diese gegenseitig beeinflussen und nur im Verbund zum Erfolg von Cloud-basierten GPM-Lösungen beitragen können. 13 der Faktoren waren Teil des aus der Literatur abgeleiteten, initialen Kodierschemas. Die Faktoren „Benutzerfreundlichkeit“, „adäquater Support“ und „Funktionalitäten in einer Umgebung eines Anbieters“ wurden, wenn auch unterdurchschnittlich oft genannt, zusätzlich erstellt. Von den 16 ermittelten Faktoren werden 14 sowohl von Anbietern als auch von Anwendern

als erfolgsrelevant wahrgenommen. Lediglich „adäquater Support“ sowie „Benutzerfreundlichkeit“ führen ausschließlich Anwender auf. Insgesamt kann somit festgehalten werden, dass sich die beiden Perspektiven hinsichtlich der genannten Faktoren und deren relativen Häufigkeit ähneln. Das bedeutet, dass Anbieter, zumindest in Summe, die Anforderungen und Wünsche der Anwender in hohem Maße abdecken.

In Tabelle 4 werden die erfolgsrelevanten Faktoren von Cloud-basierten GPM-Lösungen jeweils kurz beschrieben:

Erfolgsrelevante Faktoren	Beschreibung
bessere Kostenkontrolle / Kostenreduktion	Durch das Cloud-Bereitstellungsmodell sparen sich Anwender die hohen Anschaffungskosten für die zugrundeliegende Infrastruktur. Zudem bezahlen sie nach dem Pay-per-Use-Prinzip nur für diejenigen Ressourcen, die sie tatsächlich auch verbrauchen. Dadurch können Anwender die anfallenden Kosten besser kontrollieren und gezielt reduzieren.
geringer Implementierungsaufwand	Die Implementierung einer GPM-Lösung übernimmt zum größten Teil der Anbieter. Kunden sparen dadurch finanzielle und personelle Ressourcen ein, die bei einer eigenständigen Umsetzung und Integration nötig wären.
bedarfsgerechte Bereitstellung / Skalierbarkeit / flexible Anpassung	Anwender können die Ressourcen der Anbieter in der Cloud nutzen. Damit wird ihnen die eigenständige Beschaffung der benötigten Ressourcen erspart, wodurch die Lösung deutlich schneller in Betrieb genommen werden kann. Zudem können die bereitgestellten Ressourcen bedarfsgerecht angepasst werden (Skalierbarkeit).
orts- und zeitunabhängiger Zugriff / Mobility	Gemäß der Cloud-Computing-Charakteristika können die GPM-Lösungen zu jeder Zeit von jedem Ort aus verwendet werden, sofern ein Internetzugang vorhanden ist. Dabei ist von großem Vorteil, dass die Prozesse auch auf mobilen Endgeräten eingesehen und verwaltet werden können.
umfassende Sicherheit / Datenschutz / Standards / Compliance	Die Sicherheit und der Datenschutz werden häufig als Herausforderungen bei Cloud-basierten GPM-Lösungen angesehen. Anbieter können allerdings zahlreiche Standards und technische Verfahren zur Gewährleistung einer sicheren, konformen GPM-Lösung nutzen. Die Einhaltung dieser Faktoren ist elementar.
kontinuierliche Updates / Bereitstellung neuester Technologien	Die bereitgestellten Cloud-GPM-Lösungen werden durch den Anbieter mit Updates versorgt. Den Anwendern stehen somit kontinuierlich Verbesserungen und Erweiterungen zur Verfügung, wobei für sie keine zusätzlichen Aufwände entstehen.
schnellerer ROI	Kurze Implementierungszeiten, Kosteneinsparungen bei der Ressourcenbezahlung und schnelle Optimierungsergebnisse der Prozesse in der Cloud führen potenziell zu einem schnelleren ROI.
Zuverlässigkeit / hohe Verfügbarkeit / Robustheit der Umgebung	Der Anbieter nutzt große, sichere Rechenzentren, um die bereitzustellende Infrastruktur zu managen. Er hat dabei verschiedene Möglichkeiten dafür zu sorgen, dass die GPM-Lösung für die Anwender stets verfügbar ist.
kollaboratives Arbeiten	Durch die Nutzung einer Cloud-basierten GPM-Lösung können die Prozessbeteiligten, beispielsweise bei der Modellierung von Prozessen, zusammenarbeiten. Dadurch lassen sich Aufgaben schneller und effizienter innerhalb einer zentralen Umgebung umsetzen.
Steigerung der Performance / Effizienz / Produktivität	Die Bereitstellung einer GPM-Lösung in der Cloud wirkt sich positiv auf die Performance, Produktivität und Effizienz des GPM aus. Das Management von GPM-Aufgaben wird vereinfacht, wodurch der Ressourceneinsatz und die Ressourcenauslastung sowie Durchlaufzeiten optimiert werden können.

kurze Markteinführungszeit	Die GPM-Lösung kann schneller in die Unternehmenslandschaft der Anwender integriert werden als bei einer On-Premise-Variante. Somit können Anwender die GPM-Lösung bereits nach in kurzer Zeit nutzen.
Transparenz	Die Erkenntnisse und Einsichten, die zu Prozessen gewonnen und über Applikationen dargestellt werden, können in der Cloud nicht nur mit wenigen Experten geteilt werden, sondern sind für alle Berechtigte einsehbar. Folglich wird die Transparenz innerhalb einer Organisation gesteigert, womit ein größeres Verständnis für die Prozessabläufe geschaffen wird.
zentralisiertes Echtzeitprozessmanagement	Da durch eine Cloud-Lösung die Prozessdaten den Anwendern in Echtzeit zur Verfügung stehen, können diese kontinuierlich den Status der Prozesse überwachen und gegebenenfalls auf auftretende Probleme zeitnah reagieren.
Funktionalitäten in einer Umgebung eines Anbieters	Sämtliche Funktionalitäten für das Management der Geschäftsprozesse erhalten Anwender aus einer einzigen Suite. Dadurch können Pflege- und Anbindungsaufwände reduziert werden.
adäquater Support	Da es in bei der Nutzung einer GPM-Lösung auf der Anwenderseite zu Problemen kommen kann, ist es wichtig, dass Anbieter einen zuverlässigen Support anbieten, der den Anwendern bei ihrer Problembewältigung hilft.
Benutzerfreundlichkeit	Cloud-basierte GPM-Lösungen müssen einfach und nutzerfreundlich zu bedienen sein. Die Benutzerfreundlichkeit betrifft alle enthaltenen Funktionalitäten. Die Lösungen sind so intuitiv zu gestalten, dass Anwender ohne ausführliche, vorherige Schulung starten können.

Tab. 4: Beschreibung der ermittelten erfolgsrelevanten Faktoren

5 Schlussbetrachtung

Innerhalb einer qualitativ-empirischen Querschnittanalyse konnten zu 36 Cloud-basierten GPM-Lösungen aus Kunden- und Anbietersicht insgesamt 16 erfolgsrelevante Faktoren identifiziert werden. 13 der Faktoren waren Bestandteil des aus der Literatur abgeleiteten, initialen Kodierschemas. Die Faktoren „Benutzerfreundlichkeit“, „adäquater Support“ und „Funktionalitäten in einer Umgebung eines Anbieters“ wurden ergänzt. Von den 16 Faktoren werden 14 sowohl von Anbietern als auch von Anwendern als erfolgsrelevant wahrgenommen. Lediglich „adäquater Support“ und „Benutzerfreundlichkeit“ führen ausschließlich Anwender auf. Die relative Häufigkeit der Nennungen der einzelnen Faktoren variiert zwischen beiden Perspektiven etwas. Übergreifend wurden folgende Faktoren besonders häufig genannt:

- bedarfsgerechte Bereitstellung / Skalierbarkeit / flexible Anpassung
- orts- und zeitunabhängiger Zugriff / Mobility
- bessere Kostenkontrolle / Kostenreduktion
- geringer Implementierungsaufwand
- kontinuierliche Updates / Bereitstellung neuester Technologien
- Steigerung der Performance / Effizienz / Produktivität

Bei der Studie konnten im Bereich Cloud-basierter GPM-Lösungen folgende Trends ausgemacht werden: Zunächst lässt sich für die Erstellung von Prozessmodellen, Benutzeroberflächen und Reports erkennen, dass GPM-Lösungen vermehrt mit der Low-/No-Code-Development-Technologie ausgestattet werden. Dadurch soll die Generierung dieser Komponenten ohne oder nur mit geringen Programmierkenntnissen durch Endbenutzer verhältnismäßig aufwandsarm umgesetzt werden können. Dieser Gedanke hat bereits in anderen Domänen wie der Erstellung von Apps Einzug gehalten [FF20]. Des Weiteren wird vermehrt auf die Ermöglichung von kollaborativem Arbeiten (orts- und zeitunabhängig) gesetzt. Auch stehen zunehmend, was die Prozessautomatisierung, das Prozessmonitoring und die Prozessanalyse angeht, die roboterisierte (RPA) oder digitale Prozessautomatisierung (DPA) und die Künstliche Intelligenz im Fokus. Diese Technologien können dazu beitragen, die Durchlaufzeiten zu reduzieren, den Ressourceneinsatz zu minimieren und somit die Prozesskosten als Ganzes zu senken.

Neben den Cloud-basierten GPM-Lösungen konnten bei der Untersuchung sogenannte BPaaS-Angebote identifiziert werden. Diese ermöglichen Kunden eine ganzheitliche Auslagerung von Prozessen. Grundsätzlich bieten sich dafür standardisierte Supportprozesse an, die bei vielen Anwendern sehr ähnlich ablaufen und deshalb keine Anpassungen erfordern. Dadurch entfällt für Kunden ein Großteil der GPM-Aufgaben (Tabelle 1) [Ac11]. BPaaS-Lösungen werden für die verschiedensten Geschäftsbereiche angeboten. Nachdem derartige Angebote in dieser Untersuchung bewusst ausgeklammert wurden, könnten sich zukünftige Forschungsarbeiten mit dem BPaaS-Markt und den dortigen erfolgsrelevanten Faktoren näher beschäftigen.

Der Nutzen der Untersuchungsergebnisse für Forscher und Praktiker ist mehrschichtig: Anwender erhalten Unterstützung bei der Anbieter- beziehungsweise Toolauswahl, indem sie die abgeleiteten erfolgsrelevanten Faktoren in ihren Auswahlprozess einbeziehen können. Anbieter können von den Ergebnissen dahingehend profitieren, dass sie ihre Cloud-basierten GPM-Lösungen entsprechend den Kundenwünschen gestalten und bewerben können. Im Verbund nennen die Anbieter zwar die Aspekte auf ihren Internetauftritten, die in der Wahrnehmung der Kunden wichtig sind, aber kein Anbieter in der Stichprobe deckt die Gesamtmenge der Faktoren ab. Forschern bietet die Untersuchung einen geeigneten Startpunkt für weitere Analysen auf dem Gebiet der Cloud-basierten GPM-Lösungen (siehe nachfolgende Absätze).

Trotz des strukturierten Vorgehens bei der Recherche und der Auswertung weist diese Untersuchung gewisse Limitationen auf. Die erste Limitation betrifft die Größe der gezogenen Stichprobe des Tool-Markts (Anbieterperspektive). Die relativ geringe Trefferzahl lässt sich darauf zurückführen, dass nur diejenigen Lösungen aufgenommen wurden, zu denen als erfolgsrelevant einstufbare Faktoren auf den jeweiligen Hersteller-Webseiten gefunden werden konnten. Nachfolgestudie könnten versuchen, den Markt umfassender zu erforschen. Zweitens ist festzuhalten, dass die Ergebnisse nur so valide und reliabel wie die auf den Tool-Hersteller-Webseiten veröffentlichten Informationen und die Kundenberichte sind. Einerseits kann bei den Kundenberichten nicht ausgeschlossen werden, dass diese nicht in allen Fällen objektiv und unabhängig erstellt wur-

den. Andererseits ist davon auszugehen, dass Anbieter gezielt versuchen, sich in einem positiveren Licht darzustellen. Eine allgemeine und vergleichende Evaluation der erfolgsrelevanten Faktoren kann dieser Beitrag damit nicht leisten. Die Häufigkeit der Nennungen kann lediglich eine erste Indikation für deren Bedeutung sein. Zukünftige Forschungsarbeiten könnten deshalb versuchen, im Rahmen von Fallstudien objektivere erfolgsrelevante Faktoren zu erheben. Die Ergebnisse dieser Studie können dazu als Referenz herangezogen werden. Drittens konnte, entgegen der ursprünglichen Planung, aufgrund der oftmals anbieterneutralen Aussagen von Anwendern kein direkter, Software-Tool-spezifischer Abgleich mit den Anbieterinformationen vorgenommen werden.

Vorliegende Untersuchung stellt insgesamt einen Schritt bei der systematischen Erforschung erfolgsrelevanter Faktoren von Cloud-basierten GPM-Lösungen dar. Um die bisherigen Ergebnisse zu evaluieren und gegebenenfalls zu erweitern, ist eine multiple Fallstudie geplant, bei der sowohl Kunden- als auch Anbietervertreter von wirtschaftlich erfolgreichen Unternehmen zu erfolgsrelevanten Faktoren befragt werden sollen. Interessant ist dabei, inwieweit und warum Faktoren, die auf dem Internetauftritt eines spezifischen Anbieters nicht aufgeführt beziehungsweise von Unternehmensvertretern nicht genannt werden, dennoch auf das Softwarewerkzeug zutreffen und umgekehrt. In dem Zusammenhang kann auch der Frage nachgegangen werden, inwieweit sich die Softwarewerkzeuganbieter basierend auf ihren Lösungen beziehungsweise ihrem Geschäftsmodell als Ganzes in Gruppen zusammenfassen lassen.

Am Ende kann festgehalten werden, dass einzelne der ermittelten erfolgsrelevanten Faktoren auf andere Domänen übertragen beziehungsweise als weitgehend SaaS-generisch angesehen werden können. Diese Studie trägt somit zum Fortschritt der sich noch in einem vergleichsweise frühen Stadium befindlichen Forschung zu Erfolgsfaktoren bei SaaS-Services Ganzes [F118] bei. Die Autoren dieses Beitrags [FL22] haben kürzlich eine Metastudie über Studien zu Erfolgsfaktoren von SaaS-Services durchgeführt und damit die bislang weitestgehend voneinander losgelösten Forschungsarbeiten verglichen und integriert. Der zuvor herrschende Zustand hat den Forschungsfortschritt gehemmt.

Andererseits legt diese Studie nahe, dass die Erfolgsforschung auf Ebene der Rolle des SaaS-Anbieters, wie sie aktuell in Cloud-Computing-Ökosystemmodellen zu finden ist [FLS20], zu grobgranular ist. SaaS-Anbieter sollten wie in diesem Beitrag stattdessen auf einzelne Domänen heruntergebrochen und gesondert untersucht werden.

Danksagung

Die Autoren danken Herrn Michael Ellinger für seine wertvolle Unterstützung bei der Erhebung und Auswertung der Daten.

Literatur

- [Ac11] Accorsi, R.: Business Process as a Service: Chances for Remote Auditing. IEEE 35. Annual Computer Software and Applications Conference Workshops, München, 2011.
- [AL20] Ahmad, T.; Van Looy, A.: Business Process Management and Digital Innovations: A Systematic Literature Review, *Sustainability*, 12(17), S. 1-29, 2020.
- [CS17] Carrillo, A.; Sobrevilla, M.: BPM in the Cloud: A Systematic Literature Review, *Software Engineering*, arXiv preprint, S. 1-7, 2017.
- [FF20] Friedl, T.; Floerecke, S.: Wissenschaftliche App-Generatoren: Eine Option für die Wirtschaft?, *Wirtschaftsinformatik & Management*, 12(5), S. 370-375, 2020.
- [FL22] Floerecke, S.; Lehner, F.: Meta-Study of Success-Related Factors of SaaS Providers based on a Cloud Computing Ecosystem Perspective. In Baumann, S.: *Handbook on Digital Business Ecosystems: Strategies, Platforms, Technologies, Governance and Societal Challenges*. Edward Elgar, Cheltenham, 2022.
- [F118] Floerecke, S.: Success Factors of SaaS Providers' Business Models – An Exploratory Multiple-Case Study. 9. International Conference on Exploring Service Science, Karlsruhe, 2018.
- [FLS20] Floerecke, S.; Lehner, F.; Schweikl, S.: Cloud Computing Ecosystem Model: Evaluation and Role Clusters, *Electronic Markets*, 2020. DOI: 10.1007/s12525-020-00419-2
- [Fo20] Fortune Business Insights: Business Process Management (BPM) Market Size, Share & Industry Analysis, By Component, By Deployment Type, By Enterprise Size, By Functionality, By Industry and Regional Forecast, 2019-2026, 2020. <https://www.fortunebusinessinsights.com/business-process-management-bpm-market-102639>.
- [Ga21] Gartner: Gartner Magic Quadrant for Intelligent Business Process Management Suites, 2021. <https://www.gartner.com/en/documents/3899484/magic-quadrant-for-intelligent-business-process-management>.
- [GD20] Gabsi, H.; Drira, R.; Ben Ghézala, H. H.; Ducasse, S.: From Business Process to Cloud Application. IBIMA 2020 – International Business Information Management Association Conference, Sevilla, 2020.
- [GFA18] Gunawan, W.; Fajar, A. N.; Alianto, H.: Developing Cloud-Based Business Process Management (BPM): A Survey. 2. International Conference on Computing and Applied Informatics, Medan, 2018.
- [Gz20] Gzik, T.: Business Process as a Service – A Systematic Literature Review. In Hermes, M. et al.: *Towards Industry 4.0 – Current Challenges in Information Systems*. Springer, Cham, S. 163-182, 2020.
- [Ha10] Hammer, M.: What is Business Process Management? In Vom Brocke, J.; Rosemann, M.: *Handbook on Business Process Management: Introduction, Methods, and Information Systems*. Springer, Berlin und Heidelberg, S. 3-16, 2010.
- [HC93] Hammer, M.; Champy, J.: *Reengineering the Corporation. A Manifesto for Business Revolution*. Harper Business, New York, 1993.

- [HL21] Helbin, T.; Van Looy, A.: Is Business Process Management (BPM) Ready for Ambidexterity? Conceptualization, Implementation Guidelines and Research Agenda, *Sustainability*, 13(4), S. 1-25, 2021.
- [LNB19] Levina, A.; Novikov, A.; Borremans, A.: BPM as a Service Based on Cloud Computing. In Murgul, V.; Pasetti, M.: *International Scientific Conference Energy Management of Municipal Facilities and Sustainable Energy Technologies EMMFT 2018*. Springer, Cham, S. 210-215, 2019.
- [LWS08] Lehner, F.; Wildner, S.; Scholz, M.: *Wirtschaftsinformatik: Eine Einführung*. Hanser, München, 2008.
- [Ma15] Mayring, P.: *Qualitative Inhaltsanalyse. Grundlagen und Techniken*. Beltz, Weinheim und Basel, 2015.
- [Ri15] Riemann, U.: Benefits and Challenges for BPM in the Cloud. *International Journal of Organizational and Collective Intelligence (IJOICI)*, 5(1), S. 32-61, 2015.
- [SK14] Schlatter, U.; Kykalová, D.; Schladitz, O.; Minonne, C.; Keller, T.: *BPM-Lösungen aus der Cloud. Potenziale, Anforderungen und Erfolgsfaktoren*. ZHAW School of Management and Law, Zürich, 2014.
- [SJ15] Schulte, S.; Janiesch, C.; Venugopal, S.; Weber, I.; Hoenisch, P.: Elastic Business Process Management: State of the Art and Open Challenges for BPM in the Cloud. *Future Generation Computer Systems*, 46, S. 36-50, 2015.
- [We12] Weske, M.: *Business Process Management*. Springer, Berlin und Heidelberg, 2012.
- [WH07] Wilde, T.; Hess, T.: *Forschungsmethoden der Wirtschaftsinformatik*. *Wirtschaftsinformatik*, 49(4), S. 280-287, 2007.
- [WHW03] van der Wil, A.; Hofstede, A.; Weske, M.: *Business Process Management: A Survey*. *International Conference on Business Process Management (BPM 2003)*, Eindhoven, 2003.

On the notion of diversity in the context of information aggregation in BPM

Clemens Schreiber ¹

Abstract: Assuming a set of diverse process variants, information aggregation is used to provide an aggregated overview of the overall process. The main challenge is to provide a condensed but accurate overview, while avoiding redundant information. Although there exist different approaches for information aggregation, none of them actually consider the diversity of the multiple process variant attributes, such as resources, cost or duration. But this hinders from gaining an overview of the process diversity, i.e. the variance of multiple process attributes. Hence, valuable information regarding resource and time variability can get lost. Diversity measures can be used to evaluate the significance of a single process variant, based on multiple attributes. This short paper introduces existing diversity measures from different research disciplines and elaborates on their application in BPM for the purpose of information aggregation.

Keywords: Business process variants; Process diversity; Delta analysis

1 Introduction

The problem of providing an aggregated overview of a set of process variants was first described by van der Aalst and Basten [Aa02] as “management-Information problem”. It refers to the challenge of providing a condensed but accurate overview of multiple process variants, while avoiding redundant or trivial information. In [Aa02] a variant or process version is defined as a workflow-net derived from another workflow-net based on an ad-hoc or an evolutionary change of structural nature. Since it is not desirable to show multiple variants of a process individually, one needs to create an aggregated overview of these variants, e.g. in the form of a unique process diagram. The sufficiency and accuracy of the information provided by the aggregated overview depends, however, on the diversity of the process, i.e. the variance of the process variant attributes. The attributes can be manifold and either refer to the control flow (e.g., number of activities, connectivity of the activities) or performance (e.g., case frequency, task durations, output quality). The more diverse two process variants are, the more information should an aggregated overview contain about these variants. An incomplete process overview could for example omit valuable information regarding the overall flexibility of a process or execution paths, which are relevant for the overall performance. Process diversity should therefore always be considered with respect to performance. Another important aspect of the process diversity is that interdependencies between different process variants might become more apparent, due to the correlation between certain process attributes.

¹ Karlsruhe Institut für Technologie, clemens.schreiber@kit.edu

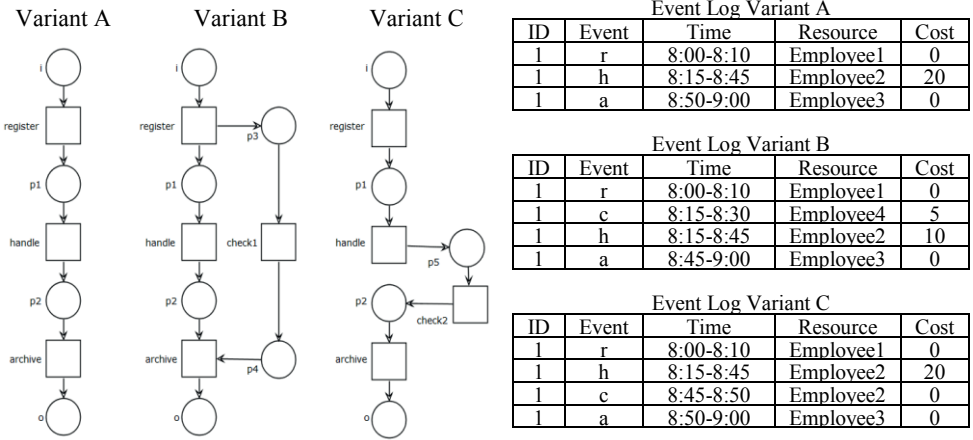


Fig. 1: Three process variants taken from [Aa02] with three respective event logs

In order to be able to consider a variety of process attributes for information aggregation, one should consider the predefined process model, as well as the execution data. Figure 1 shows three process variants taken from [Aa02]. To emphasize the importance of process attributes beyond the control-flow, the three given event logs show possible process execution data for each respective process variant. These event logs contain information regarding resources, event duration, cost and case frequency. One can see that process variant 2 contributes significantly to the overall process diversity, since 1 different resource is used and the costs are also differently distributed compared to the other variants. The two variants A and C on the other hand are much more similar. This should imply, that compared to variant A, variant B is more important than variant C for the representation of the process diversity and hence, for a more complete process overview.

Figure 2 shows two possible representations of the three process variants in one diagram, also taken from [Aa02]. The first one is based on the concept of the Greatest Common Divisor (GCD) and the second one is based on the Least Common Multiple (LCM). The GCD represents the behaviour that all variants agree upon. The Least Common Multiple is the most compact representation, which is still a subclass of all variants. For a more detailed description on how the two representations are derived see [Aa02]. The LCM does provide significantly less information regarding the overall process diversity as the GCD. Especially when one considers the process attributes costs and resources from variant B. The GCD on the other hand does not depict the execution path <r,h,a>, which could be important if the flexibility of the control flow is required as management information. Hence, this small example shows that process attributes can differ in the importance for the overall process diversity, depending on the purpose of the diversity measure. Several different diversity measures exist, which are applied in different research disciplines such as biology and economics. They each have different properties

and allow for different interpretations of diversity. In the following section, several diversity measures will be introduced and analysed in the context of information aggregation in BPM. The purpose of the diversity measure is to identify the process variants that are most significant for the representation of the overall process.

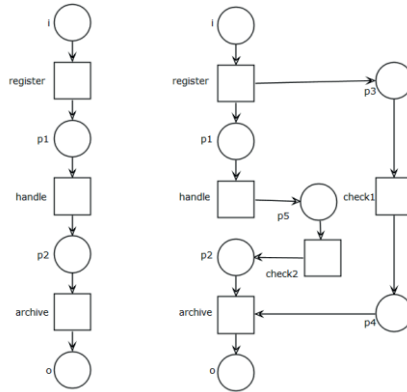


Fig. 2: The LCM (left) and the GCD (right) of the three variants in Fig. 1, taken from [Aa02]

2 Diversity Measures

Diversity can generally be defined as the totality of the dissimilarity among objects [Ne09]. Hence, the more dissimilar objects are among each other, the more diverse is their totality. Biodiversity for example aims to capture the overall dissimilarity of species, which live in a defined space. In economics diversity can refer to the overall dissimilarity among a set of products, such as the model portfolio of a car brand. However, in both cases different aspects of diversity are considered. While for biodiversity the relative occurrence of an object is important, in economics it is the characteristic features of the objects. This can be linked to a normative judgement on the purpose of diversity in its particular domain [Ba06]. In economics, product diversity is usually linked to the ability to choose between different options, in biology, biodiversity is used as an indicator for the preservation of an ecosystem. In the following 3 main approaches for the measurement of diversity among objects will be considered: entropy measures, distance measures and attribute measures. One should notice that further measures exist to capture the diversity across populations and within an object, i.e. the variance of an attribute.

Entropy measures are commonly used for the assessment of biodiversity. They can be considered as special cases of the generalized function [Ba06]:

$$f(n, p) = \sum p \log \frac{1}{p} \tag{1}$$

where n is the total number of different species and p the relative abundance of a species i . The absolute abundance of a species i is

therefore represented by a_i . The parameter α determines how much weight the diversity measure places on the abundance of the species. For $\alpha = 0$, the equation yields n , the total number of different species. In this case the relative abundance does not matter at all. A population of 1 butterfly and 99 birds would have the same diversity as a group of 50 butterflies and 50 birds. An alternative is the Shannon-Wiener entropy, which can be obtained from equation (1) by setting $\alpha = 1$:

(2)

The Shannon-Wiener entropy satisfies the following useful mathematical properties among continuous functions [Pa10]: (1) Symmetry: the calculated value is independent on how the relative abundance is distributed among the species, e.g. a population of 1 butterfly and 99 birds has the same diversity value as a population of 99 butterflies and 1 bird; (2) Maximum of equality: exists when species exist in equal proportions, i.e.

n for all i ; (3) Decomposability: the total diversity equals between category diversity plus within category diversity, where categories can be understood as disjunct subsets of species within an ecosystem. While the Shannon-Wiener entropy satisfies these properties, one main limitation of entropy measures is that they do not consider the extent of dissimilarity between objects. Hence, 50 butterflies and 50 birds could have the same diversity as 50 butterflies and 50 beetles.

Distance measures are able to take the extent of dissimilarity between objects into account. For this purpose, distance measures require a formal representation of the characteristic object attributes [Ba06]. Once these characteristic attributes are defined, the diversity is calculated based on some distance function $d(s_i, s_j)$, where s_i, s_j represent two objects of a set S . In addition to the distance function, one needs to define a mechanism to aggregate the pairwise distance between all objects in S . By simply adding up all pairwise distances, diversity would increase each time an already existing object is added to a set. Although, one should assume that adding a bird to a population of 99 birds and 1 butterfly, should rather decrease than increase diversity. An alternative approach is the recursive aggregation algorithm, developed by Weitzman [We92]: Iteratively increase the diversity value $W(S)$ by adding the distance of an object s_i to the set S , where s_i is the object with the least distance to S . In the beginning S is an empty set and one needs to define a random object to start the recursive aggregation. The main shortcoming of this approach is, that in general the result of the measure is path-dependent and does not provide a unique value. One can show that the outcome of the measure is only uniquely defined, if the pairwise distances are ultrametric, i.e. when for three possible pairwise distances between any three objects, the two greatest distances are equal [Ba06]. This is however, a strong precondition for the assessment of diversity.

Attribute measures offer a more general alternative and were first proposed by Nehring and Puppe [Ne09]. The measure solely accounts for the unique attributes represented by a set of objects and does not consider their relative abundance. This is based on the presumption that the number of distinct attributes accounts better for the diversity of a

set than the number of objects. The approach can be applied in different disciplines. The flexibility, however, comes with the cost that one needs to preselect a set of attributes and weights. The predefined attributes of the set Ω are added up by a weighting function:

$$(3)$$

where λ represents weights over the different attributes $f \in F$. As stated in [Pa10], depending on whether the weights originate from some objective criterion or subjective criterion, one either has a diversity measure or a value function. The weights should therefore reflect the purpose of the diversity analysis. An additional aspect of the applicability of the multi-attribute approach is the transferability of attributes and the separability of their functions [Pa10]. This depends on the specific assumption over diversity in the respective domain. Just as there are normative judgements needed on the purpose of diversity in biology and in economics, one needs to define what diversity is supposed to represent for the assessment of process variants.

3 Application of diversity measures in BPM

To measure how much diversity is represented by the LCM and the GCD in Fig. 2 one could define each possible execution trace of a process diagram as a single object, similar to a species of an ecosystem. Hence, there is one possible *execution trace* in the LCM and two in the GCD. For the calculation of the *entropy measure* one needs furthermore to define, how often each trace has been executed. For this purpose, following case frequencies will be assumed: $\langle r, h, a \rangle$ is executed 10 times, $\langle r, c1, a \rangle$ is executed 9 times and $\langle r, h, c2, a \rangle$ is executed 1 time. Accordingly, the relative abundance of cases and the Shannon-Wiener entropy for the LCM and the GCD can be calculated. For the calculation of the *distance measure* the Levenshtein distance is used, as described in [Wo06]. Based on the two possible traces in GCD, the outcome of the Levenshtein distance is 2. Since there are only two traces, the value does not change when the Weizman's recursive aggregation algorithm is applied. For the calculation of the *attribute measure* three attributes are assumed to be relevant, which will all be equally weighted by $\lambda=1$: number of transitions (events), number of resources and total cost options. For the LCM this results to the simple addition $3 + 3 + 1$, for the GCD the outcome equals $5 + 4 + 2$. Based on the described premises and calculations, Tab. 1 shows the calculated diversity measures for the LCM and the GCD.

	Traces	Entropy M.	Distance M.	Attribute M.
LCM	1	1	0	7
GCD	2	0.723	2	11

Tab. 1: The 4 diversity measures calculated for the LCM and the GCD in Fig. 2

One can see that only the entropy measure assigns a higher diversity value to the LCM than to the GCD. In the case, where a set only consists of one variant the entropy measure does not have much meaning. The other measures however indicate that the

GCD represents a higher variety of process attributes than the LCM. The most distinctive evaluation is given by the attribute measure, which contains the most information about the two diagrams. This small example demonstrates that existing diversity measures can be applied in BPM. But it also becomes clear that diversity measures need to be carefully selected, in order to provide a meaningful outcome. This also applies for the selection of the considered process attributes.

4 Conclusion

This short paper introduces the idea of applying diversity measures for the improvement of information aggregation in BPM. The approach can be distinguished from other aggregation approaches [Ro17] based on the diversity consideration of selected process attributes. Four common diversity measures have been introduced and tested based on two process diagrams. Several issues, such as the consideration of unfinished traces and process loops have been ignored so far. In addition, the interpretation of the diversity measure outcomes leaves a lot of room for open questions. In order to give a meaningful assessment of business process diversity one should however, consider the diversity in relation to other process aspects, such as performance and flexibility. Only then it is possible to answer whether process diversity can contribute to the improvement of the information aggregation or not. The development of meaningful diversity measures for BPM will hopefully help to answer this question in the future.

References

- [Aa02] W. Van der Aalst und T. Basten, „Inheritance of workflows: An approach to tackling problems related to change,“ *Theoretical Computer Science*, 270, 1-2, pp. 125-203, 2002.
- [Ne09] K. Nehring und C. Puppe, „Diversity,“ *The Handbook of Rational and Social Choice*, pp. 298-322, 2009.
- [Ba06] S. Baumgärtner, „Measuring the Diversity of What? And for What Purpose? A Conceptual Comparison of Ecological and Economic Biodiversity Indices,“ *SSRN*, 2006.
- [Pa10] S. E. Page, „Measuring Diversity,“ in *Diversity and Complexity*, 2. Hrsg., Princeton University Press, 2010.
- [We92] M. Weitzman, „On Diversity,“ *The Quarterly Journal of Economics*, Bd. 107. Jg., Nr. 2, pp. 363-405, 1992.
- [Wo06] A. Wombacher und M. Rozie, „Evaluation of workflow similarity measures in service discovery,“ in *In Service-Oriented Electronic Commerce, Proceedings zur Konferenz im Rahmen der Wirtschaftsinformatik 2006*. Gesellschaft für Informatik e.V., 2006.
- [Ro17] M. L. Rosa, W. Van der Aalst, M. Dumas und F.P. Milani, „Business process variability modeling: A survey,“ *ACM Computing Surveys (CSUR)*, 50(1), pp. 1-45, 2017.

Towards method- and tool-independent business process modeling

Marc Frerichs¹, Stephan Leible¹, Markus Nüttgens¹

Abstract: The last decades have brought significant developments in the field of Business Process Management (BPM). Practice and science have developed and established new approaches such as Process Mining and Robotic Process Automation. However, within these developments, new techniques, tools, formats, and models are constantly being designed that work well in their specific focus but are usually not interoperable with other applications. The diversity of the market thus also brings disadvantages. This is where we start with our research; we conceptualize and develop an architecture that builds a bridge between Process-Aware Information Systems and BPM tools using Design Science Research methodology. With this approach, we want to make a contribution towards method- and tool-independent modeling.

Keywords: business process modeling; meta model; process model transformation; tool interoperability

1 Introduction

Nowadays, most companies see Business Process Management (BPM) as a central strategic topic. The different objectives for the use of BPM are the reason why a wide variety of process modeling methods and tools are used in practice [Ju00]. Despite multiple promising approaches, it cannot be foreseen that an actual standard will ever prevail. Instead, research in the field of BPM is fueling the diversity in terms of new use cases as well as techniques. In general, this is a positive development as there are appropriate solutions for many areas of application. However, this increases the difficulty of integrating different modeling platforms and languages, although there are several reasons to do so. Riggio et al. [Ri05] summarize different motivations for the interoperability among various platforms, e.g., a single tool might not be sufficient for the whole Information System's development life cycle.

This research aims to develop methods and artifacts that enable the interoperability of the various BPM methods and tools with regard to business process modeling so that in each phase of the BPM life cycle [Du18], the advantages of the existing BPM solutions can be leveraged. One of the main motivations is to bridge the gap between Process-Aware Information Systems (PAIS), such as Enterprise Resource Planning (ERP) systems,

¹ University of Hamburg, Hamburg Research Center for Information Systems, Von-Melle-Park 5, 20146 Hamburg, Germany, firstname.lastname@uni-hamburg.de

Workflow Management Systems (WfMS), Customer Relationship Management (CRM) systems, etc., and BPM tools, e.g., for business process design, analytics, monitoring or automation, through *method- and tool-independent process modeling*.

Definition (Method- and tool-independent modeling). *Method- and tool-independent modeling describes the ability to deal with the diversity in PAIS and business process modeling tools and make these systems and tools (and the respective business process models) interoperable where applicable.*

2 Background and related work

As business process modeling has become more important, the number of modeling tools and methods has increased. The main disadvantages of the existing process representations are the lack of standardization and the variety of graphical representations available [Ro06]. The use of heterogeneous methods and tools led to various research approaches to deal with tool and method incompatibility, which we will classify in the following three areas: (1) *transformation of process models*, (2) *meta²-modeling and ontologies*, and (3) *standardization of methods and tools*.

The *transformation of process models* deals with translating models into a different notation that is more suitable for the intended purpose. Three forms of model transformation can be distinguished [Jo07]: (1) Transformation between graphically supported modeling languages, e.g., from Event-driven Process Chain (EPC) [KNS92] to Business Process Model and Notation (BPMN)²; (2) transformation of graphical process modeling notations into an interchange format, e.g., EPC Markup Language (EPML) [MN06] or BPMN Markup Language (BNML) [VZS05]; (3) transformation between different interchange formats [MN04]. Process model transformation has been extensively researched, while most approaches focus on XML-based transformation [MNN04].

The *use of (meta-)meta-modeling and ontologies* [Ho07] is another approach to deal with interoperability among process modeling languages by integrating meta models [KV11]. Karagiannis and Höfferer [KH06] state that “to integrate artifacts that are already described by a meta model, at least a [common] meta²-model is needed,” while Kühn and Murzek [KM06] show the interoperability problems that may arise and how integrated meta models can be realized. The work on ontologies regarding conceptual modeling addresses semantic integration (“tower of Babel” problem), where a Universal Ontology (UO) is proposed [O117]. The UFO (Unified Foundational Ontology) “has been employed as a basis for [...] integrating many modeling languages and standards in different domains” [Gu15].

Various *standardization efforts* were carried out in (business process) modeling that can be specifically attributed to the OMG (Object Management Group). A good example is the Unified Modeling Language (UML). The UML meta model, the Meta-Object Facility

² <https://www.omg.org/spec/BPMN/2.0/PDF/>

(MOF), is intended to bridge the gap between different meta models by creating a general basis (meta²-model). BPMN Diagram Interchange (BPMN DI) facilitates the interchange of BPMN diagrams between tools. With the Business Process Definition Metamodel (BPDM), the OMG provides a “metamodel and serialization mechanism for BPMN concepts.”³ However, according to the specification, e.g., BPMN DI cannot handle any tool-specific information, and it does not ascertain syntactical and semantical correctness. Kurz [Ku16] says that “in some cases there are ambiguities or even contradictions in the specification document” and therefore, “various tool vendors interpret parts of the specification differently.” XML Metadata Interchange (XMI) provides a standard for exchanging metadata information. The BPMN Model Interchange Working Group aims to “facilitate and demonstrate BPMN Model Interchange amongst tools that implement the standard.”⁴

Although the need for method- and tool-independent modeling, as described in sections 1 and 2, has been addressed in many ways, there is still a lack of solutions that enable method and tool interoperability. The transformation between models is a good start, but, e.g., implementations in the form of XSLT scripts are not sufficient to solve the challenges in practice. Extensive research in meta²-modeling and ontologies provides many starting points in the context of method- and tool-independent modeling that have not yet been addressed. Standardization efforts help to increase the quality and uniformity in modeling. However, in terms of business process modeling, these are often limited exclusively to the BPMN ecosystem, which also continues to face standardization challenges. Dirndorfer et al. [DFS13] show that “there are significant difficulties to do a cross-platform exchange of BPMN business process models.” In their case study, all tested BPM tools were rated “poor” or “very poor / fail” concerning their import functionality of models from other tools.

3 Research questions and methodology

The following research questions are proposed in this paper:

RQ1: What are the requirements and challenges of method- and tool-independent process modeling?

RQ2: How can interoperability be sustainably ensured with an increasing diversity of modeling methods and tools?

RQ3: How can process modeling be achieved regardless of the methods and tools used?

The proposed research follows the research approach of Design Science Research (DSR), which aims to develop different types of artifacts that contribute to the solution of an existing problem [He04]. Peffers et al.’s DSR research methodology [Pe07], which consists of six phases, is used to conduct DSR in a structured and rigorous manner. Every phase is supported by scientifically sound methods with the aim of answering the research questions described in this section:

³ <https://www.omg.org/spec/BPDM/1.0/Volume2/PDF>

⁴ <https://www.omgwiki.org/bpmn-miwg/>

- 1. Identify problem and motivate:** Define the problem and explain research relevance to answer *RQ1* through a systematic literature review.
- 2. Define objectives of a solution:** Based on this preliminary work, define and validate goals to solve the research problem.
- 3. Design and development:** Conceptualize and develop the artifacts addressing *RQ2* and needed to realize our concept briefly described in section 4.
- 4. Demonstration:** The efficacy of the created artifacts in solving the problem defined in step one is demonstrated in practical environments under real-world conditions. All single artifacts form a method- and tool-independent architecture, answering *RQ3*.
- 5. Evaluation:** Prototyping is used to design functional prototypes at an early stage as well as to test and evaluate them.
- 6. Communication:** The problem, relevance, research, and methodology will be communicated to the relevant audience through publications at conferences and journals.

4 Research agenda

We propose a concept that enables method- and tool-independent modeling. The core of this concept will be a *meta model* that is able to represent all process-related data. Furthermore, it is necessary to conceptualize and implement two generic interfaces to connect different PAIS and BPM applications: The *PAIS interface* enables the extraction of process data from PAIS, e.g., ERP systems. In addition, it should be possible to (re)import data to such systems, e.g., for business (re)engineering purposes. The *BPM tool interface* allows the exchange of data to and from BPM applications, e.g., modeling or analytics tools. Figure 1 shows a highly simplified representation of our concept.

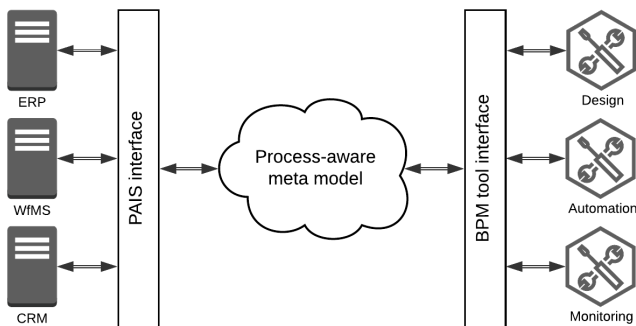


Fig. 1: Method- and tool-independent architecture

As the illustration is supposed to show, there are various challenges in implementing the concept. Beginning with the extraction of data from PAIS, it is not a trivial task to retrieve the relevant process data from these systems. The work of Gonzalez et al. [GRvdA19] seems to be a good starting point for further research. Coping with the heterogeneous underlying data models and interfaces is a challenge in itself. A meta model to be designed

will have to map the semantic and syntactic properties of the various modeling languages as well as the tool-specific features, which have not been addressed sufficiently in existing approaches. In order to ensure the universality of the proposed solution, extensibility will also be provided. The connection of the many BPM tools is heavily dependent on the openness of the respective tools. In practice, it can be seen that especially newer software offers options for importing data and connecting data sources for import purposes. However, there is often little support for exporting or extracting data. Regardless of this, our work focuses on supporting as many tools and methods as possible to achieve a high level of practical benefit and interoperability. The research agenda consists of the following steps:

1. Examination of diversity in business process modeling, especially regarding methods and tools, existing meta models or model transformation efforts (systematic literature review, semi-structured interviews),
2. conceptualization and implementation of a meta model for process-related data (reference modeling),
3. conceptualization and implementation of generic interfaces for data exchange with PAIS and BPM tools (prototyping),
4. conduction of a proof of concept (field research, case study).

The authors are aware of the fact that the proposed research project is an ambitious task. Therefore, the research agenda described should present the authors' planned contributions towards method- and tool-independent modeling but also increase awareness for this *Grand Challenge* in business process modeling.

References

- [DFS13] Dirndorfer, Max; Fischer, Herbert; Sneed, Stephan: Case Study on the Interoperability of Business Process Management Software. In: S-BPM ONE - Running Processes, pp. 229–234. 2013.
- [Du18] Dumas, Marlon; La Rosa, Marcello; Mendling, Jan; Reijers, Hajo A.: Fundamentals of Business Process Management. Springer-Verlag Berlin Heidelberg, 2nd edition, 2018.
- [GRvdA19] González López de Murillas, Eduardo; Reijers, Hajo A.; van der Aalst, Wil M.P.: Connecting databases with process mining: a meta model and toolset. Software and Systems Modeling, 18(2):1209–1247, 4 2019.
- [Gu15] Guizzardi, Giancarlo; Wagner, Gerd; Almeida, João; Guizzardi, Renata: Towards Ontological Foundations for Conceptual Modeling: The Unified Foundational Ontology (UFO) Story. Applied Ontology, 10(3-4):259–271, 2 2015.
- [He04] Hevner, Alan R; March, Salvatore T; Park, Jinsoo; Ram, Sudha: Design Science in Information Systems Research. MIS Quarterly, 28(1):75–105, 2004.

- [Ho07] Hofferer, P: Achieving Business Process Model Interoperability Using Metamodels and Ontologies. In: ECIS 2007 Proceedings. volume ECIS 2007, pp. 1620–1631, 2007.
- [Jo07] Johannsen, Florian: Transformation von Prozessmodellen: Bewertung XML-basierter Ansätze, volume 4 of BDVB-Award Geschäftsprozess- und Projektmanagement / Bundesverband Deutscher Volks- und Betriebswirte e.V. (bdvb), Fachgruppe für Geschäftsprozess- und Projektmanagement. Salzwasser-Verlag, Bremen, 2007.
- [Ju00] Junginger, Stefan; Kühn, Harald; Strobl, Robert; Karagiannis, Dimitris: Ein Geschäftsprozessmanagement-Werkzeug der nächsten Generation — ADONIS: Konzeption und Anwendungen. *Wirtschaftsinformatik*, 42(5):392–401, 10 2000.
- [KH06] Karagiannis, Dimitris; Höfferer, Peter: Metamodels in action: An overview. In: ICISOFT 2006 - 1st Int. Conf. on Software and Data Technologies, Proc. volume 1, Setúbal, Portugal, 2006.
- [KM06] Kühn, Harald; Murzek, Marion: Interoperability Issues in Metamodelling Platforms. In: *Interoperability of Enterprise Software and Applications*. pp. 215–226, 2006.
- [KNS92] Keller, G; Nüttgens, M; Scheer, A.-W.: Semantische Prozeßmodellierung auf der Grundlage 'Ereignisgesteuerter Prozeßketten (EPK)'. *Veröffentlichungen des Instituts für Wirtschaftsinformatik (IWi), Universität des Saarlandes, Heft 89*, 1992.
- [Ku16] Kurz, Matthias: BPMN Model Interchange. In: *Proc. of the 8th Int. Conf. on Subject-oriented Business Process Management*. ACM, New York, NY, USA, 4 2016.
- [KV11] Karagiannis, Dimitris; Visic, Niksa: Next Generation of Modelling Platforms. In: *Perspectives in Business Informatics Research*, pp. 19–28. 2011.
- [MN04] Mendling, Jan; Nüttgens, Markus: Transformation of ARIS Markup Language to EPML. In: *Proc. of the 3rd GI Workshop on Event-Driven Process Chain*. pp. 27–38, 2004.
- [MN06] Mendling, Jan; Nüttgens, Markus: EPC markup language (EPML): an XML-based interchange format for event-driven process chains (EPC). *Information Systems and e-Business Management*, 4:245–263, 7 2006.
- [MNN04] Mendling, Jan; Neumann, Gustaf; Nüttgens, Markus: A Comparison of XML Interchange Formats for Business Process Modelling. In: *EMISA 2004 – Informationssysteme im E-Business und E-Government*. pp. 129–140, 2004.
- [OI17] Olivé, Antoni: The Universal Ontology: A Vision for Conceptual Modeling and the Semantic Web. In: *Conceptual Modeling*, pp. 1–17. Springer, Cham, 2017.
- [Pe07] Peffers, Ken; Tuunanen, Tuure; Rothenberger, Marcus A.; Chatterjee, Samir: A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, 24(3):45–77, 12 2007.
- [Ri05] Riggio, Roberto; Ursino, Domenico; Kühn, Harald; Karagiannis, Dimitris: Interoperability in meta-environments: An XMI-based approach. In: *Advanced Information Systems Engineering. CAiSE 2005.*, pp. 77–89. Springer, Berlin, Heidelberg, 2005.
- [Ro06] Rosenkranz, Friedrich: *Geschäftsprozesse - Modell- und computergestützte Planung*. Springer-Verlag, Berlin/Heidelberg, 2nd edition, 2006.
- [VZS05] Vanderhaeghen, Dominik; Zang, Sven; Scheer, August-Wilhelm: *Interorganisationales Geschäftsprozessmanagement durch Modelltransformation*. Technical report, Veröffentlichungen des Instituts für Wirtschaftsinformatik - Saarbrücken, 2005.

Menschliche Faktoren in Workflow-Management-Systemen

Felix Holz Christoph Gibcke Sven Erdmann Michael Fellmann Birger Lantow ¹

Abstract: Workflow Management Systeme (WfMS) unterstützen Arbeitsabläufe, indem sie die Kontrolle über die Arbeitsflüsse unter Betrachtung notwendiger Ressourcen und eines Ablaufschemas ermöglichen. Der Mensch wird ebenfalls als eine solche Ressource angesehen, der die anfallenden Aufgaben erledigt. Bei der Zuweisung von Aufgaben werden menschliche Faktoren jedoch nur spärlich beachtet. Das Wohlbefinden des Personals, Interessen und Präferenzen sowie soziale Gefüge sind nur wenige Aspekte, die für eine verbesserte Arbeitsverteilung sorgen und sowohl dem Unternehmen als auch dem Mitarbeitenden zu Gute kommen können. Um den Forschungsstand zur Berücksichtigung menschlicher Faktoren in WfMS zu ermitteln, wurde eine strukturierte Literaturanalyse durchgeführt. Die analysierten Studien zeigen, dass die angemessene Zuweisung von Ressourcen mit der meisten Forschungsaktivität bedacht werden. Nichtsdestotrotz unterscheiden sich die Herangehensweisen deutlich. Die Verbesserung von Methoden und Bewertungskriterien sowie ein erhöhtes Maß an Formalisierung seien Schlüsselaspekte, um die Forschung weiter voran zu bringen.

Keywords: Workflow Management System; Business Process Management System; menschliche Faktoren; Ressourcenallokation

1 Einleitung

Die Arbeit steht im Wandel, denn sie ist geprägt vom Informationszeitalter und wird immer weiter vorangetrieben durch die Digitalisierung und technologische Entwicklungen. Die Arbeitnehmer fordern zunehmend eine flexiblere Arbeit, die unabhängig von der Zeit oder dem Ort ist. Außerdem ist erstrebenswert, dass sich die Arbeitsbedingungen ihrer aktuellen persönlichen Lebenssituation anpassen können. In klassischen Workflow Management Systemen (WfMS) wird die Arbeit nach einer vorgegebenen Ablauflogik verteilt, die über ein Prozessmodell spezifiziert wird. In solch starren Konzepten können menschliche Faktoren nur unzureichend abgebildet werden. Die fehlende Berücksichtigung des Menschen durch Kontrollmechanismen von Unternehmen können bei einigen Arbeitnehmern in Über- oder Unterforderung münden. Der Bedarf nach einer flexibleren Arbeitsorganisation ist nicht zuletzt durch die stetig komplexer werdenden Aufgaben zu begründen. So bietet z.B. ein Ticketsystem dem Einzelnen mehr Freiheiten in der Arbeitsgestaltung, wirft auf der anderen Seite aber bezüglich Effizienz und Fairness neue Fragen auf.

Um einen Überblick über den Forschungsstand gewinnen zu können, wie aktuelle Workflow-Management-Systeme menschliche Faktoren berücksichtigen, wird eine systematische

¹ Universität Rostock, Albert-Einstein-Straße 22, 18051 Rostock, Deutschland {felix.holz,christoph.gibcke,sven.erdmann,michael.fellmann,birger.lantow}@uni-rostock.de

Literaturrecherche durchgeführt. Dabei liegen insbesondere die Ermittlung neuer Ansätze zur Arbeitsverteilung und Unterstützung der Zusammenarbeit im Fokus der Untersuchung, insbesondere in welchen Teilbereichen der WfMS diese Ansätze umgesetzt werden. Der Reifegrad des Forschungsgebiets ist ebenfalls von Interesse, nicht zuletzt um Forschungspotentiale zu erheben und einen Ausblick für die zukünftige Entwicklungen zu erhalten.

In Abschnitt 2 werden die Hintergründe von Workflow Management Systemen und Business Process Management Systemen kurz aufgeführt, sowie der Faktor Mensch erläutert. Abschnitt 3 präsentiert die angewendete Forschungsmethode der Literaturrecherche zuzüglich der Forschungsfragen und der Suchstrategie. Darauf folgend wird die gefundene Literatur im Zuge des vierten Abschnittes analysiert, um die Forschungsfragen zu beantworten. Abschnitt 5 fasst die wichtigsten Erkenntnisse zusammen und gibt einen Ausblick für künftige Forschungsrichtungen.

2 Hintergrund

Workflow Management Systeme (WfMS) und Business Process Management Systeme (BPMS) dienen der Formalisierung und Unterstützung von betrieblichen Arbeitsabläufen und Geschäftsprozessen. Da die Grenzen zwischen diesen beiden Konzepten nicht eindeutig sind, wurden sie in der Vergangenheit häufig als Synonym wahrgenommen. Im Grunde kann ein WfMS Teil eines BPMS sein. Der Kerngedanke des Workflow Managements liegt in der Kontrolle von Geschäftsprozessen und der Unterstützung der Ausführung einzelner Arbeitsflüsse, unter der Beachtung von Ressourcen und Abhängigkeiten. Der Hauptzweck des WfMS liegt in der „Unterstützung von der Definition, Ausführung, Registrierung und der Kontrolle von Geschäftsprozessen.“ Es ist dafür zuständig, der Arbeitsaufgabe die richtigen Ressourcen zur richtigen Zeit zuzuordnen und die Weiterführung des Prozesses zu gewährleisten. [Re03, S.18] Das Business Process Management beschäftigt sich zusätzlich mit der grundsätzlichen Gestaltung und der Optimierung von Geschäftsprozessen, indem sie über Prozessmodelle formalisiert werden. Innerhalb des Business Process Managements sind insbesondere das Design, die Analyse, die Modellierung, die Implementation sowie die Kontrolle von Geschäftsprozessen von zentraler Bedeutung. [Re03, S.11 f.]

Jedoch scheint ein Nachteil der bestehenden Systeme dahingehend zu bestehen, dass die Ressource „Mensch“ nicht adäquat abgebildet und somit auch nur mangelhaft berücksichtigt wird. Während Konzepte wie Verantwortlichkeiten bereits Bestandteil der Systeme sind, sind im Menschen etliche Faktoren inbegriffen, die für eine verbesserte Verteilung der Arbeitslast und den optimalen Ablauf der Prozesse hilfreich sind. So bestimmen diese „menschlichen Faktoren“, wie z.B. das Wohlbefinden, die Erschöpfung oder Soziale Gefüge, unter welchen Umständen der Mitarbeiter zu welcher Art der Arbeit bereit ist. [FLW20] Sobald mehr Eigenschaften des Personals berücksichtigt werden, ist zu vermuten, dass auch die Arbeitszufriedenheit steigt. Wenn z.B. bekannt ist, dass ein Mitarbeiter produktiver in der Zusammenarbeit mit bestimmten Kollegen ist, wird eine automatische Zuweisung gemeinsamer Aufgaben positiv wahrgenommen. Insbesondere in immer flexibler werdenden

Arbeitswelten, in denen die Arbeit von hoher Komplexität, Zeitdruck und intensiver Arbeit geprägt ist, nimmt die Bedeutung der Ressource Mensch zu. [GM01]

3 Forschungsmethode

In diesem Beitrag wird eine systematische Literaturrecherche nach den Kriterien von Kitchenham [Ke07] durchgeführt. Nach Kitchenham ist die systematische Literaturrecherche ein Mittel zur Identifizierung, Bewertung und Interpretation von verfügbaren Studien, die für ein bestimmtes Themengebiet oder eine Forschungsfrage relevant sind.

Nach den Kriterien von Kitchenham werden sechs verschiedene Schritte für eine systematische Literaturrecherche definiert. Als erstes wird das vorliegende Problem des Forschungsgebiets beschrieben. Darauf aufbauend werden verschiedene Forschungsfragen für die Arbeit definiert. Im zweiten Schritt werden verschiedene Studien in einem Suchprozess anhand eines Suchstrings identifiziert. Die gefundenen Studien werden im anschließenden dritten Schritt durch Ein- und Ausschlusskriterien bewertet und relevante Studien für den vierten Schritt ausgewählt. Im vierten Schritt findet eine Datenerhebung der übrig gebliebenen Studien statt. Anschließend werden die Daten der Studien analysiert und im letzten Schritt findet eine Interpretation der Ergebnisse statt. [Ke07] Diese Schritte werden im folgenden angewandt.

3.1 Forschungsfragen

Obwohl die Strukturierung der Arbeit durch Faktoren, wie Verfügbarkeit für bestimmte Aufgaben und Verantwortlichkeiten für Produkte oder ganze Prozesse, bereits in BPMS und WfMS Verwendung findet, wird der Faktor Mensch nur sehr abstrahiert betrachtet. Dabei kann die Betrachtung weiterer Faktoren, die mit dem Menschen einhergehen (wie z.B. Präferenzen, Beziehungen, Fertigkeiten oder das Wohlbefinden) für eine Arbeitsverteilung sorgen, welche zum einen die Ressourcen optimal zuordnen kann und zum anderen auch den Mitarbeitern zu Gute kommt. Eine strukturierte Literaturanalyse wird durch Forschungsfragen getrieben, sodass sowohl der Extraktions- als auch der Analyseprozess auf die Beantwortung dieser Fragen gerichtet ist. [Ke07] Ziel dieser Untersuchung liegt darin, den derzeitigen Stand von WfMS und BPMS zu ermitteln, welche menschliche Faktoren betrachten. Von besonderem Interesse ist hierbei die Arbeitsverteilung und die Unterstützung von kooperativem bzw. kollaborativem Arbeiten sowie den Auswirkungen auf das Workflow-Management selbst. Zur Untersuchung von menschlichen Faktoren in WfMS sind demnach die Forschungsfragen in Tabelle 1 zu stellen. Mit FF1 ist festzustellen, welche Forschungsbereiche besonders häufig von den Forschern untersucht wurden. Diese Frage soll auch die Systeme abdecken, die bereits den Menschen als Ressource für die Entscheidungsfindung eingebunden haben und ist bewusst etwas allgemeiner gehalten. Ebenfalls ist die Art der Berücksichtigung menschlicher Faktoren von Interesse. FF2

ermittelt die Forschungspotentiale und Ansätze, die laut der Autoren in Zukunft verfolgt werden sollten. Mit Hilfe dieser Frage kann der aktuelle Forschungsstand und dessen Reifegrad abgeschätzt werden.

FF #	Forschungsfrage
FF1	In welchen Themenbereichen von Workflow-Managementsystemen (WfMS) bzw. Geschäftsprozess-Managementsystemen (BPMS) werden menschliche Faktoren berücksichtigt?
FF2	Welche Themen im Bereich der menschlichen Aspekte in der WfMS bedürfen nach Ansicht der Autoren weitere Forschung?

Tab. 1: Forschungsfragen

3.2 Suchstrategie

Um geeignete Studien für die systematische Literaturrecherche zu finden, wurden verschiedene Schlüsselwörter identifiziert. Dazu wurden auf Grundlage der Forschungsfragen und aus verschiedenen Ausgangsstudien Schlüsselwörter abgeleitet und anschließend verwandte Synonyme gebildet. Die gefundenen Schlüsselwörter wurden nachfolgend durch Testsuchen und Probeüberprüfungen iterativ verfeinert. Tabelle 2 beinhaltet die verwendeten Schlüsselwörter, eingeordnet in abzudeckende Themenbereiche. Das erste Thema umfasst dabei die Basisbegriffe, die verschiedene Bereiche des Workflows beinhalten. Auf diesen Basisbegriffen sollen die zu findenden Studien aufgebaut sein. Das zweite Thema umfasst die Zielgruppe von Personen, die im Workflow beschäftigt sind. Das dritte Thema geht auf die Verteilung der Arbeit der verschiedenen Personen in einem Workflow ein. Das vierte Thema umfasst die persönliche Gestaltung der Arbeit einer Person. Das letzte Thema konzentriert sich auf die weichen Faktoren einer Person, die in einem Workflow arbeitet. Alle Schlüsselwörter eines Bereichs werden mit einem „OR“ Operator verbunden und auf Titel sowie Abstracts angewandt, bei dem fünften Thema werden die Schlüsselwörter zusätzlich im gesamten Dokument gesucht. Jeder der Themenbereiche wurde mit dem „AND“ Operator verknüpft.

Mit der Anwendung des resultierenden Suchstrings in der Literaturdatenbank Scopus² konnten 101 Studien identifiziert werden. Für die Eignungsprüfung der Suchergebnisse wurden folgende Ein- und Ausschlusskriterien angewendet:

Einschlusskriterien:

- Es handelt sich um Studien, die in einer Konferenz, in Workshops, Zeitschriften oder Buchteilen veröffentlicht wurden
- Der Artikel wurde zwischen 2009-2019 geschrieben
- Der Artikel wurde in englischer oder deutscher Sprache geschrieben

² www.scopus.com

Thema	Schlüsselwörter
Basisthema	workflow scheduling, workflow technologies, workflow management, workflow system, dynamic work distribution, process-aware, workflow technology, business process, process management, workflow
Zielgruppe	human, roles, employer, worker, user
Verteilung der Arbeit	competence management, employee selection, occupation, work allocation, work distribution, resource allocation, job crafting, employee reward, human resources perspective, human capacities, human oriented, resource requirements, task execution, task performance, human abilities, human resources perspective, competence management
persönliche Gestaltung der Arbeit	personal preference, human resources, life domain balance, human abilities, work satisfaction, work motivation, work life balance, work flexib*, artificial intelligence, work satisfaction, work motivation
weiche Faktoren	empowerment, competency, ability, strength, preference, fairness, behaviour, health

Tab. 2: Auflistung der Schlüsselwörter

- Artikel enthält einen Ansatz, der menschliche Faktoren berücksichtigt

Ausschlusskriterien:

- Bei den gefundenen Artikeln handelt es sich nicht um einen Volltext
- Der Artikel unterliegt Zugangsbeschränkungen
- Der Artikel bezieht sich bei einer Ressourcenzuweisung nicht auf Menschen, sondern auch auf Maschinen
- Der Artikel beschäftigt sich mit einem Verfahren zur Ermittlung eines optimalen Bewerbers für eine Position im Unternehmen

Nach der Anwendung der Ein- und Ausschlusskriterien auf den Abstract der 101 Studien konnten 19 potenzielle Studien als relevant eingestuft werden. Durch die Analyse des Volltextes dieser Dokumente blieben abschließend acht relevante Studien für das Forschungsgebiet übrig. Um prüfen zu können, ob alle Studien für dieses Forschungsgebiet berücksichtigt wurden, wurde im Anschluss bei den acht verbliebenen Studien eine Rückwärtssuche durchgeführt. Dazu wurden deren Referenzen geprüft und passende Titel für das Forschungsgebiet identifiziert. Anschließend werden die Abstracts der Studien geprüft und bei Übereinstimmung mit den Kriterien ebenfalls der Volltext gesichtet. Nach dem Abschluss der Rückwärtssuche konnte das Studienvolumen auf 25 Studien erhöht werden.

4 Analyse der Literatur

Die 25 als relevant eingestuften Studien werden im Zuge dieses Abschnittes auf eine strukturierte Art und Weise aufbereitet und im Hinblick auf die Beantwortung der Forschungsfragen (siehe Tabelle 1) analysiert. Zunächst wird im folgenden Kapitel jedoch das allgemeine Forschungsgeschehen bzw. die Forschungslandschaft beleuchtet, welches aus dem Suchergebnis abgeleitet werden kann. Im Folgenden wird eine Konzeptmatrix nach Webster & Watson [WW02] verwendet, um die Ergebnisse strukturiert darstellen und analysieren zu können. Weitere Tabellen ergänzen die Konzeptmatrix um zusätzliche Auswertungen. Darauf folgend werden die identifizierten und extrahierten Forschungspotentiale präsentiert.

4.1 Überblick des Forschungsgeschehens

Der Suchraum wurde auf zehn Jahre (2009-2019) eingeschränkt, sodass vor allem aktuelle Lösungen die Ergebnismenge befüllen. Obwohl es sich um einen limitierten Zeitraum handelt, ist ein wachsendes Interesse an der Kombination von WfMS und menschlichen Faktoren zu erkennen. So wurden 2012 drei Dokumente verfasst; im Jahre 2016 waren es bereits sechs und 2018 wurden acht Dokumente veröffentlicht. In genau diesen Jahren zeichnet sich eine Ballung der Veröffentlichungen ab, während zwischen diesen „Hochphasen“ das Interesse kurzzeitig einfiel. Ein weiteres Zeichen für - über den gesamten Zeitraum - steigendes Interesse an der Beachtung menschlicher Faktoren ist auch durch ein erhöhtes Maß an Sponsoren zu begründen. Der Anteil der Studien, denen eine Form der Unterstützung zuteil wurden, beträgt 64%. Offenbar soll die Industrie bei der Schaffung attraktiver und flexibler Arbeitsplätze unterstützt werden, unter anderen durch die „Entwicklung Menschenzentrierter Ansätze“³. Das Interesse an einer menschen-zentrierten Ausrichtung von Systemen ausgedrückt in Fördergeldern ist deutlich in den betreffenden Publikationen zu erkennen. Dieser Umstand begünstigt, dass Dokumente einiger Autoren vermehrt auftreten.

Die 25 Studien können in 23 Primärstudien und zwei Sekundärstudien eingeteilt werden. Die Sekundärstudien umfassen eine Literaturübersicht zur Klassifizierung von Personalressourcen [Ar18b] und eine Literaturübersicht zur Taxonomie von Ressourcenallokationskriterien [AMS18]. Der Inhalt der Primärstudien ist jedoch überwiegend theoretischer Natur. 21 Dokumente (84%) befassen sich mit der theoretischen Arbeit, in denen Modelle, Methoden Frameworks oder Algorithmen entwickelt wurden sowie der Forschungsbereich analysiert und diskutiert wurde. In 14 (56%) Fällen wurde die Theorie anhand von Fallstudien und/oder Experimenten validiert. Zusätzlich konnten einzelne Dokumente gefunden werden, die entweder die Methode einer empirische Studie [Sc19] oder einer Umfrage [Ca16] genutzt haben.

³ https://ec.europa.eu/info/sites/info/files/research_and_innovation/strategy_on_research_and_innovation/documents/ec_rtd_orientations-he-strategic-plan_122019.pdf S.14

4.2 Aufteilung der Konzeptmatrix

Die Beantwortung der FF1 „In welchen Themenbereichen von WfMS bzw. BPMS werden menschliche Faktoren berücksichtigt?“ zielt darauf ab, diejenigen Aspekte der WfMS-Forschung zu ermitteln, die menschliche Faktoren berücksichtigen. Die Datenanalyse soll Schlüsse darauf zulassen können, in welchen Bereichen die Beachtung menschlicher Faktoren am sinnvollsten erscheint. Um die Ergebnisse inhaltlich strukturieren zu können, wurden sie in einer Konzeptmatrix nach Webster & Watson strukturiert. Dies erleichtert ebenfalls den Analyseschritt, sodass z.B. verstärktes Interesse in einem Themenbereich ersichtlich wird. [WW02] In Tabelle 3 sind die Dokumente nach dem Erscheinungsjahr absteigend sortiert. Zusätzlich wird das Hauptthema, mit welchem sich der entsprechende Artikel befasst, in der zweiten Spalte der Konzeptmatrix aufgeführt. Dies ist mit den recht unterschiedlichen Ansätzen zu begründen, die beschrieben wurden. Weitere Bestandteile des Inhalts der Literatur werden in vier Kategorien aufgespalten, von denen jede unterschiedliche Ausprägungen vorweisen kann.

Der thematische Bestandteil System ist in Workflow, Business Prozess und Sonstige gegliedert. Es wird damit eine Auskunft darüber gegeben, welche Art von System in dem jeweiligen Dokument behandelt wurde. Es ist zu sehen, dass nur sieben (28%) Dokumente WfMS behandeln, ganze 18 (72%) dagegen ein BPMS. Ein Grund dafür kann die Beziehung zwischen den beiden Systemen sein, da das WfMS oftmals ein Teil von BPMS ist; beziehungsweise kann ein BPMS die Funktionalität besitzen, einen Workflow zu generieren. Sonstige Systeme sind sieben mal (28%) in den Dokumenten vertreten, dazu gehören z.B. Crowdsourcing-Plattformen. Dieses Anwendungssystem weist einige Unterschiede zum WfMS oder BPMS auf, die zu Grunde liegenden Mechanismen sind dennoch ähnlich. Bei diesem System wird zur Erfüllung einer Aufgabe ein (Bearbeitungs-)Aufruf an ein Netzwerk von potentiellen Bearbeitern gestellt. Diese können daraufhin entscheiden ob sie die Aufgabe annehmen und erledigen. Beim WfMS oder BPMS ist der Bearbeiter oftmals ein Teil des Unternehmens und hat eher eine eingeschränkte Wahlmöglichkeit. Somit kommen weitere Systemkategorien in Frage, um die Behandlung eines Workflows mit menschlichen Faktoren in Einklang zu bringen, insbesondere da die dedizierten WfMS vergleichsweise selten vorkamen.

Der thematische Bestandteil Mensch ist in Individuell und Kollektiv gegliedert. Mit diesem Bestandteil ist zu ermitteln, ob das System kollaborative Unterstützungsmöglichkeiten bietet, oder das Individuum als Einzelnes betrachtet wird. Die Auswertung zeigt, dass 19 (76%) Dokumente nur das Individuum und zehn (40%) eine Gruppe von Menschen betrachten. Ein Grund dafür kann die höhere Komplexität bei der Abbildung von Gruppen sein. Neben den Faktoren der Einzelpersonen müssen zusätzlich bspw. gruppendynamische Effekte mit einbezogen werden.

Der thematische Bestandteil Perspektive ist in Organisation und Mensch gegliedert. Es soll die Frage beantworten, aus welcher Perspektive die Autoren die Thematik betrachten. Entweder aus betriebswirtschaftlicher/organisationaler oder Menschen-zentrierten Sicht.

Die Perspektive der Organisation ist bspw. geprägt von Zeit, Kosten und Leistung. Der Faktor Mensch wird hingegen in Arbeitszeit, Arbeitslast, Fähigkeiten, Erfahrung, usw. gemessen. Dem gegenüber steht eine menschen-zentrierte Perspektive in der bspw., auf die Gesundheit, Emotionen oder das Wohlbefinden Beachtung findet. Die Auswertung vom thematischen Bestandteil Perspektive zeigt, dass 21 (84%) Arbeiten eher eine organisational- und sieben (28%) eine menschen-zentrierte Perspektive einnehmen. Ein möglicher Zusammenhang kann zu den o.g. Sponsoren gefunden werden. Im Allgemeinen sind die Interessengruppen der behandelten Forschung eher Unternehmen, weshalb eine Tendenz zu einer organisational-zentrierten Perspektive besteht. Es sei denn bspw. eine staatliche bzw. überstaatliche Instanz stellt Forschungsmittel für eine „menschen-zentrierte Perspektive“ bereit.

Der thematische Bestandteil Ressource ist in Allokation und Inhalt gegliedert. Der Bestandteil soll die Frage klären, ob sich das Dokument mit der Allokation von Ressourcen oder mit deren inhaltlichen Auseinandersetzung befasst. Mit dieser Aufteilung ist zu ermitteln, ob der Fokus der Publikation eher auf die Mechanismen oder den Entwurf der Verteilung gerichtet ist oder ob diese zur Beschreibung, Konzeption oder Definition von Ressourcen bspw. dem Personal oder der Arbeitsaufgabe beiträgt.

21 (84%) Dokumente richten den Fokus auf die Verteilung von Ressourcen bspw. in Form von Arbeitsaufgaben oder Personal. Acht (32%) Dokumente behandeln die Ressource inhaltlich, bspw. welche Information über eine Ressource benötigt werden um mit dem System einen Nutzen zu generieren.

Dokument	Thema	S			M		P		R	
		Workflow	Business Prozess	Sonstige	Individuell	Kollektiv	Organisation	Mensch	Allokation	Inhalt
[Hs11]	Ganzheitliche Multi-Agenten-Systeme (HMS)	✓		✓		✓	✓		✓	
[KRR12]	Mensch-zentriertes prozessorientiertes Informationssystem (HC-PAIS)	✓			✓	✓		✓		✓
[CRR12]	Verantwortungsmanagement mit RACI-Matrizen		✓		✓		✓		✓	
[Ni12]	Bayes-Selecton-Rule(BSR)-Algorithmus		✓		✓		✓		✓	
[Ca13]	Benutzerpräferenzmodell		✓		✓			✓	✓	✓
[KDS13]	Kompatibilitätsmatrix für Akteure	✓				✓	✓	✓	✓	
[Xu14]	Personaleinsatzplanung mit Klassifikationsmodell	✓					✓		✓	
[VG15]	Dynamische Rollenauflösung		✓		✓		✓		✓	
[Wi15]	Naive-Bayes-Selecton-Rule(NBSR)-Algorithmus		✓		✓		✓		✓	
[Ca16]	Ressourcenzuweisung und -priorisierung mit Studienergebnissen von Crowdsourcing-Plattformen			✓	✓		✓		✓	✓
[SGD16]	Kontextbewusstes kollaboratives Filterungssystem		✓		✓	✓	✓		✓	
[SGK16]	Explorative Prozess-Mining-Techniken				✓	✓	✓		✓	
[UP16]	Spezifische Aufgabenmerkmale; menschliche und soziale Aspekte von Ressourcen		✓		✓	✓	✓	✓	✓	✓
[Wi16]	Zeitnahe und menschenzentrierte Ressourcenzuweisung mit dem Naive-Bayes-Modell		✓		✓		✓		✓	
[Ar16]	Prozess-Mining		✓	✓	✓		✓		✓	
[AMS17]	Ressourcen-Ranglisten		✓			✓	✓		✓	
[Ar18b]	Studie zur Personalallokation und Prozess-Mining	✓	✓	✓	✓	✓	✓	✓	✓	
[Er18]	Spezifizierung der Fleischmann Taxonomie		✓		✓			✓		✓
[AMS18]	Taxonomie von Ressourcenallokationskriterien		✓		✓	✓	✓			✓
[UPV18]	Konzept des psychologischen Profils		✓		✓		✓	✓		✓
[Ta18]	Optimale-Ressourcen-Allokations(ORA)-Modell für kooperative Aufgabenplanung		✓			✓			✓	
[Ar18a]	Rahmenwerk: Multi-Faktoren-Kriterien		✓				✓		✓	
[AAA18b]	Ressourcenbewusste Geschäftsprozess-Simulation (RBPSim)	✓	✓	✓	✓		✓		✓	
[AAA18a]	Ressourcenbewusste Geschäftsprozess-Simulation (RBPSim), Verfeinerung der Beziehungen zwischen den Workflow-Ressourcenmustern	✓	✓	✓	✓		✓		✓	
[Sc19]	Aufgabenempfehlungssystem			✓	✓		✓		✓	✓
Summe:		7	18	7	19	10	21	7	21	8

Tab. 3: Hauptthema und inhaltliche Bestandteile der relevanten Dokumente, kategorisiert nach System (S), Mensch (M), Perspektive (P) und Ressource (R).

4.3 WfMS-Themenbereiche mit Berücksichtigung menschlicher Faktoren

Die Einteilung innerhalb der Konzeptmatrix erreicht eine formelle Strukturierung, der eigentliche Inhalt lässt sich jedoch weiter verfeinern. Tabelle 4 zeigt thematische Verdichtungen, die während der Volltextanalyse aufgetreten sind. Den größten Anteil mit 14 (56%) Studien behandelt die „Zuweisung von Ressourcen“, wobei unterschieden werden kann, ob Aufgaben dem Personal gegeben werden oder ob das Personal Aufgaben zum Bearbeiten erhält. Mit zehn Studien (40%) in diesem Segment beschäftigt sich der Großteil mit der Zuweisung des Personals an bestimmte Aufgaben. Dabei werden verschiedene Konzepte angewendet, wie z.B. die Erstellung eines psychologischen Profils [UPV18] oder die Erstellung eines Benutzerpräferenzmodells [Ca13]. Drei (12%) Studien beschäftigen sich mit der „Empfehlung von Aufgaben“, „Zuordnung von Teams“ oder die „Simulation von Geschäftsprozessen“ unter Beachtung der Ressourcen sind mit jeweils zwei (8%) Studien vorhanden. Im ersten Segment behandeln die Studien eine optimale Kooperation und Kompatibilität zwischen Akteuren, das zweite hingegen die Anreicherung des BPSim-Standards um weitere Ressourcenspezifikationen. Die restlichen Studien verteilen sich thematisch, wobei sich diese grob wie in der Konzeptmatrix (Tabelle 3) anhand der Ressource (Allokation oder Inhalt) und der Perspektive (Organisation oder Mensch) einordnen lassen.

Forschungsaktivität	Dokumente
Zuweisung von Ressourcen	[Ar18b], [KRR12], [Er18], [Ca13], [SGK16], [AMS18], [UPV18], [UP16], [Ta18], [Wi15], [Xu14], [Wi16], [Ni12], [Ar16]
Empfehlung von Aufgaben	[Sc19], [SGD16], [Ar18a]
Zuordnung von Teams	[KDS13], [AMS17]
ressourcenbewusste Geschäftsprozess-Simulation	[AAA18b], [AAA18a]
Ansatz zur Rollenauflösung	[VG15]
Unterstützung des Ressourcenmanagements	[Ca16]
Beschreibung von Arbeitsabläufen	[Hs11]
Verwaltung von Personalverantwortung	[CRR12]

Tab. 4: Forschungsaktivitäten der Dokumente

4.4 Forschungspotentiale

In diesem Abschnitt wird eruiert, welche Themen laut der Autoren weiterer Forschung bedürfen. Tabelle 5 zeigt die zentralen Aussagen der Studien, welche die Autoren über künftige Forschung getroffen haben. Ein Überblick über empfohlene Bestandteile für „Human-Centered Process Aware Information Systems“ (HC-PAIS) wurde bereits von [KRR12] gegeben. Die zukünftigen Forschungsthemen sind dabei unterteilt in „Einbindung des Benutzers in das Design“, „Benutzerunterstützung“, „Vorschriften, Risiko, Verwaltung“, „Datenschutz“, „Darstellung“ und „Auswertung“. Obwohl der Beitrag aus dem Jahr 2012 stammt, ist dieser auch heute noch wegweisend für die weitere Erforschung des Themas. Aus der Gesamtschau der zukünftigen Forschungsthemen kann abgeleitet werden, dass

es weiterer Konzepte bedarf, um den Menschen in einem System abzubilden. Es scheint noch keine Einigkeit darüber zu herrschen, welche Aspekte bei der Verteilung von Personal einen Nutzen generieren. So wird noch in Dokumenten aus dem Jahr 2018 bspw. die Identifizierung der Haupttreiber bzw. der Schlüsselaspekte für zukünftige Forschung im Bereich der Personalallokation als notwendig angegeben [Ar18b]. In einem weiteren Beitrag aus diesem Jahr wird von „speziellen Fähigkeiten“ geschrieben, die benötigt werden um einen klaren Nutzen aus dem System zu ziehen [AMS18]. In einem weiteren Beitrag wird gefordert die Wirksamkeit von Ressourcenzuweisungskriterien zu überprüfen [Ta18]. Werden die Ressourcen als Aufgaben verstanden, so wird aus dem jüngsten Beitrag ein interessanter zukünftiger Forschungsbereich aufgezeigt, um die Probleme bei der Abbildung der Personalressource zu vermeiden. Dort sollen semantische Ähnlichkeitsmerkmale von Aufgabenbeschreibungen genutzt werden, um ein Aufgabempfehlungsschemata zu generieren [Sc19].

Ein weiterer wichtiger Aspekt bei der Betrachtung der zukünftigen Forschungsthemen liegt in dem Reifegrad der Technologie. Die bislang untersuchten Ansätze scheinen noch auf einem einfachen, grundlegenden Niveau zu sein. So wird in [Ca13] und [AMS17] explizit die Unterstützung „komplexer Szenarien“ angesprochen. Obwohl zwischen den Arbeiten mehrere Jahre liegen, scheint das Problem „eine komplexe Anwendung zu unterstützen“ noch gegenwärtig zu sein. Ein Grund für die mangelnde Unterstützung komplexer Szenarien kann in [Ar18b] gefunden werden. Diese ist aus dem Jahr 2018 und nennt den Bedarf an Algorithmen, Werkzeugen und Kriterien zur Bewertung von Ressourcen und Prozessdomänen. Es erweckt den Anschein, als seien grundsätzliche Ansätze zu der Verteilung von Ressourcen in Form von Personal oder Aufgaben erforscht worden. Allerdings scheint der Reifegrad der Technologie noch unzureichend zu sein, um die verschiedenen Anforderungen zur Integration des Faktors Mensch zu berücksichtigen. Erschwerend in dieser Situation ist die Erhebung, der Umgang und die Verarbeitung von personenbezogenen Daten, insbesondere beim Datenschutz. Eine Lösung für sensible Daten kann bspw. Edge-Computing bieten, wo die persönlichen Daten im Besitz des Eigentümers verbleiben und trotzdem verarbeitet werden können. Es wird deutlich, dass die Komplexität von Systemen drastisch steigt, sobald der Faktor Mensch integriert werden soll.

Insgesamt lassen sich die künftigen Forschungsthemen in drei grobe Richtungen einteilen, um bessere Ergebnisse erzielen zu können. Zum ersten besteht der Wunsch nach mehr *Formalisierung*. Somit sollen menschliche Fähigkeiten und Entwicklungsziele [KRR12], die Zusammenarbeit selbst [KDS13], Aufgaben [UP16] oder Kontextinformationen [SGK16] genauer spezifiziert werden, um die Verfügung stehenden Ressourcen genauer verarbeiten zu können. Weiterhin besteht in [AMS18] der Bedarf nach der Formalisierung der Ressourcenzuweisungskriterien. Dies entspricht dem zweiten Forschungspotential, der *Verbesserung und Erweiterung der Bewertungskriterien* für die Zuweisung von (menschlichen) Ressourcen. So sprechen z.B. [Ar16] und [Ar18a] von der Erweiterung des Ressourcenwürfels um weitere Dimensionen. Das dritte und höchste Forschungspotential liegt in der *Verbesserung bzw. Erweiterung des Methodenportfolios* zur Bewertung und Allokation der Ressourcen.

Hervorzuheben ist, dass in [SGD16] eine zu hohe Menge an benötigten Daten bemängelt wurde, um ein passables Ergebnis zu erzielen. Da viele unterschiedliche Ansätze verfolgt wurden, ist zu schließen, dass die Entwicklung eines Standards zur Berücksichtigung menschlicher Faktoren in WfMS noch in entfernter Zukunft liegt. Um dies zu bekräftigen, ist nur in einem Paper von einer tatsächlichen Evaluation der ermittelten Konzepte die Rede [AMS18].

5 Fazit

Das Forschungsthema „Wie der menschliche Faktor in Workflow-Management-Systemen berücksichtigt wird“ unterliegt einem schwankendem, aber dennoch steigendem Forschungsinteresse. Ein Grund für die Schwankungen kann die relativ geringe Anzahl der Forschenden sein. Entsprechende Förderprogramme haben sich jedoch als ein Mittel gezeigt, die Forschung für mensch-zentrierte WfMS zu unterstützen, um die starren Strukturen an die flexibler werdende Arbeitswelt anzupassen. Die in diesem Beitrag durchgeführte Literaturrecherche konnte die Forschungslandschaft kartieren und Erkenntnisse in Bezug auf den aktuellen Forschungsstand generieren sowie Zielrichtungen und Forschungspotentiale identifizieren. Die thematische Ausrichtung von der Mehrheit der Dokumente wurde für den Kontext der BPMS verfasst, ist auf das Individuum bezogen, hat eine organisational-zentrierte Perspektive und behandelt die Verteilung und Zuweisung von Ressourcen. Ebenfalls scheinen sich Crowdsourcing-Plattformen zu eignen, insbesondere was Empfehlungen von individuellen Aufgaben anbelangt; dies weicht von der Idee der zentralen Prozesskontrolle und Ressourcenzuordnung ab und überlässt dem Mitarbeiter mehr Entscheidungsgewalt zur Arbeitsgestaltung. Bereiche der kollaborativen Aufgabenfindung können somit zukunftsweisende Aspekte bieten, um WfMS zu erweitern. Die Literaturanalyse zeigt weiterhin, dass thematische Schwerpunkte in der Ressourcenzuweisung, der Zuordnung von Teams sowie der Empfehlung von Aufgaben vorherrschen. Davon zu unterscheidende Themen werden in der Literatur nur vereinzelt abgedeckt. Eine einheitliche Vorstellung zur Berücksichtigung menschlicher Faktoren existiert somit nicht. Dies schlägt sich ebenso in der Menge unterschiedlicher Algorithmen und Herangehensweisen nieder. Es sei jedoch anzumerken, dass lediglich eine Literaturdatenbank (Scopus) verwendet wurde, wenngleich die Ergebnismenge durch die Anwendung einer Rückwärtssuche deutlich angereichert werden konnte.

Ein offenkundiger Problembereich scheint der geringe Reifegrad der Technologien zu sein, welche die Komplexität möglicher Anwendungsszenarien einschränkt. Dies ist auch damit begründet, dass nur ein Autor eine Evaluation als notwendig ansah. Dieser Eindruck wird durch die dürftige Integration bzw. Abbildung des Faktors „Mensch“ verstärkt. Eine treffende Formalisierung der Ressourcen wäre für bessere Einbindung des Personals ebenfalls unerlässlich. Einige Autoren sehen weiterhin den Kontext und die Situation der Menschen als eine weitere sinnvolle Ergänzung zur Ressourcenallokation [SGK16]. Durch den starken Fokus der gefundenen Literatur auf theoretische Arbeiten und die mangelnde Erprobung in

Dokument	Forschungsthemen
[KRR12]	- Entwicklung von Konzepten menschlicher Fähigkeiten wie z.B. Erfahrung, Fertigkeiten und Qualifikation, um diese dann in das System zu implementieren. - Außerdem sollten auch die Entwicklungsziele der Nutzer berücksichtigt werden.
[CRR12]	- Die Nutzung der RASCI-Matrizen auf das Ressourcenmanagement erweitern.
[Ni12]	Eine dynamische Version der Bayes'schen Auswahlregel (BSR), die es erlaubt in einem laufenden Arbeitsprozess die Ressourcen zu wechseln. Bspw. bei umsatzstarken Ressourcen die ab einem bestimmten Zeitpunkt im Prozess nicht mehr notwendig bzw. in einem anderen Prozess dringender benötigt werden.
[Ca13]	Das Ressourcenmodell soll erweitert werden, um komplexere Fälle abzudecken.
[KDS13]	Es bedarf einer genaueren Modellierung der Zusammenarbeit, insbesondere der beeinflussenden Faktoren.
[Xu14]	Die halbautomatische Methode der Workflow-Personalzuweisung soll durch die Berücksichtigung der Beziehungen zwischen den Zusweisungsprozeduren und Wichtungparameter der Beziehungen verbessert werden.
[Wi15]	Es soll der Ressourcentransfer zwischen Aktivitäten berücksichtigt werden.
[Ca16]	Eine breitere Klassifizierung von Crowdsourcing-Plattformen wird angestrebt.
[SGD16]	Die speicherbasierte- soll durch eine modellbasierte-kollaborative Filterung ersetzt werden, die robuster auch mit wenig verfügbaren Daten arbeitet.
[SGK16]	Es wird eine Taxonomie für Prozessausführungsprotokolle gefordert, um die Ableitung von Kontextinformationen zu verbessern.
[UP16]	Es sollen die Aufgaben (individuell oder kollaborativ) jedes organisatorischen Prozesses charakterisiert werden, um menschliche Ressourcen zu finden, die in der Lage sind diese auszuführen.
[Wi16]	Es soll der Ressourcentransfer zwischen Aktivitäten berücksichtigt werden.
[Ar16]	Es sollen neue Dimensionen in den Ressourcenwürfel aufgenommen werden, um die Analyse zu verbessern.
[AMS17]	- Es sollen komplexere Szenarien unterstützt werden. - Die Bestimmung der Teams soll mittels der Erstellung von Heuristiken umgesetzt werden, um eine schnellere und günstige Berechnung umzusetzen.
[Ar18b]	- Bei der Personalallokation im BPM ist Forschung zur Identifizierung der Haupttreiber bzw. der Schlüsselaspekte notwendig. - Außerdem Algorithmen, Werkzeuge und Kriterien zu Bewertung der Ressourcen und Prozessdomänen.
[Er18]	- Die Risiken welche mit den Aufgaben verbunden sind, sollen berücksichtigt werden. - Neben allgemeinen Faktoren wie Fähigkeiten, Erfahrungen, Präferenzen und Berechtigungen, sollen insbesondere spezielle Fähigkeiten aufgenommen werden, um den Nutzen des Systems zu gewährleisten.
[AMS18]	- Es gilt die Wirksamkeit der Ressourcenzuweisungskriterien zu überprüfen. - Formalisierung der Ressourcenzuweisungskriterien.
[Ta18]	Es wird die Planung der Ressourcenallokation für kooperative Aufgaben aus der Perspektive der Ressourcenauslastung angestrebt.
[Ar18a]	- Einführung neuer Dimensionen in den Ressourcenwürfel. - Die Einbeziehung der Kontrollfluss-Perspektive.
[Sc19]	Mittels semantischer Ähnlichkeitsmerkmale aus der Aufgabenbeschreibung, sollen Aufgabenempfehlungsschemata erstellt werden.

Tab. 5: Ausblick und künftige Forschungsthemen der Studien.

der Praxis, begründet sich ein weiteres Forschungspotential im Bereich der in der Praxis eingesetzten WfMS und dessen Evaluation. Schlussendlich befindet sich die Arbeitswelt im Wandel zu einer deutlich flexibleren Umgebung, sodass auch die Arbeitsunterstützung flexibler spezifiziert werden sollte. Zur Beachtung weicher Faktoren sind erste Konzepte bekannt, eine Ausarbeitung dieser ist jedoch notwendig.

Literatur

- [AAA18a] Afifi, N.; Awad, A.; Abdelsalam, H. M.: Extending BPSim based on workflow resource patterns. 2018.
- [AAA18b] Afifi, N.; Awad, A.; Abdelsalam, H. M.: RBPSim: A resource-aware extension of BPSim using workflow resource patterns. In: CEUR Workshop Proceedings. Bd. 2072, S. 32–39, 2018.
- [AMS17] Arias, M.; Munoz-Gama, J.; Sepúlveda, M.: A multi-criteria approach for team recommendation. 2017.
- [AMS18] Arias, M.; Munoz-Gama, J.; Sepúlveda, M.: Towards a taxonomy of human resource allocation criteria. 2018.
- [Ar16] Arias, M.; Rojas, E.; Munoz-Gama, J.; Sepúlveda, M.: A framework for recommending resource allocation based on process mining. 2016.
- [Ar18a] Arias, M.; Munoz-Gama, J.; Sepúlveda, M.; Miranda, J. C.: Human resource allocation or recommendation based on multi-factor criteria in on-demand and batch scenarios. *European Journal of Industrial Engineering* 12/3, S. 364–404, 2018.
- [Ar18b] Arias, M.; Saavedra, R.; Marques, M. R.; Munoz-Gama, J.; Sepúlveda, M.: Human resource allocation in business process management and process mining: A systematic mapping study. *Management Decision* 56/2, S. 376–405, 2018.
- [Ca13] Cabanillas, C.; García, J. M.; Resinas, M.; Ruiz, D.; Mendling, J.; Ruiz-Cortés, A.: Priority-based human resource allocation in business processes. 2013.
- [Ca16] Cabanillas, C.: Exploring human resource management in crowdsourcing platforms. 2016.
- [CRR12] Cabanillas, C.; Resinas, M.; Ruiz-Cortés, A.: Automated resource assignment in BPMN models using RACI matrices. 2012.
- [Er18] Erasmus, J.; Vanderfeesten, I.; Traganos, K.; Jie-A-Looi, X.; Kleingeld, A.; Grefen, P.: A method to enable ability-based human resource allocation in business process management systems. 2018.

- [FLW20] Fellmann, M.; Lambusch, F.; Weigelt, O.: Human Energy in Organizations: Theoretical Foundations and IT-Based Assessment. In (Harris, D.; Li, W.-C., Hrsg.): Engineering Psychology and Cognitive Ergonomics. Mental Workload, Human Physiology, and Human Energy. Springer International Publishing, Cham, S. 127–140, 2020.
- [GM01] Green, F.; McIntosh, S.: The intensification of work in Europe. Labour Economics 8/2, S. 291–308, 2001, ISSN: 09275371.
- [Hs11] Hsieh, F. .-.: Collaborative workflow management in holonic multi-agent systems. 2011.
- [KDS13] Kumar, A.; Dijkman, R.; Song, M.: Optimal resource assignment in workflows for maximizing cooperation. 2013.
- [Ke07] Keele, S. et al.: Guidelines for performing systematic literature reviews in software engineering, Techn. Ber., Technical report, Ver. 2.3 EBSE Technical Report. EBSE, 2007.
- [KRR12] Kabicher-Fuchs, S.; Recker, J.; Rinderle-Ma, S.: Workshop on Human-Centric Process-Aware Information Systems: HC-PAIS. 2012.
- [Ni12] Nisafani, A.; Wibisono, A.; Kim, S.; Bae, H.: Bayesian Selection Rule for Human-Resource Selection in Business Process Management Systems. The Journal of Society for e-Business Studies 17/, S. 53–74, 2012.
- [Re03] Reijers, H. A.: Design and Control of Workflow Processes: Business Process Management for the Service Industry. Springer, Berlin und Heidelberg, 2003, ISBN: 9783540366157.
- [Sc19] Schnitzer, S.; Neitzel, S.; Schmidt, S.; Rensing, C.: Results of a survey about the perceived task similarities in micro task crowdsourcing systems. 2019.
- [SGD16] Sindhgatta, R.; Ghose, A.; Dam, H. K.: Context-aware recommendation of task allocations in service systems. 2016.
- [SGK16] Sindhgatta, R.; Ghose, A.; Khanh, H.: Context-Aware analysis of past process executions to aid resource allocation decisions. 2016.
- [Ta18] Tan, W.; Zhao, L.; Xie, N.; Tang, A.; Hu, X.; Tang, S.: Methods for optimal resource allocation on cooperative task scheduling in cross-organizational business process. 2018.
- [UP16] Uahi, R.; Pereira, J. L.: Task allocation in Business Processes supported by BPMS: Optimization perspectives. In: Iberian Conference on Information Systems and Technologies, CISTI. Bd. 2016-July, 2016.
- [UPV18] Uahi, R.; Pereira, J. L.; Varajão, J.: Improving work allocation practices in business processes supported by BPMS. 2018.
- [VG15] Vanderfeesten, I.; Grefen, P.: Advanced dynamic role resolution in business processes. 2015.

- [Wi15] Wibisono, A.; Nisafani, A. S.; Bae, H.; Park, Y. .-.: On-the-fly performance-aware human resource allocation in the business process management systems environment using Naïve Bayes. 2015.
- [Wi16] Wibisono, A.; Nisafani, A. S.; Bae, H.; Park, Y. .-.: A dynamic and human-centric resource allocation for managing business process execution. *International Journal of Industrial Engineering : Theory Applications and Practice* 23/4, S. 270–282, 2016.
- [WW02] Webster, J.; Watson, R. T.: Analyzing the past to prepare for the future: Writing a literature review. *MIS quarterly*/, S. xiii–xxiii, 2002.
- [Xu14] Xu, R.; Liu, X.; Xie, Y.; Yuan, D.; Yang, Y.: A Gaussian fields based mining method for semi-automating staff assignment in workflow application. In: *ACM International Conference Proceeding Series*. S. 178–182, 2014.

Workshop: Digitale Kompetenz, Digital Literacy, Digital Skills

Digitale Mündigkeit fördern in IT-nahen Studiengängen

Entwicklung von Learning Outcomes am Beispiel eines Studiengangs

Dorina Gumm¹, Sonja Beer², Stephan Schumacher³

Abstract: In einem Projekt an der TH Lübeck untersuchen wir, was die Förderung Digitaler Kompetenzen im Sinne einer *Digitalen Mündigkeit* für unsere Informatik- und IT-nahen Studiengänge konkret bedeuten und wie eine Verankerung in den Ausbildungsprozess erfolgen kann. In diesem Paper legen wir zunächst verschiedene Perspektiven auf Digitale Kompetenzen dar, begründen, welchem Verständnis wir in dem vorgestellten Projekt folgen, und beschreiben bisherige Problemstellungen für die Umsetzung in Studiengängen. Unter Berücksichtigung des Kontextes legen wir dann das Vorgehen des Projektes dar und diskutieren, welche Erkenntnisse zur Förderung Digitaler Mündigkeit abgeleitet werden können und welche Bedeutung diese Erkenntnisse für die Konzeptbildung zu Digitalen Kompetenzen haben können.

Keywords: Digitale Mündigkeit, Future Skills, IT-nahe Studiengänge, Curriculumsentwicklung

1 Digitale Kompetenzen: Überblick und Positionierung

Die Förderung „digitaler Kompetenzen“ in Schulen und Hochschulen ist schon lange in Wissenschaft und Bildungspolitik ein Thema, und ihre Diskussion hat durch die Coronapandemie 2020 nochmals einen intensiven Schub erhalten. Im Rahmen der „Beauftragung für Digitalisierung und Lehre“ am Fachbereich „Elektrotechnik und Informatik“ wurden verschiedene Handlungsfelder identifiziert, um die Lehre im Kontext von Digitalisierung zu unterstützen. Einer der damit verbundenen Aufgabenbereiche ist es, *Digitale Mündigkeit* der Studierenden aktiv zu fördern, u.a. durch eine strukturierte Verankerung im Curriculum, um diese Förderung nicht dem Zufall oder dem sporadischen Interesse einzelner Lehrender zu überlassen⁴.

Dies bedeutet, über die „allgemeinen Kompetenzrichtziele“ zur Umsetzung des Europäischen Qualifikationsrahmens [Sc12, S. 6] und über die im KMK-Kompetenzrahmen abstrakt beschriebenen Anforderungen hinaus [KMK16, S. 15;42], studiengangspezifische Kompetenzprofile auszdifferenzieren und in Form von Learning Outcomes zu konkretisieren.

¹ Technische Hochschule Lübeck, FB Elektrotechnik und Informatik, Mönkhofer Weg 239, 23562 Lübeck, dorina.gumm@th-luebeck.de

² Gleiche Institution, FB Maschinenbau und Wirtschaft, sonja.beer@th-luebeck.de

³ Gleiche Institution, FB Elektrotechnik und Informatik, stephan.schumacher@th-luebeck.de

⁴ Das Projekt wurde auch für das Dachprogramm „LehreN“ 2020 ausgewählt und gefördert [Gu20].

Zielsetzung für unser Projekt ist es, ganzheitlich neue, durch Digitalisierung erforderliche oder gewünschte Kompetenzen und Lernziele zu adressieren, in unseren Studiengängen systematisch zu fördern und im Curriculum zu verankern. Eine solchermaßen kompetenzorientierte Studiengangentwicklung ist von hoher Relevanz, da sich durch die fortschreitende Digitalisierung der Arbeitswelt aktuell neue bzw. zu antizipierende Anforderungsbereiche für Akademiker:innen ergeben. Digitale Kompetenzen können unmittelbar die Employability unserer Absolvent:innen betreffen. Daher gilt es, diese neuen Kompetenzanforderungen fach- und berufsfeldspezifisch zu beschreiben und für die Lehr-Lern-Gestaltung konkrete Learning Outcomes abzuleiten, vgl. [Sc12, S.22].

Bisherige Definitionen digitaler Kompetenzen eignen sich nur bedingt, um ein spezifisches Kompetenzprofil für einen Studiengang zu erfassen. Der Begriff „Digitale Kompetenzen“ kann auf sehr unterschiedliche Fähigkeiten, Wissensinhalte, Reflexions- und Bewertungsgegenstände und Kontexte verweisen: Beispielsweise geht es um das Anwendungs-Know-How bezüglich der Kommunikationsmedien (z.B. „Medienhandeln“ [HSB20]), um die Nutzbarmachung digitaler Technologien für konkrete Wirtschaftszweige [Ko20], um Datenkompetenz [LT20], um die Risiken gewisser Medien und Anwendungen bzgl. der Sicherheit [Mo19, RS20], um die Nutzung digitaler Lernplattformen und digitaler Formate für Lernmaterialien [Pe20, Pf16] oder auch um allgemeine „Future Skills“ wie Sense-Making oder adaptives Denken [DDG20, Se20]. Neben diesen Schwerpunkten ist auch zunehmend das verantwortliche Handeln mit digitalen Werkzeugen bzw. im digitalen Raum in den Fokus gerückt; dies drückt sich in Kompetenzbeschreibungen wie „Digitale Souveränität“ [KHP17, GI20], „Digitale Mündigkeit“ [Be18, Ho19] oder „Orientierungskompetenz“ [Ro18] aus. Der Begriff „Digitale Mündigkeit“ adressiert sowohl Bereiche der technischen als auch der sozialen Kompetenz bzw. Future Skills (nach [Be18] sind das Technical, Privacy, Information, Social und Civic Literacy) mit dem Fokus auf Nutzung, Gestaltung und Wirkung digitaler Technologien im gesellschaftlichen Kontext. Damit erfasst dieser Begriff weitaus umfänglicher als andere, in einer digitalen Umwelt zu interagieren, digitale Werkzeuge konstruktiv, selbstbestimmt und unabhängig (souverän) nutzen zu können oder auch gesellschaftliche Wirkung zu erzielen. Hier werden also gerade auch motivationale Orientierungen, (Wert-)Haltungen, Selbstreflexion des Handelns und Entscheidens sowie „anschlussfähige Kommunikation“ im Sinne des gesellschaftlichen Diskurses einbezogen. Dieses „Bündel“ von Einstellungen, Fähigkeiten, Wissensbestandteilen usw. entspricht damit der Auffassung von Schaper [Sc12] und Ferrari [Fe13] für akademische Kompetenzen, wie sie von Hochschulen vermittelt werden sollen.

Unter den genannten Begriffen eignet sich der Begriff der *Digitalen Mündigkeit* für unser Vorhaben am besten, als er es ermöglicht, ein ganzheitliches Verständnis für die Wechselwirkungen zwischen Technologien und deren Gestaltungs- und Wirkungsmacht zu entwickeln. Er kann damit soziale, rechtliche, ethische, politische als auch ökonomische Dimensionen digitaler Innovationen in den Blick nehmen und so eine kritische Reflexion digitaler Technologien bezüglich unterschiedlicher Einsatzbereiche fördern (insbesondere i.S.v. Chancen und verantwortungsvoller Gestaltung, nicht nur bzgl. Risiken). Allerdings

ist er mit dieser Umfänglichkeit nicht ohne weiteres für die Lehrentwicklung operationalisierbar.

Im Hinblick auf eine kompetenzorientierte Studiengangentwicklung sind daher die spezifischen Anforderungsbereiche zu berücksichtigen, in denen unsere Absolvent:innen ihre *Digitale Mündigkeit* benötigen werden, um nicht nur Lösungen hoher bzw. wissenschaftlicher Qualität zu entwickeln, sondern auch befähigt werden zu „anschlussfähiger Kommunikation“ über entsprechende Lösungen sowie zur „Selbstregulation und Reflexion des eigenen problemlösungs- und erkenntnisgeleiteten Handelns“ [Sc12, S. 29].

2 Problembeschreibung

Typische Informatik- und IT-nahe Curricula an Fachhochschulen – mit ihrem typischen Fokus auf die Ausbildung für die berufliche Praxis – adressieren soziale, gesellschaftliche oder gar gesellschaftspolitische Reflexionen und Folgenabschätzungen von Technologien meist nur am Rande, in Wahlpflichtfächern oder je nach Interesse der jeweiligen Lehrenden in einzelnen Modulen. Während die Vorteile bzgl. Medienkompetenz, Data Literacy oder auch Sicherheitskompetenz für den Berufsalltag auf der Hand liegen, scheinen Vorteile der einstellungs- und gesellschaftsbezogenen Kompetenzbereiche (kritische Beurteilung, Technikfolgenabschätzung, netzpolitische Fragestellungen) noch nicht so bewusst und werden als Querschnittsthema lieber Schlüsselkompetenzen zugeordnet, so dass die Entwicklung einer diesbezüglichen Employability und Teilhabe eher den Studierenden überantwortet wird, vgl. [Sc12, S. 21]. Dabei gibt es bereits Ansätze für die Informatik-Ausbildung, beispielsweise die Forderung, „folgenantizipierendes Design“ in die Informatik-Ausbildung zu integrieren [Ro18].

Die Übersetzung von Digitalkompetenzen oder Digitaler Mündigkeit in konkrete Learning Outcomes und didaktische Szenarien – insbesondere für Kontexte, in denen von recht hohen Fachkenntnissen bezüglich digitaler Technologien ausgegangen werden kann – wird in der bisherigen Literatur kaum beschrieben. Zum einen ist in vielen Publikationen zu beobachten, dass die geforderten Kompetenzen auf nicht-technisch ausgebildete Personen zielen (z.B. Schüler:innen, auf Erwachsene ohne oder mit wenig technischer Vorbildung oder auf (hauptsächlich) nicht-technische Studiengänge wie bei [BG18, Da21, Me19, KHP17, LT20]). Zum anderen werden die konkreten Lernziele bspw. nicht weiter ausgeführt [Se20], bleiben allgemein [Eh19]; es geht um studienbezogene Digitalkompetenz bezüglich der Lernplattformen (z.B. der Onboarding-Prozess für Studierende bei [Li20]); die Vorschläge sind auf zu spezifische Einzelbeispiele bezogen (z.B. [Ro18]) und oft bleibt unklar, wie die Kompetenzen und Lernziele tatsächlich in Hochschul-Curricula integriert werden sollen [Eh19, Se20].

Eine konsequente, modulübergreifende Verankerung im Curriculum steht bisher der Auffassung gegenüber, Schlüsselkompetenzen ließen sich isoliert vermitteln, bspw. in einem gesonderten (Wahl-)Modul. Daher könnte die Idee entstehen, den „Schlüsselkompetenz-

Anteil“ der Digitalen Mündigkeit auszugliedern und ohne fachspezifischen Bezug zu präsentieren. Ein solches Vorgehen wäre aber aus lernpsychologischer Sicht wenig nachhaltig [Sc12, S. 21]. Für eine übergreifende Konzeption spricht auch, dass die Verwertungsinteressen der Unternehmen mitbetrachtet werden sollten [Ro18], um dem Ziel von Hochschulbildung nachzukommen, Studierende den Eintritt in Berufs- und Handlungsfelder zu ermöglichen, also ihre „Employability“ zu fördern.

3 Pilotprojekt für den Studiengang „EKS“

Aus der vorgestellten Herausforderung bezüglich digitaler Kompetenzen ergaben sich für den Fachbereich folgende Fragestellungen und Zielsetzungen:

1. Was bedeutet die Förderung von *Digitaler Mündigkeit* in Informatik- und IT-nahen Studiengängen? Ziel ist es, diejenigen Kompetenzen und Lernziele zu identifizieren, die sich auf die Digitalisierung insgesamt beziehen, aber über die Kernfachkompetenzen der genannten Studiengänge hinausgehen, nämlich auch i.S.v. Einstellungen, Werthaltungen und Persönlichkeitsbildung, welche ein selbstreflektiertes, verantwortliches und gesellschaftsbezogenes Handeln betreffen können.
2. Wie kann eine entsprechende kompetenzorientierte Weiterentwicklung in Fach- und Technischen Hochschulen erfolgen? Ziel ist es, als Best-Practice-Beispiel eine Umsetzungsstrategie für die Implementierung des Querschnittsthemas *Digitale Mündigkeit* in Curricula zu erarbeiten und in ersten Schritten zu erproben.

Da eine studiengangübergreifende Entwicklung im Rahmen dieses Projekts nicht durchführbar war, wurde der Studiengang „Elektrotechnik und Kommunikationssysteme“ (EKS) ausgewählt, um diese Fragen explizit für einen Studiengang zu untersuchen.

3.1 Vorgehen

Die Erarbeitung eines gemeinsamen Verständnisses von Kompetenzprofilen und abgeleiteten Learning Outcomes erfordert ein behutsames Einbeziehen des Kollegiums sowie eine intensive Kommunikation [Sc12, S. 38]. Die Verankerung von Querschnittsthemen im Curriculum unterliegt vielen Rahmenbedingungen: Curricula sind bereits festgelegt und jahrelang erprobt, sodass neue Impulse mit vorhandenen Inhalten und Zeitplanungen konkurrieren. Änderungen des Curriculums sind mit größerem Aufwand verbunden, sowohl strukturell als auch für die inhaltliche Lehr-/Lern-Gestaltung. Die zeitlichen Ressourcen sind sowohl auf Seiten der Studierenden als auch der Lehrenden sorgfältig ge- und verplant und bieten daher geringe Flexibilität für Änderungen, sodass Änderungsideen sehr sorgsam geprüft werden müssen, bevor sie umgesetzt werden. Dabei sind auch organisationskulturelle Gegebenheiten zu berücksichtigen, um in der Hochschule Akzeptanz und Mitgestaltungswillen zu erzeugen [Eu16]. Auch sind hochschulinterne Gremien und viele Einzelpersonen mit unterschiedlichen Rollen zu berücksichtigen. Innovationen

können nicht „von oben“ eingeführt werden, Innovator:innen müssen daher mit der Autonomie der Beteiligten umgehen und „stimulierende Umgebungen“ schaffen [Br16]. Aus diesen Gründen haben wir für dieses Projekt einen partizipativen Ansatz [vU14] gewählt. Das Vorgehen bestand aus folgenden Schritten:

1. Erfassung von Lerninhalten und Kompetenzbereichen

Da unsere Studiengänge in digitalen Technologien ausbilden, muss bei der Betrachtung von *Digitaler Mündigkeit* von diesem Wissensstand ausgegangen werden. Ausgangspunkt war die Frage, welche zusätzlichen Kompetenzen für Absolvent:innen durch die zunehmende Digitalisierung insgesamt relevant geworden sind, im Studium aber noch nicht oder kaum adressiert werden. Dazu wurden alle Lehrenden des Fachbereichs eingeladen; teilgenommen hat etwa ein Dutzend (ca 30%). In einem Workshop wurden Vorschläge für Themen und grundlegende Lerninhalte bzw. Lernziele gesammelt und vorläufig geclustert. Dabei zeigte sich, dass schon für die Studiengänge innerhalb eines Fachbereichs von unterschiedlichen Voraussetzungen ausgegangen werden muss, da Studierende jeweils etwas andere Vorkenntnisse mitbringen und etwas andere technische Grundkenntnisse vermittelt bekommen. Im zweiten Teil wurden erste Ideen zur Verankerung solcher Themen diskutiert. Im Anschluss erfolgte eine Auswertung und Strukturierung der Ergebnisse sowie eine Spiegelung an vorhandener Literatur, um eine geeignete zusammenfassende Begrifflichkeit zu finden. Einzelne Themen aus dem Workshop wurden in Learning Outcomes umformuliert (immer im Hinblick darauf, dass es um Lernergebnisse geht, die nicht zu den primären, ohnehin in den IT-nahen Studiengängen vermittelten Lernergebnissen gehören). Der erarbeitete Gesamtvorschlag wurde in weiteren unterschiedlichen Gesprächskontexten mit Lehrenden thematisiert, diskutiert und angepasst, sodass auch Kolleg:innen, die nicht am o.g. Workshop teilgenommen haben, sich mit ihren Sichtweisen einbringen konnten.

2. Bestandsaufnahme im Pilotstudiengang

Für den Studiengang EKS wurden 10 Module identifiziert, für die mit den jeweiligen Lehrenden separate Workshops durchgeführt wurden, mit dem Ziel zu erkennen, a) inwieweit das Modul bereits Aspekte von Digitaler Mündigkeit fördert bzw. b) in welchen Bereichen es Handlungsansätze und -bedarfe gibt, sowie c) welche organisatorischen Handlungsmöglichkeiten favorisiert werden. Für jedes Modul wurde ein strukturierter Steckbrief erstellt, in welchem mögliche Themen zur Förderung der Digitalen Mündigkeit sowie deren Verankerungsstatus (von erprobt bis zur Idee) sowie Verankerungsformen festgehalten sind. Die Steckbriefe wurden in Anlehnung an die Delphi-Methode in mehreren Feedbackschleifen mit den Lehrenden erstellt. Die Workshops wurden auch genutzt, um unser Verständnis von Digitaler Mündigkeit und damit verbundenen Lernergebnissen zu hinterfragen und zu schärfen, sowie um die Lehrenden für diese kompetenzorientierte Lehre zu sensibilisieren, ihnen Impulse für individuelle bzw. modulspezifische Bezüge zu dem Kompetenzprofil zu geben und diese im weiteren Verlauf berücksichtigen zu können. Dadurch konnten wiederum wiederum weitere Lehrende in den Prozess einbezogen werden, die im Schritt 1 noch keine Gelegenheit hatten, sich einzubringen.

3.2 Umgang mit unterschiedlichen Perspektiven und Interessen

Mit dem Ziel, die Förderung digitaler Mündigkeit im Studium zu verankern, sind eine Reihe von Veränderungen verbunden, die sich – je nach Ausgestaltung des Prozesses – auf einzelne Module, das Curriculum eines Studiengangs oder auf Formalien wie Modulhandbücher und Studienordnungen beziehen können. Widerstände im Kollegium sind daher naheliegend: Lehrende könnten z.B. befürchten, mit zusätzlicher Arbeitslast konfrontiert zu werden, durch zusätzliche Regelvorgaben in ihrer Gestaltungsfreiheit eingeschränkt zu werden oder dass ihre eigenen Sichtweisen auf ihre Lehre, Module oder generell auf die Ziele der Ausbildung in Frage gestellt werden. Ein reflektierter Umgang mit Widerständen sollte daher in einem solchen Vorhaben hohe Aufmerksamkeit erhalten [Eu16]. Im Rahmen dieses Projektes wurde dies insbesondere durch eine „einladende Zielsetzung“ und eine wertschätzende Gestaltung der Workshops erreicht:

Einladende Zielsetzung

Digitale Mündigkeit zu fördern erfordert sowohl inhaltliche als auch didaktische Ideen, die nicht bei allen Kolleg:innen vorausgesetzt werden können. Außerdem können sich auch nicht alle Lehrenden mit dieser Thematik insoweit identifizieren, dass sie sich diese als Bestandteil ihrer eigenen Lehre vorstellen können. Zudem besteht die Zielsetzung des Projekts zunächst darin, im Kollegium eine Vorstellung vom Kompetenzprofil zu entwickeln; Lösungen für eine didaktische Umsetzung müssen sich also nicht von vornherein an konkreten Evaluationszielen bzw. Kriterien der Kompetenzmessung bewähren.

In den Workshops ging es daher vor allem darum, die Lehrenden zum Austausch einzuladen und ihre konkreten Vorstellungen kennenzulernen, um diese im Prozess berücksichtigen zu können. Es ging explizit nicht darum, Vollständigkeit bzgl. aller Module oder aller Learning Outcomes zu erzielen, alle Lehrenden zu überzeugen oder ihnen Vorgaben für ihre Lehre vorzusetzen. Stattdessen sollten einige „Innovator:innen“ [Eu16] gefunden werden, die an der Umsetzung interessiert sind, und „Zaudernden“ auch ihren berechtigten Raum gelassen werden.

Wertschätzende Gestaltung der Workshops

Digitale Mündigkeit umfasst ein breites Spektrum an Zielen, Kompetenzen und Lernergebnissen. Es ist nicht zu erwarten, dass selbst die interessiertesten Lehrenden für das ganze Spektrum ein gleichermaßen hohes Interesse mitbringen oder dass konkrete Lernergebnisse zur eigenen fachlichen Spezialisierung bzw. zu ihren Modulhalten passen. Um hier Widerständen bzw. auch Missverständnissen bzgl. der Zielsetzung des Pilotprojekts vorzubeugen, wurden die Beteiligten im Sinne der einladenden Zielsetzung als Expert:innen ihrer eigenen Module und für ihre Lehrgestaltung wahrgenommen und angesprochen.

Für die Annäherung an und Sensibilisierung für die Thematik wurden sie dazu ermuntert, ihre eigene „Beziehung“ zur Thematik zu formulieren, nämlich inwiefern sie dem bisher

erarbeiteten Verständnis von Digitaler Mündigkeit und den Learning Outcomes folgen können und ob sie Ideen zur Ergänzung oder Änderung haben – letztere wurden jeweils direkt aufgenommen und integriert. Auch ihre eigene Motivation und der eigene Fokus auf Digitale Mündigkeit standen im Fokus, und so ging es vor allem darum, ihre Module und ihre eigenen Ideen zu reflektieren, statt auf ein vorgegebenes Regelwerk zu bestehen.

3.3 Zwischenergebnisse

Im bisherigen Projektverlauf konnten bereits folgende Ergebnisse erzielt werden:

- Gemeinsames Verständnis eines Kompetenzprofils für *Digitale Mündigkeit* und für zugehörige Learning Outcomes.
- Identifikation von bereits vorhandenen Ansätzen zur Kompetenzvermittlung sowie von Potenzialen in Bezug auf *Digitale Mündigkeit*.
- Workshop-Leitfaden mit Begleitmaterial, der darauf zielt, *Digitale Mündigkeit* mit konkreten Learning Outcomes zu verbinden und im Diskurs darüber Lehrende für die kompetenzorientierte Weiterentwicklung ihrer Module zu gewinnen.
- Steckbriefvorlage zur systematischen Darstellung der Workshop-Ergebnisse sowie eine Matrix-Vorlage für eine modulübergreifende Ergebnisdarstellung.

4 Zentrale Erkenntnisse und Diskussion

In diesem Abschnitt reflektieren wir insbesondere jene Projekt-Ergebnisse, die sich auf unsere Auffassung von *Digitaler Mündigkeit* sowie die identifizierten Anforderungsbereiche beziehen. Sie sind im Überblick in Abbildung 1 dargestellt.

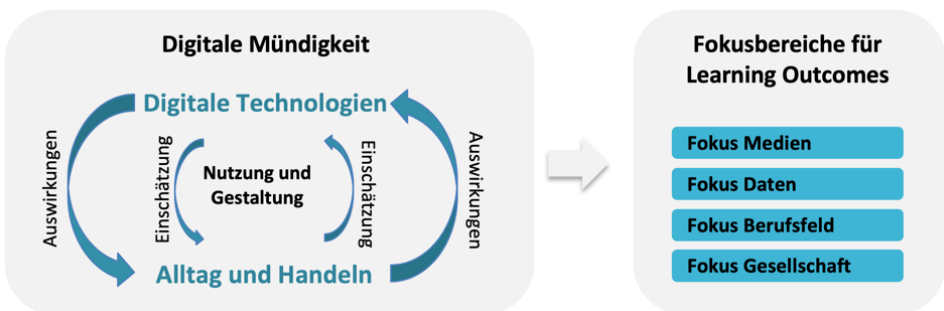


Abbildung 1: Digitale Mündigkeit: Verständnis und Learning Outcomes

4.1 Auffassung von Digitaler Mündigkeit

In den durchgeführten Workshops zeigten sich Spannungsfelder, die auch die allgemeinen öffentlichen Diskussionen zur Digitalisierung widerspiegeln: Einerseits bezieht sich das, was mit neuen Inhalten erreicht werden soll, vor allem darauf, digitale Technologien zu reflektieren und in den sozialen Kontext zu stellen. Andererseits war es vielen Teilnehmer:innen wichtig, ein Ohnmachtsgefühl oder gar eine Technik-Distanziertheit zu vermeiden, und somit die Risiken, die sich durch digitale Technologien ergeben könnten, nicht zu stark in den Vordergrund zu stellen. Einerseits wurden die gesellschaftlichen Aspekte hervorgehoben, andererseits sollten vor allem berufspraktisch verwertbare Fähigkeiten im Fokus stehen. Auch die soziale Verantwortung wurde herausgestellt, die bei Personen mit besonders tiefen Kenntnissen und Gestaltungsmöglichkeiten in Bezug auf digitale Technologien erwartet werden – mithin jene, die wir ausbilden.

Aus diesen unterschiedlichen Sichtweisen wurde eine Auffassung von *Digitaler Mündigkeit* entwickelt, die diese Sichtweisen als Gesamtkonzept statt als Gegensätze aufgreift:

Digitale Mündigkeit bedeutet in der Lage zu sein, die aktuelle **digital-technologische** Situation und zukünftige Entwicklungen bezüglich ihrer **Chancen**, aber auch ihrer **Risiken** sowie ihrer **Auswirkungen** auf unseren (**gesamtgemeinschaftlichen**) **Alltag** oder unser **Handeln** einzuschätzen und diese verantwortungsvoll für sich **und andere** zu nutzen und zu gestalten.

Die Begriffe *Chancen* und *Risiken* verdeutlichen, dass mit digitalen Technologien allgemein positive oder negative Effekte hervorgerufen werden können. Der Begriff *Auswirkungen* zielt darauf, dass Veränderungen möglicherweise erst einmal neutral oder je nach Sichtweise oder Betroffenheit unterschiedlich eingeschätzt werden können. Digitale Technologien wirken sich auf unseren *individuellen Alltag* aus, rufen aber auch *gesamt-gesellschaftliche* Veränderungen hervor. Weiterhin sind sie nicht nur Elemente unserer Umgebung, sondern wirken auf *unser eigenes Handeln*. Digitale Mündigkeit bedeutet also, all diese Facetten der Veränderungen und Kontextebenen in den Blick nehmen zu können mit dem Ziel, digitale Technologien vor diesem Hintergrund *einschätzen* (d.h. reflektieren und einordnen) zu können. Aus dieser Reflexion heraus ist ein *verantwortungsvolles Handeln* möglich und notwendig bei der *Nutzung* und *Gestaltung* digitaler Technologien.

Damit umfasst diese Definition nach unserer Auffassung des Kompetenzbegriffs im Kontext digitaler Technologien neben Fachwissen und anderen kognitiven Aspekten eben auch Einstellungen, Motivation, Volition und personale, soziale und kommunikative Schlüsselkompetenzen [Sc12, S. 13]. Diese sind jeweils studiengangspezifisch zu reflektieren, in Form von Kompetenzprofilen zu beschreiben und durch geeignete Lern-/Lehr-Formate zu vermitteln.

4.2 Fokusbereiche für Learning Outcomes im IT-Fachkontext

Aus den in den Workshops gesammelten Vorschlägen für Lerninhalte und Lernziele wurden 15 konkrete Learning Outcomes entwickelt. Sie stellen aus Sicht der beteiligten Lehrpersonen eine gute Arbeits- und Diskussionsgrundlage dar. Da diese Learning Outcomes aus der jeweiligen Fachperspektive heraus entstanden sind, sind sie im Sinne eines studienangangsspezifischen Kompetenzprofils wesentlich spezifischer und weniger umfangreich als generische, allgemeine Kompetenzbeschreibungen, etwa „Medien-, Datenkompetenz“ oder „Future Skills“. Insofern sind die identifizierten Learning Outcomes auch nicht deckungsgleich mit allgemeineren Kompetenzstrukturmodellen wie z.B. bei [Be18]. Stattdessen ergaben sich aus einer Clustering der Workshop-Ergebnisse vier Fokusbereiche, in die sich die fach- bzw. berufsspezifischen Anforderungsfelder einordnen lassen. Sie stellen den konkret identifizierten Handlungsbedarf für den Studiengang dar.

Fokus Medien

Handlungsbedarf im Fokusbereich Medien wird insbesondere bezüglich Kommunikation, Arbeitsorganisation und Informationsbeurteilung gesehen. Es geht weniger um die Technical Literacy, welche in unserem Kontext als sehr hoch angesehen wird: Studierende kennen eine große Auswahl an digitalen Kommunikationsmitteln und können diese sicher bedienen. Die Learning Outcomes beziehen sich stattdessen vor allem auf die Beurteilung digitaler Kommunikationsmittel für konkrete Arbeitskontexte (Social Literacy bzgl. dieser Tools) oder auf die Beurteilung von Informationen bzgl. ihres Verbreitungsmediums und der Angemessenheit für den fachlichen Kontext (Information Literacy).

Fokus Daten

Handlungsbedarf im Fokusbereich Daten wird insbesondere gesehen bzgl. ihres Auftretens, ihrer Nutzbarkeit und damit verbundener Auswirkungen, ihrer Plausibilität sowie bezüglich mit Daten verbundener Fragen zu Privatsphäre, Sicherheit und Selbstschutz. Zweifelsohne liegt hier ein fachlicher Kern von Informatik- und IT-nahen Studiengängen. Durch die steigende gesellschaftliche Relevanz von IT für die Gesellschaft wird aber auch eine steigende Verantwortung für IT-Gestalter:innen wahrgenommen und ein möglichst tiefes Verständnis eingefordert. Im Sinne der *Digitalen Mündigkeit* geht es hier also um vertiefte Technical Literacy, fokussiert auf die genannten Fragestellungen – und damit weit über das hinaus, was z.B. bei [Be18] unter Technical Literacy diskutiert wird.

Fokus Berufsfeld

Handlungsbedarf im Fokusbereich Berufsfeld wird v.a. gesehen in Bezug auf die Veränderungen von Berufsbildern und Arbeitsfeldern sowie von Marktmechanismen und Geschäftsmodellen, welche durch digitale Technologien bzw. deren rasante Weiterentwicklung bewirkt werden. Die Relevanz für die Employability unserer Absolvent:innen wird darin gesehen, dass nicht nur die Anforderungen an die Fachkompetenz besonders volatil geworden sind, sondern auch die gesellschaftliche Verantwortung der Gestalter:innen dieser Technologien im öffentlichen Diskurs einen größeren Stellenwert einnimmt. Dieser

Fokusbereich wird in der Literatur noch kaum thematisiert (Ausnahme ist hier z.B. [Ro18]), auch wenn sich Überschneidungen zur Civic Literacy erkennen lassen.

Fokus Gesellschaft

Handlungsbedarf im Fokusbereich Gesellschaft wird insbesondere gesehen bzgl. der Reflexion digital-gesellschaftlicher oder digital-politischer Themen aus technischer und gesellschaftlicher Perspektive, sowie bezüglich der eigenen technischen und damit auch gesellschaftlichen Gestaltungsspielräume. Dies wird vor dem Hintergrund betrachtet, dass die (Weiter-)Entwicklung digitaler Technologien von gesellschaftlichen bzw. politischen Entwicklungen genauso beeinflusst wird wie umgekehrt gesellschaftliche und politische Entwicklungen durch digitale Technologien (siehe dazu auch [Ro18]). Dieser Fokusbereich weist große Ähnlichkeiten zur Civic Literacy auf, soll hier aber konkrete Beispiele von Wechselwirkungen sowie Beurteilungskriterien und Leitlinien in den Blick nehmen und v.a. auch die Gestaltungskompetenz aus diesem Blickwinkel stärken.

5 Zusammenfassung und Ausblick

In diesem Beitrag haben wir gezeigt, wie eine konkrete Interpretation von *Digitaler Mündigkeit* im Kontext von Informatik- und IT-nahen Studiengängen aussehen kann. Um die fachspezifische Wirklichkeit der Lehrenden in die Ausgestaltung dessen zu integrieren, was „Förderung der *Digitalen Mündigkeit*“ im konkreten Studiengang bedeuten kann, wurde ein partizipatives Vorgehen gewählt. Im Pilotstudiengang zeigte sich, dass diese erarbeitete Auffassung bis auf wenige anzupassende Details übernommen werden konnte. Es ist daher anzunehmen, dass sie auch anderen IT-nahen Studiengängen als Ausgangspunkt zur Entwicklung studiengangspezifischer Kompetenzprofile dienen kann.

Ferner zeigte sich in der Diskussion über Verankerungsmöglichkeiten, dass die mit *Digitaler Mündigkeit* verbundenen Learning Outcomes sehr unterschiedlich bzgl. ihrer (Berufs-)Relevanz eingeschätzt wurden. Da Fach- und Technische Hochschulen eine sehr praxisnahe und berufsorientierte Ausbildung verfolgen, sollte in Zusammenarbeit mit Unternehmen geprüft werden, inwieweit sich *Digitale Mündigkeit* als direkter Kompetenzvorteil für Absolvent:innen bei der Jobsuche entwickeln kann. Eine Integration der unternehmerischen Perspektive in den Entwicklungsprozess ermöglicht die praxisorientierte Ausrichtung der identifizierten Learning Outcomes und kann dadurch der kompetenzorientierten Weiterentwicklung des Curriculums einen Schub verleihen.

6 Literaturverzeichnis

- [Be18] Beck, R. et.al.: Digitale Mündigkeit. Eine Analyse der Fähigkeiten der Bürger in Deutschland zum konstruktiven und souveränen Umgang mit digitalen Räumen. Forschungsbericht, Nationales E-Government Kompetenzzentrum, TUM, 2018. https://www.researchgate.net/publication/325756267_Digitale_Mundigkeit_Eine_Analyse_der_Fahigkeiten_der_Burger_in_Deutschland_zum_konstruktiven_und_souveranen_Umgang_mit_digitalen_Raumen, Stand: 23.02.2021
- [BG18] Braun, M; Gumm, D.: Schulen in eine digitalisierte Welt leiten: Vorschläge des Projektes „Chaos macht Schule“ zur Gestaltung von Schulentwicklung. In: Schulverwaltung Spezial, Ausgabe 04/2018.
- [Br16] Braun, S.; Peus, C.; Frey, D.; Knipfler, K.: Leadership in Academia: Individual and collective approaches to the quest of creativity and innovation. Emerald Group Publishing Limited, Bingley, 2016, pp. 349-365.
- [Da21] DATAx, Data Literacy Education, Projekt der Leuphana Universität Lüneburg, <https://www.leuphana.de/universitaet/entwicklung/lehre/projekte/datax.html>, Stand: 22.07.2021
- [DDG20] Davies, A.; Devin, F.; Gorbis, M.: Future Work Skills 2020. Institute for the Future for the University of Phoenix Research Institute. https://www.iftf.org/uploads/media/SR-1382A_UPRI_future_work_skills_sm.pdf, Stand: 24.2.2021
- [Eh19] Ehlers, U.-D.: Future Skills und Hochschulbildung „Future Skill Readiness“. In (Hafer, J.; Mauch, M.; Schumann, M. Hrsg.): Teilhabe in der digitalen Bildungswelt. 2019, Medien der Wissenschaft, Band 75. Waxmann, Münster; New York. S. 37-48
- [Eu16] Euler, D.: Gestaltung von Veränderungsprozessen im Rahmen der pädagogischen Hochschulentwicklung. In (Brahm, T. et al. Hrsg.): Pädagogische Hochschulentwicklung, Springer Fachmedien Wiesbaden, 2016.
- [Fe13] Ferrari, A.: DIGCOMP: A framework for developing and understanding digital competence in Europe. Publications Office of the European Union, 2013,
- [GI20] GI: Arbeitspapier Schlüsselaspekte Digitaler Souveränität, 2020. https://gi.de/fileadmin/GI/Allgemein/PDF/Arbeitspapier_Digitale_Souveraenitaet.pdf, Stand: 24.02.21.
- [Gu20] Gumm, D.: Digitale Mündigkeit aktiv fördern. Projektbeschreibung im Dachprogramm Lehre^N, Jahrgang 2020, https://lehrehochn.de/wp-content/uploads/2019/12/projekt_gumm.pdf, Stand: 23.02.2021.
- [Ho19] Hoffmann, C. P. et.al.: Dimensionen digitaler Mündigkeit und politische Beteiligung im Netz. In (Engelmann, I.; Legrand, M.; Marzinkowski H. Hrsg.): Politische Partizipation im Medienwandel, Berlin, 2019 S. 79-99.
- [HSB20] Hofhues, S.; Schiefner-Rohs, M.; Brahm, T. (Hrsg.): Studierende – Medien – Universität. Einblicke in studentische Medienwelten. Waxmann Verlag, Münster New York, 2020.
- [KHP17] Kranich, L.; Hauth, P; Pols, A.: Kompetenzen für eine Digitale Souveränität. Studie im Auftrag des Bundesministeriums für Wirtschaft und Energie (BMWi), Berlin, 2017.

- [KMK16] KMK: Strategie der Kultusministerkonferenz „Bildung in der digitalen Welt“. Dezember 2016. https://www.kmk.org/fileadmin/Dateien/pdf/PresseUndAktuelles/2017/Strategie_neu_2017_datum_1.pdf, Stand: 15.7.2021
- [Ko20] Kollmann, T.: Die Digital Skills. In: Digital Leadership: Grundlagen der Unternehmensführung in der Digitalen Wirtschaft. Springer Fachmedien, Wiesbaden.
- [Li20] Limpinsel, I. L. et.al.: All-in-One – Hochschulkontexte studierendengerecht und digital gestalten durch einen ganzheitlichen Onboarding-Prozess. In: [HSB20], S. 229-235
- [LT20] Ludwig, T.; Thiemann, H.: Datenkompetenz – Data Literacy. In: Informatik Spektrum (2020) 43:436–439.
- [Me19] Meyer-Guckel, V. et.al.: Future Skills: Strategische Potentiale für Hochschulen. Diskussionspapier 3 vom Stifterverband für die Deutsche Wissenschaft e.V. https://dspace.library.uu.nl/bitstream/handle/1874/380523/future_skills_strategische_potenziale_fuer_hochschulen_5_.pdf, Stand: 26.02.2021.
- [Mo19] Morisco, R. M.: Medienkompetenz und IT-Sicherheit. In (David et.al. Hrsg): INFORMATIK 2019, Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2019 631-642.
- [Pe20] Persike, M.: Blended Learning als Spielfeld für Learning Analytics und Educational Data Mining. In: (Friedrichsen M., Wersig W. Hrsg.): Digitale Kompetenz. Synapsen im digitalen Informations- und Kommunikationsnetzwerk. Springer Gabler, Wiesbaden, 2020, S. 111-128.
- [Pf16] Pfau, W. et.al. Hrsg.: Teaching Trends 2016. Digitalisierung in der Hochschule: Mehr Vielfalt in der Lehre. Münster ; New York : Waxmann 2016.
- [Ro18] Rolf, A.: Digitale Bildung für Informatiker. In: Informatik Spektrum (2018) 41:250–256.
- [RS20] Regitz, C., Schaar-Goldapp, A.: Digitale Souveränität in Zeiten einer Pandemie. In: GI20, S. 20-21.
- [Sc12] Schaper, N. et.al.: Fachgutachten zur Kompetenzorientierung in Studium und Lehre. HRK Projekt nexus, 2012.
- [Se20] Seidel, T.: Hochschul(aus)bildung im Zeitalter der Digitalisierung – Ziele und Kompetenzanforderungen. In: Friedrichsen M., Wersig W. (Hrsg.) Digitale Kompetenz. Synapsen im digitalen Informations- und Kommunikationsnetzwerk. Springer Gabler, Wiesbaden, 2020, S. 135-141.
- [vU14] von Unger H.: Partizipative Forschung. Springer VS, Wiesbaden, 2014.

EPINetz: Exploration of Political Information Networks

John Ziegler¹, Alexander Brand², Julian Freyberg¹, Tim König², Wolf Schünemann²,
Marina Walther¹, Michael Gertz¹

Abstract: Different societal challenges, such as information overload, emerge due to the digital transformation of the media landscape. This also demands new competencies from citizens that often lack the means to contextualize arguments or actors and to understand their interrelationships in complex topics. The EPINetz project is an approach to bridge the outlined skills gap by developing an appropriate political information system. It provides access to political news collected from multiple data sources, including social media, and offers various network exploration capabilities. Different entities such as political actors or topics are extracted from collected data and shown within their respective contexts modelled as weighted and time-varying information networks. Thereby, interested citizens and especially schoolchildren can discover current political topics and understand relationships between relevant entities.

Keywords: Digital Literacy; Information Networks; Political Information Systems

1 Introduction

How to approach and convey complex topics of social and political importance? How to enable citizens ranging from schoolchildren to adults to develop and improve their media competence and digital literacy? In an increasingly complex media landscape, many citizens lack the appropriate skills.

The EPINetz project aims to develop a Web-based platform that allows users to explore political information networks. Rather than focusing only on a few select media outlets, the platform integrates data from various publicly accessible (German) sources, including Twitter and news outlets. The project has the following objectives: (1) Integrate politically relevant information from different sources and make it accessible through a unified information retrieval system. (2) Extract and visualize entities such as actors (e.g., politicians) and topics, as well as their relationships in a temporal-sensitive way. By applying NLP-based information extraction methods, collected documents are mapped to interactively explorable information networks where nodes represent entities and edges time-varying and weighted relationships. This differs significantly from traditional search engines where users are simply provided with a list of documents barely offering any contextualized view.

¹ Heidelberg University, Institute of Computer Science, Im Neuenheimer Feld 205, 69120 Heidelberg, Germany; ziegler@informatik.uni-heidelberg.de, freyberg@stud.uni-heidelberg.de, m.walther@stud.uni-heidelberg.de, gertz@informatik.uni-heidelberg.de

² Hildesheim University, Institute of Social Sciences, Universitätsplatz 1, 31141 Hildesheim, Germany; alexander.brand@uni-hildesheim.de, tim.koenig@uni-hildesheim.de, wolf.schuenemann@uni-hildesheim.de

The remainder of this paper is structured as follows: In Section 2, we give a summary of related projects, contrast them with the EPINetz approach, and outline relevant concepts of digital literacy. This is succeeded by a description of the EPINetz platform in Section 3, along with some sample user stories. Finally, we conclude the paper with a summary and outlook on the project's roadmap in Section 4.

2 Background

In this section, we first give an overview of related work, contrast the same with the approach taken in the EPINetz project and, secondly, outline our understanding of digital literacy.

2.1 Related Work

To focus on the core functionality of the platform, we narrow down the scope of related work to the topic of political information systems. In the past, different projects have aimed at building similar tools. These approaches differ from EPINetz in several ways. First and foremost, the Media Cloud [MC21] project offers a platform for general-purpose media analysis. Despite its diverse and very sophisticated capabilities, such as topic mapping and source management, it does not offer any of the network-based exploration functionality provided by the EPINetz project. This also applies to the Vox Civitas tool [Di11] which was designed to analyze social media content around broadcast events. Similarly, the European Media Monitor [EMM21] provides basic statistics, trending topics and activity detection to the end-user, but again lacks sufficient exploration functionality. Furthermore, information from social media is not taken into account. This is also true for the TopExNet [SAG19] tool, even though it offers network-based exploration capabilities for entity relationships extracted from news articles. Probably most similar to our work is the LeadLine system [Do12]. It offers a visual analytics system for events extracted from news articles and social media. In contrast, EPINetz focuses on the less general domain of political information and more specifically targets the German media landscape.

2.2 Digital Literacy

EPINetz aims to improve both general and domain-specific digital literacy. As to our general conception of digital literacy, there is no canonical terminology to build on, but rather a great variety of concepts being used in the field [DD09]. The concept of literacy is generally oriented towards comprehensive understandings of citizen education and participation in the digital age. First, our concept emphasizes the informational skills that individual users need to develop and apply a critical understanding of their digital information ecosystems and lifeworlds. Thus, we can build on Jones-Kavalier's and Flannigan's definition of digital

literacy as “a person’s ability to perform tasks effectively in a digital environment [...] Literacy includes the ability to read and interpret media (text, sound, images), to reproduce data and images through digital manipulation, and to evaluate and apply new knowledge gained from digital environments.” [JF06] Beyond this fundamental concept, recent developments in digitalization need to be reflected. Therefore, we include skills and evaluative capacities concerning datafication, algorithmic filtering or machine learning that have been termed data literacy elsewhere. [PS20] Beyond processing and presenting information in context, EPINetz allows a ‘look behind the scenes’ of datafication by providing opportunities to observe and practice data-scientific methods. All in all, our basic conception is compatible with strategies and frameworks issued by governance actors such as the European Union [Vu16] and the German Conference of Education Ministers [KMK16] as it takes up key components defined in those frameworks, mostly covering the competence area “information and data literacy”, but also citizen engagement and a critical understanding of media in a digital world. As to our domain-specific orientation, it is important to add that in contrast to many other studies and projects on politics and policy using computational social science methodology, we do not operate in an exclusively data-driven way. EPINetz instead accounts for the pre-structuration of policy fields that resonates in public debates even when mediated in digital environments.

3 EPINetz Platform

In the following, we describe the main components of the EPINetz platform. In Section 3.1, we give a description of the different data sources and outline the data processing in Section 3.2. Finally, in Section 3.3, we detail the construction of the different types of information networks for text analysis and exploration purposes. Potential user stories are given in Section 3.4. Figure 1 serves as an overview of how the different components interact with each other.

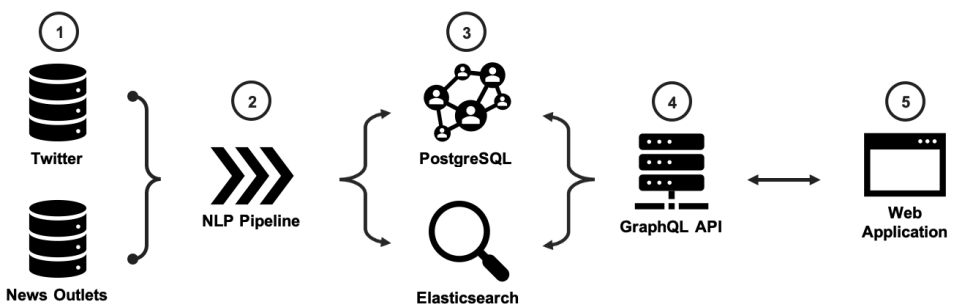


Fig. 1: Simplified overview of the architecture of the EPINetz platform.

1. Data is collected from multiple sources, such as Twitter and news outlets. 2. Textual information of the data is processed by the NLP pipeline. 3. Data is stored in PostgreSQL for network generation and Elasticsearch for information retrieval. 4. A GraphQL API is

used as a unifying data access layer. 5. Information can be explored by the end-user via the Web application.

3.1 Data Sources

A primary objective of the EPINetz platform is to provide users access to political topics and actors from different media sources. For this, we resort to social media, more precisely Twitter, as well as close to 100 major German news outlets, e.g., Spiegel, FAZ, Bild. News article collection is restricted to the general topics of politics, economy, finance, and society. Furthermore, the collected Twitter data is based on a curated list of Twitter accounts of close to 2,000 German politicians and political organizations. Twitter data is collected through the new Twitter API³, providing access to both historic information and live data. Currently, our data pool contains around 12 million tweets and 8 million news articles.

We try to avoid filter bubbles by providing data on parliamentarians and news outlets across the political spectrum. Through network exploration and information retrieval capabilities the user can have a look at multiple and potentially diverse aspects of a political debate. It is important to note that already biased media coverage might be propagated to the platform as we take the raw data without any filtering or sampling.

3.2 Data Processing

Both news articles and tweets are processed via dedicated NLP pipelines, primarily using spaCy⁴, for named-entity extraction (e.g., persons, organizations, locations). For tweets, hashtags and user mentions are also extracted. By now, we refer to topics as densely connected sets of keywords and actors as occurring in our information networks (see Section 3.3). As illustrated in Figure 1, the processed data is managed via two systems, a PostgreSQL database used as the basis for the construction of our information networks and Elasticsearch for different information retrieval tasks.

To distinguish between different policy fields, e.g., public health and migration, we develop "policy parsing" as a novel method for the identification of such fields. This involves a mixture of supervised and unsupervised learning techniques, namely a pre-informed, enriched keyword-based grouping in the tagged dossiers of the Bundeszentrale für politische Bildung (Federal Agency for Civic Education) and clustering on node-embedded representations on the aforementioned information networks.

³ Twitter API v2: Early Access, <https://developer.twitter.com/en/docs/twitter-api/early-access>, retrieved 28-06-2021

⁴ spaCy · Industrial-strength Natural Language Processing in Python, <https://spacy.io/>, retrieved 28-06-2021

3.3 Information Networks and Timelines

With fine-grained information about text data accessible in PostgreSQL, different types of information networks and timelines can be constructed. Some of these networks are built as soon as new documents come into the system while other networks are constructed on the fly in response to user requests issued via the platform interface. Such networks show, for example, the Twitter users mentioned by a user, the hashtags used by a user, or co-occurrences of hashtags, named entities, and keywords. To reflect the temporal dimension, all relationships are assigned time-encoding attributes, enabling the analysis and contrasting of entities and relationships in a time-sensitive manner.

3.4 User Stories

A fictitious user could use the EPINetz platform to keep up with ongoing political debates. By providing an aggregated view based on data from different sources, the end-user will less likely suffer from information overload. Statistical information and timelines of the collected data might serve as potential entry points to currently important topics. Once the user decided on a subset of entities, network exploration capabilities allow to grasp relationships in an intuitive manner (see Figure 2b). Usage scenarios could then be as simple as exploring topics and arguments over time, or more complex, e.g., contrasting actors and arguments on different media outlets.

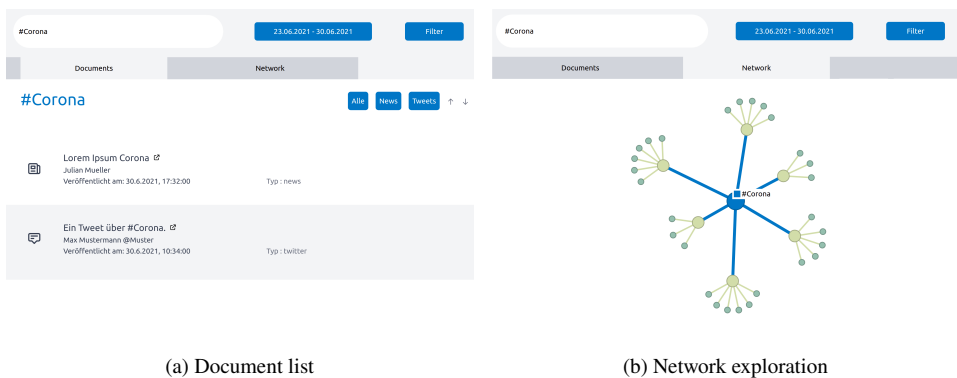


Fig. 2: Information can be retrieved as a list of documents or explored as a network.

A prototypical example of an information network based on the hashtag “#Corona” is shown in Figure 2b. Adjacent nodes might be co-occurring hashtags or Twitter users employing them frequently. Figure 2a shows a corresponding fictional list of documents.

4 Conclusion and Ongoing Work

With the steady increase of news outlets and channels in the media landscape, it will become more and more difficult for the public and citizens to deal with the amount and complexity of information, especially aspects related to political topics and debates. In the EPINetz project, we aim to address this problem by providing users with a Web-based platform that allows them to search, explore, and contrast information surrounding political actors, topics, and debates at different scales, with data uniformly integrated from diverse sources, including social media. In this paper, we have outlined how network-like structures provide a suitable means to explore actors and topics in context and over time while informing users from where and how the information has been obtained. Underlying EPINetz is a large collection of German news outlets as well as comprehensive Twitter datasets related to German politicians and organisations.

We are currently realizing the first front-end features for the platform, tailored to use cases suitable for a wide range of users, including schoolchildren, to explore the data. In particular, together with partner institutions from the education sector, we design more specific use cases and user stories as well as evaluation and quality criteria, the latter targeted towards schools with a focus on digital literacy in the landscape of political topics and debates.

Acknowledgements

We thank the Klaus Tschira Stiftung gemeinnützige GmbH for funding EPINetz. The project's website can be found at www.epinetz.de.

References

- [DD09] van Deursen, A. J.; van Dijk, J. A.: Using the Internet: Skill related problems in users' online behavior. *Interacting with Computers* 21/5-6, pp. 393–402, June 2009.
- [Di11] Diakopoulos, N.; Naaman, M.; Yazdani, T.; Kivran-Swaine, F.: Social media visual analytics for events. In: *Social Media Modeling and Computing*. Springer, pp. 189–209, 2011.
- [Do12] Dou, W.; Wang, X.; Skau, D.; Ribarsky, W.; Zhou, M. X.: LeadLine: Interactive visual analysis of text data through event identification and exploration. In: *2012 IEEE Conference on Visual Analytics Science and Technology (VAST)*. Pp. 93–102, 2012.
- [EMM21] European Media Monitor, retrieved 23-06-2021, 2021, URL: <https://emm.newsbrief.eu/overview.html>.

-
- [JF06] Jones-Kavalier, B.; Flannigan, S. L.: Connecting the Digital Dots: Literacy of the 21st Century, <https://er.educause.edu/articles/2006/1/connecting-the-digital-dots-literacy-of-the-21st-century>, retrieved 30-06-2021, 2006.
- [KMK16] KMK, Kultusministerkonferenz: Bildung in der digitalen Welt: Strategie der Kultusministerkonferenz, https://www.kmk.org/fileadmin/pdf/PresseUndAktuelles/2018/Digitalstrategie_2017_mit_Weiterbildung.pdf, retrieved 28-07-2021, 2016.
- [MC21] Media Cloud, retrieved 23-06-2021, 2021, URL: <https://mediacloud.org/>.
- [PS20] Pangrazio, L.; Sefton-Green, J.: The social utility of ‘data literacy’. *Learning, Media and Technology* 45/2, pp. 208–220, 2020.
- [SAG19] Spitz, A.; Almasian, S.; Gertz, M.: TopExNet: Entity-Centric Network Topic Exploration in News Streams. In: *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining. WSDM '19*, Association for Computing Machinery, Melbourne VIC, Australia, pp. 798–801, 2019.
- [Vu16] Vuorikari, R.; Punie, Y.; Carretero, S.; Van den Brande, L.: *DigComp 2.0: The Digital Competence Framework for Citizens: Update Phase 1: the Conceptual Reference Model*, <https://ec.europa.eu/jrc/en/printpdf/150698>, retrieved 28-07-2021, 2016.

Grundkonzepte des Maschinellen Lernens für die Grundschule – Algorithmen, Biases, Generalisierungsfehler

Ute Schmid¹ Anja Gärtig-Daug² Linda Müller³ Alexander Werner⁴

Abstract: Da digitale Technologien zunehmend mehr Lebens- und Arbeitsbereiche durchdringen, gewinnt die Förderung digitaler Kompetenzen bereits ab der Grundschule an Bedeutung. Durch entsprechende Bildungsangebote sollen Kinder beim Erschließen, Verstehen und Mitgestalten der digitalen Welt unterstützt werden. Hierbei ist wichtig, nicht nur digitale Handlungskompetenzen auf der Anwendungsebene zu fördern, sondern Kinder beim Verstehen der zugrundeliegenden Funktionsprinzipien zu unterstützen. Ein grundlegendes Verständnis der algorithmischen Konzepte, die den digitalen Anwendungen zugrundeliegen ist hierfür zentral. Zunehmend beinhalten digitale Anwendungen Komponenten, die auf Ansätzen der Künstlichen Intelligenz (KI) basieren. Kinder begegnen solchen Systemen etwa in Form von intelligenten Sprachassistenten bei Fernsehern oder beim Entsperren von Smartphones mittels Gesichtserkennung. Damit das kindliche Verständnis von KI-Technologien gefördert werden kann, muss das komplexe Thema für Kinder in didaktisch reduzierter Form erfahrbar gemacht werden. Wir haben hierzu ein Spiel entwickelt, mit dem Kinder grundlegende Konzepte des maschinellen Lernens wie regelbasierte Algorithmen (Entscheidungsbaumverfahren) und neuronale Netze (Perzeptron) spielerisch erkunden können. Neben einem Grundverständnis von Lernalgorithmen werden Kinder für kritische Aspekte wie Biases und Generalisierungsfehler sensibilisiert. Durch Rückbezug des Spiels auf KI-Algorithmen im Alltag sollen die Kinder zur Reflektion der Möglichkeiten und Grenzen von KI angeregt werden.

Keywords: Künstliche Intelligenz; Maschinelles Lernen; Bias; Generalisierungsfehler; Primarbereich

1 Einleitung

Künstliche Intelligenz (KI) erfährt seit einigen Jahren viel Aufmerksamkeit. Das lang etablierte Forschungsgebiet der Angewandten Informatik befasst sich mit der Entwicklung von Algorithmen zur Lösung von Problemen, die Menschen im Moment noch besser lösen können [Ri83]. KI-Methoden werden immer dann notwendig, wenn ein Problem nicht mit Standardalgorithmen der Informatik lösbar ist. Das ist insbesondere dann der Fall, wenn Probleme so komplex sind, dass deren vollständige Bearbeitung nicht effizient möglich ist, oder wenn das Problem und seine Lösung nicht genau beschreibbar sind. Für den ersten Fall werden heuristische Verfahren der Suche und Inferenz eingesetzt, im zweiten Fall kommt

¹ Universität Bamberg, Kognitive Systeme und Forschungsgruppe Elementarinformatik (FELI), An der Weberei 5, 96045 Bamberg, ute.schmid@uni-bamberg.de

² Universität Bamberg, FELI und Zentrum für Lehrerinnen- und Lehrerbildung Bamberg (ZLB): Kompetenzzentrum Digitales Lehren und Lernen, anja.gaertig-daug@uni-bamberg.de

³ Universität Bamberg, FELI, linda-margareta.mueller@stud.uni-bamberg.de

⁴ Universität Bamberg, FELI und LS Empirische Bildungsforschung, alexander.werner@uni-bamberg.de

maschinelles Lernen zum Einsatz. Einen Überblick über Methoden und Themengebiete der KI gibt zum Beispiel [GSB20].

Infolge der zunehmenden Durchdringung aller Lebens- und Arbeitsbereiche mit digitalen Technologien wird es als Aufgabe der Schule erachtet, Kinder beim Erschließen, Verstehen und Mitgestalten der digitalen Welt zu unterstützen. Für einen souveränen Umgang mit digitalen Medien wird es als wichtig erachtet, nicht nur – oft produktbezogene – Bedien- und Anwendungskompetenzen zu fördern, sondern Kindern ein Verständnis von informatischen Konzepten zu ermöglichen, die den Anwendungen zu Grunde liegen [Dö16; GWS19]. Längst haben intelligente Spielsachen und smarte Geräte Einzug in die kindliche Lebenswelt genommen. Kinder begegnen dem Thema KI in Medien und im Alltag auf unterschiedliche Art, etwa in Form von Roboterhunden, die auf Befehle hören und unterschiedliche Personen erkennen, intelligenten Sprachassistenten bei Fernsehern oder beim Entsperren von Smartphones mittels Gesichtserkennung. Kinder zeigen großes Interesse an diesen Funktionen und haben Spaß daran, die Grenzen dieser Systeme auszutesten. Für sie (wie vermutlich auch für die meisten Jugendlichen und Erwachsenen) ist jedoch nicht ersichtlich, warum ihr Roboterhund nur manchen Befehlen folgt, sie auf eine Anfrage eine sinnvolle Antwort erhalten, auf eine andere hingegen nicht, oder warum das eigene Smartphone auch entsperrt wird, wenn der Bruder in die Kamera blickt, ohne dass seine biometrischen Daten in den Einstellungen hinterlegt wurden.

Die Begegnung mit KI-Technologien beeinflusst die Vorstellung von Kindern zu Eigenschaften, Möglichkeiten und Kontrollierbarkeit von KI [BC08; VTV20]. Je mehr Erfahrung Kinder mit KI-Technologien sammeln und umso intensiver sie sich mit den zugrunde liegenden Algorithmen auseinandersetzen können, desto realistischer schätzen sie smarte Geräte ein. So können sie eher zwischen menschlicher und künstlicher Intelligenz unterscheiden [BC08; WPB19]. Damit das kindliche Verständnis von KI gefördert werden kann, muss das komplexe Thema geeignet aufbereitet und erfahrbar gemacht werden. Um Kindern bereits im Vor- und Grundschulalter den Aufbau richtiger mentaler Modelle zu ermöglichen und Misskonzeptionen zu vermeiden, eignen sich Spiel- und Erfahrungsmaterialien, die einen Blick in die Blackbox KI ermöglichen sowie zum Austausch und zur Reflektion der Erfahrungen innerhalb der Klasse sowie der Familie anregen [Dr18; SLR19].

Angesichts der vielfältigen Ansätze und Forschungsrichtungen innerhalb der KI stellt sich die Frage, welche Themen für Grundschul Kinder besondere Relevanz haben und einer kindgerechten Vermittlung zugänglich sind. Orientierungspunkt hierbei können auf der einen Seite die kindliche Lebens- und Erfahrungswelt [Ir18] und auf der anderen Seite die fünf „Big Ideas“ der KI darstellen [To19]. Demzufolge sollte Kindern das Verständnis ermöglicht werden, dass (1) Computer ihre Umwelt nur über Sensoren wahrnehmen können, (2) ihr Denken auf (vereinfachten) Repräsentationen der Umwelt basiert, die mittels Algorithmen manipuliert werden können, um neue Informationen aus bereits Bekanntem zu generieren, (3) sie zum Lernen große Mengen an Daten benötigen, die ihnen (zumeist) von Menschen bereitgestellt werden müssen, (4) die menschliche Interaktion mit intelligenten Agenten eine große Herausforderung darstellt, da menschliche Sprache, Gefühle oder Handlungsabsichten

nur bedingt erkannt werden können und (5) KI-Anwendungen sowohl positive als auch negative Wirkungen auf die Gesellschaft haben können. Dies kann erreicht werden, indem Kinder KI-Aspekte in bekannten Anwendungen entdecken und sich selbst mit einem KI-System (zum Beispiel zur Klassifikation von Bildern) auseinandersetzen.

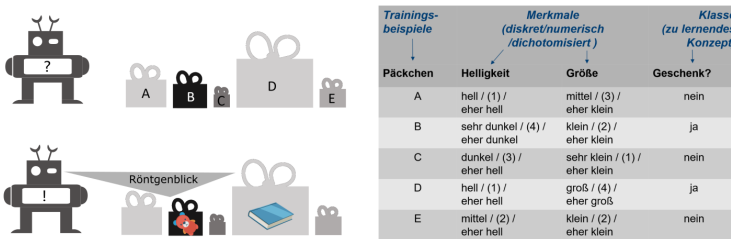
Die aktive Beschäftigung fördert neben der Motivation, sich mehr mit dem Thema auseinanderzusetzen, auch die Fähigkeiten im Bereich Computational Thinking [SL20]. Neben der Anknüpfung an die kindlichen Erfahrungen mit KI sollte auch die aktive Auseinandersetzung mit der Thematik in den Fokus von Lerneinheiten gestellt werden. Hier bieten sich Tätigkeiten an, für die die Kinder die notwendigen Fähigkeiten selbst besitzen, welche aber eben auch von intelligenten Systemen ausgeführt werden können. Anhand der Beobachtung der eigenen Herangehensweise an Aufgaben erwerben Kinder einen eigenen Zugang zum Verständnis der Simulation menschenähnlichen Verhaltens von KI-Systemen [HS19]. Zur Förderung von *AI Literacy* empfiehlt es sich deshalb, grundlegende Konzepte ab Kindergarten oder der Grundschule mit steigendem Schwierigkeits- und Abstraktionsniveau in die schulische Bildung einzubinden [Ka16; To19]. Im Allgemeinen dienen multimodale und spiel-basierte Lernansätze sowie herausfordernde, aber nicht überfordernde Lerninhalte und -materialien dazu, das kindliche Interesse und Verständnis von KI zu fördern [HS19].

Das komplexe Thema KI haben wir seit 2017 für Kinder und Jugendliche ab zwölf Jahren aufgearbeitet [SWS19]: Neben maschinellem Lernen wird in [SWS19] logikbasiertes Schlussfolgern als zweiter grundlegender Methodenbereich eingeführt und Spielen, Sprechen und Fühlen als wichtige Anwendungsgebiete vorgestellt. Um ein Verständnis aufzubauen, wie KI-Algorithmen funktionieren, wurden für alle Themen vereinfachte Umsetzungen in Python schrittweise entwickelt und konkrete Vorschläge zu algorithmischen Erweiterungen gegeben. Für den Bereich maschinelles Lernen wurde ein Anwendungsbeispiel konzipiert, das sich sowohl für das Lernen aus Tabellendaten als auch für das Lernen aus Bilddaten eignet und anhand dessen sowohl symbolische als auch neuronale Ansätze des maschinellen Lernens umgesetzt werden können. Passend zu den in [SWS19] ausgearbeiteten Themen existieren auch erste didaktisch reduzierte analoge Materialien, um Menschen ohne Informatikkenntnisse [SLR19] sowie jüngeren Kindern einen spielerischen, erfahrungsbasierten Zugang zum Thema KI zu geben.⁵ Hier bauen wir auf unsere bereits für andere grundlegenden Themen der Informatik und des algorithmisch-logischen Denkens entwickelten Materialien auf [GWS19; Kö20].⁶

In diesem Beitrag stellen wir die für maschinelles Lernen entwickelten Materialien vor, die sich an Grundschulkindern der 3. und 4. Klasse richten. Im nächsten Abschnitt wird das Lernproblem beschrieben, anhand dessen grundlegende Konzepte und verschiedene Algorithmen des maschinellen Lernens eingeführt werden können. Danach werden Entscheidungsbaumlernen sowie Perzeptronlernen für diese Beispieldomäne demonstriert. Im vierten Abschnitt zeigen wir, wie für das Verständnis und die Beurteilung wichtige

⁵ https://wehrfritz.com/de_DE/wehrfritz-digital-starter-maschinelles-lernen--367487

⁶ https://wehrfritz.com/de_DE/themenmagazin/bildungsbereiche/mathematik-informatik-naturwissenschaft-und-technik/medienkompetenz/digitalstarter



Trainingsbeispiele	Merkmale (diskret/numerisch /dichotomisiert)		Klasse (zu lernendes Konzept)
Päckchen	Helligkeit	Größe	Geschenk?
A	hell / (1) / eher hell	mittel / (3) / eher klein	nein
B	sehr dunkel / (4) / eher dunkel	klein / (2) / eher klein	ja
C	dunkel / (3) / eher hell	sehr klein / (1) / eher klein	nein
D	hell / (1) / eher hell	groß / (4) / eher groß	ja
E	mittel / (2) / eher hell	klein / (2) / eher klein	nein

Abb. 1: Das Lernproblem: Ein Roboter bekommt 5 Päckchen präsentiert, von denen einige ein Geschenk enthalten. Aus der Erfahrung mit diesen Beispielen soll er eine Hypothese aufbauen, welche Eigenschaften Päckchen haben, in denen ein Geschenk zu finden ist. Für neue Päckchen kann er dann entscheiden, ob es sich lohnt, sie zu öffnen oder nicht. Für verschiedene Lernalgorithmen werden unterschiedliche Darstellungen der Werte (z. B. Ziffern oder grobe Kategorien) genutzt.

Eigenschaften sowie grundsätzliche Grenzen von maschinellem Lernen anhand des Beispiels verdeutlicht werden können. Abschließend bewerten wir unser Konzept kritisch und geben einen Ausblick auf geplante weiterführende Arbeiten.

2 Ein induktives Lernproblem: Vielversprechende Päckchen

Die am meisten eingesetzte Art von maschinellen Lernverfahren sind sogenannte überwachte Verfahren (*supervised learning*). Hier geht es darum, aus gegebenen Trainingsbeispielen $D \subset X$ eine Funktion $f : X \rightarrow Y$ zu approximieren, die für eine meist unendliche Menge von Instanzen X eine passende Ausgabe Y liefert. Ist Y zweiwertig, handelt es sich um ein Problem des Konzeptlernens; ist Y eine Menge von diskreten Werten, handelt es sich um Klassifikationslernen; ist Y ein metrischer Wert, so handelt es sich um Regressionslernen. Im Folgenden beschränken wir uns auf Konzeptlernen. Beim überwachten Lernen wird für jedes Trainingsbeispiel die korrekte Ausgabe mit angegeben. Man sagt, die Daten sind gelabelt. Ist die Ausgabe eindeutig bekannt, so spricht man von *ground truth*. Das ist für viele visuelle Alltagskonzepte der Fall. Beispielsweise können wir bei Tieren oder deren Abbildungen eindeutig entscheiden, ob es sich um eine Katze handelt oder nicht. Die meisten Ansätze des maschinellen Lernens sind für Tabellendaten – als Merkmalsvektoren – definiert: Jede Instanz wird als Vektor von Werten für vorgegebene Merkmale beschrieben. Viele Probleme liegen in Form von Tabellendaten vor. Möchte man aber zum Beispiel ein Konzept aus Bilddaten lernen, muss man zunächst Merkmale definieren und die Merkmale dann aus den Bildern extrahieren. Aktuell finden vor allem Ansätze Beachtung, bei denen eine solche Vorverarbeitung von Daten nicht notwendig ist. So kann man etwa mit Faltungsnetzen (*convolutional neural networks*) direkt aus Bildern lernen.

Das Lernproblem, das wir als Beispiel verwenden, um maschinelles Lernen in der Grundschule einzuführen, ist in Abbildung 1 dargestellt. Ein Roboter – oder auch die Kinder selber

– erhalten einige Päckchen und sollen raten, ob ein Geschenk darin ist. Das zu lernende Konzept ist also „Päckchen, das ein Geschenk enthält“. Gegeben sind fünf Trainingsbeispiele. Die Päckchen unterscheiden sich in Größe und Helligkeit der Verpackung. Die Merkmale sind mit diskreten Werten oder alternativ als fünf Zahlenwerte angegeben. Die fünf Beispiele decken nicht alle möglichen Merkmalskombinationen ab. Werden Größe wie Helligkeit mit jeweils vier verschiedenen Werten angegeben, so gibt es 16 verschiedene Päckchen. Würden die Werte metrisch angegeben, existieren prinzipiell unendlich viele Möglichkeiten, wie ein Päckchen aussehen kann. In der letzten Spalte der Tabelle wird die Klasse angegeben, also, ob das Päckchen ein Geschenk enthält oder nicht.

Wir haben bewusst eine Beispieldomäne konstruiert, bei denen auch für uns nicht direkt ersichtlich ist, zu welcher Klasse eine Instanz gehört. Würde man maschinelles Lernen etwa mit Tierbildern vermitteln wollen, so wäre es sehr schwer, Kindern überhaupt zu vermitteln, warum hier ein Lernproblem vorliegt, weil sie ja direkt sehen, um welches Tier es sich handelt. Bevor konkrete Lernalgorithmen eingeführt werden, können die Kinder versuchen, das Lernproblem selber zu lösen. Sie können Vermutungen anstellen, welche Arten von Päckchen (Kombination von Merkmalsausprägungen) wohl ein Geschenk enthalten und dies an neuen Päckchen testen. Dies kann man mit aus Pappe gestanzten Bildern von Päckchen umsetzen, auf deren Rückseite ein Röntgenblick ins Päckchen möglich ist. Dieses Lernproblem ist passend zu der Lebenswelt der Kinder, da diese bei feierlichen Anlässen auch oft versuchen, den Inhalt des Geschenks anhand verschiedener Eigenschaften wie der Größe zu erraten.

3 Grundlegende Algorithmen des Maschinellen Lernens: Perzeptron und Entscheidungsbaum

Will man nun, gegeben die Trainingsbeispiele in Abbildung 1, eine Funktion approximieren, die bei ungesehenen Päckchen vorhersagen kann, ob ein Geschenk enthalten ist, muss man Muster entdecken, die helfen, die beiden Arten von Päckchen möglichst gut zu unterscheiden. Ein einfacher Lernalgorithmus, der bereits in den 1940er Jahren entwickelt wurde, ist das Perzeptron. Das Perzeptron ist ein idealisiertes Neuron und entspricht einer linearen Funktion. Die in Abbildung 2 dargestellte Grundform setzt mathematische Konzepte voraus, die deutlich über die Primarstufe hinausgehen. Es wird eine gewichtete Summe der Eingabewerte berechnet. Auf das Ergebnis wird eine Schwellwertfunktion angewendet. Ist der Wert größer null, so wird angenommen, dass die aktuelle Instanz zum Konzept gehört (1). Der Perzeptron-Lernalgorithmus ändert iterativ die Gewichte so, dass die Vorhersagen für die Trainingsbeispiele immer besser zutreffen. Das Gewicht w_0 wird dabei konstant mit 1 multipliziert und modelliert die Verschiebung der gewichteten Summe von Null.

Als für die Grundschule geeignete Vereinfachung kann hier mit einer festen Zahl von zwei Merkmalen gearbeitet werden. Die Gewichte werden als Wichtigkeit des Merkmals bezeichnet (siehe Abb. 2). Lernen meint also, zu erkennen, welches Merkmal wie wichtig ist, um das Konzept „enthält Geschenk“ vorherzusagen. Sagt das Perzeptron vorher, dass ein

Geschenk enthalten ist, dies ist aber nicht der Fall, sind die Gewichte zu hoch. Sie können erniedrigt werden, indem die Werte der Eingaben von denen der Gewichte subtrahiert werden (für $i = 1 \dots n$: $w_i \leftarrow w_i - x_i$). Umgekehrt werden die Gewichte erhöht, wenn falsch vorhergesagt wurde, dass kein Geschenk enthalten ist. Falls Subtraktion zu schwierig ist, können die Werte auch durch Erhöhung und Erniedrigung um einen festen Wert (zum Beispiel 1) geändert werden. Die Beispiele in Abbildung 1 wurden so konstruiert, dass nach wenigen Anpassungen alle Beispiele korrekt klassifiziert sind und dass die gelernten Gewichte auch für weitere Pakete korrekte Vorhersagen machen. Mögliche Werte für die beiden Gewichte sind 1 und -2. Um negative Zahlen zu vermeiden, kann statt auf „größer als Null“ ein höherer Wert angegeben werden.

Ein klassisches neuronales Netz kann auf Grundlage des Perzeptrons hergeleitet werden: Ändert man die Schwellwertfunktion zu einer differenzierbaren Funktion (beispielsweise eine Sigmoidalfunktion) und schaltet zwischen die Eingabeknoten und die Ausgabe weitere Perzeptrons, so entsteht ein Multi-Layer Perzeptron. Hier werden die Gewichte basierend auf partiellen Ableitungen (*backpropagation*) verändert. Hier könnte man die Kinder

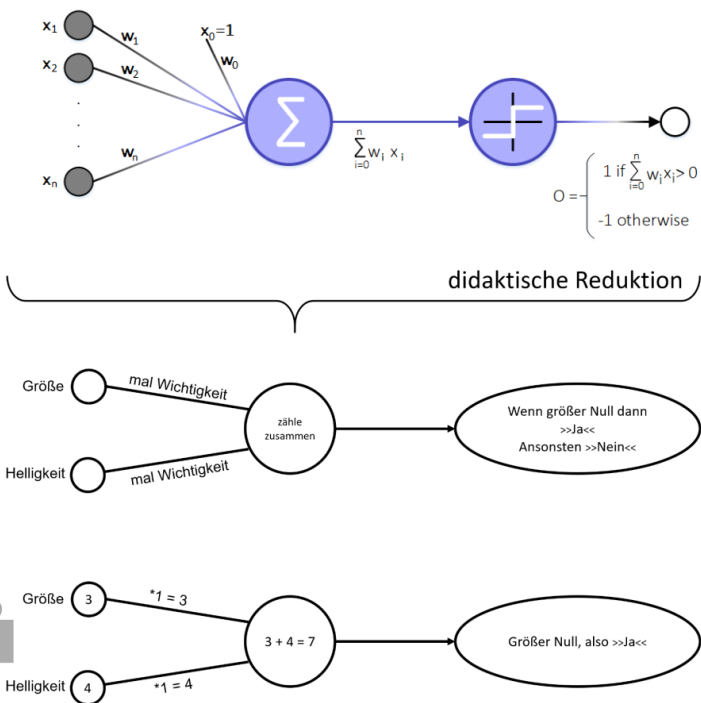


Abb. 2: Perzeptron in seiner allgemeinen Form (oben), in der didaktisch reduzierten Variante (mitte) und mit einer exemplarischen Verrechnung möglicher Merkmalswerte in der didaktisch reduzierten Form (unten).

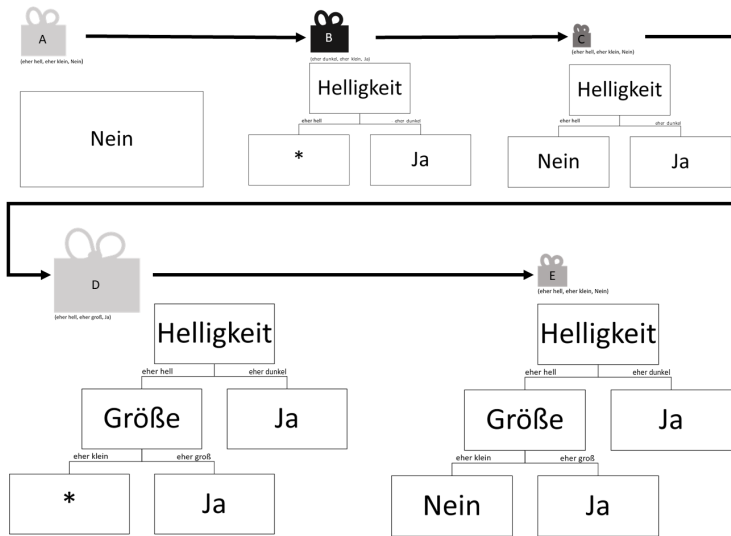


Abb. 3: Durch die serielle Prüfung der einzelnen Trainingsbeispiele, kann schrittweise ein Entscheidungsbaum erstellt werden.

ein Netz aus mehreren Knoten malen lassen und erläutern, dass das Lernen auch hier durch Veränderung der Gewichte passiert, das Rechnen aber komplizierter wird. Mit dem Perzeptron haben die Kinder Lernen als Berechnen von Wichtigkeiten kennengelernt. Dieses Prinzip liegt allen statistischen Ansätzen des maschinellen Lernens zugrunde. Hier können Beziehungen zu Alltagserfahrungen der Kinder, etwa zur Gesichtserkennung beim Entsperren von Smartphones, hergestellt werden.

Alternativ kann Lernen als explizites Konstruieren von Entscheidungsregeln modelliert werden. Der erste Entscheidungsbaumalgorithmus wurde in den 1970er Jahren entwickelt. Moderne Weiterentwicklungen wie Random Forests gehören bis heute zu den am meisten in der Praxis verwendeten Lernverfahren. Um das Beispiel einfach zu halten, werden die Merkmalsausprägungen im Päckchenbeispiel in zwei grobe Wertebereiche eingeteilt (siehe Abb. 1). Entscheidungsbaumlernen funktioniert allgemein aber auch für größere Wertemengen und auch für metrische Werte. Die meisten aktuell verwendeten Algorithmen betrachten die gesamte Trainingsmenge auf einmal. Wir verwenden einen einfachen inkrementellen Algorithmus (CAL2), der auf dem auch bei anderen Regel-Lernverfahren genutzten Prinzip des *sequential covering* basiert. Ausgewählte Zwischenschritte und der resultierende Baum werden in Abbildung 3 dargestellt.

Begonnen wird mit einem Baum, der aus einem einzigen Knoten besteht, der als Klassenausgaben ein Symbol für „unbekannt“ (zum Beispiel einen „*“) liefert. Das erste Päckchen wird präsentiert. Die Rückmeldung ist ein *. Also schauen wir in das Päckchen und sehen,

dass kein Geschenk enthalten ist. Das Fragezeichen wird durch den konstanten Wert „Nein“ überschrieben, was der übergeneralisierten Annahme entspricht, dass Päckchen keine Geschenke enthalten (links oben in Abb. 3). Das zweite Päckchen enthält jedoch ein Geschenk, also muss der Baum korrigiert werden. Das erste Merkmal wird genutzt und wir vermuten nun, dass eher dunkle Päckchen ein Geschenk enthalten. Das dritte Paket widerspricht der Erfahrung nicht. Die Hypothese ist nun, dass sehr dunkle Päckchen ein Geschenk enthalten, eher helle nicht. Das vierte Päckchen ist allerdings eher hell und enthält ein Geschenk. Der Entscheidungsbaum muss wieder korrigiert und erweitert werden. Das zweite Merkmal Größe wird anstelle der falschen Klassenentscheidung eingefügt (unten rechts in Abb. 3). Am Ende wurde der als letztes gezeigte Entscheidungsbaum aufgebaut, der für alle fünf Trainingsbeispiele und die entsprechend konstruierten Testpäckchen korrekt vorhersagen kann, ob ein Päckchen ein Geschenk enthält. Im Kontext von NLP (*Natural Language Processing*) werden häufig solche regelbasierten Ansätze zur Klassifikation hergestellt, etwa um Fakenews zu erkennen. Hier kann für die Kinder ein Bezug zu Anwendungen von Sprachassistenten hergestellt werden.

4 Was man über maschinelles Lernen wissen sollte: Biases und Generalisierungsfehler

Grundlegende Konzepte, mathematische Grundlagen und Algorithmen des maschinellen Lernens werden sowohl in allgemeinen Lehrbüchern zur KI als auch in speziellen Lehrbüchern eingeführt. Einen niedrigschwelligen Einstieg ermöglicht [KLR19]. Dort sind auch die oben besprochenen Algorithmen und deren Hintergrund anschaulich dargestellt.

Wenn Kinder zunächst selbst versuchen, zu schließen, in welchen Päckchen wohl ein Geschenk vorhanden ist, erleben sie ein Grundprinzip des induktiven Lernens – nämlich, dass aus Beispielen generalisierte Konzepte fehlerhaft sein können. Wendet man ein Verfahren des maschinellen Lernens auf einen Echtweltdatensatz an, so behält man eine Teilmenge der Daten zurück. So kann man für das Gelernte Modell abschätzen, wie die Klassifikationsleistung – die prädiktive Performanz des Modells – auf bislang ungesehenen Daten in der Zukunft sein wird. Ein Modell, das zu 99% korrekte Vorhersagen macht, macht bei jeder hundertsten Instanz einen Fehler. Dies kann Kindern sehr gut am Erkennen von Verkehrszeichen oder Ampeln veranschaulicht werden: Jedes hundertste Mal wird ein Stoppschild oder eine rote Ampel übersehen (*miss*) oder ein anderes Schild bzw. eine andere Ampelfarbe fälschlicherweise als Stoppschild oder rot erkannt (*false alarm*). Gemeinsam mit den Kindern kann überlegt werden, was in diesen Szenarien passieren kann. Zudem kann diskutiert werden, in welchen Bereichen solche Fehler schlimme Konsequenzen haben können, beispielsweise in den Bereichen Verkehr oder Medizin, und in welchen Bereichen weniger schlimme Folgen zu erwarten sind, beispielsweise bei Empfehlungssystemen. Als Konsequenz sollte gezogen werden, dass maschinelles Lernen Computer und Roboter dazu befähigt, Aufgaben ähnlich intelligent zu lösen wie wir Menschen, aber, dass sie dann auch Fehler machen – wie wir Menschen auch.

Induktives Lernen ist – egal ob bei Mensch oder Computer – nur möglich, wenn das Lernsystem einen induktiven Bias aufweist. Dieser – typischerweise implizite – Bias ist Voraussetzung dafür, dass generalisiert werden kann. Bei Menschen wird der Bias sichtbar, wenn Kinder erste Konzepte erwerben und etwa das schon erworbene Konzept einer Katze auch auf ein Eichhörnchen anwenden oder die regelmäßige Form der Vergangenheitsbildung auch auf unregelmäßige Verben übertragen („gegangen“). Die Schattenseite unserer enormen Fähigkeit zum induktiven Lernen sind Vorurteile und Stereotype wie „Jungen sind schlecht in Deutsch, Mädchen sind schlecht in Mathematik“. Anders als KI-Systeme verfügen Menschen über die Fähigkeit der Meta-Kognition. Wenn man uns auf ein Vorurteil hinweist, können wir versuchen, dem entgegenzuwirken, auch wenn es nicht immer gelingt.

Eine zweite Art von Bias ist der sogenannte *sampling bias*, der sich auf die Auswahl der Daten, mit denen gelernt wird, bezieht. Wird maschinelles Lernen auf vorhandene Daten angewendet, ohne deren Verteilung auf die verschiedenen Zielklassen zu beachten, kann dies zu unbeabsichtigter Diskriminierung oder zum „Klugen Hans Effekt“ führen. Im zweiten Modell scheint es so, als ob das Modell beispielsweise Pferde erkennen kann, aber die Klassifikation erfolgt aufgrund einer zufällig korrelierten Information wie etwa dem Bildhintergrund. Wird das Modell auf ungesehene Daten angewendet, bei denen diese Information fehlt, kann es die Klasse nicht mehr erkennen. Unfaire Modelle durch Diskriminierung sowie Nutzung irrelevanter Information basieren auf Überanpassung an die Trainingsdaten (*overfitting*). Um Kindern dieses Problem erfahrbar zu machen, kann beispielsweise eine neue Trainingsmenge von Päckchen gezeigt werden, bei der bei allen Päckchen, die ein Geschenk enthalten, eine andere Hintergrundfarbe gegeben ist. Bei den Testdaten haben dann alle Päckchen den gleichen Hintergrund.

Menschen neigen dazu, ihren eigenen Fähigkeiten auch KI-Systemen zuzuschreiben. Beispielsweise verbinden viele Menschen mit der Bezeichnung maschinelles Lernen ähnliche Eigenschaften, wie sie menschliches Lernen aufweist. Insbesondere wird angenommen, dass Lernen grundsätzlich lebenslang (inkrementell) erfolgt, während die meisten maschinellen Lernansätze zunächst aus vielen Daten ein Modell aufbauen und das Modell nachfolgend nicht mehr verändert wird. Menschliches Lernen ist zudem nie rein datengetrieben, bereits erworbenes Wissen wird genutzt, um die Plausibilität von induktiven Schlüssen zu überprüfen. In vielen Bereichen können Menschen aus sehr wenigen Beispielen generalisieren. Dies gilt für symbolische Bereiche wie grammatikalisches Wissen, aber auch für komplexe visuelle Bereiche. Diese Eigenschaften können durch Kontrastierung mit den eingeführten Lernverfahren Perzeptron und Entscheidungsbaum herausgearbeitet werden.

5 Ausblick

Wir haben die Päckchen-Domäne vorgestellt, die geeignet ist, verschiedene Algorithmen und grundlegende Konzepte des maschinellen Lernens auf spielerische Art für den Primarbereich zu vermitteln. Dabei gehen wir von der Perspektive der Informatikdiaktik aus, basierend auf der Annahme, dass digitale Kompetenz nur dann sinnvoll aufgebaut werden kann, wenn

die Funktionsweise der den Anwendungen zugrundeliegenden Algorithmen verstanden wird. Digitale Kompetenz in diesem Sinne zeichnet sich dadurch aus, dass die Nutzung komplexer Systeme auf einem tieferen konzeptuellen Verständnis basiert und damit eine bessere Einschätzung und Bewertung des Systemverhaltens und in der Folge ein souveräner und selbstbestimmter Umgang mit digitalen Medien ermöglicht wird. Gerade bei komplexen Anwendungen mit KI-Komponenten besteht die Gefahr eines zunehmenden Digital Divide. Es ist zu hoffen, dass gezielte schulische Bildungsangebote, die sich nicht auf höhere Bildung beschränken, helfen können, dem entgegenzuwirken. Die Erprobung und empirische Evaluation der vorgestellten Materialien wird im digitalen Lehr- und Lernlabor für Vor- und Grundschule der Otto-Friedrich-Universität Bamberg (FELI-Lab) erfolgen.

Literatur

- [BC08] Bernstein, D.; Crowley, K.: Searching for signs of intelligent life: An investigation of young children's beliefs about robot intelligence. *The Journal of the Learning Sciences* 17/2, S. 225–247, 2008.
- [Dö16] Döbeli Honegger, B.: *Mehr als 0 und 1: Schule in einer digitalisierten Welt*. Bern, 2016.
- [Dr18] Druga, S.; Williams, R.; Park, H. W.; Breazeal, C.: How smart are the smart toys? children and parents' agent interaction and intelligence attribution. In: *Proc. of 17th ACM Conference on Interaction Design and Children*. S. 231–240, 2018.
- [GSB20] Görz, G.; Schmid, U.; Braun, T., Hrsg.: *Handbuch der Künstlichen Intelligenz*. de Gruyter, 2020.
- [GWS19] Gärtig-Daug, A.; Werner, A.; Schmid, U.: Wie funktioniert das? – Informatische Konzepte in der Vor- und Grundschule spielerisch begreifen und anwenden. In (A. Pasternak, Hrsg.): *INFOS 2019. GI*, S. 377, 2019.
- [HS19] Ho, J. W.; Scadding, M.: Classroom Activities for Teaching Artificial Intelligence to Primary School Students. In: *Proc. Int. Conf. on Computational Thinking Education, CoolThink, Hong Kong*. S. 157–159, 2019.
- [Ir18] Irion, T.: Wozu digitale Medien in der Grundschule? Sollte das Thema Digitalisierung in der Grundschule tabuisiert werden? *Grundschule aktuell* 142/, S. 3–7, 2018.
- [Ka16] Kandlhofer, M.; Steinbauer, G.; Hirschmugl-Gaisch, S.; Huber, P.: Artificial intelligence and computer science in education: From kindergarten to university. In: *2016 IEEE Frontiers in Education Conference (FIE)*. IEEE, S. 1–9, 2016.
- [KLR19] Kersting, K.; Lampert, C.; Rothkopf, C., Hrsg.: *Wie Maschinen lernen – Künstliche Intelligenz verständlich erklärt*. Springer, 2019.

- [Kö20] Köhler, K.; Schmid, U.; Weiß, L.; Weitz, K.: Pixel & Co. - Informatik in der Grundschule. Westermann, 2020.
- [Ri83] Rich, E.: Artificial Intelligence. McGraw-Hill, New York, 1983.
- [SL20] Shamir, G.; Levin, I.: Transformations of computational thinking practices in elementary school on the base of artificial intelligence technologies. In: Proc. of EDULEARN20 Conference. Bd. 6, S. 1596–1605, 2020.
- [SLR19] Seegerer, S.; Lindner, A.; Romeike, R.: AI Unplugged – Wir ziehen Künstlicher Intelligenz den Stecker. In (Pasternak, A., Hrsg.): INFOS 2019. Bd. P-288. LNI, Gik, S. 325–334, 2019.
- [SWS19] Schmid, U.; Weitz, K.; Siebers, M.: Künstliche Intelligenz selber programmieren für Dummies Junior. Wiley, 2019.
- [To19] Touretzky, D.; Gardner-McCune, C.; Martin, F.; Seehorn, D.: Envisioning AI for K-12]: What should every child know about AI? In: Proc. of the AAAI Conference on Artificial Intelligence. Bd. 33. 01, S. 9795–9799, 2019.
- [VTV20] Vartiainen, H.; Tedre, M.; Valtonen, T.: Learning machine learning with very young children: Who is teaching whom? International Journal of Child-Computer Interaction 25/, S. 100182, 2020.
- [WPB19] Williams, R.; Park, H. W.; Breazeal, C.: A is for artificial intelligence: the impact of artificial intelligence activities on young children’s perceptions of robots. In: Proc. of CHI Conference on Human Factors in Computing Systems. S. 1–11, 2019.

Datenbewusstsein im Kontext digitaler Kompetenzen für einen selbstbestimmten Umgang mit datengetriebenen digitalen Artefakten

Lukas Höper,¹ Carsten Schulte¹

Abstract: Im Alltag spielen datengetriebene digitale Artefakte eine immer größere Rolle, sie bestimmen und leiten uns in vielen Alltagssituationen. Bei der Interaktion mit diesen werden massenhaft Daten erhoben, generiert und etwa durch die Verwendung von Technologien des maschinellen Lernens für verschiedene Zwecke automatisiert verarbeitet. Lernende haben jedoch oft kein adäquates Verständnis, wo, wie und wozu Daten von ihnen erhoben und verarbeitet werden, sind sich dieser Prozesse der Erhebung und Verarbeitung von Daten also oft nicht bewusst. Wir schlagen das Konzept *Datenbewusstsein* vor, womit die Rolle der Daten im Alltag der Lernenden, also gerade in den Interaktionen mit datengetriebenen digitalen Artefakten, in den Mittelpunkt gestellt wird. Das Konzept Datenbewusstsein ordnet sich als Teil der nötigen digitalen Kompetenzen in der heutigen und zukünftigen digitalen Welt ein.

Keywords: Digitale Kompetenzen; Informatische Allgemeinbildung; Datenbewusstsein; datengetriebene digitale Artefakte

1 Einleitung

In diesem Beitrag stellen wir das Konzept *Datenbewusstsein* vor, was auf eine Befähigung zum selbstbestimmten Umgang mit persönlichen Daten und *datengetriebenen digitalen Artefakten* (ddA) abzielt. Was das heißt, soll zunächst an einem Beispiel illustriert werden:

Beim Registrieren für einen Onlinedienst begegnen Nutzer*innen oft Captchas zum Ausfüllen; diese „Completely Automated Public Turing test to tell Computers and Humans Apart“ sollen sicherstellen, dass sich wirklich ein Mensch anmeldet. Dazu werden Nutzer*innen aufgefordert, etwa auf Bildern Dinge zu identifizieren – z.B. Verkehrszeichen. Durch das Anklicken dieser Bilder werden Daten erstellt, die zum Identifizieren der Nutzer*innen als Menschen genutzt werden. Neben diesem primären Zweck dienen die Daten ebenfalls in einer Zweitverwertung als Trainingsdaten für Bilderkennungen (vgl. [Ah11]). Inzwischen gibt es weitere Captchas, bei denen etwa lediglich eine Checkbox zum Bestätigen von „Ich bin kein Roboter“ angeklickt werden soll. Dessen Funktionsweise ist etwas unklar, man geht davon aus, dass zuvor erhobene Daten zum Browsingverhalten oder auch die charakteristischen Bewegungen der Maus genutzt werden, um Mensch und Technik auseinanderzuhalten.

¹ Universität Paderborn, Didaktik der Informatik, Fürstenallee 11, D-33102 Paderborn, {lukas.hoeper,carsten.schulte}@upb.de

Es werden also Daten implizit erhoben und verwendet, dessen sich Nutzer*innen nicht unbedingt bewusst sind. Weitergehend könnte überlegt werden, ob die möglicherweise implizit erhobenen Mausbewegungs-Daten gespeichert und zweitverwertet werden.

Dieses Beispiel der Captchas illustriert drei wesentliche Aspekte von Datenbewusstsein, die wir fördern möchten: (1) Daten werden explizit (also eher bewusst) eingegeben bzw. erhoben, es gibt aber auch implizit (oft unbewusst) erhobene Daten; (2) erhobene Daten werden oft unmittelbar für primäre Zwecke verarbeitet und verwendet, sie können aber auch für sekundäre Zwecke zweitverwertet werden; (3) Datenbewusstsein zielt darauf ab, in Interaktionen mit digitalen Artefakten den Blick vom eigentlichen Handlungsziel (hier: die Registrierung) absehen zu können, und auf die explizit und implizit erhobenen Daten sowie auf mögliche primäre und sekundäre Verwendungszwecke zu achten. Lernende müssten also befähigt werden, in Interaktionen mit einem ddA die Aufmerksamkeit auf die Rolle der Daten und Datenverarbeitungsprozesse lenken sowie das jeweilige ddA auf mögliche relevante Datenerhebungs- und Datenverarbeitungsprozesse hin untersuchen zu können.

2 Digitale Kompetenzen für den selbstbestimmten Umgang mit datengetriebenen digitalen Artefakten

Bevor wir das Konzept Datenbewusstsein weiter ausdifferenzieren und anschließend Hinweise zur didaktischen Umsetzung anhand eines kurzen Beispiels geben, werden wir in diesem Abschnitt das Konzept sehr knapp in verwandte Arbeiten etwa aus der Informatik, Mathematik oder Medienpädagogik einordnen.

Die reflektierende Betrachtung der Interaktion mit digitalen Artefakten, früher sagte man eher Informatiksysteme, kann auch als „informatischer Aufgabenbereich der Medienbildung“ [Sc01] verstanden werden, etwa wenn webgestützte Medienangebote betrachtet werden. Einerseits wird deutlich, dass informatische Bildung und Medienbildung Gemeinsamkeiten, Überlappungen und Schnittstellen haben, andererseits werden die verschiedenen Betrachtungen aber nebeneinander gelegt. Das ist im Grunde auch in neueren Debatten wiederzufinden, wie etwa zum Dagstuhl-Dreieck [Ge16], in dem die drei Seiten verschiedenen Ebenen grob zugeordnet werden – der Informatik eher die technologische Perspektive und der Medienpädagogik eher die gesellschaftlich-kulturelle Perspektive. Der Ansatz Datenbewusstsein nimmt ebenso die Interaktion mit Artefakten (wie implizit im Dagstuhl-Dreieck, und noch stärker im Frankfurt-Dreieck (vgl. [Ge19])) als Ausgangspunkt, erlaubt und erfordert aber unseres Erachtens durch den bewusst eng gefassten Bezug auf die Rolle der Daten und Datenflüsse in diesen Interaktionen, die verschiedenen Betrachtungsebenen an einem kohärenten Aspekt auch tatsächlich einzubeziehen und zu integrieren.

Durch den Fokus auf Daten hat das Konzept ebenso Überschneidungen zu anderen Konzepten, wie Data Literacy oder Statistical Literacy. Ersteres kann wie folgt verstanden werden: „Data Literacy umfasst die Fähigkeiten, Daten auf kritische Art und Weise zu sammeln, zu managen, zu bewerten und anzuwenden“ [SKR21]. Es geht unter anderem darum, Daten zu

sammeln und mit diesen kritisch reflektierend eine Handlung durchzuführen. Datenliteraten können also kompetent mit Daten umgehen und unter anderem Datenanalysen durchführen. Statistical Literacy als verwandtes Konzept zu Data Literacy umfasst etwa Fähigkeiten, statistische Zusammenhänge interpretieren und kritisch evaluieren sowie diskutieren und kommunizieren zu können [Ga02]. Es geht also um das Untersuchen einer Fragestellung mit statistischen Methoden, Interpretieren von Daten und Beantworten der Fragestellung, wozu verschiedene Statistikenkenntnisse nötig sind. Für die Erklärung der Zusammenhänge dieser Ansätze zu anderen Literacies verweisen wir an dieser Stelle auf [HP20].

Heutzutage nutzen viele digitale Artefakte oder Medien Technologien und Methoden des maschinellen Lernens, mit denen massenhaft Daten erhoben, generiert und automatisiert verarbeitet werden können. Maschinelles Lernen spielt bei der alltäglichen Nutzung von ddA eine große Rolle. Im Kontext der informatischen Allgemeinbildung schlagen [Te20] aufbauend auf Data Literacy den Ansatz Data Agency vor, womit – neben dem Verstehen von Daten – die Fähigkeit, 'information flows' aktiv zu kontrollieren und zu manipulieren und bei dessen Nutzung weise und ethisch zu handeln, betont wird. Es geht also um aktives Handeln mit und Gestalten von Machine Learning basierten Technologien. Im Kontext der Media Literacy argumentieren mehrere dieser Autor*innen dafür, dass für die Förderung einer Media Literacy ebenfalls etwa Technologien des Trackings oder Methoden des maschinellen Lernens vermittelt werden sollten [Va19]. Maschinelles Lernen als Teil der künstlichen Intelligenz spielt beim Ansatz AI Literacy eine besondere Rolle. Nach einem Literaturreview über AI Literacy definieren [LM20] AI Literacy als „a set of competencies that enables individuals to critically evaluate AI technologies; communicate and collaborate effectively with AI; and use AI as a tool online, at home, and in the workplace“. Es geht also vor allem um einen kompetenten Umgang mit KI Technologien.

Diese Ansätze greifen für Lernende verschiedener Altersstufen Kompetenzen auf, etwa um Daten zum Gewinnen einer Erkenntnis zu untersuchen oder um Technologien mit künstlicher Intelligenz zu nutzen und zu erkennen. Damit die Schüler*innen in eigenen Interaktionen mit einem ddA selbstbestimmt handeln können, ist jedoch vor allem ein Bewusstsein über die Erhebung und Verarbeitung persönlicher Daten während der Nutzung eines ddA notwendig. Bei einer Studie fanden [Te20] heraus, dass viele Schüler*innen zwar aktiv mit ddA interagieren, aber nur wenig Verständnis davon hatten, wo, wie und wozu Daten über sie erhoben und verarbeitet werden, sie sind sich dieser Prozesse der Erhebung und Verarbeitung von Daten also oft nicht bewusst. Die Autor*innen konnten damit in ihrer Studie ähnliche Berichte von [PS19] bestätigen. Sofern sich dieses Ergebnis auch in weiteren Studien bestätigen ließe, stellt sich die Frage, wie eine selbstbestimmte Interaktion mit ddA möglich ist, wenn sich die Schüler*innen der Datenerhebungs- und Datenverarbeitungsprozesse nicht bewusst sind. Unsere These ist, dass derzeit die informatische Allgemeinbildung eine Lücke zwischen dem allgemeinen und abstrakten Verstehen der Bedeutung von Daten und dem alltäglichen Interagieren mit diversen digitalen Artefakten besteht. Denn in diesen Interaktionen steht das Ziel der Nutzung im Vordergrund, und die in einer Black Box ablaufenden datengetriebenen Prozesse bleiben oft verborgen oder stehen

nicht im Mittelpunkt der Aufmerksamkeit. Die vorherigen Ansätze haben oft ähnliche Zielvorstellungen, etwa eine Mündigkeit im Umgang mit Daten oder digitalen Artefakten zu fördern. Im Kontext einer informatischen Allgemeinbildung in Schulen soll das Konzept Datenbewusstsein verschiedene Ziele dieser Ansätze verbinden.

3 Das Konzept Datenbewusstsein

Bei der Entwicklung eines Datenbewusstseins im Kontext der informatischen Allgemeinbildung soll eine situationsspezifische Aufmerksamkeit für die Erhebung und Verarbeitung persönlicher Daten in Interaktionen mit ddA entwickelt werden. Die wesentlichen Aspekte von Datenbewusstsein werden im Folgenden ausdifferenziert und mit einem Beispiel für eine didaktische Umsetzung illustrieren.

3.1 Datengetriebene digitale Artefakte (ddA) aus dualen Perspektiven

In einem Beitrag zu Data Agency beschreiben [Te20] die aktuellen Veränderungen durch die Nutzung von datengetriebenen Algorithmen, zum Beispiel aus dem Bereich des maschinellen Lernens, als *data-driven turn of technology*, was auch Einfluss auf informatische Allgemeinbildung hat. Ähnlich beschreiben [SFN18] eine nötige Verschiebung des Fokus in der informatischen Bildung auf die Thematisierung der Erhebung, Verarbeitung und Organisation von Daten. Beiden Argumentationen liegt die Veränderung des Verstehens von digitalen Artefakten zugrunde: herkömmliche digitale Artefakte mit klassischen Algorithmen können durch Verstehen des Algorithmus verstanden werden; das Verstehen von ddA mit datengetriebenen Algorithmen erfordert jedoch auch ein Verständnis von Daten und deren Rolle bei diesen Artefakten. So kann etwa Bias oder Fairness in einem ddA erst verstanden werden, wenn die Rolle der Daten und Datenmodelle erkannt wird. Zum Verstehen eines ddA reicht somit eine alleinige Betrachtung eines Algorithmus nicht mehr aus, es muss auch die Rolle der Daten berücksichtigt werden. Bei der Interaktion mit einem ddA spielen vor allem die eigenen Daten inmitten derer anderer eine wesentliche Rolle, welche bewusst gemacht werden sollte.

Um auf das Beispiel der Captchas zu Beginn dieses Beitrags zurückzukommen: Bei der Interaktion mit einem Captcha werden Daten erhoben, die zum Zwecke der Unterscheidung von Mensch und Maschine verwendet sowie zu sekundären Zwecken zweitverwertet werden. Im Gegensatz zu diesem Beispiel wäre die Eingabe eines Passwortes bei einem Login, wobei ebenfalls Daten relevant sind, kein ddA. Der Unterschied ist, dass sich das ddA selbst oder eine Rückmeldung dessen nicht durch die Eingabe der Daten verändert, es also nicht datengetrieben ist. Ein weiteres Beispiel für ein ddA, welches durch die Verarbeitung erhobener Daten beeinflusst wird, ist etwa das Erstellen einer personalisierten Startseite bei Netflix durch einen Empfehlungsdienst, wozu unter anderem zuvor erhobene Interaktionsdaten verarbeitet werden [AB15]. Für das Konzept Datenbewusstsein ist das

Identifizieren von relevanten Kontexten und Inhalten nötig, sodass ein ddA als eine Art Erklärmodell dienen kann, womit geeignete exemplarische Beispiele für den Unterricht gefunden werden können. Zentral ist, dass bei der Nutzung eines ddA Daten erhoben und verarbeitet werden, um etwa für Nutzer*innen ein potentiell interessantes Angebot anzuzeigen.

3.2 Interaktion zwischen Nutzer*in und ddA

Im Alltag interagieren Schüler*innen mit verschiedenen ddA, wobei Daten erhoben und für verschiedene Zwecke verwendet und verarbeitet werden. Dabei spielt die Interaktion eine wesentliche Rolle, etwa bei der Beobachtung des Nutzungsverhaltens. Ein Bewusstsein für die Erhebung und Verarbeitung von Daten erfordert gerade eine Aufmerksamkeit für diese Prozesse, sodass es nicht reicht, nur das ddA selbst zu betrachten. Vielmehr sollte die Interaktion mit einem ddA als Betrachtungsgegenstand verstanden werden. Eine Interaktion wird nach [SB18] von einer Interaktionskette bestimmt, die als aneinandergereihte Interaktionsschritte – bestehend aus Eingabe, Verarbeitung und Ausgabe – verstanden werden kann. Nutzer*innen tätigen in einer solchen Interaktion eine Eingabe, die anschließend durch das ddA verarbeitet wird, wodurch nachfolgende Handlungen des Nutzenden beeinflusst werden. So eine Beeinflussung findet sich etwa beim Attention Engineering oder auch Behaviour Engineering wieder (vgl. [Va19]), wobei das Verhalten von Nutzer*innen etwa durch Empfehlungsdienste möglichst beeinflusst werden soll. Dadurch entstehen Wechselwirkungen in einem System, welches von [SB18] als hybrides Interaktionssystem beschrieben wird, bestehend aus den beiden Akteuren – Nutzer*in und ddA – und der Interaktion zwischen diesen. Auf nähere Ausführungen dazu wird an dieser Stelle auf [SB18] verwiesen, es sei lediglich festzuhalten, dass die Interaktion mit einem ddA als Interaktionssystem zu verstehen ist, in dem sich Mensch und ddA gegenseitig beeinflussen – sowohl auf einer individuellen als auch gesellschaftlichen Ebene. Exemplarisch könnte dafür im Kontext von Empfehlungsdiensten das Entstehen von Filterblasen aufgegriffen und reflektiert werden.

3.3 Explizite und implizite Erhebung von persönlichen Daten

Der Erhebung und Generierung von Daten in Interaktionen mit ddA müssen sich die Nutzer*innen nicht unbedingt bewusst sein. Dazu lässt sich zwischen verschiedenen Arten der Datenerhebung differenzieren, wie etwa von [OE14] beschrieben: *provided data* als Daten, die bewusst und aktiv durch die Nutzer*innen erstellt werden; *observed data* als Beobachtungs- und Metadaten, deren Erhebung dem Einzelnen nicht bewusst sein muss; *derived data* als Daten, die aus anderen bestehenden Daten generiert wurden; *inferred data* als Daten, die durch wahrscheinlichkeitsbasierte analytische Prozesse generiert werden, wie etwa Vorhersagewerte für das Verhalten des Nutzers. Eine ähnliche Unterscheidung ist auch bei [PS19] wiederzufinden. Es kann also wie folgt unterschieden werden:

(1) *explizit* erhobene Daten, wie etwa durch aktive Eingabe durch Nutzer*innen (v.a. provided data), oder (2) *implizit* erhobene oder generierte Daten, wie etwa durch Beobachtung oder Tracking oder durch Datenverarbeitungsprozesse (observed, derived und inferred data). Beim Nutzen einer Suchmaschine werden beispielsweise durch die Bestätigung einer Eingabe von Suchbegriffen (explizit erhobene Daten) ebenfalls implizit Daten – z.B. ein Zeitstempel – erhoben. Wichtig zu begreifen ist hier vor allem die implizite Erhebung von Daten, so können etwa durch die Verarbeitung einfacher bereits erhobener Daten auch neue umfangreichere Daten generiert werden, zum Beispiel durch Relationierung dieser Daten mit derer anderer Nutzer*innen. Damit die Erhebung und Verarbeitung von Daten bei der Interaktion mit einem ddA verstanden oder auch bewertet werden kann, müssen sowohl explizit als auch implizit erhobene Daten erkannt sowie bewusst werden, wie und wozu diese Daten verwendet und verarbeitet werden könnten.

3.4 Verarbeitung erhobener Daten zu primären und sekundären Zwecken

Für die Verarbeitung erhobener oder generierter Daten können durchaus verschiedene Zwecke verfolgt werden, wobei insbesondere die Bedeutungszuweisung und der jeweilige Kontext der Verarbeitung und Verwendung relevant ist. [Zu18] beschreibt die Verarbeitung von Daten zum Ermitteln eines „Verhaltensüberschusses“, wodurch basierend auf Daten über das vorherige Verhalten Vorhersagen für das zukünftige Interesse und Verhalten ermittelt werden. Damit können verschiedene Zwecke verfolgt werden, wie etwa das Erzielen wirtschaftlicher Gewinne. Für die automatisierte Datenverarbeitung können Methoden des maschinellen Lernens zum Erstellen von Datenmodellen oder auch einfache Datenverarbeitungsschritte zur Informationsgewinnung, wie etwa die Data Moves nach [Er18], genutzt werden. Mithilfe dieser zum einen komplexen und zum anderen einfachen Datenverarbeitungen können Profile von Nutzer*innen (Datenmodelle) erstellt werden. Aus Sicht der Medienbildung setzt sich [Is16] mit solchen digitalen Personenprofilen im Kontext einer informationellen Selbstbestimmung und der strukturalen Medienbildung auseinander. Solche digitalen Personenprofile werden durch Zusammenführung von erhobenen Daten automatisiert erstellt, wobei diese Daten eine Nutzerin bzw. einen Nutzer anhand von Merkmalen abbilden [Is16]. Ähnlich charakterisieren [BK15] den *digitalen Doppelgänger*. Im Konzept Data Agency wird von [Te20] etwa der Begriff „digital self“ aufgegriffen, welches zum Beispiel an sich und inmitten derer anderer erkannt und reflektiert werden solle. Diese Personenprofile oder digitale Doppelgänger können für verschiedene Zwecke verwendet werden. [Is16] führt etwa die problematische Anwendung von Personenprofilen aus, die etwa nicht nur im Sinne einer subjektiv oft positiv beurteilten personalisierten Werbung genutzt werden, sondern dessen Nutzung auch zu Benachteiligungen führen kann. Zusammenfassend lässt sich etwa ein digitaler Doppelgänger mit folgenden wesentlichen Charakteristika beschreiben: Ein digitaler Doppelgänger (1) bezieht sich auf eine konkrete Person; (2) stellt eine Person anhand spezifischer und begrenzter Merkmale dar – ist also kein vollständiges Abbild; (3) ist inmitten der digitalen Doppelgänger anderer Nutzer*innen zu verstehen; (4) wird durch Zusammensetzen – vor allem implizit – erhobener Daten

konstruiert und (5) kann für verschiedene primäre und sekundäre Zwecke genutzt werden. Im wesentlichen können dabei zwei Zwecke beschrieben werden: (1) Verwendung und Verarbeitung von Daten, die für das Betreiben von Funktionen des ddA technisch notwendig sind (*primäre Zwecke*) und (2) womit weitergehende, wie etwa kommerzielle Zwecke, verfolgt oder Weiterentwicklungen eines ddAs untersucht werden können (*sekundäre Zwecke*). Für ein Datenbewusstsein ist das Bewusstmachen verschiedener Verwendungs- und Verarbeitungszwecke von Daten bedeutsam, wozu für einen kompetenten Umgang auch das Anstellen von eigenen Vermutungen gehört.

3.5 Ziele und Inhalte für das Konzept Datenbewusstsein

Mit Datenbewusstsein wird darauf abgezielt, dass Schüler*innen ein Bewusstsein über die Erhebung und Verarbeitung von persönlichen Daten in Interaktionen mit ddA entwickeln. Dazu folgt als Ziel einer Befähigung, in Interaktionen mit ddA die explizite und implizite Erhebung von persönlichen Daten und deren Verarbeitung zu primären und sekundären Zwecken erkennen und bewerten zu können. Mit Bezug zur referenzierten Studie von [Te20] werden in der Tabelle 1 wesentliche Aspekte von Datenbewusstsein zu den Fragen „Wo, wie und wozu werden persönliche Daten erhoben und verarbeitet?“ zugeordnet.

	Wo?	Wie?	Wozu?
Erhebung von Daten	Interaktion zwischen Mensch und ddA	explizit und implizit	primäre und sekundäre Zwecke
Verarbeitung von Daten		v.a. Data Moves und Maschinelles Lernen	

Tab. 1: Tabellarische Übersicht relevanter Aspekte des Konzepts Datenbewusstsein

Für ein selbstbestimmtes und kompetentes Interagieren mit ddA ist ein Verständnis davon nötig, wo, wie und wozu persönliche Daten erhoben und verarbeitet werden. Dazu sind gerade Interaktionssysteme bestehend aus den beiden Akteuren – Nutzer*in und ddA – und der Interaktion zwischen diesen relevant. Bei diesen Interaktionen werden Daten massenhaft explizit und implizit erhoben und generiert. Diese Daten werden zu primären oder sekundären Zwecken verwendet und verarbeitet. Dabei werden in der Regel Datenmodelle erstellt, etwa digitale Doppelgänger, die zur Verarbeitung neuer Daten genutzt werden können. Schüler*innen sollten sich dieser Erhebung und Verarbeitung von Daten in einer konkreten Interaktion mit einem ddA bewusst sein und diese erkennen können, um in diesen Interaktionssystemen bewusst und kompetent handeln zu können. Dazu gehört auch ein grundlegendes Verständnis von Datenarten und Schritten zur Verarbeitung von Daten, insbesondere zur Erstellung eines digitalen Doppelgängers. Dabei ist es notwendig sowohl Relevanz als auch die Architektur eines ddAs zu betrachten und zu kombinieren, um etwa Auswirkungen oder Wechselwirkungen verstehen und erklären zu können. Dadurch sollten Schüler*innen dazu befähigt werden, in diesen Interaktionen bewusst zu handeln

und eigene Handlungsmöglichkeiten zu bewerten und zu entwickeln. Diese Facetten von Datenbewusstsein werden in Abbildung 1 zusammenfassend dargestellt.

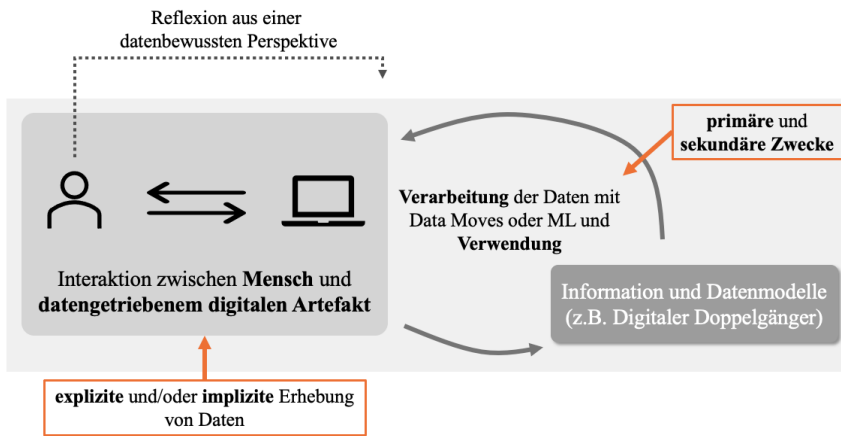


Abb. 1: Modell zur Beschreibung der verschiedenen Facetten von Datenbewusstsein

3.6 Förderung von Datenbewusstsein im Informatikunterricht

Zur weiteren Illustration der Förderung von Datenbewusstsein etwa im Informatikunterricht, wird im Folgenden ein Unterrichtsbeispiel zu Standortdaten im Mobilfunknetz skizziert. Bei der Nutzung des Mobilfunknetzes, etwa beim Telefonieren, werden verschiedene Daten explizit und implizit erhoben – unter anderem sind dies Standortdaten (vgl. [Wi12]). Diese Daten hat ein deutscher Politiker zusammen mit Zeit veröffentlicht². Dieses Beispiel für eine Interaktion mit einem ddA greifen wir in einer Unterrichtsreihe für den Informatikunterricht in den Jahrgangsstufen 5 und 6 wie folgt auf: Zunächst wird im Unterricht der Aufbau und die Funktionsweise des Mobilfunknetzes mithilfe eines Videos und eines Puzzles motiviert und grob erarbeitet. Dieses Interaktionssystem kennen die Schüler*innen bereits aus dem Alltag, wozu nun erarbeitet wird, welche Daten explizit und implizit erhoben werden. Für die Nutzer*innen werden die Standorte der verbundenen Basisstationen gespeichert und zum Herstellen einer Mobilfunkverbindung genutzt (primärer Zweck). Dann stellt sich die Frage, wozu diese Standortdaten zweitverwertet werden könnten – losgelöst von tatsächlichen Zweitverwertung von Mobilfunkanbietern. Mithilfe einer erstellten Web-Anwendung explorieren die Schüler*innen die gegebenen Standortdaten einer Person und sollen dabei durch Filtern und Visualisieren der Standortdaten möglichst viele Informationen über die ihnen unbekannte Person ermitteln und einen Steckbrief erstellen (digitaler Doppelgänger). Abschließend reflektieren die Schüler*innen die Erhebung und Verarbeitung von Standortdaten im Kontext des Mobilfunknetzes und in weiteren Kontexten, in denen ebenfalls Standortdaten erhoben werden.

² <https://www.zeit.de/digital/datenschutz/2011-02/vorratsdaten-malte-spitz>, Stand: 22.07.2021

4 Diskussion und Ausblick

Mit dem Konzept Datenbewusstsein schlagen wir einen Ansatz vor, wie Schüler*innen zu einem selbstbestimmten Umgang mit persönlichen Daten und datengetriebenen digitalen Artefakten (ddA) befähigt werden könnten. Ein Ziel von Datenbewusstsein ist die Fähigkeit, sich in einer Interaktion mit einem ddA selbst im Rahmen dieses Interaktionssystems wahrzunehmen und zu erkennen, welche Daten explizit und implizit erhoben werden und zu welchen primären und sekundären Zwecken diese genutzt oder verarbeitet werden könnten – dafür sollen sie ein Bewusstsein entwickeln. Insbesondere gehört dazu auch das Wahrnehmen und Entwickeln einer Aufmerksamkeit für die Rolle eigener Daten in Interaktionen mit ddA. Unsere These ist, dass auf ein Datenbewusstsein auch eine Datenmündigkeit folgen könnte, ein mündiges Interagieren mit ddA könnte jedoch für die informatische Allgemeinbildung ein ambitioniertes Ziel sein, sodass Datenbewusstsein zunächst operationalisiert als Konzept für die Unterrichtsentwicklung dienen kann. Die Aufnahme des Konzepts Datenbewusstsein im Bereich „Informatik, Mensch und Gesellschaft“ im Kernlehrplan für Informatik in den Jahrgangsstufen 5 und 6 von Nordrhein-Westfalen im Jahr 2021 zeigt bereits, dass das Konzept für die Beschreibung und Entwicklung von Informatikunterricht hilft. Diesem Beitrag folgend werden wir für Datenbewusstsein entwickelte Unterrichtsmaterialien³ weiter erproben und beforschen, um das Konzept Datenbewusstsein weiter auszudifferenzieren.

Literatur

- [AB15] Alvino, C.; Basilio, J.: Learning a Personalized Homepage, Netflix Technology Blog, Stand: 21.07.2021, 2015, URL: <http://techblog.netflix.com/2015/04/learning-personalized-homepage.html>.
- [Ah11] von Ahn, L.: Massive-scale online collaboration, TEDxCMU, Stand: 21.07.2021, 2011, URL: https://www.ted.com/talks/luis_von_ahn_massive_scale_online_collaboration?utm_campaign=tedsread&utm_medium=referral&utm_source=tedcomshare.
- [BK15] Bode, M.; Kristensen, D.: The digital doppelgänger within. A study on self-tracking and the quantified self-movement. In (Canniford, R.; Bajde, D., Hrsg.): Assembling Consumption. Routledge, 2015.
- [Er18] Erickson, T.; Finzer, B.; Reichsman, F.; Wilkerson, M.: Data Moves: One Key to Data Science at the School Level. In: Looking Back, Looking Forward. IASE, Kyoto, Japan, 2018.
- [Ga02] Gal, I.: Adults' Statistical Literacy: Meanings, Components, Responsibilities. International Statistical Review 70/1, S. 1–25, 2002.
- [Ge16] Gesellschaft für Informatik e.V.: Dagstuhl-Erklärung. Bildung in der digitalen vernetzten Welt, Stand: 22.07.2021, 2016, URL: <https://dagstuhl.gi.de/dagstuhl-erklaerung>.

³ <https://unterrichtsmaterial-ddi.cs.upb.de/Datenbewusstsein>

- [Ge19] Gesellschaft für Informatik e.V.: Frankfurt-Dreieck zur Bildung in der digital vernetzten Welt, Stand: 22.07.2021, 2019, URL: <https://dagstuhl.gi.de/frankfurt-dreieck>.
- [HP20] Hosman, L.; Pérez Comisso, M. A.: How Do We Understand “Meaningful Use” of the Internet? Of Divides, Skills and Socio-Technical Awareness. 18/3, S. 461–479, 2020.
- [Is16] Iske, S.: Medienbildung im Kontext digitaler Personenprofile. In (Verständig, D.; Holze, J.; Biermann, R., Hrsg.): Von der Bildung zur Medienbildung. Springer Fachmedien Wiesbaden, Wiesbaden, S. 257–280, 2016.
- [LM20] Long, D.; Magerko, B.: What is AI Literacy? Competencies and Design Considerations. In: 2020 CHI. CHI '20, ACM, Honolulu, HI, USA, 2020.
- [OE14] OECD: Summary of the OECD Privacy Expert Roundtable „Protecting Privacy in a Data-driven Economy: Taking Stock of Current Thinking“. DSTI/ICCP/REG(2014)3, 2014.
- [PS19] Pangrazio, L.; Selwyn, N.: ‘Personal data literacies’: A critical literacies approach to enhancing understandings of personal digital data. *New Media & Society* 21/2, S. 419–437, 2019.
- [SB18] Schulte, C.; Budde, L.: A Framework for Computing Education: Hybrid Interaction System: The Need for a Bigger Picture in Computing Education. In: *Koli Calling '18*. ACM, Koli, Finland, 2018.
- [Sc01] Schulte, C.: Informatische Aufgabenbereiche der Medienbildung. *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung* 4/, S. 1–16, 2001.
- [SFN18] Shapiro, R. B.; Fiebrink, R.; Norvig, P.: How machine learning impacts the undergraduate computing curriculum. *CACM* 61/11, S. 27–29, 2018.
- [SKR21] Schüller, K.; Koch, H.; Rampelt, F.: *Data-Literacy-Charta*, Stifterverband, Version 1.2, Berlin, 2021.
- [Te20] Tedre, M. et al.: Machine Learning Introduces New Perspectives to Data Agency in K-12 Computing Education. In: 2020 FIE. IEEE, S. 1–8, 2020.
- [Va19] Valtonen, T.; Tedre, M.; Mäkitalo, K.; Vartiainen, H.: Media Literacy Education in the Age of Machine Learning. *Journal of Media Literacy Education* 11/2, 2019.
- [Wi12] Wicker, S. B.: The Loss of Location Privacy in the Cellular Age. *Communications of the ACM* 55/8, S. 60–68, 2012.
- [Zu18] Zuboff, S.: *Das Zeitalter des Überwachungskapitalismus*. Campus Verlag, Frankfurt am Main, 2018.

Workshop: Für Informatik
begeistern - vom Kindesalter bis
zum Abitur
(Informatik-Begeisterung 2021)

BOB3 – Ein kleiner Roboter zum Einstieg in Informatik und Programmieren lernen

Digitale Mikrolerneinheiten in Kombination mit klassischen Unterrichtsmaterialien

Katja Bach¹, Nils Springob²

Abstract: In diesem Beitrag wird das Informatikprojekt BOB3 für Schülerinnen und Schüler ab der Sekundarstufe I vorgestellt. Es wird zunächst kurz auf die Entstehung eingegangen, anschließend werden die vorhandenen Lehr- und Lernmaterialien des Projekts für den Schulalltag beschrieben. Es wird eine Kombination aus digitalen, interaktiven Mikrolerneinheiten und klassischen Arbeitsmaterialien vorgestellt. Abschließend wird ein kurzes Fazit gezogen

Keywords: Einstieg in die Digitale Welt, Einstieg in Informatik, textuelle Programmierung, Sekundarstufe.

1 Einleitung

BOB3 ist ein Informatikprojekt, das zwischen 2015 und 2017 zusammen mit Lehrkräften und Schülerinnen und Schülern (im Folgenden SuS) aus der Städteregion Aachen entstanden ist. Es wurde ein Konzept entwickelt, das einen Einstieg in die Informatik und die Grundlagen der Programmierung für SuS ab Klasse 5 im Klassenverband ermöglicht. Um die Lernerfahrung anhand eines *physischen Objekts* realistisch zu gestalten, wurde der kleine Roboter BOB3 als programmierbares Mikroprozessorsystem entwickelt (Abb. 1), der von den SuS wahlweise auch selbst zusammengebaut werden kann. Ein Modellprojekt von Januar bis Juni 2017 mit sieben Schulen, 14 Lehrerinnen und Lehrern, 142 Schülerinnen und 81 Schülern zeigte, dass der textuelle Einstieg in die Programmierung mit BOB3 in der Schulpraxis im regulären Klassenverband mit bis zu 30 SuS sehr gut umsetzbar ist. In den Folgejahren wurde die Unterrichtsreihe um weitere digitale Mikrolerneinheiten, automatisierte Leistungsüberprüfungen zur Sicherung des Lernerfolgs, klassische Arbeitsblätter und eine Lernstandsübersicht für die Lehrkräfte mit selbst-erstellbaren Kapiteln und ein Klausurmodul erweitert.

¹ Am Denkmal 8, 52223 Stolberg, katja.bach@bob3.org

² Am Denkmal 8, 52223 Stolberg, nils.springob@bob3.org



Abb. 1: Mikroprozessorsystem BOB3

2 OER-Unterrichtsmaterial

Die klassischen Unterrichtsmaterialien umfassen 14 Arbeitsblätter als OER-Material, die in informatische Grundkonzepte einführen und frei im Unterricht verwendet werden können. Die behandelten Themengebiete sind ‚Robotik‘, ‚Syntax und Semantik von Quellcode‘, ‚Mikrocontroller-Programmierung‘, ‚Sequenzen‘, ‚if-else-Kontrollstruktur‘, ‚Hardwarekomponenten und Sensorik‘, ‚Softwarebibliothek/Methoden und Funktionen‘, ‚for-Schleife‘, ‚Variablen‘, ‚switch-case-Kontrollstruktur‘, ‚Funktionen‘, ‚Funktionen mit/ohne Parameter‘, ‚Funktionen mit/ohne Rückgabewert‘, ‚while-Schleife‘, ‚break-Anweisung‘ und ‚Messprinzip und Anwendung des IR-Sensors‘. Die Arbeitsblätter liefern Theorieteile, konkrete Beispiele und Aufgaben für die Schüler. Drei aufeinander aufbauende Modulbeschreibungen (AB 1-4, AB 5-8, AB 9-14) liefern Anregungen für konkret durchführbare Unterrichtseinheiten, die jeweils passende Mikrolerneinheiten mit den OER-Materialien kombinieren. Um die Modellbildung zu den Konzepten der Programmierung seitens der SuS voranzubringen, werden bei den Arbeitsblättern Erklärungen, Analogien und Beispiele eingesetzt. Bei elementaren Konzepten wie Variablen und Schleifen werden Analogiemodelle verwendet, um eine effiziente Verankerung der Konzepte zu bewirken. In der Psychologie wird die „Analogiebildung als elementarer Mechanismus des Denkens angesehen“ [Ha08].

2.1 Analogiemodell zur Einführung des Konzepts ‚Variable‘

Um ein anschauliches Analogiemodell für das Konzept einer Variablen im informatischen Sinne zu entwerfen, wurde auf einen kleinen Notizzettel mit Bleistift und Radiergummi als Sinnbild zurückgegriffen. Mit diesem Modell ließ sich im neunten Arbeitsblatt zur Einführung des Variablenbegriffs der volatile Charakter des Wertes einer Variablen herausstellen und gleichzeitig andeuten, dass immer nur ein Wert gespeichert werden kann.

Die SuS lernen mit dem Arbeitsblatt zunächst, dass eine Variable ein Speicherort für Zahlen, Zeichen oder sonstige Daten ist und dass jede Variable einen Namen und einen

Datentyp hat. Zur besseren Vorstellung wird jeweils ein Quellcodebeispiel und eine bildhafte Darstellung der Variablen nebeneinander gezeigt (Abb.2).

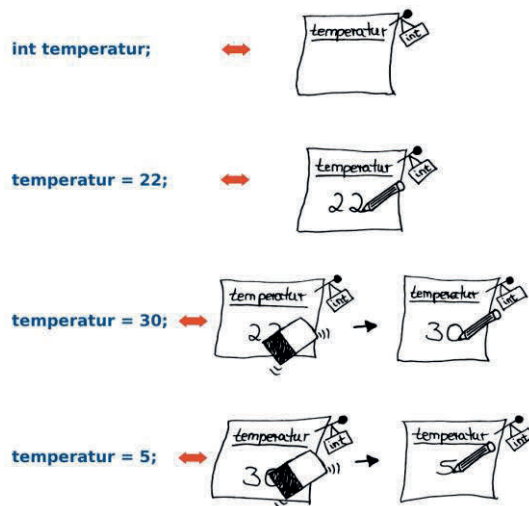


Abb. 2: Analogiemodell

Das für die SuS zunächst schwer zu verstehende Prinzip der Veränderlichkeit wird ebenfalls durch Abb. 2 veranschaulicht. Die SuS lernen, dass eine Variable immer nur genau einen Wert abspeichert und dass dieser Wert beliebig oft geändert werden kann. Der Datentyp wird zur Einführung vorerst nur erwähnt und als Attribut der Variablen behandelt. Im Analogiemodell wird zunächst eine Integer-Variablen mit dem Namen „temperatur“ deklariert. Anschließend wird diese Variable mit dem Wert 22 initialisiert. In zwei weiteren Schritten bekommt die Variable nacheinander den Wert 30 und anschließend den Wert 5 zugewiesen. Jeder dieser Schritte wird durch den Programmcode und ebenfalls durch eine zugehörige Grafik dargestellt und somit für die SuS greifbar und verständlich. Das Arbeitsblatt lässt sich gut mit den entsprechenden digitalen Mikrolerneinheiten kombinieren. Durch die wiederkehrende Anwendung in konkreten Programmierbeispielen wird nach und nach ein gutes Verständnis des Konzepts bei den SuS erreicht.

2.2 Einführung des Konzepts ‚Funktion‘ anhand eines hypothetischen Fallbeispiels

Das zwölfte Arbeitsblatt thematisiert das Konzept und die Verwendung von Funktionen. Die SuS lernen das Prinzip an einem hypothetischen Beispiel aus ihrer Erlebniswelt kennen, indem eine Funktion zum Pfannkuchen backen definiert wird (Abb. 3). Sie lernen, dass Programme mittels Funktionen in einzelne Teilbereiche bzw. Teilprobleme

gegliedert werden. Jede Funktion übernimmt dabei eine bestimmte Aufgabe und kann selber auch wieder in verschiedene Teilbereiche zerlegt werden.

Dieses Prinzip wird im ersten Beispiel des Arbeitsblatts veranschaulicht, indem eine Funktion „backeEinenPfannkuchen()“ definiert wird, die ihrerseits eine Reihe weiterer Funktionen aufruft. Die SuS lernen, dass das Hauptproblem, einen Pfannkuchen zu backen, in einzelne Teilprobleme zerlegt werden kann. Eins dieser Teilprobleme ist die Funktion „warteBisBraun()“, die wiederum zwei weitere Funktionen zur Lösung des Teilproblems aufruft.

```
void backeEinenPfannkuchen () {
    fettePfanneEin();
    mixeTeig();
    tuePortionInPfanne();
    warteBisBraun();
    drehePfannkuchenUm();
    warteBisBraun();
    serviereDenPfannkuchen();
}

void warteBisBraun () {
    while (guckUnterPfannkuchen () != braun) {
        warte (1 min);
    }
}
```

Abb. 3: Programmcode

In den abschließenden Aufgaben müssen die SuS unter Beweis stellen, dass sie sowohl die Definition und Verwendung als auch die korrekte Syntax bei der konkreten Implementation von Funktionen verstanden haben. Zusätzlich bekommen die SuS die Aufgabe, die Funktion „mixeTeig()“ selbstständig sinnvoll durch Zerlegung in Teilprobleme zu definieren.

3 Digitale Mikrolerneinheiten

„Im Informatikunterricht ist die Problematik, dass Schüler [und auch Schülerinnen] unterschiedlich rasch lernen und unterschiedlich rasch vorankommen, gravierender als in anderen Fächern.“ [HNR07] Aus dieser Überlegung heraus wurden die Lerneinheiten so entwickelt, dass sie von den SuS in individueller Geschwindigkeit bearbeitet werden können. Die ‚digitale‘ Wissensvermittlung geschieht über einzelne Mikrolerneinheiten (MLE), die jeweils in durchschnittlich fünf typisierte Abschnitte (Information, Programmieraufgabe, Multiple-Choice-Abfrage) unterteilt sind. Abb. 4 zeigt die Organisation von 102 Mikrolerneinheiten in 11 thematische Kapitel.

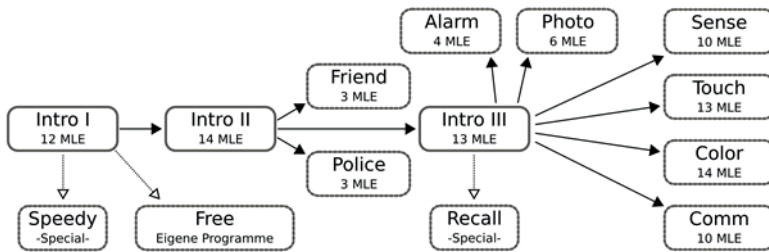


Abb. 4: Struktur der digitalen Lerneinheiten

Die Kapitel Intro I bis III vermitteln informatische Grundkonzepte und sollen von allen SuS bearbeitet werden, die Vertiefungskapitel sind zur inneren Differenzierung konzipiert. Die Lerneinheiten beinhalten dabei fundamentale Ideen der Informatik nach Andreas Schwill [Sc93] aus den Bereichen Algorithmisierung, strukturierter Zerlegung und Sprache (Syntax, Semantik).

4 Fazit

Es ist gelungen ein motivierendes Informatikprojekt für Einsteiger zu entwickeln, welches inzwischen von weit über 10000 Schülerinnen und Schülern der Sekundarstufe intensiv genutzt wurde. Der Unterricht mit BOB3 zeigt, dass die SuS den Themenbereich Informatik als interessant und erstrebenswert empfinden.

Literaturverzeichnis

- [Ha08] Haider, M.: Der Stellenwert von Analogien für den Erwerb naturwissenschaftlicher Erkenntnisse. Dissertation, Regensburg, 2008.
- [HNR07] Hartmann, W.; Näf, M.; Reichert, R.: Informatikunterricht planen und durchführen. Springer-Verlag, Berlin Heidelberg, 1., korrig. Nachdruck, S. 84, 2007.
- [Sc93] Schwill, A; Fundamentale Ideen der Informatik. ZDM – Zentralblatt für Didaktik der Mathematik, Heft 1, S.20-31, 1993.

Die BYTE Challenge – One bit at a time

Stefan Hildebrand¹, Carolin Neumann²

Abstract: Die BYTE Challenge ist ein digitaler Wettbewerb für Schüler*innen aus ganz Deutschland, der Begeisterung für Informatik, Informationstechnische Grundlagen sowie deren gesellschaftliche Bedeutung vermitteln und auf die Reduzierung bestehender Disparitäten im MINT-Bereich hinwirken soll.

Keywords: Wettbewerb; Begeisterung; Digitale Lehre; Schüler*innen; Informatik für alle

1 Einleitung

Während die Nutzung digitaler Kommunikationsmittel zur Aufrechterhaltung des Unterrichts in der Pandemie dringend erforderlich geworden ist, ist die informationstechnische Grundbildung in Deutschland sowohl bei den Schüler*innen, als auch bei den Lehrer*innen bei Weitem nicht einheitlich. Die Vermittlung von Informationstechnischen Grundlagen (ITG) sowie Informatik erfolgt in den Bundesländern abhängig von Schulform und Klassenstufe als eigenständiges Fach, als Teil anderer Fächer oder gar nicht. Während Schleswig-Holstein Informatik flächendeckend als Pflichtfach in der Sekundarstufe I einführt [Wi21], wird es z.B. in Berlin abhängig vom verfügbaren Lehrpersonal als Wahlpflicht-Fach angeboten, nur ITG wird als Pflichtfach unterrichtet [Sc20]. Dabei herrscht ein Mangel an Informatiklehrer*innen [Mu17].

Weiterhin bestehen dauerhaft Disparitäten in der Bildung und Auswahl von Studienfächern. In der Informatik beträgt der Frauenanteil unter den Studienanfänger*innen nicht einmal 20% (WS 2011/12 sowie WS2018/19) [St19, S. 240]

Entsprechend dieser Ausgangssituation ist unser Ziel, Schüler*innen für Technik und Informatik zu begeistern und digitale (Grund-) Kompetenzen gleichberechtigt an Schüler*innen ohne oder mit geringen Vorkenntnissen zu vermitteln und somit auf den Abbau bestehender Disparitäten hinzuwirken. Wir streben dabei eine Chancengleichheit in der Informatik-Bildung durch einen inklusiven Ansatz an. Wir beschränken unsere Bestrebungen deswegen nicht auf bestimmte, mutmaßlich benachteiligte, Gruppen, sondern erarbeiten Angebote, die allen Schüler*innen offen stehen. Diese sollen aber besonders attraktiv und hilfreich für diejenigen sein, die nicht bereits umfassende Förderung und Vorkenntnisse im Informatik-Bereich erhalten haben.

¹ Gesellschaft für Informatik e.V., Hochschulgruppe Berlin-Brandenburg, Anna-Louisa-Karsch-Straße 2, 10178 Berlin, Deutschland stefan.hildebrand@byte-challenge.de

² Gesellschaft für Informatik e.V., Hochschulgruppe Berlin-Brandenburg, Anna-Louisa-Karsch-Straße 2, 10178 Berlin, Deutschland info@byte-challenge.de

Als Format haben wir dazu im Mai 2020 einen komplett digital organisierten Wettbewerb gegründet – die BYTE Challenge. Diese findet zum ersten Mal vom 18. März bis zum 18. Juni 2021 statt, die Teilnehmenden kamen in diesem Jahr aus Klassenstufe 6-10 an Integrierten Oberschulen, Mittel- und Realschulen sowie Gymnasien und einer Schule für Gehörlose. Die ursprüngliche Zielgruppe war Klasse 7-10 an diesen Schultypen.

2 Organisation

Der Wettbewerb ist vollständig ehrenamtlich unter dem Dach der Hochschulgruppe Berlin-Brandenburg der GI ausgerichtet. Die Preise und die IT-Infrastruktur werden durch Sponsoren ermöglicht, die Inhalte werden mit zahlreichen Kooperationspartnern erstellt. Dazu gehören die Interviewpartner*innen unserer Online-Live-Seminare, gemeinsam verwendete Kursmaterialien sowie Input und Feedback zur Ausgestaltung der Kurse.

3 Der Wettbewerb – Die Herausforderung

Über drei Monate hinweg mit einem angepeilten Zeitaufwand von 4 h pro Woche bekommen die Teilnehmenden zunächst Wissen und Fertigkeiten vermittelt, soweit möglich, ohne spezifische Vorkenntnisse vorauszusetzen.

Kurs-Struktur Dazu haben wir Kurse in einem Lernmanagementsystem (LMS) erstellt, überwiegend im Video-Format. Die Sicherung der Ergebnisse erfolgt anschließend durch automatisch bewertete Aufgaben, als Quiz sowie durch offene Praxisaufgaben, bei denen abschließend eigene Spiele in Scratch selbstständig entwickelt werden. Die Bewertung der Spiele erfolgt durch das BYTE Team anhand eines eigenen Bewertungsbogens.

2021 wurden 20 Kurse mit äußerst unterschiedlichem Umfang aus den Themengebieten *Coding (in Scratch)* (3 Phasen – Einführung, angeleitetes Programmieren, selbstständig ein Spiel programmieren), *LIVE-Seminare* (8 Kurse à max. 1,5 h Interview/ Vortrag) und *Technik-Vertiefung* (25 Kurse Text-Bild zum Selbststudium) angeboten. 2022 sollen die Kurse jeweils ca. 15 Minuten dauern. Sie werden in drei Niveaustufen angesiedelt, wobei das inhaltliche, strukturelle und sprachliche Niveau zusammenpassen sollen. Um Barrieren abzubauen, wird sich das Einstiegs-Niveau an einfacher Sprache orientieren, besonders kurze Sinneinheiten enthalten und mittels Feedback von Sonderpädagog*innen gestaltet werden.

Nach unserer Erfahrung werden Video-Formate wesentlich besser verstanden und lieber verwendet, als Text-Bild-Kombinationen, wobei letztere aber kein festes Tempo vorgeben. Daher sollen 2022 flächendeckend untertitelte Videos für die Vermittlung von Stoff eingesetzt werden, ergänzt um Text-Bild-Transkripte. Auch das Material der LIVE-Seminare von 2021 wird in dieser Form verarbeitet. Zusätzlich sollen regelmäßige (mehrmals wöchentlich) Sprechstunden im Videokonferenz-Format angeboten werden.

Offene, aber zeitaufwendige Aufgaben wie das Erstellen von Scratch-Spielen haben verhältnismäßig wenige Rückläufer, diese jedoch mit beeindruckender Ergebnis-Qualität. Passend zum Niveau der Kurse und bei weitgehender Automatisierung der Korrekturen soll 2022 eine Mischung geschlossener, halb-offener und soweit möglich offener Aufgabentypen angeboten werden.

Preise Wir wollen keine womöglich demotivierende Konkurrenz unter den Teilnehmenden provozieren. Die Herausforderung in der BYTE Challenge besteht daher darin, die gestellten Aufgaben rechtzeitig und korrekt zu bearbeiten. Dadurch sammeln die Teilnehmenden Punkte in unserer „Währung“ \$Byte, die sie anschließend in unserem Byte-Store in Preise umtauschen und sich nach hause liefern lassen können. Bei den Preisen haben wir die Wünsche der Teilnehmenden abgefragt und neben einem Programmier-Lernspiel (Human Resource Machine) Oberbekleidung, Powerbanks, USB-Sticks, Schmuck, Geschirr und Naschereien ins Programm aufgenommen.

Einbindung von Lehrer*innen Wir haben dieses Jahr mehrfach die Anfrage erhalten, ganze Klassen/ Kurse geschlossen anmelden zu können. Auch insgesamt haben wir sehr positives Feedback von den Lehrer*innen bekommen und erfahren, dass sie die BYTE Challenge in den Unterricht einbinden möchten. Um ihre Schüler*innen kontinuierlich unterstützen zu können, wünschten sie sich mehrfach einen Zwischenstand des Bearbeitungsfortschritts sowie eine Abschluss-Urkunde mit den erreichten Punkten.

Während das Datenschutzkonzept 2021 die individuelle Einwilligung aller Eltern vorsah, streben wir an, 2022 durch Umgestaltung des Datenschutzkonzeptes zusätzlich direkt als Auftragsverarbeiter für die Schulen agieren zu können. Dann können die Lehrer*innen auch direkten Zugriff auf die Ergebnisse ihrer jeweiligen Schüler*innen erhalten und unmittelbar bei der Betreuung im Rahmen des Wettbewerbs unterstützen.

4 Auswertung

Die vorläufigen Anmeldestatistiken zeigen, dass unsere Teilnehmenden 2021 aus allen Bundesländern Deutschlands kommen, überwiegend 12-14 Jahre alt sind, keine Scratch- und Wettbewerbs-Erfahrung haben und durch ihre Lehrer*innen auf uns aufmerksam wurden. Das am Ende der LMS-Kurse erbetene Feedback der Teilnehmenden fällt auf der Skala „Super“, „Geht so“, „Nicht so gut“ und „Es gibt noch Dinge zu verbessern“ überwiegend positiv aus.

Der Anteil derjenigen, die die Aufgaben als zu schwer empfanden, stieg bei den fortgeschrittenen Programmier-Aufgaben (2. Teil von Phase 2 und Phase 3) ohne Video-Anleitung jedoch sprunghaft an. Vor allem bei den Online-Seminaren wurde mehrfach genannt, dass die Videos zu lang waren.

Eine abschließende Auswertung wird allerdings erst nach dem Ende des Durchlaufs 2021 möglich sein.

5 Ausblick – Digitalisierung der Digital Natives

Nach unseren Erfahrungen aus dem Durchlauf in 2021 bestehen drastische Unterschiede in der digitalen Vorbildung unserer diesjährigen Teilnehmenden. Unsere Zielgruppe in 2022 sind Schüler*innen von Klassenstufe 6-13 an allen Schulformen in Deutschland. Auch Schüler*innen mit Förderbedarf und körperlichen Einschränkungen möchten wir dabei einbinden.

Während laut der Rückmeldung von Lehrer*innen der Umgang mit digitalen Endgeräten, Dateien, E-Mail und Webseiten an Gymnasien i.A. gut funktioniert, weisen Real- und Hauptschüler*innen und vereinzelt auch deren Lehrkräfte hier wesentliche Defizite auf.

Um der Niveau-Spreizung im Umgang mit der Technik sowie ggf. aus dem ITG-/ Informatik-Unterricht zu begegnen, ist für 2022 ein Programm aus rund 100 Kursen à 15 Minuten aus fünf verschiedenen Themengebieten (Digital Skills, Grundlagen der Informatik, Technik-Vertiefung, MINT & Schnittstellen, Orientierungskurse) geplant. Davon sollen die Teilnehmenden ca. 30 Kurse belegen, die entweder sie selbst oder deren Lehrer*innen auswählen können.

Digital Skills erklären den grundlegenden Umgang mit digitalen Werkzeugen und das soziale Verhalten im Internet. **Grundlagen der Informatik** vermitteln Themen, die ähnlich auch im Informatikunterricht auftreten, von Computerbestandteilen über Scratch-Programmierung bis zu Algorithmen und Komplexität. **Technik-Vertiefung** gehen auf aktuelle Themen und Anwendungsgebiete der Informatik wie Künstliche Intelligenz und Virtuelle Realität ein. **MINT & Schnittstellen** zeigt die Verbindungen von Informatik in andere Disziplinen wie Wirtschaft, Jura, Pharmazie, Meteorologie usw. auf. Die **Orientierungskurse** zeigen die Möglichkeiten nach dem Schulabschluss auf, z.B. Universität, Fachhochschule, Ausbildung und Freiwilligendienste.

Abhängig von den Möglichkeiten des LMS soll die bewertete Ergebnis-Sicherung für diese Kurse über reine Quizzes hinausgehen.

Literaturverzeichnis

- [Mu17] Mueller, Dorothee: Berufswahl Informatiklehrkraft. 2017.
- [Sc20] Schwarz, Richard: , Informatikunterricht in den Bundesländern Deutschlands im Jahr 2020, 2020.
- [St19] Stanat, Petra; Schipolowski, Stefan; Mahler, Nicole; Weirich, Sebastian; Henschel, Sofie: IQB-Bildungstrend 2018 Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I im zweiten Ländervergleich. Waxmann Verlag GmbH, 2019.
- [Wi21] Wilkens, Andreas: Heise Online. May 2021.

EduInf - Education in Informatics

Erik Marx,¹ Anne Hamann,¹ David Baberowski,¹ Christin Nenner,¹ Antje Kazimiers,¹
Thiemo Leonhardt,¹ Nadine Bergner¹

Abstract: Das Schülerlabor EduInf der Professur für Didaktik der Informatik der Technischen Universität Dresden strebt die Begeisterung von Kindern und Jugendlichen und die gleichzeitige Vermittlung von Informatikkompetenzen an, indem eine Vielzahl von Workshops für Schüler:innen aller Altersstufen angeboten wird. In diesem Beitrag werden Angebote, Ziele und Besonderheiten des EduInf vorgestellt. So soll ein Einblick gegeben werden, wie im Rahmen des EduInf Nachwuchsförderung, Ausbildung von zukünftigen Lehrkräften der Informatik und Forschung im Bereich Informatikdidaktik miteinander verzahnt werden.

Keywords: Schülerlabor, Schüler:innen-Workshops, Teaching Lab, Informatische Bildung, Informatische Grundkompetenzen, Lehrkräftebildung, Lehramtsstudium

1 Aufgaben und Selbstverständnis

Die Professur für Didaktik der Informatik der Technischen Universität Dresden betrachtet es als wichtige Aufgabe, aufzuzeigen, wie interessant, vielseitig und gesellschaftlich relevant Informatik ist. Dieses Ziel wird im Rahmen der Lehrkräftebildung und Forschung im Bereich der Fachdidaktik Informatik verfolgt. Zusätzlich existieren außerschulische Lernangebote für alle Altersgruppen, die unter dem Label *EduInf - Education in Informatics* zusammengefasst und organisiert sind. Im Lernlabor EduInf werden Kinder, Jugendliche und junge Erwachsene in Workshops für die Informatik begeistert und Lehramtsstudierende erhalten die Möglichkeit praktische Lehrerfahrung zu sammeln. Im Mittelpunkt steht die Vermittlung einer realistischen Vorstellung des Fachgebiets und das Wecken von Interesse für die Informatik im Allgemeinen. Auf diese Weise sollen Schüler:innen motiviert werden, ein Studium oder eine Ausbildung im Bereich Informatik für sich zu erwägen. So wird zusätzlich ein Beitrag zur Fachkräftegewinnung geleistet. Dabei wird besonders die Mädchenförderung [Fr18], unter anderem über mädchenorientierte Workshops unter der Leitung von weiblichen Rolemodels, fokussiert.

¹TU Dresden, Professur für Didaktik der Informatik, Nöthnitzer Str. 46, 01187 Dresden, Deutschland {erik.marx|anne.hamann|david.baberowski|christin.nenner|antje.kazimiers|thiemo.leonhardt|nadine.bergner}@tu-dresden.de

2 Formate und Inhalte

Da durch die Angebote des EduInf verschiedene Zielgruppen angesprochen werden sollen, werden sowohl kurze Online-Workshops, ganztägige Präsenzangebote für komplette Schulklassen als auch mehrtägige Feriencamps für einzelne Interessierte angeboten. Thematisch deckt das EduInf ein breites Spektrum, von der Programmierung über 3D-Modellierung bis hin zu künstlicher Intelligenz, ab, wobei die einzelnen Inhalte jeweils für das jeweilige Format aufbereitet werden. Für alle Angebote gilt gleichermaßen, dass ein lernerzentrierter und aktivierender Ansatz verfolgt wird, bei dem ein selbstgesteuerter, kollaborativer Lernprozess angestrebt wird, in dem Lehrpersonen als Lernbegleiter:innen auftreten.

2.1 Formate

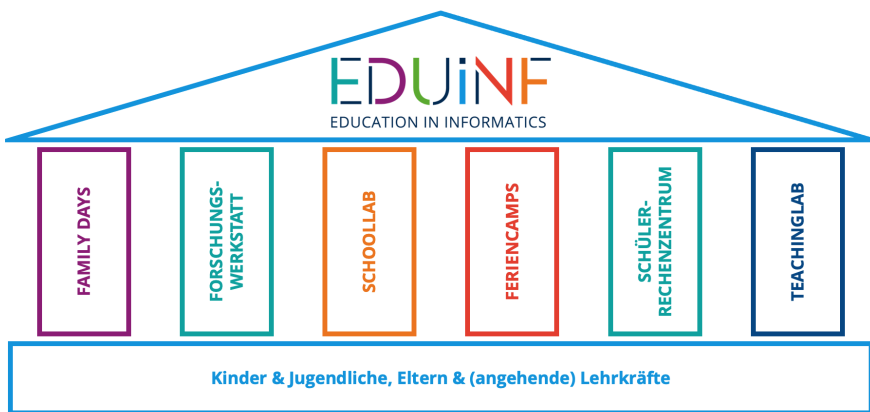


Abb. 1: Formate der Angebote des EduInf

In den *FamilyDays* ist ein „Hineinschnuppern“ ohne Vorkenntnisse für die ganze Familie möglich. Kinder, Eltern und Großeltern entdecken dabei gemeinsam spielerisch die Welt der Informatik, probieren sich beispielsweise in der Roboterprogrammierung und erforschen, wie das Internet funktioniert.

Die *Forschungswerkstatt Informatik* richtet sich direkt an Grundschulklassen, die spielerisch einen ganzen Tag in die Welt der Informatik eintauchen und sich beispielsweise mit der Sortierung, der Codierung, dem Internet und der Roboterprogrammierung beschäftigen. Das *Schoollab* lädt Schulklassen von der Grundschule bis zum Abitur zu einem Tag Informatik ein. In dem Schülerlabor mit flexiblen Möbeln, mobilen Arbeitsgeräten wie Laptops und Tablet und einer großen Auswahl an Lernmaterialien werden diverse Workshops zu verschiedenen Themen angeboten.

In *Feriencamps* können die Kinder und Jugendlichen ihr Interesse an der Informatik in ein- oder mehrtägigen Workshopreihen vertiefen und detaillierter als in normalen Workshops in ein Thema eintauchen. Zeitgleich erhalten die Schüler:innen einen Einblick in die

universitäre Lehre und Forschung. So wird auch für Schüler:innen aus ländlichen Gebieten ein Angebot geschaffen, wenn sonst nicht die Möglichkeit besteht, das EduInf außerhalb der Ferien zu besuchen.

Das **Schülerrechenzentrum** konzentriert sich auf die Talent- und Spitzenförderung in den Bereichen Informatik und Elektrotechnik. Hier werden Halbjahres- und Jahreskurse für verschiedene Niveaustufen angeboten, an denen die Jugendlichen regelmäßig einmal wöchentlich teilnehmen. Das Kursangebot reicht vom Einsteig in die Programmierung, über die Android-App Programmierung bis hin zur Robotik.

Die sechste Säule des EduInf bildet das **TeachingLab**, welches eng verzahnt mit der Lehramtsausbildung ist. Hier wird zukünftigen und aktiven Lehrkräften die Möglichkeit geboten, den Unterricht der Zukunft mitzugestalten und zu erproben. Beispielsweise werden innovative Ansätze für den Informatikunterricht erarbeitet, direkt mit Schulklassen erprobt und im Anschluss als freie Materialien (OER) zur Verfügung gestellt.

Zusätzlich zu diesen regulären Formaten besteht für Studierende die Möglichkeit Räume und Ausstattung des EduInf im Rahmen von Abschlussarbeiten z.B. zur Erprobung von Lernszenarien zu nutzen. Auf diese Weise entstehen regelmäßig neue Materialien, Workshops und Tools für den Einsatz im EduInf wie auch im Schulunterricht.

2.2 Inhalte

Die Angebote des EduInf umfassen mehrstündige bis mehrtägige Workshops (online oder in Präsenz), in denen die Teilnehmer:innen im Sinne des außerschulischen Lernens interessengesteuert und handlungsorientiert aktuelle Themen der Informatik entdecken [Jo09; Ov06]. Das aktuelle Themenangebot des EduInf erstreckt sich von einem blockbasierten ProgrammierEinstieg mit Scratch² oder Calliope mini³, über Projekte mit Mikrocontrollern bis zur Grafikprogrammierung mit der Programmiersprache Python unter der Nutzung der Online-Plattform CS Academy⁴ sowie komplexere Projekte mit dem App Inventor⁵. Daneben existieren weitere Materialien zum Thema KI und maschinellem Lernen. Ein Workshop zu 3D-Druck sowie Lernspiele für Multitouch-Displays und VR-Brillen befinden sich gerade in der Entwicklung. Besonderer Fokus liegt auf dem „be-greifbar machen“ der Themen, durch vielfältige *Hands-on-Materialien* [KTW07].

Neben den betreuten Workshopangeboten werden innerhalb des EduInfs auch eigenständig nutzbare Onlineangebote, wie Lernspiele, entwickelt, wodurch Interessierte auch zeit- und ortsunabhängig informatisch gefördert werden können.

Bei der Entwicklung der Konzepte und Materialien wird besonderer Fokus auf schülernahe Kontexte gelegt, um möglichst viele Kinder und Jugendliche, insb. auch die Mädchen, bei ihren eigenen Interessen abzuholen. Alle Materialien werden in Durchführungen mit Kindern und Jugendlichen evaluiert, beständig weiterentwickelt und als OER über die

² <https://scratch.mit.edu/>

³ <https://calliope.cc/>

⁴ <https://eu.academy.cs.cmu.edu/>

⁵ <http://appinventor.mit.edu/>

Webpräsenz⁶ des EduInf für Lehrkräfte, Studierende und Schüler:innen zur Verfügung gestellt.

3 Fazit und Ausblick

Ziel des EduInf ist es Kinder und Jugendliche für die Informatik zu begeistern, sie an diese zukunftsrelevante Disziplin heranzuführen, ihnen ihr Informatiktalent bewusst zu machen und dieses nachhaltig zu fördern. Gleichzeitig soll ein Verständnis dafür geschaffen werden, was Informatik eigentlich ist und mit Vorurteilen aufgeräumt werden. Ein langfristiges Ziel des EduInf ist es, vor allem mehr junge Frauen für die Informatik zu gewinnen und insgesamt mehr begeisterte Nachwuchs-Informatikerinnen und Informatiker auszubilden. Gleichzeitig erhalten Lehramtsstudierende die Gelegenheit im geschützten Raum des Teaching Labs Erfahrungen für den Unterricht von Morgen zu sammeln. Die kontinuierlichen Angebote ermöglichen außerdem das fortlaufende Weiterentwickeln und Evaluieren von Materialien und Formaten. Im EduInf bietet sich so die Möglichkeit Forschung, Ausbildung und Nachwuchsförderung auf eine einzigartige Weise zu verbinden.

Literatur

- [Fr18] Friedrich, J.-D.; Hachmeister, C.-D.; Nickel, S.; Peksen, S.; Roessler, I.; Ulrich, S.: Frauen in Informatik: Welchen Einfluss haben inhaltliche Gestaltung, Flexibilisierung und Anwendungsbezug der Studiengänge auf den Frauenanteil? Centrum für Hochschulentwicklung, 2018.
- [Jo09] Johannes, W.: Forschendes Lernen: Lernen im „Format“ der Forschung. *Journal Hochschuldidaktik* 20/2, S. 4–7, Sep. 2009.
- [KTW07] Klahr, D.; Triona, L. M.; Williams, C.: Hands on what? The relative effectiveness of physical versus virtual materials in an engineering design project by middle school children. *Journal of Research in Science Teaching* 44/1, S. 183–203, 2007.
- [Ov06] Overwien, B.: Informelles Lernen—zum Stand der internationalen Diskussion. *Informelles Lernen im Jugendalter. Vernachlässigte Dimensionen der Bildungsdebatte* 2/, S. 35–62, 2006.

⁶ <https://tu-dresden.de/inf/eduinf>

Das I in MINT von Anfang an

Bildungsangebote zur informatischen Bildung der Stiftung „Haus der kleinen Forscher“

Christine Günther ¹, Christin Nenner ²

Abstract: Informatische Bildung spielt bereits in der Grundschule eine wichtige Rolle, denn Informatiksysteme begegnen Grundschulkindern häufig in ihrem Alltag. Die Kinder bei der Beantwortung ihrer Fragen zu Themen der Informatik nicht allein zu lassen, stellt Pädagoginnen und Pädagogen in Kita, Hort und Grundschule vor die Herausforderung Themen der Informatik in ihre pädagogische Arbeit zu integrieren. In diesem Beitrag wird das Fortbildungsangebot der Stiftung „Haus der kleinen Forscher“ im Bereich der informatischen Bildung vorgestellt. Das Angebot gibt pädagogischen Fach- und Lehrkräften vielfältige Anregungen für ihre pädagogische Praxis und Hintergründe zu informatischer Bildung.

Keywords: Informatische Bildung, Kita, Grundschule, Fortbildung, Blended Learning, forschendes Lernen

1 Das Bildungsangebot der Stiftung „Haus der kleinen Forscher“ zum Thema informatische Bildung

In unserer digital vernetzten Welt kommen schon die Jüngsten mit informatischen Phänomenen, Gegenständen und Situationen in Kontakt. Als Teil ihrer Lebenswelten sollte ihnen das Entdecken und Erforschen dieser Inhalte schon in der Kita und der Grundschule ermöglicht werden. Die Stiftung „Haus der kleinen Forscher“ engagiert sich daher im Bereich der frühen informatischen Bildung und bietet sowohl für Pädagoginnen und Pädagogen in Kita und Hort als auch für Lehrkräfte der Grundschule spezielle Fortbildungen zu diesem Thema an.

Die Bildungsangebote der Stiftung leisten einen Beitrag zum aktuellen Querschnittsthema der Digitalisierung, indem sie es pädagogischen Fach- und Lehrkräften ermöglichen, sich dem Thema der informatischen Bildung niedrigschwellig und aktiv zu nähern und sich selbst als Lernende zu erfahren. Die Fortbildungen setzen auf Vernetzung und Austausch und ermutigen dazu, miteinander zu lernen und mit Kindern gemeinsam die Welt von heute und morgen zu erforschen. Pädagogische Fach- und Lehrkräfte entdecken Themen wie Verschlüsselung und Datenverarbeitung, Programmieren und Robotik. Sie stellen

¹ Stiftung „Haus der kleinen Forscher“, Abteilung Forschung und Entwicklung, Rungestraße 18, 10179 Berlin
christine.guenther@haus-der-kleinen-forscher.de

² Technische Universität Dresden, Professur Didaktik der Informatik, Nöthnitzer Straße 46, 01187 Dresden,
christin.nenner@tu-dresden.de, <https://orcid.org/0000-0002-5230-4343>

Verbindungen zu den Lebenswelten der Kinder her und erfahren, wie sie die Kinder in ihren Bildungsprozessen der informatischen Bildung begleiten können - und zwar mit und ohne Computer. Die Fortbildungen und pädagogischen Materialien bieten dazu fachliche Grundlagen und Praxisideen.

Die Konzeptionierung der Angebote fußt auf langjähriger Fundierung. Dabei bildet u. a. die Expertise zu Zielen und Gelingensbedingungen für informatische Bildung im Elementar- und Primarbereich [Be18], die von Fachexpertinnen und -experten der informatischen Bildung für die Stiftung „Haus der kleinen Forscher“ angefertigt wurde, die Grundlage für die Konzeptionierung und die konkrete Ausgestaltung der Fortbildungen.

Diese Expertise liefert unterschiedliche Empfehlungen. Für die Konzeptionierung der Bildungsangebote wurde z. B. die Kombinationen aus den folgenden Inhalts- und Prozessbereichen angeregt, da sie besonders vielversprechend für gelingende informatische Bildung in Elementar- und Primarbereich sind:

- Interaktion mit und Exploration von Informatiksystemen
- Repräsentieren/Darstellen und Interpretieren von Daten und Information
- Modellieren und Implementieren von Algorithmen und Programmen
- Nachdenken über und Bewerten des Zusammenhangs von Informatik, Mensch und Gesellschaft

Außerdem wurde für die Begleitung des forschenden Lernens der Kinder in Informatik analog zum etablierten Forschungskreis für die Naturwissenschaften der Informatikkreis entwickelt [Fr19]. Er unterscheidet zwei Prozesse, die für die informatische Bildung grundlegend sind: Das Erkunden bestehender Informatiksysteme und das Gestalten eigener digitaler Artefakte [Sc17]. Das informatische Forschen ist dabei in sechs Schritte aufgeteilt, bei dem sich Phasen des Denkens und Handelns abwechseln [St17].

Die Bildungsangebote der Stiftung zur Informatischen Bildung konzentrieren sich außerdem auf drei Kategorien von Zugängen zur Informatik: Unplugged, softwarebasiert oder über kindgerechte Robotiksysteme [Br19]. Diese Zugänge sind besonders geeignet für das Gelingen früher informatischer Bildung.

Beim Ansatz „Unplugged“ geht es darum, dass Kinder Phänomene der digitalen Welt **ohne** digitale Geräte wie z. B. Computer erforschen. Den Schwerpunkt des „Unplugged“-Ansatzes [BWF15] bilden Erkenntnisse darüber, dass und wie sich gedankliche Prozesse automatisieren lassen.

Der zweite empfohlene Zugang zur Informatik ist das softwarebasierte Programmieren mit kindgerechten Programmierumgebungen: das Coding. Es gibt Programmierumgebungen wie bspw. ScratchJr, die bereits von sehr jungen Kindern ohne Schreib- und Lesekenntnisse genutzt werden können, da die entsprechenden Befehle durch Symbole

dargestellt werden. Das Programmieren wird als eine kreative Tätigkeit erlebt, denn die Kinder können damit eigene, digitale Produkte erschaffen.

Die dritte, für die Kinder besonders motivierende Zugangsweise zur Informatik stellen kindgerechte Robotiksysteme dar. Denn die Mädchen und Jungen können mit ihrer Hilfe direkt in der physisch erfahrbaren Welt programmieren. Die Robotiksysteme geben ihnen außerdem direkte Rückmeldung über den Erfolg ihrer Programmierung.

Diese und weitere Empfehlungen der Expertinnen und Experten liefern eine wesentliche Grundlage dafür, wie informatische Bildung in Kita, Hort und Grundschule gestaltet werden kann und wurden in den verschiedenen Bildungsangeboten der Stiftung umgesetzt.

2 Die Fortbildung „Informatik entdecken – mit und ohne Computer“

Mit dem ersten Bildungsangebot "Informatik entdecken – mit und ohne Computer" betrat die Stiftung „Haus der kleinen Forscher“ 2017 pädagogisches Neuland. Die eintägige Fortbildung behandelt die Bedeutung von Informatik im Alltag und bietet den Teilnehmenden viele Praxisideen für die Bildungsarbeit mit Kindern in der Altersgruppe von 3-10 Jahren. Gemeinsam mit den Kindern können sie z. B. erforschen, wie Computer Zahlen sortieren, wie sich mit Hilfe einer algorithmischen Vorgehensweise Spiele gewinnen lassen oder wie man Nachrichten verschlüsselt. Die Fortbildung "Informatik entdecken – mit und ohne Computer" ist so angelegt, dass Einrichtungen keine digitalen Endgeräte einsetzen müssen, um eigene Lerngelegenheiten zu gestalten. In der Fortbildung können die pädagogischen Fach- und Lehrkräfte den Einsatz von Programmierumgebungen und Robotiksystemen jedoch ausprobieren und reflektieren.

3 Die Blended-Learning-Fortbildung „Informatische Bildung“ für Grundschullehrkräfte

Das zweite Bildungsangebot der Stiftung zur informatischen Bildung wurde im Projekt PRIMA! entwickelt. In diesem Projekt bearbeitet die Stiftung grundlegende Fragen zu MINT-Bildung im Grundschulunterricht und entwickelt dazu passende Fortbildungen für Lehrkräfte. Bisher existieren Blended-Learning-Fortbildungen zu den Themen Energiebildung und informatische Bildung.

Die Fortbildungen von PRIMA! sind als Blended-Learning-Fortbildungen konzipiert. Online-, Präsenz- und Praxiserprobungsmodule greifen über einen Zeitraum von etwa zehn Wochen ineinander und betonen jeweils die Stärken der verschiedenen Modi. Präsenztage fördern persönlichen Austausch und praktisches Ausprobieren. Onlineteile ermöglichen zeitsouveränes Lernen und kollaboratives Arbeiten. Zu reflektieren,

inwiefern dieses Prinzip zu einer effektiven Fortbildung führt, und die Fortbildungen iterativ zu verbessern, gehört zu den Kernaufgaben im Projekt. So wurde bspw. die bundesweite Pilotierung der Fortbildung im Pre-Post-Design evaluiert. Fachwissen, fachdidaktisches Wissen, Einstellungen, Motivation und Selbstwirksamkeit waren dabei die Themenbereiche der Evaluation. Bei den Lehrkräften konnten signifikante Zuwächse in den Bereichen der Selbstwirksamkeit, des Fachwissens und fachdidaktischen Wissens verzeichnet werden.

4 Literaturverzeichnis

- [BWF15] Bell, T., Witten, I., Fellows, M., 2015. cs unplugged, verfügbar unter <http://csunplugged.org/books/>
- [Be18] Bergner, N., Köster, H., Magenheim, J., Müller, K., Romeike, R., Schulte, C. & Schroeder, U.: Zieldimensionen informatischer Bildung im Elementar- und Primarbereich. In Stiftung Haus der kleinen Forscher (Hrsg.), Frühe informatische Bildung – Ziele und Gelingensbedingungen für den Elementar- und Primarbereich. Wissenschaftliche Untersuchungen zur Arbeit der Stiftung „Haus der kleinen Forscher“ (Band 9). Verlag Barbara Budrich, Opladen, Berlin, Toronto, 2018.
- [Br19] Brünger, K., Franke-Wiekhorst, A., Griffiths, K., Günther, C., Radtke, M.: Informatische Bildung für Kinder im Kita- und Grundschulalter – ein Konzept zum entdeckenden und forschenden Lernen für die Praxis In: GDSU-Journal, Heft 9, 2019, S. 106-117.
- [Fr19] Franke-Wiekhorst, A., Günther, C., Brünger, K., Magenheim, J., Romeike, R.: „Der Informatikkreis“: Kinder von drei bis zehn Jahren beim Forschen in Informatik begleiten – ein methodisch-didaktisches Material. In: GDSU-Journal, Heft 9, 2019, S. 10-21.
- [Sc17] Schulte, C., Magenheim, J., Müller, K., Budde, L.: The Design and Exploration Cycle as Research and Development Framework in Computing Education. In: Beiträge zur EDUCON; Global Engineering Education Conference, IEEE, 2017.
- [St17] Stiftung „Haus der kleinen Forscher“ Informatikkreiskarte, 2017.

InfoLab Saar - Schülerlabor Informatik der Universität des Saarlandes

Nadine Kirsch,¹ Kerstin Reese^{1,2}, Verena Wolf¹

Abstract: Das „InfoLab Saar - Schülerlabor Informatik an der Universität des Saarlandes“ ist 2018 auf Initiative von Prof. Verena Wolf gegründet worden, um die Informatik an saarländischen Schulen zu stärken. Neben Events für Schüler*innen bietet das InfoLab Saar auch Veranstaltungen für Studierende und Fortbildungen für Lehrkräfte an. In diesem Beitrag werden Ausstattung und Angebot des InfoLab Saar beschrieben. Das Feedback der Teilnehmenden wird aufgezeigt und ein Ausblick auf geplante Projekte gegeben.

Keywords: Schülerlabor; Lehrkräftebildung; Informatische Bildung für alle Lehramtsstudierenden; Online-Kurse; Grundschule; Sekundarstufe 1; Sekundarstufe 2; Berufliche Schulen

1 Ziele des InfoLab Saar

Das InfoLab Saar möchte die Informatik an saarländischen Schule stärken. Um das Ziel zu erreichen wird auf mehreren Ebenen angesetzt. Neben dem Erarbeiten von Inhalten werden alle Akteure angesprochen. Das InfoLab bietet für Lehramtsstudierende Seminare mit informatischen Inhalten an, für Lehrkräfte Lehrkräftefortbildungen und für Schüler*innen Kurse im Labor, in der Schule und online. Bei der Etablierung des Pflichtfachs Informatik unterstützt das InfoLab die Qualifizierung von Lehrkräften.

2 Das InfoLab Saar an der Universität des Saarlandes

Für das InfoLab Saar steht ein großer und gut ausgestatteter Raum an der Universität zur Verfügung. In dem Raum können Gruppen mit bis zu 32 Personen arbeiten (siehe Abb. 1). Die Einrichtung ist flexibel gestaltet worden und erlaubt verschiedene Unterrichtsformen. Es stehen Laptops, Tablets und verschiedene Hardwaresysteme in ausreichender Anzahl zur Verfügung. Pandemiebedingt konnte der Raum leider bisher kaum genutzt werden. Das InfoLab Saar bietet seit Frühjahr 2020 Online-Kurse an.

¹ Universität des Saarlandes, Didaktik der Informatik, Campus E1 3, 66123 Saarbrücken, Deutschland, s8nakirs@stud.uni-saarland.de, [kerstin.reese|verena.wolf]@uni-saarland.de

² Das diesem Beitrag zugrundeliegende Vorhaben wird im Rahmen der gemeinsamen „Qualitätsoffensive Lehrerbildung“ von Bund und Ländern mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 01JA2035 gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autorinnen und Autoren.



Abb. 1: Raum des InfoLab Saar

3 Zielgruppen

Das InfoLab Saar versucht an mehreren Stellen, die informatische Bildung zu fördern.

- Schulklassen können das InfoLab an der Universität besuchen oder Mitarbeiter*innen des InfoLab fahren an die Schulen.
- Das InfoLab Saar unterstützt die Durchführung von Seminaren für Lehramtsstudierende, in denen informatische Bildung vermittelt wird.
- Für Lehrkräfte aller Schulformen werden Lehrkräftefortbildungen angeboten.
- Ebenso werden Kurse für Privatpersonen angeboten.

4 Informatische Themen

Da das InfoLab Saar Angebote für alle Jahrgänge von der Grundschule (meist ab Lesekompetenz) bis zum Abitur macht, sind die Themen sehr vielfältig. In 2020 wurden

Kurse zu Themen angeboten, die sich ohne Probleme online unterrichten lassen: Scratch für Kinder ab 10 Jahren, Python ab der 6. Klasse, Künstliche Intelligenz und Ethik ab der 8. Klasse. Die Online-Kurse dauern pro Einheit in der Regel 60 Minuten und beinhalten eine oder zwei Einheiten. Seit Frühjahr 2020 hat das InfoLab Saar mehr als 60 Online-Events durchgeführt. Eine Übersicht der Kurse findet sich auf der Webseite des InfoLab: infoLab.cs.uni-saarland.de. Vor der Online-Phase in 2020 hat das InfoLab Saar auch Kurse zu Calliope mini, Lego Mindstorms und Dronenprogrammierung mit Python angeboten. Zielgruppen waren hier Grundschulen, Förderschulen, Gymnasien, Gemeinschaftsschulen und berufliche Schulen. Wenn es wieder möglich ist, werden diese Themen wieder aufgegriffen. Weitere Hardware steht zur Verfügung: Ozobots, Arduino, Lego Wedo, Bob-3, Dash, ... Ebenso wird angeboten, mit Unterstützung des InfoLab-Teams an regelmäßig stattfindenden Events teilzunehmen, z.B. Jugendwettbewerb Informatik, Informatik Biber, „Codeweek“ oder „Hour of Code“. Ein Kurs mit Stationen zum Thema „Was ist Informatik?“ ist online getestet worden. Die Verbindung der Informatik mit anderen Schulfächern, z.B. Englisch oder Chemie, soll weiter ausgebaut werden.

5 Vernetzung vor Ort

Über das Projekt MoDiSaar (siehe modisaar.de), das Teil der vom Bundesministeriums für Bildung und Forschung geförderten „Qualitätsoffensive Lehrerbildung“ ist, ist das InfoLab Saar mit verschiedenen Fachbereichen der Universität vernetzt und steht im Austausch mit anderen Schülerlaboren. Es gibt regelmäßigen Kontakt und gemeinsame Veranstaltungen mit dem Ministerium für Bildung und Kultur, dem Wirtschaftsministerium, der Begabtenförderung und dem Landesinstitut für Pädagogik und Medien. Im Rahmen der „Codeweek“ findet jedes Jahr ein gemeinsames Event mit der Firma SAP in St. Ingbert statt. Zum CoderDojo Saar und den Makerspaces HackSaar und FabLab Saar findet häufiger Austausch statt und führt zu neuen Projekten.

6 Bewerbung des InfoLab Saar und seiner Angebote

Da Schulen in 2020 viel Unsicherheit erlebt haben, waren sie für das InfoLab Saar kaum erreichbar. Daher hat das InfoLab Saar sein Angebot für Privatpersonen erweitert. So konnten trotzdem Kinder und Jugendliche erreicht und Material erprobt werden. Um auch online präsent zu sein, besitzt das InfoLab Saar von Gründung an eine Webseite, über die Events angekündigt und dokumentiert werden und das InfoLab erreichbar ist: infoLab.cs.uni-saarland.de. Mit dem Start der Online-Kurse kam ein Newsletter dazu, dessen Empfängerliste stetig wächst und mittlerweile mehr als 400 Kontakte umfasst. Die Veranstaltungen des InfoLab Saar werden auch über die sozialen Medien beworben, jedoch ist der Newsletter deutlich erfolgreicher. Meist sind die Veranstaltungen 24 Stunden nach Versand des Newsletters ausgebucht. Schwieriger ist es aktuell, Schulen und Lehrkräfte zu erreichen. Hier werden im nächsten Schuljahr die erprobten Materialien verstärkt beworben und von den guten Erfahrungen damit berichtet.

7 Feedback der Teilnehmenden

Die Erfahrungen bei der Durchführung von Kursen für Kinder und Jugendliche sind durchweg positiv. Die Kinder sind motiviert und folgen den Kursinhalten konzentriert und erreichen meist das Kursziel. Da die Kurse meist der erste Kontakt mit informatischen Inhalten sind, wird viel Wert auf eine gute Stimmung im Kurs gelegt und dass die Kinder am Ende ein Produkt erstellt haben, das sie verstehen und das funktioniert. Die Kinder sollen stolz sein, auf das, was sie erstellt haben. Im Zweifel wird der Inhalt reduziert, was aber nur sehr selten nötig ist. Der Wechsel zu Online-Kursen hat am Erfolg nichts geändert. Die Gruppengröße musste reduziert werden, da es in einer Videokonferenz schwieriger ist, zu erkennen, ob das Tempo des Kurses angemessen ist. Die Teilnehmer*innen haben nach unseren Beobachtungen Spaß und erreichen das gewünschte Kursziel. Mit einem Online-Kurs werden Kinder und Jugendliche aus dem ganzen Saarland erreicht, was als Vorteil bewertet wird. Als Einrichtung der Universität des Saarlandes sollte das Saarland der Einzugsbereich des InfoLab sein. Die Angebote für Kinder bis ca. zur 6. Klasse sind sehr stark nachgefragt. Jugendliche ab der 7. Klasse sind etwas schwerer zu erreichen und geben im Kurs wenig Feedback. Vermutlich sind diese Beobachtungen aber mit dem Alter der Teilnehmer*innen zu begründen sind.

8 Ausblick

Das InfoLab Saar wird weiter daran arbeiten, die Informatik an saarländischen Schulen zu stärken. Es ist schön, wenn wieder Kurse im Labor an der Universität angeboten werden können. Ebenso sollen wieder Schulen besucht werden, um zu zeigen, dass informatische Themen in jeder Klassenstufe altersgerecht unterrichtet werden können und auch unterrichtet werden sollten. Im nächsten Schuljahr sollten wieder mehr Lehrkräftefortbildungen angeboten werden. Das Seminar „Informatische Bildung“ für angehende Lehrkräfte verschiedener Schulstufen wird weiterhin angeboten. Da die Online-Kurse stark nachgefragt werden und ihre eigenen Vorteile haben, werden sie nicht ganz aufgegeben werden. In den Kursen soll nach der Wissensvermittlung am Anfang, mehr freie Zeit für kreative Arbeit der Teilnehmenden eingeplant werden. Mehr Angebote für die Sekundarstufe 1 passen gut in die Entwicklung der Schulen im Saarland. Informatik soll an allen weiterführenden Schulen im Saarland Pflichtfach werden. Im September 2021 startet eine Qualifizierungsmaßnahme für das Fach Informatik von Klasse 5 - 10, für die sich jede saarländische Lehrkraft einer weiterführenden Schule bewerben kann. Mittelfristig könnten Informatiklehrkräfte für die Sekundarstufe 1 an der Universität grundständig ausgebildet.

INGA – Informatik Garten: Eine Initiative zur Förderung von Informatikkompetenzen im Kindergarten

Mayr Franziska¹, Waldvogel Bettina², Schumann Nicole³, Rüdüsüli Cornelia⁴

Abstract: Mit der Initiative „Informatik Garten (INGA)“ werden Brücken gebaut zwischen der Systematik der Informatik-Fachdidaktik und altersgerechten, fächerübergreifenden, spielerischen und individualisierenden Lernangeboten für den Kindergarten. Ab Frühjahr 2021 wurden erste INGA-Boxen an den Pädagogischen Hochschulen Schaffhausen und Thurgau entwickelt und im Rahmen des von der Hasler-Stiftung finanzierten INGA-Projektes [Ha2021] in mehreren Kindergärten erprobt. Konzeptionelle Überlegungen zur Entwicklung der INGA-Boxen und erste Erfahrungen werden präsentiert.

Keywords: Computational Thinking, Informatikkompetenzen, Kindergarten⁵.

1 Einleitung

INGA-Boxen sind kindgerecht ausgestaltete Kisten der Grösse ca. 40 x 60 cm mit Informatik-Rätseln und einer kleinen Schatztruhe mit Zahlenschlössern und sollen der Förderung von Informatikkompetenzen auf der Kindergartenstufe dienen. Mit einer interaktiven Hörspielanleitung tauchen die Kinder in die Rahmengeschichte ein:

Hallo miteinander. Ich bin Harry der Hase, ich wohne ganz in der Nähe des Kindergartens. [...] Die Schatztruhe ist [...] abgeschlossen, und ich muss nun verschiedene Rätsel lösen, um die richtigen Codes herausfinden [...]. Könnt ihr mir helfen, liebe Kinder?

Die Kinder müssen nun selbständig die Rätsel lösen, um damit die Zahlenschlösser zu öffnen und das Geheimnis der Schatztruhe zu lüften.

Mit den INGA-Boxen werden die Informatikkompetenzen basierend auf den Vorgaben des Deutschschweizer Lehrplan 21 gefördert [DEDK2016]: Datenstrukturen (ordnen und sortieren), Algorithmen (formale Anleitungen befolgen) und Informatiksysteme (Geräte

¹ Pädagogische Hochschule Thurgau, Unterer Schulweg 3, CH-8280 Kreuzlingen, franziska.mayr@phtg.ch

² Pädagogische Hochschule Schaffhausen, Ebnatstrasse 80, CH-8200 Schaffhausen, bettina.waldvogel@phsh.ch

³ Pädagogische Hochschule Schaffhausen, Ebnatstrasse 80, CH-8200 Schaffhausen, nicole.schumann@phsh.ch

⁴ Pädagogische Hochschule Schaffhausen, Ebnatstrasse 80, CH-8200 Schaffhausen, cornelia.ruedisueli@phsh.ch

⁵ In der Schweiz bildet der Kindergarten den ersten Teil der obligatorischen Volksschule, darum werden die Begriffe Schule, Unterricht und Lehrpersonen verwendet. Die Kinder sind in der Regel 4 bis 6 Jahre alt.

bedienen). Die Aufgaben und Rätsel in den INGA-Boxen decken die Bereiche Algorithmen und Datenstrukturen ab. Die Bedienung der interaktiven Hörspielanleitung fördert Kompetenzen im Bereich von Informatiksystemen.



Abb. 1: Blick von oben in eine INGA-Box: Die Schatzkiste mit den Zahlenschlüssern. Die Materialien und Hinweise zu den vier Informatikaufgaben sind je in einer Ecke der INGA-Box zu finden.

2 Fachlicher Hintergrund der Informatik

Bei der Entwicklung der INGA-Boxen werden sowohl theoretische Grundlagen in Hinblick auf das Lehren und Lernen im Kindergarten als auch fachdidaktische Aspekte der Informatik zusammengeführt. Ein besonderer Fokus liegt auf dem kreativen Problemlösen und den spielerischen Zugängen zu individualisierten Lernangeboten rund um Informatikthemen.

Der amerikanische Psychologe Jérôme Bruner forderte, dass sich der Unterricht an den Strukturen der Fachwissenschaften orientieren soll [Br60]. Diese sollen helfen, das Gelernte zu vernetzen und zu strukturieren. Für die Kindergartenstufe bietet sich der Zugang zu den fachwissenschaftlichen Strukturen der Informatik über das Computational Thinking an [Wi06], insbesondere auch mit „unplugged“ Methoden, ohne digitale Geräte [BWF15]:

- Dekomposition: Zerlegung eines komplexen Problems in Teilprobleme.
- Mustererkennungen: Wiederholende Muster und Ähnlichkeiten erkennen.
- Abstraktion: Fokussierung auf das Wesentliche. Weglassen von Details.

- Algorithmus: Schritt-für-Schritt Lösung oder Regeln zum Finden der Lösung.

Mit den Aufgaben und Rätseln in den INGA-Boxen werden diese vier Ansätze des Computational Thinking umgesetzt, welche die im Lehrplan 21 geforderten Bereiche Algorithmen und Datenstrukturen abdecken.

3 Kindgerechtes Lehren und Lernen im Kindergarten

Die INGA-Boxen bieten nicht nur Material für die Förderung der Informatik-Kompetenzen, sondern stellen wichtige Verbindungen zu anderen Fach- und überfachlichen Kompetenzen, wie z.B. Deutsch, Natur, Mensch und Gesellschaft bzw. Sozialkompetenzen oder Feinmotorik.



Abb. 2: Strukturierte Anordnung im "Karottenfeld" als Datenstruktur Array (links). Problem bestehend aus einem Zahlencode (Mitte), Entwicklung von Lösungsstrategien im Team (rechts).

Die teilweise abstrakten Konzepte des Computational Thinking lassen sich anhand konkreter Alltagshandlungen und Spielgelegenheiten besonders gut veranschaulichen und «begreifbar» machen. Bereits Piaget (zit. nach [Sa02]) weist dem Spiel und der Exploration der Kinder für die Entwicklung der kreativen Imagination eine wichtige Rolle zu. Dadurch werden auch überfachliche Kompetenzen gefördert.

Mit Blick auf den inklusiven Unterricht soll das Lernen am gemeinsamen Gegenstand [Fe05] bereits ab Schuleintritt fokussiert werden. Alle befassen sich demnach mit der gleichen Problemstellung zentral sind jedoch die individuellen Aneignungsprozesse des einzelnen Kindes und ihre persönliche Lernentwicklung.

4 Begleitforschung und erste Erfahrungen mit den INGA-Boxen

Die Idee des Projekts besteht darin, dass Aufgaben auf der Basis vom Computational Thinking gesammelt, adaptiert oder neu entwickelt und in Hinblick auf einen systematischen Kompetenzaufbau zu einer stufengerechten, motivierenden INGA-Box arrangiert werden. Parallel zur Erstellung der INGA-Boxen findet eine qualitative

Begleitforschung statt, mit welcher der Erwerb von Informatikkompetenzen unmittelbar beobachtet und analysiert wird. Ziel ist es, zu erfahren, inwiefern die Kinder mit diesem Lernarrangement tatsächlich fachspezifische Informatikkompetenzen erwerben und erweitern können.

Im Frühling 2021 wurden vier Zweierteams bei ihrer erstmaligen Begegnung mit einer INGA-Box gefilmt. Das erhobene Datenmaterial wird qualitativ, mittels Videosequenzanalyse, ausgewertet.

Eine erste Analyse des Videomaterials zeigt folgende Erkenntnisse:

- Die Kinder arbeiten motiviert, konzentriert und ausdauernd an der INGA-Box.
- Mithilfe der interaktiven, in der Anleitung integrierten Tipps und individueller Unterstützung der Erhebungsperson gelingt es allen Zweierteams, die Schatztruhe zu öffnen.
- Die Kinder weisen sehr unterschiedliche Problemlösestrategien auf.
- Die Mehrschichtigkeit der Aufgaben ist für die Kinder anspruchsvoll.
- Im Verlauf des Arbeitsprozesses gelingt es den Kindern zunehmend, fokussiert auf das Ziel (Öffnen der Schatztruhe) hinzuarbeiten.

Weitere thematische INGA-Boxen werden entwickelt und in Schulklassen der Kindergarten- sowie Primarstufe eingesetzt, damit die Begleitforschung vorangetrieben werden kann.

Literaturverzeichnis

- [BWF15] Bell, T., Witten, I. H., & Fellows, M.: Computer Science Unplugged. An enrichment and extension programme for primary-aged students. <https://csunplugged.org/>, 2015.
- [Br60] Bruner, J.: The Process of Education. A Landmark in Educational Theory. Harvard Univ. Press, 1960.
- [DEDK16] Deutschschweizer Erziehungsdirektorenkonferenz (D-EDK): Lehrplan 21, Luzern, 2016.
- [Fe05] Feuser, G.: Behinderte Kinder und Jugendliche zwischen Integration und Aussonderung. 2. Auflage. Wissenschaftliche Buchgesellschaft, Darmstadt, 2005.
- [Ha21] Hasler Stiftung. Projekt Informatik-Garten: Förderung von Informatikkompetenzen im Kindergarten», <https://haslerstiftung.ch/>, 2021-2022.
- [Sa02] Saracho, O.: Young Children's Creativity and Pretend Play. Early Child. Development and Care, 172(5), S. 431-438, 2002.
- [Wi06] Wing, J. M.: Computational thinking. Communications of the ACM, 49(3), 33–35, 2006.

InfoSphere – Schülerlabor Informatik

Ein außerschulischer Lernort für Schülerinnen und Schüler aller Altersstufen und Schulformen

Annabell Brocker¹, Christina Schramm¹, Ulrik Schroeder¹

Abstract: Ein Schülerlabor ist ein außerschulischer Lernort, der Schülerinnen und Schülern einen zielgruppengerechten Zugang zu aktuellen (MINT)-Themen über den schulischen Betrieb hinaus ermöglichen soll. Das InfoSphere fokussiert dabei die informatischen Themengebiete und deren Relevanz in Alltagssituationen. Hauptanliegen ist dabei, den Schülerinnen und Schülern einen selbstständigen und explorativen Zugang zu neuen und spannenden Themen zu bieten und sie durch individuelle Betreuung auf diesem Weg zu unterstützen. Konkret bietet das InfoSphere begleitete Präsenz- und Online-Workshop an und stellt darüber hinaus (Selbst-)Lernmaterialien als OER bereit.

Keywords: Schülerlabor, außerschulischer Lernort, Informatik, Workshops, Selbstlernangebote, spielerisch.

1 Das InfoSphere

Das InfoSphere – Schülerlabor Informatik ist ein seit 2010 an der RWTH Aachen verankerter außerschulischer Lernort, der Informatik erlebbar macht, indem Schülerinnen und Schülern ab Klasse 3 eine große Auswahl an verschiedenen informatischen Themengebieten spielerisch nähergebracht. Hierzu zählen beispielsweise die (blockbasierte/textuelle) Programmierung, der Umgang mit und das Verständnis von Mikrocontrollern und dem Internet, die App- und Spiele-Entwicklung oder auch die Kryptographie. Dabei sind die Angebote des Schülerlabors für alle Interessierten kostenlos. Es werden sowohl Kurse für ganze Schulklassen als auch Wochenendkurse bzw. Ferienworkshops für individuelle Gruppen angeboten. Das allgemeine Angebot umfasst Präsenz-Module, Online-Module sowie (Selbst-)Lernmaterialien. [Be16]

Pandemiebedingt hat sich das InfoSphere im letzten Jahr verstärkt auf die Entwicklung von Online-Modulen sowie Selbstlernangeboten konzentriert. Die Selbstlernangebote sind unter anderem als Leitprogramme konzipiert, welche einerseits von Schülerinnen und Schülern zuhause selbstständig bearbeitet, andererseits als Grundlage bzw. Ergänzung des Informatikunterrichtes von Lehrkräften eingesetzt werden können. Des Weiteren werden Moodle-Kurse angeboten, welche ebenfalls zum selbstständigen Arbeiten als auch für den Einsatz in der Schule geeignet sind. Zusätzlich wurde, um die Reichweite auch in Zeiten

¹ RWTH Aachen, Lehr- und Forschungsgebiet Informatik 9, Ahornstr. 55, 52074 Aachen, {a.brocker, schramm, schroeder}@informatik.rwth-aachen.de

physischer Distanz zu erweitern, die Präsenz des Schülerlabors InfoSphere in seinen Social-Media-Kanälen (Instagram Twitter, Facebook) ausgebaut.

Während der Durchführung von (Präsenz- oder Online-) Modulen werden die Schülerinnen und Schüler unter der Betreuung von engagierten studentischen Hilfskräften (meist Informatiklehramtsstudierende) durch die Arbeitsaufträge und informatischen Themengebiete begleitet. Dabei steht die selbstständige Auseinandersetzung der Teilnehmenden mit den Arbeitsmaterialien im Vordergrund. Zumeist werden diese im Team bearbeitet, damit zusätzlich die kooperative Kompetenz gefördert wird.

Das Schülerlabor hat sich verschiedenen Netzwerken angeschlossen. Einerseits ist es Mitglied in deutschlandweiten Verbänden wie LernortLabor², andererseits Mitglied in regionalen Verbänden wie bspw. zdi³ oder eXploregio.net⁴. Darüber hinaus beteiligt sich das InfoSphere regelmäßig an Veranstaltungen seitens der RWTH Aachen, wie bspw. verschiedenen Formen von Schüleruniversitäten. Durch die gute Vernetzung und die Kooperation mit vielfältigen Partnern können neue innovative Konzepte erarbeitet und in das Schülerlabor integriert werden.

2 go4IT! – Mädchenförderung

Das im InfoSphere verankerte Projekt „go4IT!“ richtet sich speziell an Mädchen der Jahrgangsstufen 6 bis 8 und verfolgt das Ziel, das Selbstvertrauen dieser Mädchen in der Nutzung und Gestaltung von Technik zu fördern sowie das Interesse für die Beschäftigung mit Technik und Programmierung anzuregen. [Le15]

Dazu werden zunächst sogenannte Basis-Workshops angeboten, in welchen die Mädchen Roboter (Lego Mindstorms NXT 2.0⁵ oder B-O-B-3s⁶) zunächst eigenständig zusammenbauen/löten und diese anschließend programmieren. Die Basis-Workshops finden an zwei aufeinanderfolgenden Tagen während der regulären Unterrichtszeiten direkt in den Schulen statt.

Darauf aufbauend werden regelmäßig in den Schulferien zusätzliche, dreitägige Aufbau-Workshops angeboten, in deren Rahmen Mädchen, welche bereits an einem Basis-Workshop teilgenommen haben, vertiefende Programmierkonzepte sowie Projekte verfolgen.

² LernortLabor – Bundesverband der Schülerlabore, <https://www.lernortlabor.de/>, Stand 18.05.2021

³ Zukunft durch Innovation, <https://www.zdi-portal.de/>, Stand 18.05.2021

⁴ eXploregio.net, <http://www.explore.io.net/de/index.htm>, Stand 18.05.2021

⁵ LEGO Mindstorms NXT 2.0, <https://www.lego.com/de-de/service/buildinginstructions/8547>, Stand 19.05.2021

⁶ B-O-B-3, <https://www.bob3.org/de/>, Stand 19.05.2021

3 Angebote für Lehrkräfte und Lehramtsstudierende

Das Schülerlabor richtet sich nicht nur an Schülerinnen und Schüler, sondern auch an Lehrkräfte sowie Lehramtsstudierende.

Alle Arbeitsmaterialien können als PDF von der Webseite frei heruntergeladen werden. Darüber hinaus können Lehrkräfte das InfoSphere formlos kontaktieren, um die Rohmaterialien der einzelnen Angebote zu erhalten und diese bspw. für den Unterricht anzupassen und dann einzusetzen. Da sämtliche Materialien unter der „Creative Commons Namensnennung - Weitergabe unter gleichen Bedingungen 4.0 International Lizenz“⁷ (CC-BY-SA) lizenziert sind, können diese jederzeit unter Beibehaltung der gleichen Lizenz verändert werden.

Lehramtsstudierende der Informatik erhalten im Rahmen des Projektes LeBiAC bzw. Lehr-Lern-Labor Informatik⁸ die Möglichkeit, schon früh im Studium Lehrerfahrungen mit Schülerinnen und Schülern zu sammeln. Den direkten Kontakt erhalten sie bspw. bereits im 4. bzw. 6. Fachsemester. Hier sollen die Studierenden bereits existierende Arbeitsmaterialien des InfoSpheres altersgerecht erweitern, verbessern bzw. verändern und diese im praktischen Einsatz erproben.

4 Statistiken und Erfahrungsberichte

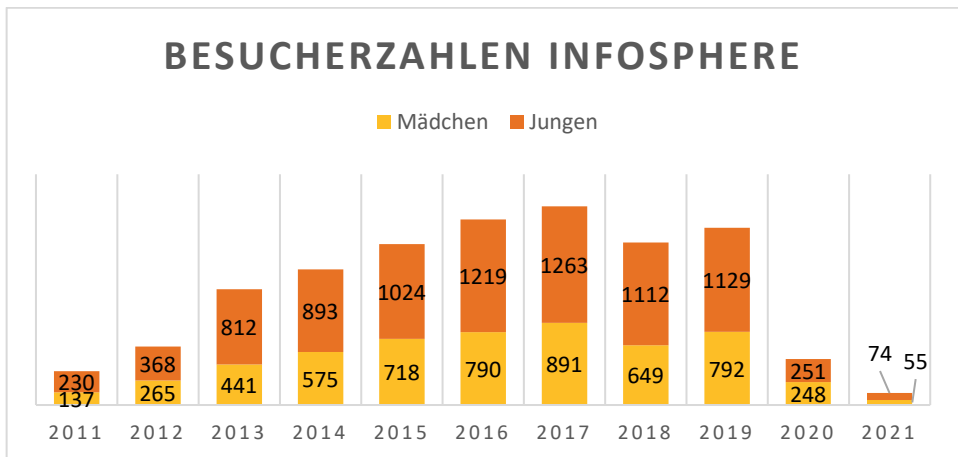


Abb. 1: Teilnehmer*innenzahlen von (Präsenz- und Online-) Modulen des InfoSpheres und go4IT! (beinhaltet nicht die Anzahl der Downloads bzw. Anfragen der Selbstlernangebote).

⁷ Attribution-ShareAlike 4.0 International (CC BY-SA 4.0), <https://creativecommons.org/licenses/by-sa/4.0/>, Stand 18.05.2021

⁸ LeBiAC, http://www.lebiac.rwth-aachen.de/?page_id=213, Stand 18.05.2021

Das InfoSphere konnte seit 2011 bis zum Beginn der Pandemie meist ein Wachstum seiner Besucherzahlen verzeichnen. Dies liegt einerseits an der stetigen Vergrößerung des Angebotes, andererseits an wachsenden Kooperationen. Pandemiebedingt sind die Besucherzahlen in den Jahren 2020 und 2021 stark gesunken, da der gesamte „klassische“ Schülerlaborbetrieb zunächst gestoppt werden musste. Anstelle der Durchführung von Modulen wurde diese Zeit vor allem für die Entwicklung von Selbstlernangeboten und Online-Modulen genutzt. Abbildung 1 beinhaltet nicht die Anzahl der Downloads bzw. Anfragen der Selbstlernangebote, sondern ausschließlich die Besucherzahlen der Präsenz- und Online-Module.

Die Umstellung von Präsenz- auf Online-Betrieb hat dem Schülerlabor neue Möglichkeiten eröffnet. Die Reichweite der online verfügbaren Module konnte auf das gesamte Bundesgebiet ausgeweitet werden, sodass nicht nur Schülerinnen und Schüler aus der Region an den InfoSphere-Workshops teilnehmen können. Darüber hinaus sind die Durchführungen der Online-Module im Klassenverband unabhängig von der aktuell möglichen Unterrichtsform. So ist es nicht relevant, ob die Schülerinnen und Schüler sich im Distanz-, Hybrid- oder Präsenzunterricht befinden. Die Zusammenarbeit kann entweder vor Ort in der Schule erfolgen oder aber über das Feature der Breakout-Sessions des verwendeten Videokonferenzsystems. Bei der Konzeption wurde darauf geachtet, die technischen Hürden möglichst gering zu halten: Voraussetzungen sind lediglich ein PC/Laptop, eine stabile Internetverbindung sowie ein Mikrofon/Headset. Weitere digitale Medien, wie bspw. Tablet/Smartphone oder Mikrocontroller, werden durch Simulatoren ersetzt. Zudem wurden die Plattformen, auf denen die Schülerinnen und Schüler arbeiten, so gewählt, dass die Kinder und Jugendliche nicht unbedingt einen eigenen Account erstellen müssen. Allerdings zeigte sich in der Praxis für die Arbeit des InfoSphere-Teams, dass die Vorbereitung und Durchführung von Online-Modulen einen größeren personellen Aufwand erzeugt, da administrative Dinge (bspw. das Einholen des Einverständnisses der Eltern zur Nutzung des Videokonferenzsystems oder die Bereitstellung von Arbeitsblättern) ebenso wie die Betreuung und die Problemlösung über die Distanz während des Moduls deutlich komplexer sind.

Zukünftig wird das InfoSphere die digitalen Formate beibehalten und regelmäßig anbieten. Dennoch werden diese Formate einen Präsenz-Betrieb, sobald dieser wieder möglich ist, nur ergänzen, da der direkte Kontakt zu den Lernenden wie auch zu den Lehrkräften als essentiell angesehen wird.

Literaturverzeichnis

- [Be16] Bergner, N.: Konzeption eines Informatik-Schülerlabors und Erforschung dessen Effekte auf das Bild der Informatik bei Kindern und Jugendlichen., 2016.
- [Le15] Leonhardt, T.: Etablierung eines begabungsfördernden Lernumfeldes für Mädchen im Bereich Informatik., 2015.

Die intoMINT App als Einstieg in die Informatik

Korinna Bade¹, Stephanie Böhnke¹, Grit Marschik¹, Tobias Scheidat¹

Abstract: So vielseitig Interessen und Charaktere von Schüler:innen sind, so facettenreich sind auch die Bereiche von MINT und ganz besonders auch die Informatik. Die große Vielfalt erschwert es Schüler:innen, den Überblick zu bewahren, ihr eigenes Potential zu erkennen und weiterzuentwickeln. Die intoMINT-App setzt genau dort an und richtet sich gezielt an Mädchen und junge Frauen. Mit ihr werden verschiedene MINT-Projekte und -Experimente an einer Stelle gebündelt. Schülerinnen können so aus einem Pool an vielfältigen MINT-Aktivitäten wählen, welche ihre Neugier und Interesse wecken, und so über einfache Schritt-für-Schritt-Anleitungen das Interesse an MINT-Themen finden. Die Themenvielfalt in der Informatik erstreckt sich beispielsweise von 3D-Modellierung und Dechiffrierungsverfahren bis hin zu umfangreichen Programmierprojekten mit einem Einplatinencomputer. So lassen sich schnell Erfolge erzielen und das Selbstbild stärken.

Keywords: Mädchen, Mitmach-Angebote, Bildungsapp, außerschulische Angebote.

1 Motivation

Mit der intoMINT-App sollen Mädchen ab Klasse 8 spielerisch an Mathematik, Informatik, Naturwissenschaften und Technik (MINT) herangeführt werden. Laut einer von Microsoft durchgeführten Studie verlieren Mädchen ab dem Alter von etwa 16 Jahren [Mi17], das Interesse an technischen Berufen, welches oft nicht zurückkehrt. Gründe dafür sind u.a. mangelnde Vorbilder, fehlende positive Erfahrungen oder ein diffuser Überblick über Bereiche und Einsatzmöglichkeiten im späteren Berufsleben. Das führt zu Fachkräftemangel und einer niedrigen Frauenquote [In20]. Dagegen wollen wir mit der 2019 erschienenen intoMINT-App² wirken. Mit kleinen Projekten und Experimenten können Mädchen ab der 8. Klasse kritik- und klischeefrei verschiedene MINT-Bereiche durch eigenständiges Nachahmen der Schritt-für-Schritt-Anleitungen erproben und darüber geleitet Berufs- und Studienmöglichkeiten erkunden. Durch App-gestütztes individuelles Feedback erhalten sie Zuspruch und lernen, dass auch Fehlschläge positive Auswirkungen haben können. Stand Juni 2021 wurde die App über 2.750-mal installiert. Im Folgenden stellen wir die App kurz vor, erläutern insbesondere die Informatik-orientierten Projekte und zeigen, wie der breitere thematische Fokus der App dazu beitragen kann, Mädchen für die Informatik zu begeistern.

¹ Hochschule Anhalt Fachbereich Informatik und Sprachen, Lohmannstraße 23, 06366 Köthen, Deutschland, korinna.bade@hs-anhalt.de, stephanie.boehnke@hs-anhalt.de, grit.marschik@hs-anhalt.de, tobias.scheidat@hs-anhalt.de

² www.intomint.de

2 Die intoMINT 4.0 App

Der Kern der intoMINT-App (Abb. 1 links) sind die so genannten MINT-Aktivitäten - fachspezifische Projekte und Experimente, die mit einer bebilderten Schritt-für-Schritt-Anleitung zum Nachmachen einladen und zwar ganz real und nicht rein virtuell. Begleitet wird die eigentliche Anleitung durch Hinweise zur Vorbereitung, Hintergrundwissen, einen Quiz, ein passendes Berufsbild oder Studienfach sowie ein Feedbackformular, mit dem die Nutzerinnen die Aktivität bewerten können. Zudem können die Nutzerinnen ihre Ergebnisse per Foto oder Text hochladen und erhalten dafür Punkte in der App (Mintis). Diese repräsentieren den App-Fortschritt und können später gegen virtuelle Gegenstände im virtuellen MINT-Labor eingetauscht werden. Die eingereichten Ergebnisse werden manuell und individuell bewertet und durch das positive Hervorheben von Erreichtem die Nutzerinnen weiter bestärkt. Über dieses App-Design nähern sich die Nutzerinnen also über spannende Projekte den verschiedenen MINT-Themen. Das eigentliche Fachgebiet steht zunächst nicht im Vordergrund, sondern das Projektthema. Durch das Selbermachen und das Feedback lernen die Nutzerinnen ihre Stärken kennen. Mithilfe der zusätzlichen Informationen wird der Bogen zum Fachgebiet sowie zu Studium und Beruf geschlagen und damit Stück für Stück Fachinteresse und Berufsorientierung aufgebaut.

Unter den aktuell 139 vorhandenen MINT-Aktivitäten befinden sich 19 spezifische Informatikprojekte, sowie 35 Informatik-nahe Projekte. Diese sind entweder Einzelthemen wie z.B. zu Kryptowährung, Chiffrierverfahren oder Chatbots oder thematisch übergreifend, so dass sich die Nutzerinnen mit mehreren Aktivitäten zu einem Thema befassen können. Dazu gehören verschiedene kleine Projekte zum Basteln und Programmieren mit dem BBC micro:bit mit Online-Entwicklungsumgebung für Programmier-einsteigerinnen, wo neben dem eigentlichen Programmieren auch die handwerkliche Kreativität eine Rolle spielt, sowie die 3D-Modellierung und Animation, wo u.a. die räumliche Vorstellungskraft sowie der Umgang mit Vektoren / Achsen geschult wird.

Da die App einen möglichst niedrighschwelligem Zugang bieten soll, werden die Projekte mit so wenig kostenintensiven Materialien durchgeführt wie möglich. Bei den Informatikprojekten nutzen wir daher ausschließlich kostenfreie Software, wie etwa Blender (www.blender.org) für die 3D-Modellierung und Animation oder Dialogflow (dialogflow.cloud.google.com) für Chatbots. Bei den handwerklichen Projekten, als deren Basis wir den BBC micro:bit verwenden, lassen sich Zusatzkosten nicht vermeiden. Über gesponserte Starterkits, die wir in begrenzter Stückzahl kostenfrei verschenken, unterstützen wir besonders motivierte Nutzerinnen zusätzlich. Die micro:bit-Projekte wurden insbesondere für Programmieranfängerinnen ausgelegt und nutzen daher die micro:bit-eigene MakeCode-Webseite (makecode.microbit.org) unter Verwendung der Blockprogrammierung sowie JavaScript. Das Heranführen an das Programmieren folgt den Leitfragen: „Was machen wir?“ und „Warum machen wir das so?“ Damit wird die Denkweise beim Programmieren trainiert. Neben reinen Programmier-Projekten, wie „Die Macht des Zufalls“, „Erstelle einen Kompass“ oder „Thermometer - Den Temperaturen auf der Spur“, gibt es auch Projekte, bei denen ein Endprodukt entsteht, wie z.B.



Abb. 1: links: Screenshots intoMINT-App; rechts: Eindruck einer Teilnehmerin

„Erstelle ein Bewässerungssystem für deine Blume“ oder „Erstelle dir dein eigenes Tamagotchi“. Zudem gibt es zeitintensivere Projekte, bei denen anstelle von vorgefertigten Elektrobauteilen auf Haushaltswaren, wie Alufolie, Papier, Kleber und Zahnstocher, zurückgegriffen wird und auch Kreativität und Bastelgeschick gefordert ist. Zu diesen Projekten gehören „Das Papierklavier“ und „Baue deinen eigenen Controlller“. Bei einigen Projekten entstehen reale Spiele, wie z.B. bei „Spiele Schere-Stein-Papier gegen den Micro:bit“, „Der heiße Draht“ oder „Programmiere eine einfache Form von Space Invaders“. Die Programmier-Projekte variieren im zeitlichen Umfang (von ca. fünf Minuten bis zwei Stunden) und in der Komplexität, dargestellt als Schwierigkeitsgrad und Zeitaufwand in der Projektübersicht. Geschult werden einfachste elektrotechnische Grundlagen, der sichere Umgang mit Hard- und Softwareanwendungen und natürlich der visuellen Programmierumgebung. Durch die Vielfalt der Programmierprojekte ist eine breite Varianz an Problemstellungen gegeben, mit denen die Lösungskompetenz trainiert und die Logik hinter einer Programmiersprache gefestigt wird. Informatik-nahe Projekte runden das Angebot ab und schulen u.a. den Umgang mit Zahlen und Zahlensystemen. Neben dem klassischen Projekt zum Binärsystem, werden auch das Dezimal-, Quadratzahlen- sowie das Duodezimalsystem im Rahmen einer MINT-Aktivität praktisch erläutert. Nützliches und anwendbares Wissen wird zudem auch zu Themen wie „Fibonacci-Zahlen“, der „Zahl Pi“, vedische Mathematik, Tangram, Maßeinheiten, Quadratur des Kreises und Zufall in der App angeboten. Aus dem Technik-Genre bietet die App u.a. Projekte an, mit denen das räumliche Vorstellungsvermögen und das Wissen um technische Prinzipien erläutert wird wie z.B. das Thema Brücken- oder Turmbau.

Die App steht über die App-Stores für Android und iOS zur Verfügung. Auf dieser Basis wurde mit der intoMINT Challenge von März bis August 2020 ein Event durchgeführt, um durch den Wettbewerbscharakter und damit verbundene Preise die Nutzung

anzuregen. Dem Aufruf sind knapp 700 Mädchen aus ganz Deutschland gefolgt. Dabei haben 31% der aktiven App-Nutzerinnen im Schnitt 6,9 Informatik- oder Informatik-nahe Projekte bearbeitet. Die dabei durchgeführte Evaluierung mit In-App-Befragung, Onlinefragebogen und Telefoninterviews ergab, dass bei über 70% der Mädchen, dank der App, Interesse an einem MINT-Beruf geweckt wurde. Mit Fokus auf die Informatik konnten 57% der Nutzerinnen für einen Beruf interessiert werden. An der in Abb. 1 exemplarisch aufgeführten Rückmeldung lässt sich erkennen, dass durch die App auch Mädchen an Informatik herangeführt werden konnten, die daran vorher noch kein Interesse hatten. Da die App fachlich breiter ausgerichtet ist, sind die Berührungsängste geringer und der Kontakt mit der Informatik kann sich „ungeplant“ ergeben.

3 Fazit

Mit der intoMINT-App ist es Mädchen und jungen Frauen möglich, sich mit MINT und im Speziellen der Informatik vielseitig und langfristig zu beschäftigen. Gleichzeitig erlangen sie so einen Einblick in die verschiedenen Bereiche und Berufe, was ihnen hilft, sich besser auf ihre Zukunft vorzubereiten. Das individuelle Feedback und der Kontext der Challenge wecken Motivation und bringen sie dazu, sich dauerhaft mit den Themen auseinander zu setzen. Gamification-Elemente, wie das geplante MINT-Labor, Badges (siehe Abb. 1) und Punkte, verleihen zusätzlich das Gefühl, etwas geschafft zu haben. Die ursprünglich für den Einsatz in der Freizeit konzipierte App lässt sich dank ihrer inhaltlichen Vielfalt auch im Schulunterricht im Rahmen bestimmter Formate sehr gut einsetzen. Gerade die eigenständige Beschäftigung mit den Aktivitäten, die auch hervorragend in kleinen Gruppen erarbeitet werden können, bietet sich für eine Auflockerung im Unterricht oder auch für ganze Projekttage oder nachmittägliche Arbeitsgemeinschaften an. Die Neuauflage der intoMINT-Challenge oder vergleichbarer Eventformate ist auch für die Zukunft geplant.

Das Verbundvorhaben „intoMINT 4.0: Begeisterung und Engagement für MINT wecken und ausbauen – Entwicklung einer App zur Berufs- und Studienorientierung für Schülerinnen im MINT-Bereich“ der Hochschule Anhalt und der Otto-von-Guericke-Universität Magdeburg wird mit Mitteln des Bundesministeriums für Bildung und Forschung (BMBF) unter den Förderkennzeichen 01FP1623 und 01FP1625 gefördert. Für den Inhalt sind die Autor:innen verantwortlich.

4 Literatur

- [In20] Institut der deutschen Wirtschaft Köln: MINT-Herbstreport 2020: MINT-Engpässe und Corona-Pandemie: kurzfristige Effekte und langfristige Herausforderungen, 2020.
- [Mi17] Microsoft: Why Europe's girls aren't studying STEM, Whitepaper, 2017.

Make Up Your MINT: Multimediale Popup-Lernlabore

Korinna Bade¹, Stephanie Böhnke¹ und Andrea Hirschel¹

Abstract: Die richtige Studienwahl, die Spaß bringt, finanzielle Freiheiten eröffnet, Sicherheit verspricht und eine individuelle Karrieregestaltung ermöglicht – wer will das nicht? Insbesondere mit einem Informatik-Studium kann all das erreichbar sein, denn Digitalisierung spielt in jedem Lebensbereich eine Rolle und die „Grand Challenges der Informatik“² bieten eine Menge Herausforderungen, die von den Fachkräften von heute und morgen gemeistert werden müssen. Der Weg zum passenden Studium ist jedoch oft steinig, die Entscheidung schwierig und Vorbilder Mangelware. Unser MINT-Orientierungsangebot „Make up your MINT“ unterstützt junge Frauen in dieser Findungsphase. In Online-Lernlaboren können sie ihre Berufsvorstellungen festigen und mit praktischen Erfahrungen untersetzen.

Keywords: MINT-Interessensförderung für Mädchen; Lernlabor; online; außerschulisch

1 Einleitung

Fördern, schulen, Vorbild sein und dabei Perspektiven zeigen: das sind die Ansätze, die im vom Europäischen Sozialfonds geförderten Projekt „Make up your MINT“ seit 2015 in vielerlei Formaten für Schülerinnen ab der 10. Klasse umgesetzt werden. Und über allem steht der Anspruch, divers, klischeefrei und begeistert an einen Bereich heranzuführen, der bis heute als „Männerdomäne“ und „nerdig“ stark stereotypenbelastet ist: nämlich u.a. an die Informatik. Dort, wo es Schulen an Zeit und Ressourcen mangelt und dort, wo Eltern aufgrund fehlender Infrastruktur im ländlichen Raum kaum Möglichkeiten haben, ihren Töchtern Förderung im Gebiet der Informatik zu ermöglichen, setzt das Projekt an. 2015 startete die Gleichstellungsinitiative „Make up your MINT“ mit dem Anspruch, Vorbilder sichtbar zu machen, mediale Aufmerksamkeit für die Unterpräsenz von Frauen in MINT zu schaffen und junge Frauen letztlich sowohl praktisch als auch interaktiv an MINT-Themen heranzuführen und über längere Zeit hinweg zu begleiten. Erfahrungen aus der ersten Projektphase (2015 – 2018) [HB21] haben bereits gezeigt, dass eine große Barriere in den großen Distanzen zu finden ist, welche zur Wahrnehmung außerschulischer Angebote von Schülerinnen und Eltern im ländlichen Raum überbrückt werden müssen. Auch aufgrund des Lehrermangels und damit fehlenden Arbeitsgemeinschaften und Wahlpflichtkursen³ sollte das Angebot möglichst flexibel und bestenfalls (auch in Einzellektionen) innerhalb der Schule nutzbar sein.

¹ Hochschule Anhalt, Fachbereich Informatik und Sprachen, Lohmannstraße 23, 06366 Köthen, Deutschland
korinna.bade@hs-anhalt.de, stephanie.boehnke@hs-anhalt.de, andrea.hirschel@hs-anhalt.de

² <https://gi.de/grand-challenges>

³ <https://www.news4teachers.de/2020/07/lehremangel-bildungsminister-streicht-unterricht-zusammen/>

Daraus entstanden die Online-Lernlabore. In verschiedenen Kursen mit multimedial aufbereiteten Lernmaterialien können Teilnehmerinnen in eigener Geschwindigkeit tiefer in ein Teilthema der Informatik einsteigen. Das Selber-Machen, Anwendungsbezug, niedrige Zugangsvoraussetzungen, schnelle Erfolgserlebnisse und individuelle Betreuung kennzeichnen das Angebot. Eine Lernmanagementumgebung bietet ein interaktives Lernerlebnis, persönliche Betreuung wirkt motivierend, Präsenz in Social-Media-Kanälen gewinnt Aufmerksamkeit bei der Zielgruppe.

2 Ausgestaltung der Lernlabore

Zentrales Element der Kurse sind Anleitungsvideos⁴, in denen Schritt-für-Schritt das Projekt vorgeführt wird. Die Art der Videos ist angelehnt an typische Tutorialvideos für Do-It-Yourself-Projekte, wie sie vielfältig auf YouTube zu finden sind. Durch die Einbettung der Videos in die Lernplattform Moodle kann jedoch zusätzlich eine interaktive Lernumgebung gestaltet werden, die durch das bloße Einstellen von Videos auf YouTube nicht erzielt wird. Da YouTube aber als Plattform von unserer Zielgruppe stark frequentiert wird, beginnen unsere Kurse genau dort. Mit Hilfe eines Werbevideos⁵ machen wir auf den Kurs aufmerksam und stellen dessen Inhalte kurz vor. Der eigentliche Lernlabor-Kurs findet wie in Abb. 1 skizziert innerhalb der Moodle-Umgebung statt.

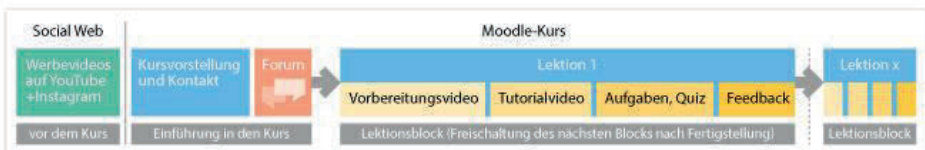


Abb. 1: Kursablauf

Die eigentlichen Inhalte eines Kurses werden in einzelne Lektionen gegliedert. Diese werden nacheinander freigeschaltet, so dass die Schülerinnen zu Beginn des Kurses nicht überfordert sind und neugierig bleiben – sie bewegen sich quasi immer ein Level weiter. Jede Lektion enthält ein Tutorialvideo und wird um weitere lernförderliche Materialien ergänzt. Ein Vorbereitungsvideo stellt sicher, dass alle Materialien für die Lektion im Vorfeld vorbereitet werden. Jede Lektion verfolgt ein kleineres Teil-Lernziel, auf das hingearbeitet werden soll. Mit Abschluss jeder Lektion kann somit direkt ein Lernfortschritt und Erfolgserlebnis durch die Teilnehmerin wahrgenommen werden, was zum Durchhalten ermutigt. Am Ende fügen sich diese Einzelschritte zu einem großen Gesamtbild zusammen, wie z.B. einem kleinen, eigenen Animationsfilm oder verschiedene Anwendungsprojekte mit unterschiedlicher Komplexität, die „benutzt“ werden können und Nutzerinteraktionen über verschiedene Kanäle wie Töne, Licht und

⁴ Ein Beispiel aus dem Blender-Kurs: <https://www.youtube.com/watch?v=XbwDax8FVlk>

⁵ Ein Beispiel aus dem Blender-Kurs: https://www.youtube.com/watch?v=ch0R_TQFZ4

Bewegung ermöglichen. Von Beginn an werden detaillierte Schritt-für-Schritt-Anleitungen gegeben, die selbst von völlig Unerfahrenen Punkt für Punkt nachgemacht werden können. In Folgelektionen werden dann in den Aufgabenstellungen die wachsenden Kenntnisse berücksichtigt und erlauben in Teilen die selbstständige Auseinandersetzung mit dem jeweiligen Thema. Weiterführende Aufgabenstellungen regen zur selbstständigen Weiterentwicklung und Vertiefung der eigenen Fertigkeiten an und geben Raum für eigene Kreativität. Ein ergänzender Feedbackbogen hilft, die Lernlabore kontinuierlich zu verbessern. Insgesamt gliedert sich der Aufgabenbereich in „Wiederholung“, „Selbstständiges Anwenden“, „Quiz“ und „Feedback“. Die Kursleiterin schaut sich die über Moodle hochgeladenen Ergebnisse an und gibt individuelle Rückmeldung sowie Optimierungsvorschläge. Fragen werden umgehend via Social Media, Moodle, Mail oder das Forum beantwortet.

Insgesamt sollen die Kurse nicht nur unterschiedliche Informatik-Themen, sondern auch unterschiedliche Zugänge über einen Themenmix mit Anknüpfung an die Realität und verschiedene Interessen bedienen. Jede Teilnehmerin soll ein geeignetes Projekt finden können, das ihr Impulse für die tiefere Auseinandersetzung mit der Informatik setzt. Aktuell gibt es drei verschiedene Lernlabor-Kurse zu Informatik-Themen: „Wundersame 3D-Welten mit Blender“, „Mi, Ma, Micro:bit – steuern, messen, regeln und spielen mit dem Minicomputer“ und „Erstelle dein eigenes Spiel mit Unity 3D“. Kurse zu weiteren Themen werden kontinuierlich aufgebaut. Titel und Inhalte sind auf das Interesse von jungen Frauen ausgerichtet, setzen bei Bekanntem an oder greifen einen für Mädchen relevanten Kontext auf. Dr. Sylvia Jahnke-Klein beschreibt es so: „Wenn es möglich ist, sollten also das Lebendige oder die Wirkungen auf das Lebendige einbezogen werden (Anwendungen in der Medizin, Bezug zum eigenen Körper, Umweltschutz, historische und gesellschaftliche Bezüge, ...).“ [JK13] So wird im Blender-Kurs z.B. ein Mensch und ein Baum modelliert. Im Micro:bit-Kurs wird z.B. das Lebendige aufgegriffen: Unter der Fragestellung „Hat meine Pflanze Durst?“ bietet sich ein praktisches Anwendungsbeispiel mit Technik- und Informatikkomponenten sowie Bezug zur Lebenswelt, das darüber hinaus die Interdisziplinarität des MINT-Bereichs an einem Projekt vereinfacht visualisiert. Aber auch Alltagsgegenstände wie etwa ein Thermometer oder Kompass sind gut greifbare Beispiele, die mit Hilfe des Einplatinencomputers Micro:bit sowie Unterstützung der Programmierung durch vorgefertigte Bausteine schnelle Erfolgserlebnisse garantieren. Werden Materialien für die Durchführung der Projekte benötigt, werden diese den Teilnehmerinnen kostenfrei leihweise zur Verfügung gestellt.

Teilnehmerinnen können jederzeit mit dem Kurs starten. Es gibt keine festen Termine oder Lerngruppen. Die Kurse sind als begleitete Selbstlernkurse gedacht, in denen sich die Teilnehmerinnen selbstständig durch das Material arbeiten, aber jederzeit bei der Kursleiterin Unterstützung erhalten können. Über das integrierte Forum ist der Austausch mit anderen Teilnehmerinnen möglich. Auch die Geschwindigkeit, mit der der Kurs absolviert wird, ist variabel. In der Konzeption ist eine Lektion pro Woche mit max. 2h Aufwand angestrebt.

3 Ergebnisse und Fazit

Seit Mitte 2019 konnten 83 Teilnehmerinnen in Online-Kursen erreicht werden. Rückblickend zeigt sich, dass nicht alle Teilnehmerinnen den Kurs auch beenden, da er gewisses Durchhaltevermögen und eine eingehendere Beschäftigung fordert. Nicht jede junge Frau findet ihre Bestimmung durch das Projekt. Auf der anderen Seite können aber Interessen bestärkt und Berufs- und Studienwege gefestigt werden. Zudem leistet „Make up your MINT“ gerade auch in Kooperation mit Schulen einen wichtigen Beitrag zur MINT-Interessensförderung. Dazu lassen sich die Online-Lernlabore auch in (mono-educative) Arbeitsgemeinschaften oder Projekte an Schulen integrieren. Dies schafft einen weiteren vorteilhaften Rahmen, um Schülerinnen die Teilnahme zu ermöglichen.

Insgesamt ermöglichen die aufeinander aufbauende Lerneinheiten sowie die Unterstützung durch die Kursleiterin schnell positiv bestärkende Erfolgserlebnisse. Der hohe Anwendungsbezug sowie Ergebnisse „zum Anfassen“ machen Fortschritt schnell sichtbar und fördern das Dran-Bleiben in Folgelektionen. Didaktisches Prinzip ist das Lernen an praktischen Beispielen und Wiederholung gepaart mit bedarfsgerechter Hilfestellung: also anwendungsorientiertes und geführtes Lernen je nach Kenntnisstand der jeweiligen Teilnehmerin. Der große Vorteil des Online-Angebotes ist die orts- und zeitunabhängige Nutzungsmöglichkeit, wodurch prinzipiell eine große Zielgruppe erschlossen werden kann. Dem entgegen steht die niedrigere gefühlte Verbindlichkeit und die Schwierigkeit, in Kontakt mit den Teilnehmerinnen zu bleiben, damit diese auch über längeren Zeitraum dabeibleiben. Aus diesem Grund erproben wir derzeit ergänzende Live-Veranstaltungen (in Präsenz oder online), entweder als Einstieg in die Lernlabore oder als motivierenden Impuls während der Teilnahme im Kurs, z.B. zum Austausch untereinander oder mit der Kursleiterin. Sowohl Kursinhalte als auch Durchführungsmodalitäten werden anhand des Teilnehmerinnenfeedbacks sowie gemachter Beobachtungen kontinuierlich weiterentwickelt und neue Kurse ergänzt. So werden z.B. schwierige Stellen in den Anleitungen durch Video-Overlays ergänzend erklärt.

Förderhinweis: Das Projekt Make Up Your MINT wird gefördert vom ESF und dem Land Sachsen-Anhalt. Für den Inhalt der Arbeit sind die Autorinnen verantwortlich.






4 Literaturverzeichnis

- [HB21] Hirschel, Andrea; Bade, Korinna: LESSONS LEARNED: Erfahrungen aus drei Jahren Gleichstellungsarbeit im Projekt Make Up Your MINT, Hochschule Anhalt, 2021, <https://dx.doi.org/10.25673/35482>.
- [JK13] Jahnke-Klein, Sylvia: Benötigen wir eine geschlechtsspezifische Pädagogik in den MINTFächern? Ein Überblick über die Debatte und den Forschungsstand. Schulpädagogik heute, Heft 8, 2013.

Workshop: Hochschule 2031

Vision, Mission und Governance der Landesstrategien im Zeitalter der Hochschul-Digitalisierung


Unterschiede und Gemeinsamkeiten im Vergleich von fünf Bundesländern in Deutschland


Markus von der Heyde¹ , Armin Gerl² , Rainer Groß³ , Frank Klapper⁴, Andreas Knaden⁵, Vera G. Meister⁶ , Simone Rehm⁷ 

Abstract: Seit Jahrzehnten kooperieren Hochschulen aller Formen bei der gemeinsamen Nutzung und Erstellung von IT-Services. Zu beobachten ist ein Paradigmenwechsel von der Versorgung mit Informationstechnologien über die Bereitstellung integrierter IT-Dienste hin zur Digitalisierung der Kernprozesse, welcher diese Kooperationen stärker vorantreibt und neue initiiert. Auf Ebene der Bundesländer wird seit kurzem der Bedarf zur übergreifenden Steuerung formuliert und durch “Digitalisierungsstrategien” zum Ausdruck gebracht, da die zur Verfügung stehenden Mittel effektiv und effizient eingesetzt werden müssen. Diese strategischen Aussagen besitzen erhebliche Parallelitäten zu strategischen Formulierungen aus der Wirtschaft. Deshalb erscheint es angemessen, die dort verwendeten Begriffe, am Beispiel des Business Motivation Models (BMM), auf die Hochschullandschaft zu übertragen. Bestehende Beispiele aus fünf Landesstrategien werden zusammengefasst und im Hinblick auf die kommunizierte Vision, Mission und Governance analysiert und verglichen. Durch Analyse und Vergleich wird deutlich, dass der Wandel von IT-Strategie zu Digitalisierungsstrategien dort gelingen kann, wo im kooperativen Dialog zwischen den Verantwortlichen für die Kernprozesse Forschung und Lehre und denjenigen, die Informationsversorgung gestalten, eine potentiell transformierende Wirkung entfaltet werden kann.

Keywords: Digitalisierung, Hochschulen, Kooperation, Vision, Mission, Strategie, Governance.


¹ vdH-IT, D-99425 Weimar, info@vdh-it.de,  0000-0002-6026-082X


² Universität Passau, D-94032 Passau, armin.gerl@uni-passau.de,  0000-0001-9991-4539

³ Technische Hochschule Nürnberg, D-90121 Nürnberg, rainer.gross@th-nuernberg.de,  0000-0002-6876-060X

⁴ Universität Bielefeld, D-33615 Bielefeld, frank.klapper@uni-bielefeld.de

⁵ Universität Osnabrück, D-49069 Osnabrück, aknaden@uni-osnabrueck.de

⁶ Technische Hochschule Brandenburg, D-14770 Brandenburg an der Havel, vera.meister@th-brandenburg.de,  0000-0002-2780-0222

⁷ Universität Stuttgart, D- 70174 Stuttgart, simone.rehm@rektorat.uni-stuttgart.de,  0000-0002-9445-6417

1 Einleitung

Unabhängig von der Hochschulform (Musik- und Kunsthochschulen, Hochschulen für angewandte Wissenschaften (HAW), Universitäten) sehen die meisten Hochschulen in der kooperativen Versorgung mit IT-Services eine unverzichtbare Möglichkeit, um Dienstleistungen in der notwendigen Spezialisierung und mit langfristig konstantem Ressourcenbedarf erbringen zu können. Seit Jahrzehnten wird der dafür wichtige Erfahrungsaustausch vor und während des produktiven Betriebs auf Landes- und auf Bundesebene gepflegt. In vielen Bundesländern wurden in den vergangenen Jahren neben bilateralen Kooperationen landesweit genutzte und zum Teil arbeitsteilig erstellte Angebote entwickelt.

Viele Stakeholder aus den Hochschulen haben bei den zuständigen Ministerien den Bedarf nach weiteren Ressourcen angemeldet, um die vielfältigen Digitalisierungsmaßnahmen zu realisieren. Da Ressourcen limitiert sind, müssen neue Wege gefunden werden, die Anforderungen kostengünstiger zu erfüllen. Unabhängig von dieser gemeinsamen Vision von Synergieeffekten der Kooperation wird ein steigender Ressourcenbedarf in der IT-Versorgung beobachtet. So konnten selbst langjährige Kollaborationen im Deutschen Forschungsnetzes (DFN) die Preise nicht mehr konstant halten⁸. Die Hochschulrektorenkonferenz (HRK) forderte eine dauerhafte Förderung der Weiterentwicklung digitaler Lehrinfrastrukturen durch Bund und Länder⁹ unabhängig von den Corona-Pandemie bedingten Investitionen.

In [He21] vergleichen von der Heyde et al. landesweit bestehende Gremien, die Modelle der Steuerung (IT-Governance) und Wege in eine nachhaltige Kooperation. Die dort gesammelten Aussagen fassen auf Basis der ebenfalls hier verwendeten Papiere den grundsätzlichen Bedarf zur Steuerung und Finanzierung der zu erbringenden IT-Services in den gleichen fünf Ländern zusammen. Es wurde dabei festgestellt, dass die Modelle der Institutionalisierung sich auf den Ebenen der Steuerung und der Realisierung der Services unterscheiden. Beide Ebenen unterliegen den gemeinsamen Wertemodellen und prägen damit die Wahl der Organisationsform.

Welche Strategien können die Hochschulen einzeln oder als Gemeinschaft dauerhaft verfolgen? Der vorliegende Beitrag vergleicht die Ansätze innerhalb von fünf Bundesländern auf der Ebene der Formulierung bzw. Formierung von Wertemodellen, die eine Zusammenarbeit langfristig prägen. Da die Zusammenarbeit nicht verordnet werden kann, ist die Ebene der kommunikativen Bilder und gemeinsamen Ziele essentiell, um die Stakeholder der ansonsten in Konkurrenz stehenden Hochschulen erfolgreich interagieren und zusammenarbeiten zu lassen. Harari formuliert in [Ha15] unter anderem die Bedeutung der Flexibilität der Sprache bei der Formulierung gemeinsamer Zielvorstellungen bei der menschlichen Kollaboration.

Die Anwendung dieses Gedankens auf die vorliegenden Kooperationsmodelle der Hochschulen bei ihrer Digitalisierung führt zum Kern dieses Beitrags: Dazu werden zunächst die Begriffe aus dem Business Motivation Model (BMM) um Aspekte der Governance

⁸ siehe [Erhöhung der Entgelte zum 1.1.2022](#)

⁹ siehe [EntschlieBung des HRK-Senats vom 8.6.2021](#)

erweitert (Kapitel 2). Der Hauptteil des Beitrags stellt Beispiele aus fünf Bundesländern vor und analysiert diese im Hinblick auf die enthaltenen Aussagen zu Vision, Mission und Governance (Kapitel 3). Abschließend werden die Ergebnisse in Verbindung mit einem Ausblick in Kapitel 4 zusammengefasst.

Die am Vergleich beteiligten Kolleg:innen aus den Bundesländern hatten weder einen Auftrag noch ein explizites Mandat zur Teilnahme, sondern haben sich innerhalb des existierenden CIO-Netzwerks zusammengefunden. Die Nennung der Länder erfolgt stets in alphabetischer Reihenfolge, ohne dass damit eine Präferenz ausgedrückt werden soll: Baden-Württemberg (BaWü), Bayern (BY), Brandenburg (BRB), Niedersachsen (NDS), Nordrhein-Westfalen (NRW).

2 Definition zentraler Begriffe

Die Beschreibung und Definition von Digitalisierungszielen an Hochschulen verwendet oftmals Begriffe, die in der Formulierung von Geschäftsplänen in der Wirtschaft ebenfalls Verwendung finden. Das Business Motivation Model (BMM, v1.3) beschreibt diese Kernbegriffe und ordnet sie konsistent in eine Hierarchie ein [Ob15, Figure 7.1]. Geschäftspläne (Business Plans) definieren, welchen Zweck ein Unternehmen verfolgt und welche Mittel es dazu einsetzt.

Die im BMM bereits definierten Begriffe wurden in den Hochschulkontext übertragen, ohne die Kernbedeutungen zu verändern. Die englischen Begriffe aus dem BMM wurden in diesem Beitrag durch weitgehend äquivalente deutsche Übersetzungen ersetzt, sind aber zur Konsistenz weiter in Klammern angegeben. Die zunächst aus dem BMM vorgestellten Begriffe dienen danach der systematischen Einordnung der vorgestellten Beispiele der Landespapiere.

Zielebenen (ends): beschreiben Eigenschaften eines Zustands, den die Organisation anstrebt.

Leitbild (Vision) zeichnet ein abstraktes Bild vom Zielzustand, der erreicht werden soll.

Resultat / Ergebnis (Result) definiert den Zielzustand konkreter.

Ziel (Goal) ist ein langfristig zu erreichendes, qualitativ beschriebenes Ergebnis.

Zielsetzung / Planziel (Objective) ist ein quantitativ beschriebenes, kurz- oder mittelfristig zu erreichendes Ergebnis.

Wege / Hilfsmittel (means): sind Handlungsoptionen, die eine Organisation wahrnimmt, um Ziele zu erreichen.

Auftrag (Mission) formuliert einen Wertekanon mit einer allgemeinen Aufforderung zur Handlung.

Vorgehensweise (Course of Action) bricht die Handlungsoptionen z.B. in der zeitlichen Dimension auf, um in kleineren Schritten verfahren zu können.

Strategie (Strategy) formuliert einen gut verständlichen Gesamtplan zur grundsätzlichen, langfristigen Umsetzung der Ziele.

Taktik (Tactic) ist im Gegensatz zur Strategie kurzfristiger orientiert und beschreibt Einzelschritte.

Maßnahme (Measure) ist ein Baustein zur Umsetzung einer Taktik und letztlich der Strategie.

Tagsordnung (Agenda) formuliert einen Plan, der in einer Abfolge abgearbeitet werden kann.

Kontrollmechanismen (controls): sind Regeln, die eine Organisation besitzt, um die Ziele mit den definierten Mitteln zu erreichen.

Kontrolle (Governance) bezeichnet das Steuerungs- und Regelsystem einer Organisation.

Entscheidung (Decision) wird auf Basis von Argumenten getroffen, wenn mehrere Handlungsoptionen innerhalb der geltenden Regeln und Gesetze zielführend sein könnten.

Zuständigkeit (Responsibility) regelt, wer eine spezifische Handlung vornimmt bzw. Entscheidung trifft.

Verantwortung (Accountability) benennt, wer für das Ergebnis der Handlungen politisch oder juristisch haftet.

Vorgabe (Directive) formuliert verbindliche Regeln für Handlungen in der Organisation.

Richtlinie (Policy) umfasst grundsätzlich und generell geltende Regeln der Organisation.

Vorschrift (Guideline) definiert im Gegensatz zur Richtlinie konkretere Vorgaben und bietet weniger Spielraum.

Die BMM-Bereiche der Beschreibung von Zielebenen einer Organisation (Ends) und den Hilfsmitteln (Means), diesen Zielzustand zu erreichen, wurden um das Begriffsfeld der Kontrollmechanismen (Controls) ergänzt. Die BMM-Struktur wurde modifiziert und so erweitert, dass die oberste Gliederungsebene drei Begriffsfelder bzw. Kategorien unterscheidet, die selbst kaum in den Diskussionen vorkommen, aber die darunter liegenden Begriffe bündeln. So werden auch die breit diskutierten Begriffe Vision, Mission und Governance eingeordnet, welche auf die Kernaspekte dieser Kategorien fokussieren, ohne die umfängliche Bedeutung aller gebündelten Begriffe zu besitzen. Es sind innerhalb der Kategorien bewusst auch Begriffe angegeben, die nicht unmittelbar bei der Analyse aufgegriffen werden. Sie sollen der Abgrenzung dienen und die Diskussionen bei der Erstellung derartiger Papiere unterstützen. Den Autor:innen ist bewusst, dass diese Begriffe in Teilen synonym verwendet werden und damit in Diskussionen möglicherweise eine Balance zwischen Verbindlichkeit und Flexibilität angestrebt wird.

Das Verständnis von Governance in diesem Beitrag basiert auf den hochschulspezifischen Vorschlägen zu IT-Governance in [Sc09] und [Wi17]. Wegen der übergreifenden Betrachtung der Wirkweisen der analysierten Papiere wird jedoch nicht auf die von Weill und Ross in [WR04] eingeführte Differenzierung in Archetypen der Entscheidungsbeteiligung und IT-Domänen Bezug genommen. Die folgenden Definitionen bleiben somit auf einer allgemeineren Verständnisebene und sind nicht für die Informationsversorgung spezifisch gewählt.

Eine Übertragung der allgemeinen Begriffe in den Kontext der Hochschulen ist notwendig, da dort spezifisch mit den Kernprozessen Forschung und Lehre umgegangen wird und Prinzipien der Wirtschaft nicht ohne Anpassung übernommen werden können, wie Wimmer in [Wi17] fordert. Wie Schünemann und Budde bspw. aus den Erfahrungen der Peer-to-Peer-Beratung des Hochschulforum Digitalisierung in [SB18] folgern, sind gerade die Lehrprozesse oftmals nicht in den Digitalisierungsstrategien ausreichend verankert, da die Lehre - wie im Grundgesetz festgehalten - frei ist und von den Professor:innen dezentral und unabhängig vorbereitet und durchgeführt wird.

3 Digitalisierungsstrategien

Da Hochschulen sich als innovative Mitgestalter der Gesellschaft verstehen, formulieren sie auch beim "Megatrend Digitalisierung" diesen Anspruch und ergreifen entsprechende Maßnahmen. Zum Ausdruck kommt dies in den letzten 1-3 Jahren in der Formulierung von Digitalisierungsstrategien in unterschiedlichen Ausrichtungen und Kontexten. So werden für die Kooperation auf Landesebene Aussagen formuliert, die nach innen (auf die Hochschulen selbst) und nach außen (zur Politik auf Ebene der Ministerien) wirken sollen.

In der Hochschullandschaft wurden in den vergangenen Jahren eine Reihe von neuen Begriffen vorgeschlagen, um damit Aspekte, Prozesse o.ä. zu bezeichnen, die im Kontext der Digitalisierung von Interessen sind. Viele Autoren beklagen zurecht eine Inflation dieser

Begriffe. In [He17] wurde eine Übersicht und Einordnung der verschiedenen Perspektiven vorgenommen.

Der folgende Abschnitt formuliert drei Perspektiven, welche die potentiell intendierten Wirkungen der strategischen Aussagen differenzieren. Anschließend werden beispielhaft Strategien aus fünf Bundesländern analysiert und verglichen. Dabei sei vorausgeschickt, dass eine Wirksamkeit nicht empirisch betrachtet, sondern lediglich verglichen wird, indem geprüft wird, welche Aussagen zu den folgenden Perspektiven passen.

3.1 Perspektiven

Um die Wirkung einer Digitalisierungsstrategie analysieren zu können, schlagen wir die Unterscheidung von drei Perspektiven vor. Die folgenden Erläuterungen werden jeweils durch den Bezug zu den oben bereits eingeführten BMM-Begriffe ergänzt.

Politische Wirksamkeit wird primär durch Formulierung des Zwecks, also der hinter den Zielen stehenden Gründe erreicht (vgl. Vision→**V**).

Einigung auf Grundwerte zielt auf die Vereinbarung, mit welchen Mitteln die Ziele erreicht werden (vgl. Mission→**M**).

Entscheidungsentlastung wird erreicht, wenn viele Mitglieder der Organisation konsistent handeln und wenige Entscheidungen getroffen werden müssen (vgl. Governance→**G**).

Die drei Perspektiven beleuchten somit die drei oberen Ebenen der BMM-Begriffe. Es wird unterstellt, dass die im Folgenden analysierten Beispielpapiere dann das Potential vergleichbar einem Geschäftsplan ausschöpfen können, wenn entsprechende Aussagen zu allen drei Ebenen enthalten sind.

3.2 Beispiele

Aus fünf Bundesländern wurden bestehende Papiere zur Hochschulkooperation für die Erstellung und Nutzung von Services der Informationstechnologie (IT) analysiert und die darin enthaltenen strategischen Kernaussagen zu **V**ision, **M**ission und **G**overnance extrahiert. Diese Kurzfassungen der Aussagen wurden nicht durch offizielle Gremien bestätigt, wohl aber die Texte, aus denen sie stammen. Die Analyse spiegelt die jeweiligen Aussagen an den zuvor erläuterten Perspektiven **V**ision, **M**ission und **G**overnance, indem die Aussagen den enthaltenen Schwerpunkten nach diesen Perspektiven zugeordnet werden.

Die Landesrektorenkonferenz (LRK) in BaWü hat in einem Positionspapier¹⁰ zum Ende 2020 strategische Kernaussagen zusammengetragen, welche den Digitalisierungsschub, der

¹⁰ siehe [Positionspapier Baden-Württemberg 2021-26 vom 14.12.2020](#)

durch die Corona-Pandemie ausgelöst wurde, aufgreift und daraus Konsequenzen zieht. Aus diesem Positionspapier wurden folgende strategische Aussagen extrahiert:

Beispiel: Baden-Württemberg (BaWü)

- M** Universitäten müssen den Wandel weiter aktiv gestalten, wenn sie **V** gestärkt aus der digitalen Transformation hervorgehen möchten.
- V** Damit das Potential der Digitalisierung für Lehre und Forschung vollständig geschöpft werden kann, **M** muss der Umstellungsprozess noch stärker didaktisch, technisch und organisatorisch begleitet werden.
- G** Gemeinsames Fundament für individuelle Profile: Wo ein gemeinsames Problem besteht, wollen wir es gemeinschaftlich lösen.
- V** Für die Digitalisierung in der Lehre gilt es einen angemessenen Mix aus Digitalität und Präsenz zu finden, in dem jedes Format seine Stärken entfaltet.
- M** Die räumlichen und technischen Infrastrukturen müssen agil an die neuen Bedürfnisse angepasst werden.
- G** Digitalisierung muss als ein Querschnittsthema über die Bereiche Forschung, Lehre und administrative Prozesse betrachtet und betrieben werden.
- G** Die Kooperationen dienen dazu, Ressourcen zu bündeln, Synergien zu schöpfen und bestimmte Standards im Land gemeinsam umzusetzen, **V** damit man sich von der Konkurrenz außerhalb von Baden-Württemberg abhebt.

Im Prozess der Erstellung der IT-Strategie der Bayerischen Hochschulen wurde 2020/2021 eine Steuerungsebene (IT-Governance) in Anlehnung an COBIT geschaffen [Ge20]. Eine Arbeitsgruppe entwickelte für zwölf Themenfelder Statements nach der Ansprache-Bewertung-Folgerung (ABF)-Systematik. Aus der Präambel der IT-Strategie wurden folgende strategische Aussagen extrahiert:

Beispiel: Bayern (BY)

- M** Die IT-Strategie der bayerischen Hochschulen definiert Handlungsempfehlungen für die fortlaufende kooperative und umfassende Erneuerung und Optimierung der IT. **V** Sie gestaltet damit gleichzeitig die Digitalisierung der Hochschulen und trägt damit zur umfassenden und fortlaufenden digitalen Transformation von Forschung und Lehre, Verwaltung, Gesellschaft und Wirtschaft bei. **G** Kern ist eine nachhaltige Abstimmung und Kooperation zur koordinierten Schaffung von Synergien zwischen den Hochschulen.
- G** Grundlage für die verbesserte Zusammenarbeit zwischen den Hochschulen ist die Weiterentwicklung der IT-Governance mit folgenden Schwerpunkten:

- Weiterentwicklung der Rolle der CIOs zur ● Stärkung der Zusammenarbeit mit den Hochschulleitungen und den Hochschulverbänden;
- Koordination und Priorisierung von Abstimmungs- und Kooperationsbedarf in bayernweiten Gremien ● gebündelt im Digitalverbund der bayerischen Hochschulen;
- Schaffung von Synergien durch ● Abstimmung gemeinsamer Vorgehensweisen, Ressourcenbündelung und -verteilung, Balance zwischen Individualbedarf und Zentralisierung, bis hin zur Einführung von hochschulübergreifenden IT-Services

In BRB wurde im Frühjahr 2021 eine Digitalisierungsagenda¹¹ erarbeitet, die auf Leitgedanken aufbaut und dann für die Handlungsfelder Studium und Lehre, Forschung, Transfer sowie Verwaltung und Infrastruktur konkrete Umsetzungsschritte ableitet.

Beispiel: Brandenburg (BRB)

Das Ministerium für Wissenschaft, Forschung und Kultur des Landes Brandenburg und die brandenburgischen Hochschulen verfolgen die Digitalisierung gemeinsam unter folgenden Leitgedanken:

- Die Hochschulen sind bestrebt, ihr Handeln stets effektiv und effizient zu gestalten, ihre Prozesse und Strukturen kontinuierlich zu verbessern und diese den aktuellen Anforderungen kontinuierlich anzupassen.
- Digitale und analoge Lern- und Arbeitsinfrastrukturen werden in Zukunft gleichermaßen wichtig sein und sich gegenseitig bereichern.
- Die brandenburgischen Hochschulen sind dem Gedanken der Openness verpflichtet, der einen grundlegenden Wandel im Denken und Handeln von Studierenden, Lehrenden und Forschenden voraussetzt.
- Digitalisierung ermöglicht zeit- und ortsunabhängiges Arbeiten, Forschen sowie Studieren und schafft somit neue Möglichkeiten für Kooperationen.
- Digitalisierung umfasst nicht nur technologische Veränderungen, sondern geht mit einem fundamentalen Kulturwandel auf verschiedenen Ebenen einher.
- Digitalisierung bietet Chancen für mehr Teilhabe.
- Die brandenburgischen Hochschulen werden die Vermittlung digitaler Kompetenzen daher noch stärker in den Blick nehmen und in den Studiengängen sowie in der wissenschaftlichen Weiterbildung Angebote entwickeln.

¹¹ siehe [Digitalisierungsagenda vom 17.05.2021](#)

G Das Themenspektrum des Zentrums der brandenburgischen Hochschulen für Digitale Transformation wird auf alle Handlungsfelder der Digitalisierung erweitert und die Governance-Struktur angepasst.

Im Jahr 2017 wurden von der LandesHochschulKonferenz (LHK) in NDS Eckpunkte der Digitalisierungsoffensive für die niedersächsischen Hochschulen verfasst und veröffentlicht¹². Die Eckpunkte formulieren zunächst die Ausgangslage und Potentiale der Digitalisierung, um dann zur Umsetzung der Digitalisierung in den Handlungsfeldern Forschung, Studium und Lehre, Informationsversorgung sowie IT-Infrastrukturen konkrete Maßnahmen bzw. Bausteine zu formulieren. Das Eckpunktepapier skizziert den inhaltlichen Fokus der aktuell entwickelten landesweiten Governance im Rahmen der Hochschule.digital Niedersachsen. Kernaussagen sind:

Beispiel: Niedersachsen (NDS)

V Die Digitalisierungsoffensive leistet einen Beitrag zur Sicherung der Wettbewerbsfähigkeit der Hochschulen in Niedersachsen.

V “Big Data” in der Forschung und digitale Technologien in der Lehre erfordern zukunftsfähige digitale Infrastrukturen, neue Kompetenzen und standortübergreifende Kooperationen.

M Das Land und die Hochschulen wirken zusammen, um die Digitalisierung an den Hochschulen in Niedersachsen wirksam und nachhaltig voranzubringen.

V Digitale Instrumente und Werkzeuge, digitalisierte Prozesse und Kommunikationsstrukturen in Forschung und Lehre sind eine Bereicherung für die Wissenschaft.

M Die Digitalisierung in der Lehre leistet einen Beitrag zur Öffnung der Hochschulen für neue Zielgruppen.

M Digitalisierung schafft die Voraussetzungen für Open Science.

G Digitalisierung ermöglicht und erfordert weitergehende Kooperationen, sowohl zwischen den Hochschulen als auch innerhalb der einzelnen Hochschulen.

G Eine nachhaltige Digitalisierung erfordert eine Weiterentwicklung der Ziele, Strategien und Strukturen der Ausbildung (Lehre, Studienprogramme, Studiengänge und die Rahmenbedingungen des Studiums).

V Ziele der Kooperationen sind eine Reduktion der zusätzlichen Kosten durch **M** eine verstärkte Arbeitsteilung und eine Ressourcenbündelung bei den Supportinfrastrukturen sowie **V** eine konsequente Vernetzung in gemeinsamen Forschungsvorhaben und Lehrformaten.

¹² siehe [Eckpunktepapier für eine Digitalisierungsoffensive vom 22.05.2017](#)

Im Sommer 2021 wurde in NRW ein Positionspapier veröffentlicht, welches für die Kernprozesse (Forschung und Lehre) und Unterstützungsleistungen (Administration und Informationstechnologien) der Hochschulen Ziele formuliert, mit denen die digitale Transformation an den Hochschulen unterstützt wird.

Beispiel: Nordrhein-Westfalen (NRW)

- ✓ Synergieeffekte durch die Kooperation über Standorte und Hochschultypen hinweg.
- ✓ Qualifikation aller Akteur:innen für ihre Aufgaben und Rollen in einer durch die Digitalisierung geprägten Welt.
- ✓ Offenheit - freie Verfügbarkeit, Verwendbarkeit und Transparenz
- ✎ Verbinden, kommunizieren und konsolidieren - Förderung des Aufbaus von Communities und Erarbeitung von Zukunftskonzepten zur Integration der vielfältigen dezentralen Angebote.
- ✎ Etablierung digitaler Infrastrukturen - Nachhaltige und nach Möglichkeit landesweite Services.
- Ⓒ Projekte müssen ein hohes Innovationspotenzial entfalten, Synergien generieren und von möglichst allen öffentlich-rechtlichen Hochschulen genutzt werden.
- Ⓒ Jede Hochschule bleibt entscheidungsbefugt und autonom in Bezug auf ihre jeweiligen Belange.

3.3 Analyse

Die Aussagen aus dem vorangegangenen Abschnitt wurden nicht alle mit dem Gedanken verfasst, den landesweiten Kooperationen eine strategische Leitlinie zu geben oder diese aus den Hochschulen heraus zu formulieren. Dennoch lassen sich die selektierten Aussagen zum einen direkt vergleichen und zum anderen unter den zuvor formulierten Perspektiven betrachten.

3.3.1 Analyse der einzelnen Papiere anhand der Perspektiven

In **BaWü** wurden zehn Aussagen identifiziert. Davon entfallen 40% auf politische Wirksamkeit (✓) und jeweils 30% auf übereinstimmende Werte (✎) und Aussagen, die dezentral zur Entlastung von Entscheidungen führen könnten und damit potentiell steuernde Wirkung (Ⓒ) haben könnten. Die drei in **BY** stark zusammengefassten Aussagen verteilen sich gleichmäßig über alle Perspektiven (✓, ✎ und Ⓒ). Eine weitergehende Analyse aller im Strategiepapier enthaltenden Kernaussagen könnte von dieser sehr vereinfachten Sicht allerdings deutlich abweichen, so dass vermutet werden kann, dass durch die Zusammenfassung

die Gleichverteilung verursacht wurde. In **BRB** bilden mit 50% aus den acht extrahierten Aussagen Formulierungen zu Werten (**M**) einen Schwerpunkt für die Umsetzung. Mit drei aus acht (knapp 40%) sind Aussagen zu den Zielen und der politischen Wirksamkeit (**V**) formuliert. Steuernde Aussagen (**G**) sind mit weniger als 15% am wenigsten vertreten. In **NDS** liegt mit rund 45% der Aussagen der Schwerpunkt auf der politischen Wirksamkeit (**V**). Ähnlich stark ist mit rund 35% der Aussagen die Umsetzung anhand der Wertmodelle (**M**) ausgeprägt. Den kleinsten Teil machen mit etwa 20% die steuernden Aussagen (**G**) aus. In **NRW** entfällt die Mehrheit mit rund 40% der Aussagen auf die politische Wirksamkeit (**V**). Mit jeweils ca. 30% der Aussagen zur Umsetzung und Maßnahmen (**M**) sowie den Kontrollmechanismen (**G**) sind die anderen Perspektiven gleich stark ausgeprägt.

3.3.2 Vergleich der Papiere der Bundesländer

Die Ebene der politischen Wirksamkeit durch Formulierungen des übergeordneten Ziels bzw. eines Leitbildes (**V**ision) wird in allen fünf Beispielen stark adressiert. In BaWü und NDS bildet dies den größten Schwerpunkt mit rund 50% der Aussagen; in den anderen Bundesländern ist dieser Bereich mit mindestens 25% der Aussagen ebenfalls stark vertreten.

Der Bereich der Umsetzung durch die Formulierung übergreifender Werte und Maßnahmen (**M**ission) steht in BRB mit dem höchsten Anteil der Aussagen im Fokus des Beispiels. Allerdings nimmt in den Papieren aller Länder die Formulierung von konkreten Umsetzungsmaßnahmen in allen betroffenen Kernprozessen einen erheblichen Anteil ein.

Die größten Unterschiede der Ansätze fallen im Bereich der Steuerung (**G**overnance) auf. Im Papier aus BY auf die Einführung von Steuerungsmechanismen und -strukturen fokussiert, stehen diese sowohl bei NDS als auch BRB eher im Hintergrund. Alle Papiere enthalten aber Aussagen zur Governance, die Leitplanken darstellen können, wenn innerhalb der Zusammenarbeit Alternativen bestehen oder Konflikte gelöst werden müssen. Diese dienen stets der Entlastung von zentralen Entscheidungsgremien, da auf Basis der gemeinsam formulierten Überzeugungen und Grundsätze dezentral, aber dennoch stets kohärent entschieden werden kann. In den Papieren von BY und NRW wird diese dezentrale Autonomie der Hochschulen ganz explizit formuliert.

4 Interpretation und Zusammenfassung

Die Kooperation der Hochschulen in den Ländern wird durch Akteure mit vielfältigen Perspektiven auf Basis einer hoch flexiblen Kommunikation und persönlichen Vertrauens gestaltet. Die Zusammenarbeit steht dabei sowohl unter dem Einfluss kontinuierlicher Veränderungen (bspw. im Bereich der seit Jahren sich verstärkenden Anforderungen der Vernetzung von Organisationen und Daten) als auch kaum planbarer Herausforderungen (bspw. die Corona-Pandemie). Insbesondere die Versorgung der Bildungseinrichtungen mit

Informationstechnologien hat sich in den vergangenen Jahren zunächst in Richtung einer verstärkten Serviceorientierung gewandelt, um nunmehr in kollaborativer Partnerschaft zwischen den Kernprozessen der Hochschulen (Lehre und Forschung) und der Realisierung mannigfaltiger informationsverarbeitender und vernetzender Prozesse auf die eigene digitale Transformation zu fokussieren. Ob und ggf. wie weit dies durch die Hochschulen selbst zu gestalten wäre, wurde bereits in [He19] diskutiert.

In diesem Beitrag wurden die auf Ebene der Länder verfassten Konzepte zum Umgang mit der Digitalisierung analysiert. Unabhängig davon, ob diese als “Eckpunkte”, “Agenda” oder “IT-Strategie” betitelt wurden, haben die Papiere den Anspruch, politische Wirkung zu erzielen, indem ein Leitbild (Vision) gezeichnet wird, zu dem eine Veränderung angestrebt wird. Um diese Zielstellung zu erreichen, formulieren die untersuchten Papiere Maßnahmen, Vorgehensweisen und Kontrollmechanismen. Diese Beschreibungen der Maßnahmen stehen oftmals in direktem Bezug zu den Kernprozessen und weniger zur klassischen IT-Versorgung. Der generelle Auftrag (Mission) zur Umsetzung der Maßnahmen ist bspw. zeitlich oder durch Abstraktionsebenen gegliedert. Die Umsetzung ist durch Vorgehensweisen, also Entscheidungsmuster, Richtlinien und einen dahinter liegenden Wertekanon vorweg genommen bzw. geprägt. Die Papiere enthalten zusätzlich explizite Aussagen zur Kontrolle (Governance), die sich dazu eignen, dass Entscheidungen konsistent und nachvollziehbar dezentral getroffen werden, indem dazu notwendige Entscheidungsstrukturen und -prinzipien benannt werden.




Diese Fallstudie der fünf Bundesländer (BaWü, BRB, BY, NDS und NRW) hat basierend auf den Begriffsdefinitionen aus dem um Governanceaspekte erweiterten Business Motivation Model den Wandel von IT-Strategie zu Digitalisierungsstrategien nachgezeichnet. Es scheint in der Kooperation der Hochschulen neben dem IT-versorgenden Charakter, der durch die Ausrichtung an den Zielen der Kernprozesse (Alignment) zum Ausdruck kommt, zunehmend um die kooperative Gestaltung der digitalen Veränderung der Kernprozesse zu gehen (Enabling). Falls diese zusätzliche, eher transformierende Wirkung auf Prozesse durch den Einsatz von digitalen Technologien tatsächlich eintritt, tragen die Konzepte mit Recht das Schlagwort “Digitalisierung” im Titel. Daher sind die tatsächlichen Veränderungen in Zukunft aufmerksam zu beobachten und zu begleiten.

Literatur

- [Ge20] Gerl, A.; von der Heyde, M.; Groß, R.; Seck, R.; Watkowski, L.: Applying COBIT 2019 to IT Governance in Higher Education. In: INFORMATIK 2020. GI Jahrestagung 2020. Bd. Back to the Future. Lecture Notes in Informatics (LNI) - Proceedings, Gesellschaft für Informatik, Bonn, Karlsruhe, Germany, S. 517–530, 28. Sep. 2020, ISBN: 978-3-88579-701-2, DOI: 10.18420/inf2020_47.
- [Ha15] Harari, Y.N.: Sapiens – A Brief History of Humankind. Vintage, 2015, ISBN: 978-0099590088.

- [Ob15] Object Management Group (OMG): Business Motivation Model, 19. Mai 2015, URL: <http://www.omg.org/spec/BMM/1.3/>.
- [SB18] Schünemann, I.; Budde, J.: Hochschulstrategien für die Lehre im digitalen Zeitalter - Keine Strategie wie jede andere!, Arbeitspapier 38, Hochschulforum Digitalisierung (HFD), Sep. 2018, S. 31, URL: https://hochschulforumdigitalisierung.de/sites/default/files/dateien/HFD_AP_Nr38%20Empfehlungen_Strategieentwicklung_WEB.pdf.
- [Sc09] Schwabe, G.: IT-Governance an Universitäten in Deutschland, Schweiz und Österreich. VM Verwaltung & Management 15/6, S. 317–325, 2009, DOI: 10.5771/0947-9856-2009-6-317.
- [He17] von der Heyde, M.; Auth, G.; Hartmann, A.; Erfurth, C.: Hochschulentwicklung im Kontext der Digitalisierung - Bestandsaufnahme, Perspektiven, Thesen. In: INFORMATIK 2017. GI Jahrestagung 2017. Bd. 275. Lecture Notes in Informatics (LNI) - Proceedings, ISSN 1617-5468, Chemnitz, Germany, S. 1757–1772, 28. Sep. 2017, ISBN: 978-3-88579-669-5, DOI: 10.18420/in2017_175.
- [He19] von der Heyde, M.; Auth, G.; Hartmann, A.; Erfurth, C.: Digital Transformation of Higher Education - Enabling Scaling Platforms. In: ICERi 2019. Bd. Proceedings of ICERI2019 Conference, Seville, Spain, S. 2107–2115, 11. Nov. 2019, ISBN: 978-84-09-14755-7, DOI: 10.21125/iceri.2019.0584.
- [He21] von der Heyde, M.; Gerl, A.; Groß, R.; Klapper, F.; Knaden, A.; Meister, V.; Rehm, S.: Gemeinsame Digitalisierung durch die Entwicklung von IT-Governance-Strukturen auf Landesebene - Vergleich von Zielen und Steuerung der Hochschul-Kooperation von fünf Bundesländern in Deutschland. Wirtschaftsinformatik & Management/, Submitted, 2021.
- [WR04] Weill, P.; Ross, J.: IT Governance: How Top Performers Manage IT Decision Rights for Superior Results. Harvard Business Review Press, Boston, 2004, ISBN: 978-1-59139-253-8.
- [Wi17] Wimmer, M.: IT-Governance an Hochschulen. Notwendigkeit, Stand und Wege zum Erfolg. Die Hochschule : Journal für Wissenschaft und Bildung, die hochschule 26/1, S. 70–82, Jan. 2017, ISSN: 1618-9671, URL: <http://nbn-resolving.org/urn:nbn:de:0111-pedocs-166375>.

WireGraph - Entwicklung eines Lernspiels mit Prosumentenumgebung zur Förderung der fachlichen und digitalen Kompetenz von Wirtschaftsstudierenden


Christian Fleiner ¹, Julian Grümmer ², Goran Zvekan ³ und Andreas Harth ⁴


Abstract: Die Europäische Kommission warnt davor, dass die Entwicklung der Digitalisierung unzureichend in der Hochschullehre beachtet wird. Vielen Absolventen:innen mangelt es an digitaler Kompetenz. Dem steht gegenüber, dass die digitalen Anforderungen auf dem Arbeitsmarkt stetig wachsen. Dies gilt auch für Berufe des Finanzbereichs. Um die Vermittlung von digitaler Kompetenz in die Lehre miteinzubinden, eignet sich das von der Europäischen Kommission vorgeschlagene Referenzmodell der digitalen Kompetenz. Es besteht allerdings ein großer Forschungsbedarf, wie das Referenzmodell explizit in die Lehre übertragen werden kann. In diesem Beitrag wird das Konzept des Lernspiels *WireGraph* vorgestellt, mit dem Studierende der Wirtschaftswissenschaften die Arbeit mit Wissensgraphen erlernen, um den Wirecard-Skandal aufzuarbeiten. Dazu wird eine Prosumentenumgebung gestellt, mit der der Kompetenzbereich *Erstellung digitaler Inhalte* des Referenzmodells vollständig und über alle Leistungsniveaus hinweg vermittelt wird. Dabei erlernen Studierende nicht nur die Grundlagen von RDF und SPARQL, sondern können diese auch gezielt für ihre Ideen und weiteres Fachwissen einsetzen.


Keywords: Audit; Big Data; DigComp; Digitalisierung; Educational Game; Interaktive Lehre; Linked Data; Prosument; Wissensgraph; Wirecard


1 Motivation

Bereits 2017 hat die Europäische Kommission auf das “Missverhältnis zwischen den Kompetenzen, die in Europa benötigt werden, und den Kompetenzen, die vorhanden sind” als einer der vier großen Herausforderungen für die Europäische Union hingewiesen. Mit dem Vermerk, dass viele Absolventen:innen eine mangelhafte digitale Kompetenz aufwiesen, die aber dringend erforderlich sei, soll insbesondere in der Hochschulbildung mit geeigneten Maßnahmen und Mechanismen reagiert werden [Eu17]. In der Arbeitswelt wird bereits seit einiger Zeit deutlich, dass eine gute digitale Kompetenz nicht nur in den

¹ Friedrich-Alexander-Universität Erlangen-Nürnberg, Technische Informationssysteme, Lange Gasse 20, 90403 Nürnberg, christian.fleiner@fau.de,  <https://orcid.org/0000-0003-4280-670X>

² Friedrich-Alexander-Universität Erlangen-Nürnberg, Rechnungs- und Prüfungswesen, Lange Gasse 20, 90403 Nürnberg, julian.gruemmer@fau.de,  <https://orcid.org/0000-0002-0810-394X>

³ Friedrich-Alexander-Universität Erlangen-Nürnberg, Rechnungs- und Prüfungswesen, Lange Gasse 20, 90403 Nürnberg, goran.g.zvekan@fau.de,  <https://orcid.org/0000-0003-3919-0326>

⁴ Friedrich-Alexander-Universität Erlangen-Nürnberg, Technische Informationssysteme, Lange Gasse 20, 90403 Nürnberg, andreas.harth@fau.de,  <https://orcid.org/0000-0002-0702-510X>

informationstechnischen Berufen vorausgesetzt wird⁵. Auch der Finanzbereich ist davon betroffen, wie beispielsweise das Berufsbild des Wirtschaftsprüfers [HS17]. Künstliche Intelligenz, Blockchain und Big Data werden als die Technologien genannt, die einen großen Einfluss auf Berufe des Rechnungswesens haben werden. [GHV19]. Das ist nachzuvollziehen, da erst durch den Einsatz dieser Technologien neue Möglichkeiten verfügbar wurden. So werden im Big Data-Kontext Wissensgraphen für Thematic Investing oder Risikoanalysen eingesetzt [Ni20]. Zudem konnten mit Hilfe von Wissensgraphen Verbindungen zwischen unautorisierten Informationen in der Panama-Papers-Affäre hergestellt werden⁶. Aber auch andere Fachrichtungen, wie zum Beispiel Humanmedizin sind davon betroffen [KFT18]. Daher muss die Forderung an die Hochschulbildung, digitale Kompetenzen gezielt zu fördern, um nachhaltige Lehre zu erreichen, beachtet werden. Im Rechnungswesen identifizierten Cockcroft und Russell sechs forschungsrelevante Themengebiete von Big Data: *Risiko, Sicherheit, Datenvisualisierung, Predictive-Analytics, Datenmanagement und Datenqualität* [CR18]. Unter Datenmanagement ist zum Beispiel die Integration mehrerer Datenbanken zu verstehen und die Vermeidung von Redundanz. Datenqualität befasst sich mit dem Wahrheitsgehalt und der Inferenzmöglichkeit. Wissensgraphen sind eine Big Data-Technologie, die Mechanismen für Datenmanagement und -qualität liefert.

Um auf diese Entwicklung zu reagieren, wird das Lernspiel *WireGraph* vorgestellt, mit dem Studierende der Wirtschaftswissenschaften die Arbeit mit Wissensgraphen erlernen, aber auch Fachwissen im Bereich Prüfungstechnik vermittelt bekommen. Die Handlungsgrundlage für das Lernspiel bildet der Wirecard-Fall. Dieser ist insbesondere durch seine Aktualität interessant und bietet zudem teilweise noch ungeklärte Handlungsstränge. Als Maßstab für die digitale Kompetenz wird das Europäische Referenzmodell der digitalen Kompetenz für Bürger (DigComp) herangezogen [CVP17]. Das Lernspiel wird in einer Prosumentenumgebung bereitgestellt, in der Studierende die Möglichkeit haben, die unterschiedlichen Leistungsniveaus des Referenzmodells zu erreichen, indem sie Fachwissen nicht nur verstehen und anwenden, sondern auch weiterentwickeln. Der explizite Transfer von DigComp auf das Lehrdesign wurde als Forschungslücke erkannt [Si18] und wird mit der Einführung der Prosumentenumgebung und deren Ausgestaltung am Beispiel des Lernspiels adressiert.

In diesem Beitrag wird zunächst die digitale Kompetenz, ihr Referenzmodell und die Prosumentenrolle der Studierendenschaft generell beschrieben. Anschließend werden aktuelle Lehr-Lern-Konzepte beschrieben, die in den Lehrveranstaltungen eingesetzt werden, in denen *WireGraph* als unterstützendes Medium zukünftig verwendet werden soll. Die Prosumentenrolle wird dabei diesen gegenübergestellt. Abschließend wird das Lernspiel *WireGraph* aufgeführt und die Ausgestaltung der Prosumentenumgebung ausformuliert.

⁵ <https://www.econstor.eu/bitstream/10419/146097/1/843867167.pdf>, letzter Zugriff am 14.04.2021.

⁶ <https://medium.com/analytics-vidhya/panama-papers-meet-knowledge-graph-c70652d98f92>, letzter Zugriff am 14.04.2021.

2 Einordnung der digitalen Kompetenz

2.1 Digitale Kompetenz - eine Definition

Es existiert keine einheitliche Definition der digitalen Kompetenz. In diesem Beitrag wird daher die Definition aus der *Empfehlung des Europäischen Rates zu Schlüsselkompetenzen für lebensbegleitendes Lernen* (2018) fortan verwendet [Ra18], die sich aus der Definition der *Computerkompetenz* [Eu06] herausbildete:

Digitale Kompetenz umfasst die sichere, kritische und verantwortungsvolle Nutzung von und Auseinandersetzung mit digitalen Technologien für die allgemeine und berufliche Bildung, die Arbeit und die Teilhabe an der Gesellschaft. Sie erstreckt sich auf Informations- und Datenkompetenz, Kommunikation und Zusammenarbeit, Medienkompetenz, die Erstellung digitaler Inhalte (einschließlich Programmieren), Sicherheit (einschließlich digitales Wohlergehen und Kompetenzen in Verbindung mit Cybersicherheit), Urheberrechtsfragen, Problemlösung und kritisches Denken. [...]

2.2 DigComp 2.1 - Referenzmodell der digitalen Kompetenz für Bürger

Das in Version 2.1 veröffentlichte Referenzmodell [CVP17] umfasst fünf Kompetenzbereiche, in denen Bürger Fähigkeiten und Kenntnisse aufweisen müssen, um in der heutigen und zukünftigen digitalen Gesellschaft selbstbestimmt agieren zu können. Die Kompetenzbereiche sollen im Folgenden kurz skizziert werden:

- **Informations- und Datenkompetenz.** Dieser Kompetenzbereich beinhaltet das Finden, Filtern, Interpretieren und das Verwalten von digitalen Inhalten, wie zum Beispiel die korrekte Verwendung einer Suchmaschine.
- **Kommunikation und Zusammenarbeit.** Hierzu zählt die bewusste Verwendung von Technologien zur Ermöglichung oder Verbesserung der zwischenmenschlichen Zusammenarbeit, wie zum Beispiel beim Einsatz von Videokonferenzanwendungen.
- **Erstellung digitaler Inhalte.** Hierzu gehört die Fähigkeit neue, digitale Inhalte zu erstellen und das Verständnis gesetzlicher Regelungen, wie zum Beispiel das (Kunst-)Urhebergesetz, im digitalen Umfeld.
- **Sicherheit.** In diesen Kompetenzbereich fällt das bewusste Schützen von Geräten, persönlichen Daten und allgemein der Privatsphäre im digitalen Umfeld.
- **Problemlösung.** Dazu gehört der Umgang mit technischen Problemen und der kreative Umgang mit Technologien, sowie die Erkennung von digitalen Kompetenzlücken zur nachhaltigen Weiterentwicklung des Referenzmodells.

Bemerkung Die in 2.1 genannte *Medienkompetenz* und das *kritische Denken* sind implizit im Referenzmodell enthalten.

In Bezug auf das Referenzmodell kann eine digitale Gesellschaft nur als nachhaltig gelten, solange deren Bürger in dieser selbstbestimmt handeln können. Daraus ergibt sich der Auftrag an eine nachhaltige Hochschullehre, zu gewährleisten, dass Studierende nach Abschluss des Studiums die erforderlichen digitalen Kompetenzen dafür besitzen. Im Positionspapier des Kompetenzzentrums *Nachhaltige Universität* der Universität Hamburg⁷ wird als Maßnahme genannt, dass Studierende die Möglichkeit erhalten und nutzen sollen, ihre Lernprozesse selbst zu gestalten. In *WireGraph* erlernen Studierende Wissensgraphen zu erstellen und abzufragen. Diese Kompetenz gehört zur Unterkategorie Programmierung und ist damit dem Kompetenzbereich *Erstellung digitaler Inhalte* zuzuordnen. Damit die Studierenden ihre Lerninhalte und -methoden allerdings selbst (mit-)gestalten können, ist es notwendig, dass das Wissen nicht nur gelernt, sondern auch angewendet wird. Wie es bereits im Referenzmodell aufgeführt wird, müssen eigene Inhalte erstellt und Ideen umgesetzt werden - aus dem konsumierenden Studierenden wird ein Prosument im digitalen Zeitalter.

2.3 Der Prosument im digitalen Zeitalter

Prosument ist ein Mischwort, bestehend aus den Begriffen *Produzent* und *Konsument*. Der Begriff *Prosumerismus* wurde erstmals im Jahre 1980 eingeführt [To80]. Im digitalen Umfeld ist derjenige Prosument, der sowohl digitale Inhalte selbst erstellt als auch Inhalte des gleichen Kontexts abrufen bzw. verwertet. Die Erstellung und Verwertung kann dabei synchron oder asynchron stattfinden.

Die Prosumentenrolle wird im Referenzmodell implizit in zwei Dimensionen adressiert. (A) Auf der Makro-Ebene beinhaltet der Kompetenzbereich *Informations- und Datenkompetenz* hauptsächlich Anforderungen, die das korrekte Konsumieren von digitalen Inhalten betrifft. Die Produzentenrolle spiegelt sich im Kompetenzbereich *Erstellung digitaler Inhalte* wider. (B) Zusätzlich erfolgt eine zweite Trennung zwischen Konsument und Produzent innerhalb jedes Kompetenzbereichs auf der Mikro-Ebene in Folge der acht Leistungsniveaus [CVP17]. So ist es erforderlich, dass jeder zunächst erprobte Methoden und Inhalte der Kompetenzbereiche versteht und anwenden kann (Konsument; Stufe 1 – 4), bevor neue Inhalte gelehrt und hinzugefügt werden (Produzent; Stufe 5 – 8). Eine Gegenüberstellung der Leistungsniveaus und der Prosumentenrolle wird in Tabelle 1 dargestellt. Die Prosumentenrolle ist dabei in jedem Leistungsniveau vertreten, doch wird sie für die Nachhaltigkeit umso relevanter, desto mehr Produzentenanteile sie einnimmt. Besonders im Hinblick auf die forschungsorientierte Lehre sollte das Erreichen des achten Leistungsniveaus angestrebt werden.

Um eine nachhaltige Hochschullehre zu gewährleisten, muss die Prosumentenrolle der Studierenden im Detail verstanden werden. Die in Abschnitt 1 erwähnte Forderung an die Hochschullehre bezieht sich auf die Vermittlung wichtiger digitaler Kompetenzen nach (A), die zwischen Studiengängen unterschiedlich gewichtet werden können. So werden

⁷ <https://www.nachhaltige.uni-hamburg.de/downloads/2015-04nachhaltigkeitskonzept-knu-team2-endfassung.pdf>, letzter Zugriff am 14.04.2021.

Informatikstudierende im digitalen Umfeld öfter die Produzentenrolle einnehmen als Medizinstudierende. Viel bedeutender für eine nachhaltige Hochschullehre ist allerdings, dass Studierende die Möglichkeit erhalten, jede gelehrte digitale Kompetenz über alle Leistungsniveaus nach (B) kennenzulernen, sodass Studierende und Absolventen:innen bei der voranschreitenden Digitalisierung mitwirken können.

Leistungsniveau	Beispiel	PR
1 – Einfache Aufgaben mit Hilfe ausführen	MP mit Supervisor schreiben/versenden	
2 – Ein. Aufgaben teilw. mit Hilfe ausführen	MP mit Rückfragen schreiben/versenden	
3 – Einfache Probleme selbstständig lösen	Formatierung-/Anhangsfunktion nutzen	K
4 – Komplexe Probleme selbstständig lösen	File-Hosting-Dienst für Anhänge nutzen	
5 – Geeignete Lösungswege kommunizieren	Instant-Messenger einführen	
6 – Besten Lösungsweg kommunizieren	Instant-Messenger für MP optimieren	P
7 – Lösungswege verbessern	API des Instant-Messengers nutzen	
8 – Kompetenzbereich erweitern	Sprach-/Texterkennung integrieren	

Tab. 1: Gegenüberstellung der Leistungsniveaus des Referenzmodells und der Prosumentenrolle im Kompetenzbereich *Erstellung digitaler Inhalte*. Als Beispiel dient ein Szenario, in dem ein Mitarbeiter ein Meetingprotokoll (MP) per Email an sein Team senden soll. (PR - Prosumentenrolle; K - Konsumentenrolle; P - Produzentenrolle).

3 Aktueller Stand der eingesetzten Lehr-Lern-Konzepte

Eine Übersicht bestehender Lehr- und Lernformate präsentiert Kuhn et al. [KFT18]. Zwar bezieht sich die Übersicht auf die medizinische Ausbildung, kann aber generisch für die Hochschullehre betrachtet werden.

Das Lernspiel wird zukünftig als unterstützendes Medium in zwei Lehrveranstaltungen eingesetzt: Eine Lehrveranstaltung über Wissensgraphen und einer Lehrveranstaltung der Prüfungstechnik. Hier werden abgesehen von Übungen, digitale Präsentations- und Videodateien eingesetzt, die für Studierende auf einer E-Learning-Plattform abrufbar sind. Zusätzlich wird das Lernspiel *AuditSim* in der Prüfungstechniklehre eingesetzt. Für diese Lehr-Lernkonzepte wird nun die Möglichkeit einer Prosumentenumgebung analysiert:

- *Digitale Präsentations- und Videodateien* werden meist von den Dozierenden bereitgestellt, wodurch Studierende vornehmlich auf der Makro-Ebene die Konsumentenrolle einnehmen. In Praktika, Seminaren und Abschlussarbeiten sind Präsentationen und somit die Erstellung digitaler Präsentationsdateien für Studierende oftmals verpflichtend. Wenn diese digitale Kompetenzausprägung gelernt wird, sollte die Erreichung des achten Leistungsniveaus angestrebt werden. Das bedeutet nicht nur die Verwendung bestehender Vorlagen, sondern auch die Erarbeitung eigener Elemente, die zum Beispiel in Form von Designs oder kreativer Elementkombinationen erfolgen können. Nach diesem Ansatz ist es also für den Studierenden nicht vorteilhaft, wenn Universitäten Präsentationsvorlagen bereitstellen, da so die Möglichkeit der Produktion entfällt.
- *Lernspiele* fördern das kritische Denken und die Fähigkeit kreativ Probleme zu lösen [Jo14]. Sie können als didaktisches Mittel in der Hochschullehre als Teil eines Kurses eingesetzt werden [Jo11]. Dabei ist wichtig, dass Lernspiele in engen Bezug zu den vermittelnden Lerninhalten stehen [ZML12]. Abgesehen von *AuditSim*, wird auch das Lernspiel *SQL Island* als repräsentatives Programmier-Lernspiel betrachtet:
 - In *AuditSim*⁸ schlüpfen die Teilnehmer in die Rolle eines Wirtschaftsprüfers. Ein virtueller Manager stellt Aufgaben und andere Ressourcen, in Form von Dokumenten oder Videos, zur Verfügung. Die Bearbeitung und Bewertung der Aufgaben finden in einem unabhängigen Textverarbeitungsprogramm statt. Dadurch ist ein Spielleiter erforderlich. Studierende sind hierbei Konsumenten der *Informations- und Datenkompetenz*. Allerdings entfällt die Informationssuche. Studierende haben keine Möglichkeit die Produzentenrolle einzunehmen. Dafür müsste es ihnen freistehen, Such- und Evaluationsmethoden selbst zu bestimmen. Neue Suchmöglichkeiten könnten durch die Anwendung von SQL oder SPARQL ermöglicht werden, durch die Datenbanken abgefragt werden können.
 - *SQL Island*⁹ ist ein browserbasiertes Lernspiel, um „SQL-Anfragen zu verstehen und zu formulieren“ [SD15]. SQL ist als Programmiersprache dem Kompetenzbereich *Erstellung digitaler Inhalte* zuzuordnen. Die Lernaufgaben beziehen sich dabei auf die *Data Manipulation Language* (DML) von SQL, mit der relationale Datensätze geändert und abgefragt werden können. *SQL Island* wird nach Aussage der Autoren an Hochschulen „in Datenbankvorlesungen als unterstützende Übung eingesetzt“ [SD15]. Das Lernziel dieses Spiels ist es Grundlagen zu vermitteln. Studierende können daher ein Leistungsniveau bis zur 3.Stufe hinsichtlich SQL-Kenntnissen erreichen. Um Prosument zu werden, müsste *SQL Island* es Studierenden zusätzlich ermöglichen, die Datenbasis zu erweitern, eigene *Quests* zu definieren und die Ergebnisse mit Kommilitonen zu besprechen.

⁸ <https://www.pw.rw.fau.de/studium-lehre/master/schlüsselqualifikation-fact/>, letzter Zugriff am 14.04.2021.

⁹ <https://sql-island.informatik.uni-kl.de>, letzter Zugriff am 14.04.2021.

4 WireGraph - Wissensgraphen spielerisch lernen

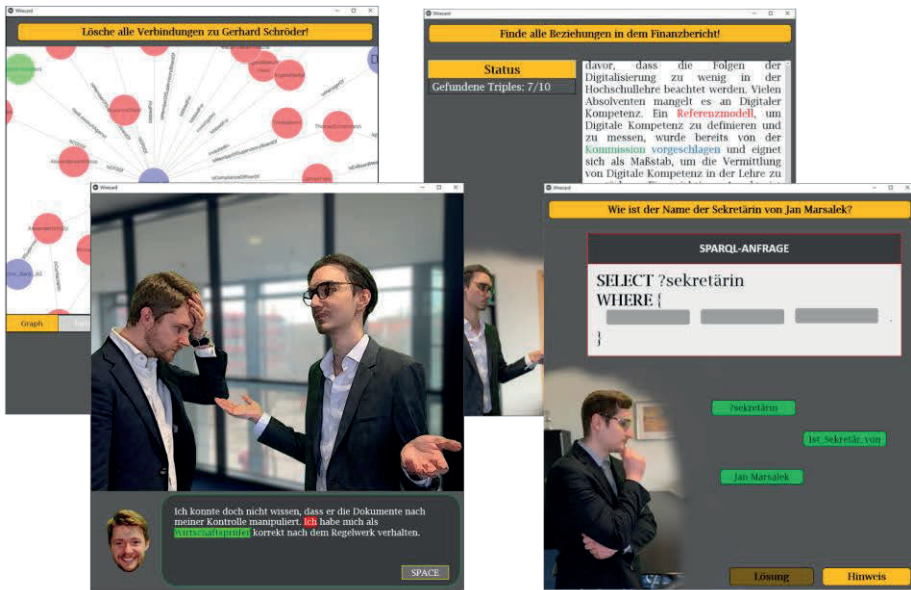


Abb. 1: *WireGraph*-Konzeptbilder zeigen (v.l.n.r.): Visuelles Graphenmodellierungswerkzeug, Dialog- und Dokumentenanalyse, um Graphen zu erweitern, SPARQL-Abfrage (Drag&Drop).

4.1 Handlung

Der in 2020 aufgedeckte Wirecard-Skandal, der bis heute noch nicht vollständig aufgeklärt ist, ist thematisch dem Finanzbereich zuzuordnen, da das in der digitalen Zahlungsabwicklung spezialisierte Unternehmen eine Fehlsomme in Höhe von 1,9 Milliarden Euro mit nicht vorhandene Treuhandkonten in Asien verschleierte¹⁰, was auch unter Studierenden bekannt sein sollte. In *WireGraph* übernimmt der Spieler die Rolle eines Investigativjournalisten, der die wichtigsten Meilensteine der Skandalauflösung durchläuft. Dies ist auch deshalb realitätsnah, da Dan McCrum, Journalist der Financial Times, einen wichtigen Beitrag zur Aufdeckung des Skandals beitrug¹¹. Die Handlung orientiert sich dabei am Krimi-Genre, u.a. auch weil dieses das populärste Genre der Belletristik ist¹². Der Spieler muss daher Personen von Interesse befragen oder Emails durchleuchten, um darauf aufbauend einen Graphen zu bauen oder den bestehenden Graphen mit gewonnenen Indizien als Filter abzufragen. Damit würde sich das Spiel, wie in einem Kriminalroman, im weiteren Verlauf des Geschehens weiterentwickeln. Zum

¹⁰ <https://www.ft.com/content/284fb1ad-ddc0-45df-a075-0709b36868db>, letzter Zugriff am 14.04.2021.

¹¹ <https://www.ft.com/content/745e34a1-0ca7-432c-b062-950c20e41f03>, letzter Zugriff am 14.04.2021.

¹² <https://www.splendid-research.com/de/studie-buecher.html>, letzter Zugriff am 14.04.2021.

einem erfährt der Spieler mehr über den Wirecard-Skandal und zum anderen wird der Spieler mit neuen Herausforderungen konfrontiert, indem die Komplexität der IT-Inhalte ansteigt.

4.2 Zielgruppe und Lernziele

Die Zielgruppe des Spiels sind Masterstudierende der Betriebswirtschaftslehre und der Wirtschaftsinformatik, da die Modulhandbücher beider Gruppen an der Autorenuniversität bereits Vertiefungen zum Thema *Big Data* beinhalten. Zudem konnten beide Gruppen bereits Fachwissen ansammeln, das ihnen helfen könnte, kreative Produktionen an dem Lernspiel vorzunehmen (wie zum Beispiel zum Thema Vertrags- oder Urheberrecht). Abschließend ermöglicht die geltende Studienprüfungsordnung beiden Gruppen fortgeschrittene Kurse im Bereich Wissensgraphen zu besuchen, sodass es ihnen offensteht, sich weiter in das Thema zu vertiefen. Die Lernziele orientieren sich an dem Referenzmodell der digitalen Kompetenz und sind wie folgt definiert:

- Teilnehmer verstehen das grundlegende Arbeiten mit Wissensgraphen.
- Teilnehmer können SPARQL-Abfragen schreiben, um Informationen auf Wissensdatenbanken, wie Wikidata¹³, abrufen zu können.
- Teilnehmer können selbstständig Graphen auf Basis von Turtle- oder RDF/XML-Ausdrücken erstellen.
- Teilnehmer verstehen den Wirecard-Skandal und ihre Zusammenhänge.
- Teilnehmer verstehen den Aufbau und die Möglichkeiten der Lernspiel-Vorlagen und können alternative Lernszenarien entwickeln.

4.3 Umsetzung der Prosumentenumgebung

Das Lernspiel wird in Python und dem Kivy-Framework programmiert und ist derzeit noch in der Entwicklung. Das Spiel ist in Szenen gegliedert, wobei jede Szene eine andere Spielmechanik aufweisen kann, wie zum Beispiel die einer Visual-Novel- oder einer Point&Click-Spielweise. Dies soll mit Abbildung 1 veranschaulicht werden. Aktuell existieren Testszenen, mit denen die erforderlichen Interaktionselemente bestätigt wurden. Jede verfügbare Spielmechanik steht durch eine Szenen-Vorlage zur Verfügung, das mit Hilfe von RDF-Ausdrücken konfiguriert wird. Damit ist gewährleistet, dass die Teilnehmer ihr konsumiertes Wissen direkt für die Produktion eigener Inhalte verwenden können. Python-Kenntnisse sind daher nicht erforderlich, können aber genutzt werden, um neue Vorlagen zu erstellen. Das Hauptprojekt, sowie Projekte der Studierenden, werden auf GitHub gehostet. Das hat den Vorteil, dass Studierende ihre Projekte als Referenz potenziellen Arbeitgebern vorlegen und eine geeignete Softwarelizenz wählen können.

¹³ <https://www.wikidata.org/>, letzter Zugriff am 14.04.2020.

Vergleichbar mit der Pygame-Projektseite¹⁴ wird eine Übersicht aller Projekte gehostet, die auf das jeweilige GitHub-Projekt verweist, damit Studierende einen schnellen Überblick über die verfügbaren Lehrmaterialien bekommen. Die Prosumentenumgebung ergibt sich daher aus verfügbaren Spielen zum Erlernen der Thematik, der Übersichtsseite mit Kommentarfunktion zum Austausch und den Szenen-Vorlagen zur Erstellung neuer Lernspiele. Dadurch deckt die Prosumentenumgebung alle vier Unterpunkte des Kompetenzbereichs *Erstellung digitaler Inhalte* ab: Entwicklung von digitalem Inhalt, Überarbeitung von digitalem Inhalt, Lizenzvergabe und Programmieren.

4.4 Didaktische Einbettung und geplante Evaluation

Um die digitale Kompetenzausprägung hinsichtlich Wissensgraphen nachhaltig und vollständig zu vermitteln, muss das Lernspiel so ausgestaltet sein, dass Studierende die acht Leistungsniveaus des Referenzmodells in linearer Reihenfolge erreichen können [CVP17]. Das bedeutet, dass zunächst ein Verständnis aufgebaut, dann Wissen bewusst angewendet und abschließend kreativ erweitert werden soll. Für die Prosumentenrolle wird also zunächst die Konsumentenrolle und anschließend die Produzentenrolle eingenommen. Um die Prosumentenumgebung für *WireGraph* abzubilden, müssen drei Phasen durchlaufen werden:

1. Konsumtion. Studierende spielen das Lernspiel.
2. Reflektion. Studierende diskutieren und evaluieren das gespielte Lernspiel mit anderen Kommilitonen.
3. Produktion. Auf Basis der Reflektionsphase erweitern bzw. verbessern Studierende das Spiel, um das Lernziel besser zu adressieren.

Die Entwicklung von *WireGraph* ist im Oktober 2021 abgeschlossen. Ab dem Wintersemester 2021/22 wird *WireGraph* in der Prüfungstechnik-Lehrveranstaltung als informatives Medium eingesetzt. Ab dem Sommersemester 2022 wird *WireGraph* auch als unterstützendes Medium im Rahmen der Wissensgraph-Lehrveranstaltung für Studierende eingesetzt. Das Lernspiel kann von den Studierenden explizit dazu genutzt werden, um das Schreiben von RDF-Ausdrücken und SPARQL-Abfragen zu üben (Konsumtion), da diese Fähigkeit für ein erfolgreiches Bestehen der Abschlussprüfung erforderlich ist. Am Ende beider Veranstaltungen wird den Studierenden ein Evaluationsbogen bezüglich *WireGraph* gegeben. Im Wintersemester 2022/23 wird Studierenden ein Praktikum angeboten, in dem sie *WireGraph* erweitern oder eigene Lernszenarien mit Hilfe der Vorlagen umsetzen (Produktion). Mit diesen Studierenden wird ein qualitatives Interview durchgeführt. Eine quantitative Analyse ist zunächst nicht vorgesehen, wird aber momentan nicht ausgeschlossen. Mit dem Wintersemester 2022/23 endet die erste Iteration.

¹⁴ <https://www.pygame.org/tags/all>, letzter Zugriff am 14.04.2021.

Folgende Forschungsfragen (FF) werden dabei für *WireGraph* aufgestellt und sollen nach der ersten Iteration beantwortet werden:

- FF1 Stellt die Prosumentenumgebung von *WireGraph* ein geeignetes Mittel dar, den Kompetenzbereich *Erstellung digitaler Inhalte* für Wissensgraphen vollständig abzubilden (Makro-Ebene)?
- FF2 Erreichen Studierende durch die Prosumentenumgebung von *WireGraph* alle acht Leistungsniveaus des Kompetenzbereichs *Erstellung digitaler Inhalte* (Mikro-Ebene)?
- FF3 Zeigen Studierende durch die Prosumentenumgebung eine höhere Bereitschaft sich mit dem Thema auseinanderzusetzen, weil sie beispielsweise digitale Inhalte (in Zusammenarbeit oder in Konkurrenz mit Kommilitonen) erstellen können (Peer-Motivation [Ha04]; Produktion)?

5 Zusammenfassung und Ausblick

Es besteht ein großer innerer und äußerer Druck für die Hochschullehre digitale Kompetenzen nachhaltig zu vermitteln. Unter anderem wird *Big Data* als zukunftsweisendes Thema für den Finanzbereich genannt, weshalb Studierende der Wirtschaftswissenschaften sich mit Technologien im Big Data-Umfeld auseinandersetzen müssen, um auf dem Arbeitsmarkt konkurrenzfähig zu bleiben. Wissensgraphen sind eine dieser Technologien und die Arbeit mit diesen soll mit dem vorgestellten Lernspiel *WireGraph* erlernt werden. Um die digitale Kompetenz einzuordnen, schlägt die Europäische Kommission ein geeignetes Referenzmodell der digitalen Kompetenz mit fünf Kompetenzbereichen und acht Leistungsniveaus vor, deren Anwendung auf Lehr-Lern-Konzepte noch Forschungslücken aufweist. Zusätzlich wird das Konzept der Prosumentenumgebung vorgeschlagen, die für das Lernspiel *WireGraph* erstmals umgesetzt wird, um die Kompetenz der Wissensgrapharbeit im Rahmen des Kompetenzbereichs *Erstellung digitaler Inhalte* vollständig und über alle Leistungsniveaus hinweg abzudecken. Dabei müssen Studierende die Prosumentenrolle einnehmen, in der sie digitales Fachwissen nicht nur konsumieren, sondern auch weiterentwickeln. Studierende aufarbeiten in *WireGraph* mit Hilfe von RDF und SPARQL den Wirecard-Skandal. Anschließend haben sie die Möglichkeit mit Hilfe der Szenen-Vorlagen eigene Ideen und Fachwissen in einem eigenen Wissensgraph-Lernspiel umzusetzen. *WireGraph* wird erstmalig im Wintersemester 2021/22 eingeführt und nach einem Jahr wird *WireGraph* für alle drei Phasen *Konsumption*, *Reflektion* und *Produktion* evaluiert worden sein. Dazu wurden drei Forschungsfragen aufgestellt, denen nachgegangen wird, um erste Ergebnisse zu erhalten und damit die Bedeutung der Prosumentenumgebung für die zukünftige Hochschullehre zu ermitteln.

Literaturverzeichnis

- [CR18] Cockcroft, S.; Russell, M.: Big Data Opportunities for Accounting and Finance Practice and Research. *Australian Accounting Review*, 28(3):323–333, 2018.
- [CVP17] Carretero, S.; Vuorikari, R.; Punie, Y.: The digital competence framework for citizens. Publications Office of the European Union, 2017.
- [Eu06] Europäisches Parlament und Rat: Empfehlung vom 18. Dezember 2006 zu Schlüsselkompetenzen für lebensbegleitendes Lernen. *Amtsblatt der Europäischen Union*, L394:10 – 18, Dezember 2006.
- [Eu17] Europäische Kommission: Mitteilung der Kommission an das Europäische Parlament, den Rat, den Europäischen Wirtschafts- und Sozialausschuss und den Ausschuss der Regionen- über eine europäische Erneuerungsagenda für die Hochschulbildung. COM, 247, Mai 2017.
- [GHV19] Gulin, D.; Hladika, M.; Valenta, I.: Digitalization and the Challenges for the Accounting Profession. *Proceedings of the ENTRENOVA - ENTERprise REsearchInNOVAtion Conference*, 5(1):428–437, Oct. 2019.
- [Ha04] Hancock, D.: Cooperative Learning and Peer Orientation Effects on Motivation and Achievement. *The Journal of Educational Research*, 97(3):159–168, 2004.
- [HS17] Henselmann, K.; Scherr, E.: Auswirkungen der Digitalisierung auf den Berufsnachwuchs in der Wirtschaftsprüfung. In (Baldauf, J.; Graszitz, S., Hrsg.): *Theorie und Praxis aus Rechnungswesen und Wirtschaftsprüfung*, S. 231–241. LexisNexis Verlag, Wien, 2017.
- [Jo11] Johnson, L. et.al.: The 2011 Horizon Report. Bericht, EDUCAUSE, 2011.
- [Jo14] Johnson, L. et.al.: Horizon report Europe: 2014 schools edition. Bericht, EDUCAUSE, 2014.
- [KFT18] Kuhn, S.; Frankenhauser, S.; Tolks, D.: Digitale Lehr- und Lernangebote in der medizinischen Ausbildung. *Bundesgesundheitsblatt-Gesundheitsforschung-Gesundheitsschutz*, 61(2):201–209, 2018.
- [Ni20] Nigam, V. et.al.: A Review Paper On The Application Of Knowledge Graph On Various Service Providing Platforms. In: 2020 10th International Conference on Cloud Computing, Data Science Engineering (Confluence). S. 716–720, 2020.
- [Ra18] Rat der Europäischen Union: Empfehlung vom 22. Mai 2018 zu Schlüsselkompetenzen für lebenslanges Lernen. *Amtsblatt der Europäischen Union*, C189:1–13, Mai 2018.
- [SD15] Schildgen, J.; DeBloch, S.: SQL-Grundlagen spielend lernen mit dem Text-Adventure SQL Island. In (Seidl, T.; Ritter, N.; Schöning, H.; Sattler, K.; Härder, T.; Friedrich, S.; Wingerath, W., Hrsg.): *Datenbanksysteme für Business, Technologie und Web (BTW 2015)*. Gesellschaft für Informatik e.V., Bonn, S. 687–690, 2015.
- [Si18] Sicilia, M. et.al.: Digital Skills Training in Higher Education: Insights about the Perceptions of Different Stakeholders. In: *Proceedings of the Sixth International Conference on Technological Ecosystems for Enhancing Multiculturality. TEEM'18*, Association for Computing Machinery, New York, NY, USA, S. 781–787, 2018.

- [To80] Toffler, A.: The third wave, Jgg. 484. Bantam books New York, 1980.
- [ZML12] Zender, R.; Moebert, T.; Lucke, U.: RouteMe - Routing in Ad-hoc-Netzen als pervasives Lernspiel. In (Desel, J.; Haake, J.; Spannagel, C., Hrsg.): DeLFI 2012: Die 10. E-Learning Fachtagung Informatik der Gesellschaft für Informatik e.V. Gesellschaft für Informatik e.V., Bonn, S. 201–212, 2012.

Distanzlabore in der ingenieurwissenschaftlichen Ausbildung

Christian Bunse¹ und Lieven Nils Kennes² und Jan-Christian Kuhr³

Abstract: Laborveranstaltungen sind vor allem in ingenieurwissenschaftlichen Studiengängen ein unverzichtbarer Bestandteil der Ausbildung. Labore vertiefen theoretisch erworbene Grundlagen, überführen sie in praxisbezogenes Handlungswissen und Motivieren Studierende zu eigenständigen Problemlösungen. Akute Probleme der Laborausbildung sind das Verhältnis von Studierenden zu Laborplätzen, Mangel an Ressourcen sowie die Verfügbarkeit von moderner Gerätetechnik. Das hier vorgestellte Konzept (*DistLab*) realisiert das bereits länger diskutierte Konzept der Distanzlabore für ausgewählte Module ingenieurwissenschaftlicher Studiengänge in zwei Säulen: Hybrid und virtuell. Ein hybrides Labor ermöglicht die Nutzung real existierender Geräte von einem räumlich entfernten Standpunkt. Ein virtuelles Labor bildet reale Geräte vollständig digital ab und ermöglicht die Versuchsdurchführung auch von zu Hause aus. Beide Konzepte ermöglichen die orts- sowie zeitungebundene Realisierung von Laborveranstaltungen und bieten dadurch einen deutlichen Mehrwert hinsichtlich Flexibilität, Qualität und selbstbestimmter Lerngeschwindigkeit. Dieser Vorteil ist so groß, dass Distanzlabore auch in Nicht-Pandemiezeiten aufrechterhalten werden sollten.

Keywords: Distanzlehre; Labore; ingenieurwissenschaftliche Ausbildung.

1 Einleitung

Labore sind ein wichtiger Ausbildungsbestandteil an Hochschulen, insbesondere in den Bereichen Mathematik, Informatik, Naturwissenschaft und Technik (MINT). Eine Laborveranstaltung erlaubt Studierenden praktische Erfahrungen zu sammeln, theoretisch erworbenes Wissen anzuwenden und selbstbestimmt zu lernen. Laborveranstaltungen sind somit ein wichtiger pädagogischer Aspekt zur Erreichung von Lernzielen.

Nicht erst seit den Digitalisierungsherausforderungen der COVID-19-Pandemie besteht ein großer Bedarf, Labore zu virtualisieren bzw. per Fernzugriff (bspw. über das Internet) verfügbar zu machen [CC95, SP00]. Die Digitalisierung bzw. Virtualisierung von Laborveranstaltungen bietet Studierenden die Möglichkeit, Labore auch außerhalb der regulären Lehrzeiten und damit zeitlich flexibel zu nutzen. Zusätzlich ermöglichen sie die Nutzung von Laboreinrichtungen für größere Studierendengruppen und somit eine effizientere Nutzung von Ressourcen.

¹ Hochschule Stralsund, Fakultät für Elektrotechnik und Informatik, christian.bunse@hochschule-stralsund.de

² Hochschule Stralsund, Fakultät für Wirtschaft, lieven.kennes@hochschule-stralsund.de,

³ Hochschule Stralsund, Fakultät für Maschinenbau, jan-christian.kuhr@hochschule-stralsund.de

Im Rahmen dieser Veröffentlichung wird mit *DistLab* ein Konzept zur Labordurchführung aus der Distanz zur Diskussion gestellt. Aufbauend auf Erfahrungen mit digitalen Werkzeugen (z.B. virtuelle Maschinen bzw. Desktop-Virtualisierung) bspw. in den Bereichen Software-Entwicklung und Data Science, soll *DistLab* die ortsungebundene, digitale Durchführung von Laborveranstaltungen in Laboren mit kostenintensiver und vor allem ortsfester Geräteausstattung in ingenieurwissenschaftlichen Studiengängen ermöglichen.

Zur curricularen Verankerung und vor allem zur Vereinheitlichung der Infrastruktur für alle Labore in der ingenieurwissenschaftlichen Ausbildung einer Hochschule muss ein solches Konzept adaptierbar, portierbar und vor allem skalierbar sein. Zusätzlich ist zur Unterstützung fremdsprachiger Studierender das System mehrsprachig zu gestalten. Zur Realisierung aller Qualitätsziele setzt das vorgestellte Konzept auf eine zentral gehostete Infrastruktur zur Unterstützung der Durchführung von Laborveranstaltungen in hybrider (Programmierung und Steuerung von real existierenden Systemen über das Netz) und in rein virtueller Form. So werden Schwächen anderer Ansätze wie bspw. der Fokus auf einzelne Experimente vermieden. Neben technischen Aspekten erfordert eine curriculare Verankerung zusätzlich die Entwicklung von Strukturen und Regeln, um im Rahmen der virtuellen Labore ergebnissichere Prüfungen und Testate durchführen zu können.

Insgesamt soll *DistLab* in der ingenieurwissenschaftlichen Ausbildung die orts- und zeitflexible Nutzung von Versuchsequipment ermöglichen. Hierdurch wird die Auslastung der Geräte verbessert und der Teilnehmerdurchsatz erhöht. Das vorgestellte Distanzkonzept wird ergänzt durch eine zentrale und skalierbare IT-Infrastruktur innerhalb der Hochschule, so dass eine Übertragbarkeit auch auf andere Laborkontexte gegeben ist. Insgesamt lässt sich so die Angebotsvielfalt der Hochschulen und die Ausbildungsqualität verbessern. Durch die zeit- und ortsunabhängige Verfügbarkeit wird gleichzeitig das selbstbestimmte Lernen der Studierenden unterstützt.

Im Folgenden wird zunächst auf existierende Ansätze und Überlegungen zum Laborbetrieb in der Distanzlehre eingegangen (Abschnitt 2). Darauf aufbauend werden die Konzeption eines generischen und skalierbaren Distanzlabors, basierend auf virtuellen und hybriden Strukturen (Abschnitt 3), sowie die Aspekte und Herausforderungen von dessen Realisierung (Abschnitt 4) betrachtet. Die curriculare Verankerung dieser Konzepte wird in Abschnitt 5 und die Evaluierungs- und Skalierungsstrategien werden in Abschnitt 6 diskutiert. Abschnitt 7 fasst die Ergebnisse zusammen und gibt einen Ausblick auf zukünftige Arbeiten.

2 Stand der Forschung und Praxis

Computer-basiertes Lernen ist in der Hochschulbildung von stetig wachsender Bedeutung, eng verbunden mit der Hoffnung, das Lernen für Studierende zeit- und ortsunabhängig zu machen, Qualität zu standardisieren und Kosten zu senken. Dieser Trend wurde durch die COVID-19-Pandemie nochmals verstärkt. In der Praxis zeigt sich

aber rasch, dass insbesondere Labore in der ingenieurwissenschaftlichen Ausbildung besonderer Aufmerksamkeit bedürfen. Ansätze wie bspw. der bwLehrpool [TS15] oder die Arbeiten von Meier et al. [MG17] zeigen, dass durch Desktop-Virtualisierung Labore in der Mathematik und in vielen Bereichen der Informatik erfolgreich orts- und organisationsunabhängig angeboten werden können. Diese Ansätze lassen sich aber nicht einfach auf ingenieurwissenschaftliche Laborveranstaltungen übertragen, in denen der Zugriff auf bzw. die Beobachtung von ortsfesten Geräten notwendig ist (bspw. Labore in der Robotik oder der Automatisierungstechnik).

Zurzeit existieren keine *standardisierten* Lösungen zur Durchführung von Laboren in der Ingenieurausbildung in Form von E-Learning-Veranstaltungen. Dies verwundert, da die Idee der Remote-Nutzung von Laborveranstaltungen bereits seit einiger Zeit diskutiert wird [CC95, SP00, RJ00, AC08, GF08]. So forderte bereits Schwarz et al. [SP00] die Einrichtung von Tele-Laboren, um Studierenden die Durchführung physikalischer Experimente über das Web zu ermöglichen.

Eine empirische Untersuchung zum Einsatz von Distanzlaboren [VP18] zeigt, dass bisherige Ansätze nur für wenig komplexe Laboraufgaben geeignet sind und dass trotz hohem Automatisierungsgrad weiterhin eine intensive Begleitung durch Lehrkräfte erforderlich ist. Post et al. [PG19] untersuchten empirische Studien zu Distanzlaboren und stellten fest, dass diese insgesamt einen positiven Effekt auf die Erreichung von Lernzielen, das erworbene konzeptuelle Wissen und die Zufriedenheit der Studierenden haben. Voraussetzung ist hierbei aber, dass die Handlungsmöglichkeiten der Studierenden in der Laborveranstaltung, zu Hause und an der Hochschule vergleichbar sind.

Rigby [RD06] analysierte verschiedene Ansätze für Distanzlabore hinsichtlich der Korrelation zwischen Übertragbarkeit (zwischen Kursen), Kosten, Verwaltungsaufwand und Zeit. Hierbei zeigte sich, dass Flexibilität und Skalierbarkeit des Konzepts ausschlaggebend für den Erfolg sind. Dies wurde in einer weiteren Studie [MH06] zu unterschiedlichen Implementierungsmodellen untersucht und bestätigt. Gravier [GF08] identifizierte Charakteristika von Distanzlaboren hinsichtlich des Mangels an Flexibilität, Skalierbarkeit, Wartbarkeit und Wiederverwendbarkeit existierender Ansätze. Distanzlabore fokussieren oft auf spezielle Geräte und Versuche mit geringer Komplexität. Yan [YL06] verweist auf den Einsatz von Web-Services um Interoperabilität zu erreichen und somit Labore „einfacher“ skalieren und portieren zu können.

Neuere Ansätze wie die von der Europäischen Union geförderten Projekte SimLab, HomeLab oder das ebenfalls von der Europäischen Union geförderte Network of Excellence „Robotics und Mechatronics“, verfolgen Ideen zur Verbreiterung des Einsatzes von Distanzlaboren mit Fokus auf Portabilität und Skalierbarkeit (<https://robolabor.ee/homelab/doku.php>). Der Fokus liegt dabei zumeist auf einzelnen Versuchen bzw. einfachen Systemen auf Basis von Mikrocontrollern sowie singulären (d.h. isoliert durchführbaren) physikalischen Experimenten. Das KUKA Robot Learning

Lab am KIT in Karlsruhe (https://www.ipr.kit.edu/projekte_2891.php) ist ein sich in der Entwicklung befindliches, aus der Ferne zugängliches Robotiklabor. Das Labor soll Studierenden und Forschern ermöglichen, über das Internet auf Roboter (KUKA LBR iiwa Roboter) im Labor zuzugreifen und ihre eigenen Projekte auszuführen. Skalierung und Portabilität gehören nicht zu den Zielen des Projekts. [BD21] zeigt die Auswirkungen von Distanzveranstaltungen in der Robotik auf das Lehr/Lernverhalten ohne Zugriff auf Geräte und ortsfeste Ressourcen auf.

Insgesamt zeigt sich ein hoher Bedarf an Infrastrukturen für Distanzlabore, aber auch, dass existierende Ansätze hohe Investitionen und ebenfalls hohe laufende Kosten (Zeit, Personalaufwand, etc.) erfordern. Ein weiteres Problem ist der Fokus auf kleine, von der Komplexität überschaubare Experimente. Distanzlabore mit Großgeräten oder frei konfigurierbaren Robotern, bspw. im Kontext des autonomen Fahrens, fehlen dagegen. Ansätze zur Realisierung von Distanzlaboren sind meist aus einem konkreten Bedarf entstanden und dienen der Realisierung eines einzelnen Labors bzw. von Laboren in einem speziellen Themenfeld. Die hierzu geschaffenen Konzepte und Infrastrukturen sind, anders als bei der Desktop-Virtualisierung, nicht einfach auf andere Kontexte bzw. Labore übertragbar. Interessanterweise ist dies einer der Hauptgründe für die bislang geringe Verbreitung von Distanzlaboren [HA07].

3 Die DistLab-Idee

Der hier vorgestellte Ansatz zur Realisierung von Distanzlaboren in der ingenieurwissenschaftlichen Ausbildung basiert auf zwei Säulen (vgl. Abb. 1): Virtuelle Zwillinge (1) und Hybrid-Labore (2). Virtuelle Zwillinge sind voll funktionsfähige digitale Abbilder realer Laborgeräte und basieren auf einem hochgenauen physikalischen Modell des ihnen zugrundeliegenden realen Gegenstücks. Sie werden idealerweise auf einer zentralen Instanz ausgeführt und sind über das Netz zugreifbar. Studierende bereiten sich von zu Hause mit Hilfe des virtuellen Zwillings auf den Versuch vor, den sie später im Labor am realen Objekt durchführen. Virtuelle Zwillinge kommen zumeist bei Einzelgeräten mit überschaubaren Freiheitsgraden zum Einsatz. Hybrid-Labore erlauben den Zugriff auf existierende Infrastruktur an einer Hochschule über das Internet. Studierende arbeiten von zu Hause aus im Labor, das sich in den Räumlichkeiten der Hochschule befindet. Dieser Lösungsansatz bietet sich vor allem für Systeme an, für die keine virtuellen Zwillinge verfügbar sind oder denen eine von vielen Faktoren (Klima, Bodenbeschaffenheit, etc.) abhängige Komplexität innewohnt.

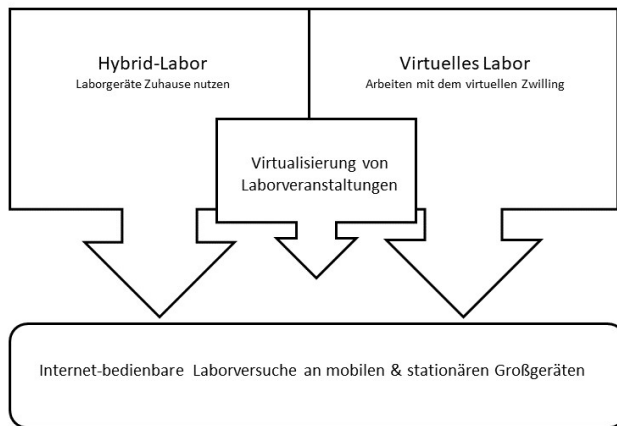


Abb. 1: Virtualisierungsstrategie

3.1 Virtuelle Zwillinge

Virtuelle Zwillinge können Abbilder von Maschinen oder von Laborgeräten sein. Ein Beispiel für Ersteres ist ein ortsfester Montageroboter, ein Beispiel für Letzteres ein Apparat zur Tank-Füllstandsregelung. Ein virtueller Zwilling hat Gemeinsamkeiten mit dem Konzept des digitalen Zwillings, unterscheidet sich von diesem jedoch in einem wesentlichen Punkt: Es findet keine Datensynchronisation zwischen realem und virtuellem Objekt statt, die den Zustand des realen Zwillings während der gesamten Lebensdauer auf sein virtuelles Gegenstück zurückspiegelt. Vielmehr ist es ausreichend, wenn eine einmal am virtuellen Objekt entwickelte Lösung auf sein reales Gegenstück übertragbar und dort funktionsfähig ist. Beide Objekte sind hinsichtlich ihres Einsatzes (Deployment) interoperabel.

Die geforderte Interoperabilität von virtuellem und realem Zwilling lässt sich gut am Beispiel einer Regelungsaufgabe erläutern, bei der die Regelstrecke als *Hardware in the Loop* (HIL) in den Regelkreis eingebunden ist. Beispielsweise realisiert das kanadische Unternehmen Quanser den HIL-basierten virtuellen Zwilling auf folgende Weise: Bei der Arbeit mit der realen Hardware finden die Schreib- und Lesezugriffe auf die Regelstrecke über eine USB-Schnittstelle statt. Wenn jedoch die Hardware durch den virtuellen Zwilling ersetzt wird, erfolgt die Interaktion mit der nunmehr *simulierten* Hardware über den Aufruf eines HTTP-basierten Dienstes, der zuvor als Programm auf dem Windows-PC gestartet wurde. Soll später die Lösung auf dem realen Objekt ausgeführt werden, wird der zu Hause entwickelte Regelalgorithmus mit in das Labor genommen und dort gegen die Hardware deployed. Dabei muss in der Schnittstelleninitialisierung der lokale Aufruf durch das USB-Interface ersetzt werden. Zukünftig soll der lokale HTTP-Aufruf durch einen remote HTTP-Zugriff auf die im Rechenzentrum der Hochschule instanziierten virtuellen Zwillinge erfolgen.

In einem virtuellen Labor arbeiten die Studierenden zunächst nicht am realen Objekt, sondern am virtuellen Zwilling. Hierdurch werden Labor-Präsenzzeiten in die häusliche Umgebung verlagert, die Verfügbarkeit der Versuche erhöht und das selbstbestimmte Lernen gefördert. Virtualisierung ist dabei im Sinne von *Blended Learning* zu verstehen. Ein virtueller Zwilling soll das Präsenzformat einer Laborveranstaltung erweitern, aber nicht vollständig ablösen. Dies kann beispielsweise so aussehen, dass Studierende den Stoff unmittelbar nach der Vorlesung am virtuellen Zwilling vertiefen und die Versuche später im Labor am echten Objekt durchführen.

3.2 Hybrid-Labore

Laborgeräte mit hoher Komplexität, hohen Freiheitsgraden (bspw. Systeme aus dem Themenfeld *Autonomes Fahren*) und flexiblen Einsatzgebieten erfordern andere Ansätze zur Virtualisierung bzw. Digitalisierung als die Nutzung von virtuellen Zwillingen. Existierende virtuelle Zwillinge und Simulatoren können die Komplexität der Bedingungen hinsichtlich Systemcharakteristika und Einsatzumgebung aufgrund der Vielzahl von Faktoren und deren gegenseitigen Abhängigkeiten nicht vollständig abbilden. Mit ihrer Hilfe erstellte virtuelle Prototypen (Hardware oder Software) müssen auf bzw. mit real existierenden Systemen evaluiert werden. Der praktische Einsatz hat gezeigt, dass die im Simulator „perfekt“ arbeitende Software in der Realität unerwartete Verhaltensweisen aufzeigt. So führen bspw. Schlupf und Unterschiede im Antrieb dazu, dass ein im Simulator geradeaus fahrender Roboter in der Realität eine Kurve fährt und daher eine Steuer- bzw. Antriebsregelung benötigt.

Zur Durchführung von Distanzlaboren mit Geräten bzw. Systemen hoher Komplexität kommt ein hybrides Konzept zum Einsatz (vgl. Abb. 2): Studierende nutzen real existierende Geräte und Einrichtungen der Hochschule von zu Hause aus über ein digitales Interface. So können bspw. autonome Fahrzeuge (im Modellformat) auf einer lokalen Aktionsfläche genutzt werden. Studierende können so Ihre Ergebnisse auf realen Maschinen, in einer realen Umgebung ausführen. Das Verhalten des Laborgeräts kann über verschiedene Videoströme beobachtet und über ein Dashboard mit Telemetriedaten kontrolliert werden. Hybrid-Labore ermöglichen somit zeit- und ortsungebundenes Lernen auch für komplexe Anwendungsfälle.

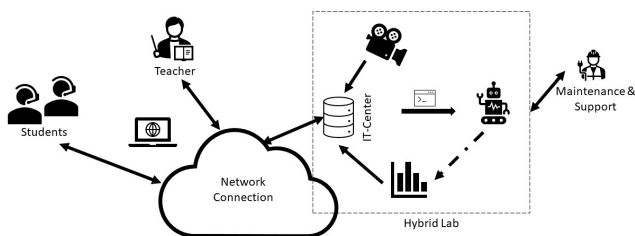


Abb. 2: Konzept – Hybrid-Labor

Die Aktionsfläche des Labors ist dem Layout des internationalen *Turtlebot* AutoRace Wettbewerbs nachempfunden. In diesem Wettbewerb müssen Roboter in einem Parcours verschiedene Herausforderungen des autonomen Fahrens bewältigen. Zusätzlich verfügt die Aktionsfläche über Ladepunkte. Roboter mit niedrigem Batteriestand fahren diese automatisiert an und stehen erst nach Aufladung wieder zur Verfügung. Nicht automatisch behebbare Probleme (bspw. Umkippen eines Roboters) müssen zu den üblichen Arbeitszeiten durch einen Laboringenieur gelöst werden. Die Architektur des Hybrid-Labors ist dabei als komponentenbasiertes System konzipiert, das in einem heterogenen Systemumfeld einsetzbar und skalierbar ist. Studierende benötigen ein internetfähiges Endgerät sowie die zur Software-Entwicklung notwendige IDE. Betrieb, Datenübertragung und Steuerung erfolgen über zentrale Systeme der Hochschule.

4 Realisierung

Distanzlabore erlauben Studierenden den Zugriff auf reale Laborgeräte über das Netz bzw. die Arbeit mit simulierten Maschinen und Geräten von zu Hause aus. Die Realisierung dieses Konzepts erfordert eine effiziente und vor allem skalierbare Infrastruktur, die auch massiven, parallelen Zugriffen gewachsen ist. Dezentrale Lösungen (jedes Labor hält eigene Kommunikations- und Interaktionsfunktionalität vor) erhöhen nicht nur Wartungsaufwände, sondern erfordern für jedes Labor separate Zugänge mit eigenen Anforderungen hinsichtlich Software und Hardware sowie der IT-Sicherheit. Im Fall des virtuellen Zwillings hat dessen Ausführung in einer lokalen Umgebung den Nachteil, dass sich bei der Vielzahl von operativen Umgebungen, die auf studentischen PCs vorliegen, häufig Installations- oder Laufzeitprobleme ergeben. Alle diese Nachteile können vermieden werden, wenn auf eine zentrale Server-Infrastruktur zurückgegriffen kann. *DistLab* setzt daher auf ein Hosting der Infrastruktur durch das Hochschulrechenzentrum (vgl. Abb. 3) und einen einheitlichen hochschulweiten Zugang beispielsweise über eine VPN-Verbindung.

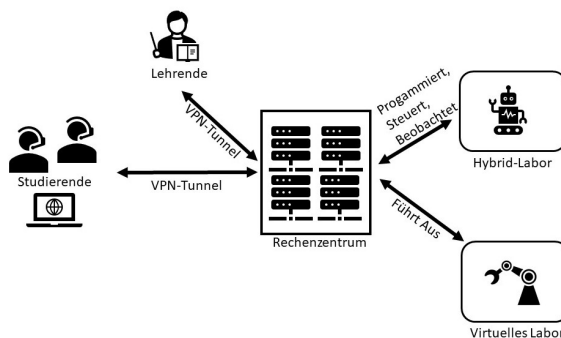


Abb. 3: *Distlab* Architektur

Die zentrale Infrastruktur ermöglicht Studierenden den sicheren und performanten

Zugang zu einzelnen Laboren. Virtuelle Zwillinge werden in einer virtuellen Maschine auf dem Server ausgeführt. Studierende erhalten Zugriff auf „ihre“ Maschine und somit exklusiven Zugang für die Dauer der Nutzung. Dort entwickelte Software steht später automatisiert im Labor zur Erprobung auf realen Geräten bereit. Bei Hybrid-Laboren wird die Kommunikation mit den realen Entitäten über die Infrastruktur realisiert. Studierende erhalten Zugriff auf eine freie Entität und Zugang zu deren Video-, Telemetrie- und Steuersystemen. Daten des hybriden Labors werden über hochschulinterne Kommunikationsstrukturen an das Rechenzentrum geliefert und von diesem für den externen Zugriff bereitgestellt.

Vorteile einer zentralisierten Infrastruktur sind neben dem vereinheitlichten Zugang, die Sicherheit der Verbindung und der ausgetauschten Daten. Weiterhin stehen zentrale Backup-Funktionalitäten zur Fehlervorsorge zur Verfügung. Studierende haben mit geeigneten Endgeräten Zugang zu allen für sie verfügbaren Laborgeräten und müssen nur wenige Anforderungen hinsichtlich Hardware und Software erfüllen. So entfällt bspw. die Installation von virtuellen Zwillingen mit spezifischen Anforderungen an Grafikhardware, die Bereitstellung und Deaktivierung von Lizenzen, sowie das Update-Management für dezentrale (studentische) Systeme. Eine zentrale Infrastruktur erscheint weiterhin in Fragen der Skalierbarkeit des Ansatzes von Vorteil. Weitere Labore einer Hochschule können, ggf. unter Erweiterung der Serverkapazitäten, vergleichsweise schnell angebunden und zugänglich gemacht werden.

5 Curriculare und hochschulweite Verankerung

Viele Arbeiten und Bestrebungen zur Etablierung von Distanzlaboren sind bislang technologiebasiert, ohne explizite Berücksichtigung pädagogischer Konzepte und moderner Lehr- oder Lernansätze. Distanzlabore können nur dann erfolgreich sein, wenn sie curricular und rechtssicher verankert sind. Konzepte für Distanzlabore müssen daher soziale Aspekte wie die Bildung von studentischen Gemeinschaften und die Unterstützung von Multi-User Interaktionen ermöglichen. So ist bspw. im Projekt bwLehrpool [TS15] neben der Virtualisierung von Lehrangeboten explizit auch die Nutzung der Infrastruktur für Leistungskontrollen und Prüfungen vorgesehen. Im Rahmen des hier vorgestellten Konzeptes soll unter anderem auf dort gemachte Erfahrungen aufgesetzt werden. Daher ist das Engagement eines mediendidaktischen Experten explizit verankert, der sich der wichtigen didaktischen Einbettung in den existierenden Lehrbetrieb widmet.

Generell lässt sich Lehre in Distanzlaboren in asynchrone und synchrone Lehr-/Lernzeiten aufteilen. Virtuelle Labore können sowohl synchron als auch asynchron hinsichtlich der Lehr- und Lernzeiten sein. Wichtig ist dabei, dass *beide* Formen für alle Studierende zugänglich sind. Virtualisierung und Digitalisierung bieten zudem den Vorteil, Labore mehrsprachig zu gestalten. Angesichts steigender Studierendenzahlen, deren Muttersprache nicht Deutsch ist, tragen virtualisierte bzw. digitalisierte Labore zur

Qualitätssteigerung bei.

Um insbesondere in asynchronen und virtuellen Laboren soziale Aspekte zur Interaktion und Gruppenbildung einfließen zu lassen, sind Distanzlabore in die an der Hochschule vorhandenen Systeme zum Distanzlernen (Moodle, ILIAS, etc.) einzubinden. Hier bietet sich ein serviceorientierter Ansatz an. Unter einer einheitlichen Oberfläche (zur Minimierung der Einarbeitung, etc.) werden die benötigten Elemente als Dienste eingebunden und verfügbar gemacht. Es ist auf jeden Fall zu vermeiden, dass Studierende mit wechselnden Werkzeugen in verschiedenen Fenstern arbeiten müssen.

Ein zentraler Bestandteil in der studentischen Ausbildung ist das Prüfen der gelernten Inhalte. Dabei müssen Prüfungsleistungen nicht nur rechtssicher gestaltet, sondern auch technisch und pädagogisch abgesichert werden. Probleme hinsichtlich Identitätskontrolle und Täuschung sind dabei zu lösen. Die handschriftliche Prüfung von Lerninhalten des Distanzlabors ist zu vermeiden. Gleichzeitig sind PC-Labore für Prüfungen tendenziell weniger geeignet, da sie oft stark in der Platzkapazität sowie der Übersicht des ganzen Raumes begrenzt und nicht in individuellen Tischen angeordnet sind. Eine Lösung ist die Nutzung mobiler Endgeräte, die sich im Eigentum der Hochschule befinden und auch in großen Kohorten die Durchführung von Präsenzprüfungen ergänzend und außerhalb der Labore erlauben. Studierenden wird ermöglicht, realitäts- und praxisnah ihr akquiriertes Wissen unter Beweis zu stellen, welches besonders im Bereich der Programmierung essenziell ist. Gleichzeitig besitzen die Prüfer die Möglichkeit, die technischen Rahmenbedingungen vorzugeben und die Prüfung auch digital auf den mobilen Endgeräten durch geeignete vorinstallierte Software zu überwachen. Diese Prüfungsstruktur ist für viele curricular verankerte Module geeignet und kann leicht auf weitere Fächer hochschulweit übertragen werden.

6 Evaluation

Das hier vorgestellte Konzept für Distanzlabore verspricht nennenswerte Vorteile wie bspw. die Egalisierung des Zugangs zu knappen Labor-Ressourcen oder die Unabhängigkeit von Lernzeit und Lernort. Zusätzlich müssen aber auch mögliche Auswirkungen auf die Qualität der studentischen Arbeit oder den benötigten Lernaufwand betrachtet werden. Zur systematischen Untersuchung des Einflusses der Laborart (remote vs. ausschließlich Präsenz) auf diese und weitere Punkte werden während der Projektlaufzeit empirische Studien durchgeführt, die den Einfluss der Laborart untersuchen. Ziel ist die Validierung der aufgestellten Hypothese, dass das *DistLab*-Konzept sich positiv auf Lernqualität, -aufwand und -zufriedenheit auswirkt.

Erste Erfahrungen mit virtuellen Zwillingen wurden im Wintersemester 2020/21 an der Fakultät für Maschinenbau der Hochschule Stralsund im Rahmen des Pflichtlabors des Moduls *Regelungstechnik* gemacht. Die zu absolvierenden Versuche wurden von allen Teams in der Kombination virtuell/real erfolgreich durchgeführt. Bei einer Kohortengröße von etwa 80 Studierenden ergab die Einzelbefragung, dass die

Kombination von virtuellem und realem Versuch durchweg positiv aufgenommen wurde. Demnach erscheint eine Weiterführung und Vertiefung des Konzeptes „Virtueller Zwillings“ auch in Post-Pandemiezeiten definitiv zielführend zu sein. Eine erste Studie des Gesamtkonzeptes soll in Form eines Quasi-Experiments durchgeführt werden, welches zwei zeitlich versetzte Labordurchläufe (geplanter Start von *DistLab* im Wintersemester 2021/22) evaluiert. Die Laborveranstaltungen betreffen zwei verschiedene studentische Kohorten und liegen zeitlich ein Jahr auseinander. Die in Quasi-Experimenten möglichen Störfaktoren müssen daher explizit berücksichtigt und diskutiert werden. Zusätzlich wird ein randomisiertes, kontrolliertes Cross-Over-Design mit nicht in die Endnote eingehenden Zwischenprüfungen verwendet, um den Einfluss der Laborart zu untersuchen [LH12].

Neben objektiven Daten wie Zeit- und Qualitätsmetriken werden auch subjektive Daten erfasst. Hierzu beantworten Teilnehmer eine Reihe von Fragen, sowohl vor als auch nach der Veranstaltung. Die Vorbefragung dient dazu, Daten hinsichtlich Vorwissen und Motivation zu erfassen. Die Nachbefragung soll Daten zur Zufriedenheit der Teilnehmer und zu möglichen Problemen erfassen. Zusätzlich soll das Verhältnis von Selbsteinschätzung und Ergebnis in Abhängigkeit von der Durchführungsart untersucht werden. So zeigte eine Studie an der Universität Magdeburg [ZH17] zu Distanzlaboren Diskrepanzen zwischen Selbsteinschätzung und tatsächlichen Ergebnissen auf.

7 Zusammenfassung und Ausblick

Im Rahmen dieser Veröffentlichung wurde das Konzept zur ortsungebundenen, digitalen Durchführung von Laborveranstaltungen in Laboren mit kostenintensiver Geräteausstattung vorgestellt. Das Konzept wird im Rahmen eines geförderten dreijährigen Projektes, beginnend im August 2021, in die Praxis umgesetzt werden. *DistLab* unterstützt sowohl vollständig virtuelle Labore, als auch Labore mit Fernzugriff auf real existierende Geräte. Das Konzept hat den Anspruch, portierbar und skalierbar zu sein. Die klassische Präsenzausbildung wird qualitativ insgesamt auf eine höhere Stufe gehoben, da Studierende intensiv und selbstbestimmt mit Laborgeräten interagieren können. Beschränkte Laborzeiten und die limitierte Verfügbarkeit von Hardware stellen kein Hemmnis bei der Wissensaneignung mehr dar.

Das *DistLab*-Konzept erlaubt die Einbindung weiterer Labore der Hochschule. Dazu können die Software-Systeme der Prototypen beider Laborformen (virtuell, hybrid) adaptiert und eingesetzt werden. Kommunikation und Hosting werden von der zentralen Infrastruktur im Rechenzentrum übernommen. Dies erlaubt auch die Nutzung der zentralen Ressourcen als Dienstleistung für andere Hochschulen. Ein weiterer Aspekt ist die enge Verzahnung des Konzeptes mit der beruflichen Praxis. Hier ist daran gedacht, Partnerunternehmen aus dem Bereich des praxisintegrierenden Studiums (Studium mit vertiefter Praxis) als Partner zu gewinnen, die über die *DistLab*-Plattform „ihren“ Studierenden kontrollierten Zugriff auf im Unternehmen verfügbare Geräte ermöglichen.

Literaturverzeichnis

- [AC08] Agrawal, JP.;Cherner, YE: A Classroom/Distance Learning Engineering Course on Optical Networking with Virtual Lab. IEEE International Conference on Computational Technologies in Electrical and Electronics Engineering. Novosibirsk, Russia, 2008.
- [BD21] Birk, A. Dineva, E., Maurelli, F., Nabor, A.: A Robotics Course during COVID-19: Lessons Learned and Best Practices for Online Teaching beyond the Pandemic, MDPI robotics Journal, 10(5), 2021.
- [CC95] Cao, Y. U., Chen, T. W., Harris, M. D., Kahng, A. B., Lewis, M. A., & Stechert, A. D.: A remote robotics laboratory on the internet. Proc. INET-95, Honolulu., 1995
- [GF08] Gravier C, Fayolle J, Bayard B, Ates M, Lardon J.: State of the art about remote laboratories paradigms: Foundations of ongoing mutations. International Journal of Online Engineering (iJOE), 2008.
- [HA07] Hasnim H, Abdullah M Z.: Remote lab generator: A software tool for lab lecturer. Proceedings of the Int. Conf. Electrical Engineering and Informatics, Institut Teknologi Bandung, Indonesia, Jun. 17-19, 2007.
- [LH12] Lemos, M., Hannig, A., Renardy, C., Kennes, L., Wolfart, S., Ohnesorge-Radtke, U., & Rafai, N.: Aachener Modell zur Vermittlung praktischer Fertigkeiten in der zahnmedizinischen Lehre. Jahrestagung der Gesellschaft für Medizinische Ausbildung (GMA), 2012.
- [MG17] Meier, K., Grüning, B., Blank, C., Janczyk, M.,Suchodoletz, D.v.: Virtualisierte wissenschaftliche Forschungsumgebungen und die zukünftige Rolle der Rechenzentren, 10. DFN-Forum Kommunikationstechnologien, GI, 2017.
- [MH06] Mougharbel I, Hajj A E, Artail H, Rimam C. Remote lab experiments models: A comparative study. International Journal of Engineering Education, 2006.
- [Od1] Odeh, S.: Building Reusable Remote Labs with Adaptable Client User-Interfaces. Journal of Computer Science and Technology, 25(5), 2010.
- [PG19] Post, L.S.; Guo, P.; Saab, N.; Admiraal, W.: Effects of remote labs on cognitive behavioral, and affective learning outcomes in higher education. Journal on Computers & Education, Volume 140, 2019
- [RD06] Rigby S, Dark M.: Designing a flexible, multipurpose remote lab for the IT curriculum. Proceedings of the 7th Int. Conf. Information Technology Education, New York, USA, 2006.
- [RJ00] Röhrig, C., Jochheim, A.: Java-Based Framework for Remote Access to Laboratory Experiments, IFAC Proceedings Volumes, 33(31), 2000,
- [SP00] Schwarz, J., Polze, A., Wehner, K., Sha, L.: RemoteLab: A Reliable Tele-Laboratory Environment, International Conference on internet Computing, 2000
- [SS11b] Seiler, S.; Sell, R.: Comprehensive Blended Learning Concept for Teaching Micro Controller Technology, Proceedings of the 4th International eLBA Science Conference, Rostock (Germany), 2011.

- [SS10] Sell, R.; Seiler, S.: Combined Robotic Platform for Research and Education. SIMPAR. Intl. Conf. on SIMULATION, MODELING, and PROGRAMMING for AUTONOMOUS ROBOTS, Darmstadt (Germany) November 15-16, 2010.
- [SS11a] Sell, R.; Seiler, S.: Integrated Concept for embedded system study, Proceedings of the 7th International Conference Mechatronic Systems and Materials (MSM 2011), Kaunas, Lithuania, 2011.
- [TS15] Trahasch, S., Suchdoletz, D.v., Münchenberg, J., Rettberg, S., Rößler, C., Wilson, M.: bwLehrpool: Plattform für die effiziente Bereitstellung von Lehr-und Klausurumgebungen, 8. DFN-Forum Kommunikationstechnologien, GI, 2015
- [VP18] Viegas, M.C.; Pavani, A.M.B.; Lima, N.; Marques, A.; Pozzo, M.I.; Dobboletta, E.; Atencia, V.; Barreto, D.; Calliari, F.; Fidalgo, A.; Lima, D.A.; Temporão, G.P.; Alves, G.: Impact of a remote lab on teaching practices and student learning. Journal on Computers & Education, Vol. 26,2018
- [YL06] Yan Y, Liang Y, Du X, Hassane H S, Ghorbani A.: Putting labs online with Web services. IT Professional, 8(2), 2006..
- [ZH17] Zug, S; Hawlitschek, A.; Krenz, T.: What are the key features of future Remote Labs? A critical evaluation of an existing one. FDIBA Conference Proceedings, 2017.

Modellierung einer digitalen Hochschule post-Corona

Einflüsse und Nachwirkungen der pandemiebedingten Transformation

Harald Gilch¹, Imke Jungermann² und Klaus Wannemacher³

Abstract: Ausgehend von verschiedenen empirischen Erhebungen zum Stand der Digitalisierung an den Hochschulen wird unter Berücksichtigung der an den Hochschulen bereits vorliegenden vielfältigen Ansätze, Konzepte, Ideen und Erfahrungen eine prototypische Beschreibung einer weiter entwickelten, umfassend digital unterstützten Hochschule vorgenommen. Die im Kontext einer bundesweiten Vollerhebung 2019 erstmals definierten Faktoren, die eine digitale Modellhochschule prägen könnten, werden mit dem neuen Sachstand abgeglichen, der sich 2020/2021 in Zusammenhang mit dem weitreichenden Digitalisierungsschub an den Hochschulen ergeben hat. Durch die erheblichen Veränderungen, zu denen es in allen Bereichen der Hochschule angesichts der erzwungenen Umstellungen im Zuge der Corona-Pandemie kam, müssen zahlreiche der Faktoren, die ursprünglich für die Modellierung einer Hochschule der digitalen Zukunft herangezogen wurden, neu gewichtet und das Modell einer umfassend digitalisierten Hochschule fortgeschrieben werden.

Keywords: Digitale Hochschule, Modellhochschule, Strategie, Forschung, Lehren und Lernen, Studienangebot, Verwaltung, Vernetzung, IT-Sicherheit, Personalentwicklung

1 Einleitung

In den letzten Jahren sind verschiedene empirische Erhebungen zum Stand der Digitalisierung an den Hochschulen durchgeführt worden [LG17; GHS19; PWC21], darunter auch eine bundesweite Vollerhebung unter den Leitungen der öffentlich-rechtlichen Hochschulen in Deutschland für die Studie „Digitalisierung der Hochschulen“ [Gi19], die das HIS-Institut für Hochschulentwicklung (HIS-HE) im Auftrag der Expertenkommission Forschung und Innovation durchgeführt hat. In dieser Studie haben die Autor:innen bereits 2019 anhand der bis dahin an den Hochschulen bestehenden vielfältigen Ansätze, Konzepte, Ideen und Erfahrungen eine umfassend digital unterstützte Modellhochschule

¹ HIS-Institut für Hochschulentwicklung e. V., Geschäftsbereich Hochschulmanagement, Gosseriede 13a, 30159 Hannover, gilch@his-he.de.

² Ebd., jungermann@his-he.de.

³ Ebd., wannemacher@his-he.de,  <https://orcid.org/0000-0003-4810-3693>.

beschrieben. Dabei wurden in erster Linie die durch eine Hochschule beeinflussbaren Elemente betrachtet, wobei vorausgesetzt wurde, dass die von den Hochschulen immer wieder thematisierten Herausforderungen im Bereich der Ressourcen sowie der rechtlichen Rahmenbedingungen intern und extern gelöst werden konnten. Die seinerzeit definierten hochschulbezogenen Einflussfaktoren auf Lehre, Forschung und Verwaltung wurden mit dem neuen Sachstand abgeglichen, der sich in Zusammenhang mit dem weitreichenden Digitalisierungsschub an den Hochschulen ergeben hat [ALK20; Bo20; De20; SEW20]. Durch die erheblichen Veränderungen, zu denen es in allen Bereichen der Hochschule angesichts der erforderlichen digitalen Umstellungen im Zuge der Bewältigung der Corona-Pandemie kam, sind jetzt zahlreiche Faktoren, die ursprünglich für die Modellierung einer Hochschule der digitalen Zukunft herangezogen wurden, neu zu gewichten. Insofern wird an dieser Stelle eine aktualisierte Version dieser idealtypischen Modellhochschule vorgestellt.

2 Modellhochschule der digitalen Zukunft

Ein Ergebnis der Studie „Digitalisierung der Hochschulen“ von 2019 war, dass es (zumindest in Deutschland) bisher keine Hochschule gibt, die in allen Bereichen durchgehend als Vorbild für eine konsistent digitale Modellhochschule dienen könnte. Dies hat sich auch seitdem – trotz der pandemiebedingten Notwendigkeit, Lehrveranstaltungen unmittelbar in virtuelle Lehr- und Lernumgebungen zu überführen und auf digitalisierte Lehre umzustellen – sowie weiterer vielfältiger Fortentwicklungen in anderen Bereichen der Hochschulen nicht prinzipiell geändert. Es lassen sich aber an unterschiedlichen Hochschulen verschiedene Best-practice-Beispiele identifizieren, die wie Puzzleteile zu solch einer prototypischen Modellhochschule zusammengeführt werden können. Abweichend von der Intention der ursprünglichen Studie, die digitale Modellhochschule ausschließlich auf Grundlage der Befunde von Gesprächen, die 2018 mittels leitfadengestützter Interviews zwecks Erhebung von Fallbeispielen an zehn Hochschulen mit unterschiedlichen Akteursgruppen geführt wurden, zu beschreiben, werden im vorliegenden Beitrag exemplarisch ergänzende Hinweise aus einer breiten Auswahl der zwischenzeitlich erschienenen Literatur zur Digitalisierung der Hochschulen aufgenommen. Auch im Rahmen dieser erweiterten Beschreibung der digitalen Modellhochschule war eine Beschränkung auf eine Auswahl von Aspekten, die den aktuellen Diskurs um die Digitalisierung der Hochschulen prägen, unumgänglich.

2.1 Hochschulinterne Voraussetzungen

In Bezug auf Strategie und Governance würde die Modellhochschule

- Digitalisierung⁴ als Mittel zur Erreichung ihrer strategischen Ziele in Forschung, Lehre und Verwaltung verstehen und hätte daher entsprechende Ziele, Programme und Maßnahmenpakete formuliert, den Stand der Digitalisierung in allen Bereichen weiter voranzutreiben. Die erfolgreich durchgeführte digitale Transformation samt entsprechender Programme und Maßnahmenpakete werden zur Profilbildung der Hochschule insgesamt genutzt und aktiv in der Außenkommunikation eingesetzt [HFD16a; DINI18];
- ihre Ausrichtung und Ziele verbindlich in einem Leitbild verankern, das Mission und Vision in Bezug auf die Digitalisierung festlegt und die Bedürfnisse der gesamten Hochschule (Forschung, Lehre und Verwaltung) berücksichtigt. Dieses Leitbild ist dabei das Ergebnis eines kooperativen Dialogs mit den unterschiedlichen Statusgruppen der Hochschule und wird stetig weiterentwickelt;
- diese strategischen Ziele in Bezug auf die Digitalisierung beständig weiterentwickeln und deren Umsetzung steuern und überwachen. Dabei erfolgen eine regelmäßige Prüfung und ein stetiger Austausch über bestehende und sich verändernde Bedarfe mit dem akademischen, administrativen und technischen Bereich. Technische Neuerungen, innovative Gestaltungsmöglichkeiten, aktuelle Kompetenzen und allgemeine Rahmenbedingungen können so laufend bedarfsorientiert genutzt bzw. angepasst werden.

Weiterhin hätte die Modellhochschule

- einen Chief Information Officer (CIO) oder auch ein CIO-Gremium eingesetzt, in dessen Verantwortungsbereich die Digitalisierung liegt und der/die auch Mitglied im Präsidium/Rektorat ist. Dieser CIO arbeitet eng mit dem zentralen Fachpersonal für IT-Fragen (Rechenzentrumsleitung, Leitungen von Servicestellen für digitale Forschung bzw. digitale Lehre, Hochschulbibliothek) zusammen.
- Über einen IT-Beraterkreis sind ferner die Nutzer:innen (Vertreter:innen aus Fakultäten, Verwaltung, Studierenden) in die Digitalisierungsprozesse einbezogen. In der Praxis kommt es ausgeprägter noch auf die Qualität der Beziehungen im Team der unterschiedlichen IT-Leitungsverantwortlichen an als auf die konkrete Umsetzung der CIO-Funktion [He18; LW14; Po17; PWC21].
- Darüber hinaus hätte die Modellhochschule zur effizienten Steuerung der Digitalisierungsaktivitäten dynamische und nachhaltige Kommunikations- und Entschei-

⁴ Im Sinne einer Definition des Rates für Informationsinfrastrukturen wird unter Digitalisierung „ganz allgemein die Umstellung der gesamten Gesellschaft auf die Verwendung von Digitaltechnologien (digitale Revolution, Digital Turn)“ verstanden [RfI16].

dungsstrukturen, in welche die oben genannten Nutzer:innen sowie relevante Support-/Serviceeinrichtungen partizipativ einbezogen sind, geschaffen [BM18; PH17].

2.2 Forschung

Für den Bereich Forschung tritt die Modellhochschule

- für Open Science, die Verbreitung und Verwendung offener Lizenzen, die Entwicklung von Regelungen für den offenen Zugang zu lizenzierten Materialien, die Anwendung offener Standards und Protokolle, die Entwicklung und den Einsatz von Open-Source-Software und die weitere Transformation des wissenschaftlichen Publikationswesens in den Open Access ein [DINI18].
- Es würden in den Fakultäten und Fachbereichen der Modellhochschule föderierte IT-Infrastrukturen für die Forschung wie die European Data Infrastructure, die European Open Science Cloud oder die Nationale Forschungsdateninfrastruktur (NFDI) rege genutzt [Ko20]. Die Hochschule richtet ihre forschungsbezogene IT-Infrastruktur auf die zu erwartenden hohen technischen Anforderungen aus, die die technischen Entwicklungen kommender Jahre u. a. im Bereich anspruchsvoller KI-Systeme mit sich bringen dürften.
- Im Hinblick auf digitale Dienste für die Wissenschaft dominiert bislang an vielen Hochschulen angesichts leichter Zugänglichkeit, Performanz und zunächst geringen Kosten die pragmatische Nutzung globaler kommerzieller Lösungen. Die Modellhochschule hingegen nutzt ihren Einfluss auf die Auswahl und Entwicklung wissenschaftlicher Informationsdienste, um ihre wissenschaftlichen Communitys mit zeitgemäßen und nachhaltigen Diensten zu versorgen (d. h. Dienste der wissenschaftlichen Infrastrukturanbieter und der Communitys selbst, darunter kollaborative Plattformen, die nicht nur Daten, sondern auch Softwarewerkzeuge, Publikationen und Expertennetzwerke zusammenführen) [Ru20; DFG21].
- Zudem wäre an der Modellhochschule selbst eine Serviceeinrichtung eingerichtet, an der Kompetenzen zu E-Research und zum Forschungsdatenmanagement (FDM) konzentriert sind. Diese Stelle schafft ein hochschulinternes Bewusstsein für die Möglichkeiten des E-Research und die Notwendigkeit des FDM und berät sowohl präventiv bei der Konzeption von Forschungsprojekten als auch kurativ bei konkreten Problemen in E-Research und FDM.

Außerdem wäre Digitalisierung auch

- Gegenstand der Forschung selbst, wobei diese zunächst die Fachdisziplinen bzw. -gebiete umfasst, durch die sich in Zusammenhang mit einer stetig wachsenden Menge an großen Datenbeständen neue Auswertungsmöglichkeiten jenseits diszip-

linärer Grenzen eröffnen, darunter beispielsweise Data Science, Künstliche Intelligenz einschließlich Robotik, Machine Learning und Deep Learning [WB21]. In diesem Zusammenhang auftretende rechtliche und soziale Fragestellungen wären dabei mit zu reflektieren.

- Digitalisierung ist aber nicht nur Gegenstand der Forschung in den Fachdisziplinen. Die Geisteswissenschaften werden an der Modellhochschule von den Digital Humanities geprägt, die Naturwissenschaften treiben Grundlagenforschung wie Biocomputing und Quantum Computing voran und die Ingenieurwissenschaften nutzen die Möglichkeiten der Digitalisierung zur Entwicklung ganz neuer Werkzeuge und technologischer Anwendungen für die Industrie 4.0 bzw 5.0.

2.3 Lehre und Lernen

Im Bereich Lehren und Lernen

- reagiert die Modellhochschule auf einen qualitativen Wandel der Nachfrage nach Studienangeboten durch neue Formen und Angebote im Bereich der (digitalisierten und hybriden) Hochschulbildung, darunter reine Online-, hybride und Blended-Learning-Angebote, vermehrt modulare Open Education-, Selbststudien- und Lifelong Learning-Angebote, kleinteilige Leistungsnachweise („Unbundling of Credentials“) sowie allgemein stärker nachgefragte Angebote mit hohem Mehrwert für die berufliche Laufbahn [Or20; Wa20]. Diese Transformation war schon vor der Corona-Pandemie begonnen worden, wird jetzt aber im Zuge der erheblichen Veränderungen angesichts der erzwungenen Umstellungen auf die komplette Online-Lehre im Zuge der Corona-Pandemie mit verstärkter Priorität vorangetrieben.
- gelten allgemeingültige Leitlinien bzw. Qualitätsstandards zur Erstellung und Anwendung digitaler Lehr- und Lernangebote. Dies garantiert eine optimale didaktische Einbettung digitaler Formate in die Lehre sowie ihre Verankerung in die Curricula. Digitale Lehre trägt so zu einer kontinuierlichen Verbesserung der Qualität der Lehre in der Breite der Studiengänge der Hochschule bei und ermöglicht qualitätsgesicherte Anerkennungs-, Anrechnungs- und Zertifizierungsverfahren [HFD16b].
- Die Modellhochschule bietet zudem ein ausgeprägtes Anreizsystem, um die Lehrenden nachhaltig an der fortschreitenden Weiterentwicklung digitaler Lehrinnovationen zu beteiligen. Ein solches Anreizsystem basiert auf einer breiten und partizipativen Kommunikation, um Potenziale digitaler Medien und Lehr- und Lernszenarien dauerhaft präsent zu halten. Ferner ist eine Anpassung der Lehrverpflichtung erfolgt, um die Beteiligung an der Weiterentwicklung von digitalen Lehrinnovationen zu honorieren. Niedrigschwellige Formate zum Austausch und zur Weiterbildung sind etabliert [Ste19].

- Auf binnenstruktureller Ebene treibt an der Modellhochschule eine zentrale Service-Einrichtung für digitales Lernen die Entwicklung und Anwendung von verschiedensten digitalen Lehrformaten voran, bietet Weiterbildungsformate zur Gestaltung digitaler Lehre an und berät und unterstützt die Lehrenden didaktisch und technisch, u. a. bei der Erarbeitung mediendidaktischer Lehr- und Lernszenarien sowie der Entwicklung digitaler Contents [Sc17]. Die spezifische Anpassung von Elementen des E-Learnings auf konkrete Anforderungen der verschiedenen Veranstaltungsformate und Fachdisziplinen ist ein zentraler Beratungsbaustein dieses Kompetenzzentrums, das in engem Austausch zur Hochschuldidaktik digitale und nicht-digitale Lehr- und Lernformate neu beschreibt und miteinander vernetzt. Auch diese Entwicklung wird durch die Corona-Pandemie, die bei den meisten Lehrenden breite Nachfrage nach Beratungsangeboten für digitalisiertes Lehren und Lernen bewirkt hat, verstärkt und mit noch höherer Priorität vorangetrieben.
- Die zentrale Service-Einrichtung für digitales Lernen bietet zugleich für Studienanfängerinnen und -anfänger Brückenkurse und Weiterbildungsangebote zur Ergänzung der digitalen Skills der Studierenden an, um unterschiedliche Kompetenzniveaus auszugleichen und möglichst allen Studierenden ein gewisses technisches Grundverständnis samt disziplinunabhängiger arbeitsmarktrelevanter Fähigkeiten zu vermitteln.
- Anknüpfend an Erfahrungen aus der Corona-Pandemie wurde an der Modellhochschule trotz aller zwischenzeitlich etablierten Möglichkeiten, sich online rasch und komfortabel zu vernetzen, zugleich die zentrale Bedeutung des Campuslebens und der persönlichen Kommunikation wiederentdeckt. Zu den Lerneffekten der pandemiebedingten Digitalsemester zählt, dass digitale Lösungen in Studium und Lehre vor allem dort eingesetzt werden, wo dieser Einsatz die jeweils sinnvollste Maßnahme darstellt bzw. die Präsenzlehre optimal unterstützt.
- Vor diesem Hintergrund spielt auch die Ausgestaltung der Lehr- und Lernräume eine wichtige Rolle in der zukünftigen Hochschule. Die Modellhochschule hat sich auf die räumlichen und technischen Bedarfe digitaler und kollaborativer Lehr- und Lernformate sowie neue Lern- und Arbeitszusammenhänge eingestellt und entsprechende Räumlichkeiten, die diese diskursiven Lehr- und Lernprozesse ermöglichen, geschaffen. So bietet die Modellhochschule Lehrräume verschiedener Größe und hochwertiger technischer Ausstattung, die die unterschiedlichsten Formate wie interaktive Seminare, Flipped Classroom und Videokonferenzen unterstützen, sowie zahlreiche Selbstlern- und Kommunikationsflächen [We20].

Auch in den Curricula

- grundsätzlich aller Studiengänge ist die Vermittlung von digitaler Kompetenz im jeweils fachspezifischen Kontext verankert. Dies betrifft sowohl die Nutzung digitaler Technologien und Methoden als auch die Inhalte der Fächer selbst, die sich

zumindest im Rahmen von Schwerpunktstudien auch mit der Entwicklung der Fächer innerhalb der digitalen Wissenschaft und Gesellschaft beschäftigt [HRK12; St20].

- Die Modellhochschule reagiert in ihrem curricularen Angebot dabei zugleich flexibler auf den chronischen Fachkräftemangel, der sich volkswirtschaftlich infolge eines fortgeschrittenen demografischen Wandels, der KI-getriebenen Automatisierung von Arbeitsprozessen und der sich ändernden Anforderungen der Industrie abzeichnet und eine Umschulung einer wachsenden Anzahl von Arbeitnehmer:innen, die aus ihren angestammten Berufsfeldern verdrängt wurden, erforderlich macht („Re-skilling“) [Ho18; WLS19].

Mit neuen Studienangeboten

- in Bereichen wie Künstliche Intelligenz, Data Science, Digitale Transformation, Business Analytics, Cyber Security, Informationsmanagement, E-Government etc. profiliert sich die Modellhochschule außerdem, um für Wirtschaft und Gesellschaft entsprechend zukunftsorientierte Lehr- und Ausbildungsangebote zu schaffen [Ma20; WRP20].

Die Möglichkeiten einer digital vermittelten ortsunabhängigen Flexibilisierung der Lehre werden auch für ein erweitertes postgraduales (Online-)Studienangebot im Bereich der wissenschaftlichen Weiterbildung für die Zielgruppe der Berufstätigen genutzt [Or20].

2.4 Verwaltung

Die Verwaltung der Modellhochschule hat ihre Geschäftsprozesse und Anwendungsfälle

- an den Anforderungen des 2017 erlassenen Gesetzes zur Verbesserung des Online-Zugangs zu Verwaltungsdienstleistungen (OZG Onlinezugangsgesetz) ausgerichtet und ihre digitalen Verwaltungsdienste umfassend erweitert und verbessert, so dass diese die vom OZG geforderten rechtlichen und technischen Anforderungen erfüllen. Die interoperable Verknüpfung der Dienstleistungsebene der Hochschule mit einem landesweiten Portalnetz samt einem einheitlichen Benutzerkonto ist OZG-konform umgesetzt. Im Bereich der rechtsgültigen und belastbaren Signaturen und Authentifizierungs- bzw. Echtheitsprüfungen sowie hinsichtlich der Usability werden die technisch bestehenden Möglichkeiten voll ausgeschöpft [Ru20].
- durchgehend und medienbruchfrei digitalisiert. Elektronische Studierendenakten, Personalakten und elektronische Rechnungsverarbeitung sind die Regel. Die entsprechenden IT-Systeme sind ausgereift, um auch die teilweise sehr speziellen Anforderungen der Modellhochschule digital zu unterstützen. Die gesamtgesellschaftliche Entwicklung in der öffentlichen Verwaltung hin zum durchgehenden eGovernment, und damit die Möglichkeit, auf die Schriftform zu verzichten, hat die Modellhochschule umgesetzt [Ru20].

- Bei dem genutzten Campus-Management-System (CaMS) konnte unter Anbindung z. B. an EMREX-Infrastrukturen erreicht werden, dass die Modellhochschule auf den Versand von beglaubigten Papierkopien oder gar Originalen von Zeugnissen und Abschlussurkunden verzichten kann. Zudem wird das CaMS für einen hochschulübergreifenden Datenaustausch mit den CaMS anderer Hochschulen genutzt, um Studierenden den Wechsel an einen anderen Studienort nach Abschluss des Bachelors zur Aufnahme eines Masterstudiums zu erleichtern [Au17].

2.5 Vernetzung und außengerichtete Maßnahmen

Die Modellhochschule hat sich weiterhin mit anderen Hochschulen auf nationaler und internationaler Ebene vernetzt

- und ermöglicht es ihren Studierenden, digitale Online-Angebote der Partner in ihr Studium zu integrieren und auf diese Weise absolvierte Module anerkennen zu lassen [BGW21].

Als Basis für die genannten digitalen Formate und Anwendungen in Forschung und Lehre verfügt die Modellhochschule über moderne und leistungsfähige IT-Infrastrukturen,

- die bezogen auf Basisdienste wie Netze, Server, Storage etc. zentral organisiert sind und von übergreifenden Servicediensten im 24/7-Support begleitet werden. Im Zuge der Corona-Pandemie wurde zudem ein breites Angebot an digitalen, überwiegend barrierefreien Kommunikations- und Kollaborationstools etabliert (wie z. B. Videokonferenzsysteme) und in die bestehende IT-Landschaft integriert, das von Lehrenden und Studierenden gleichermaßen genutzt wird.
- Mit analytischem Blick auf die Geschehnisse in den Corona-Semestern sind Rahmenbedingungen in der Struktur und den Abläufen der Hochschule nachjustiert worden, z. B. „Nutzungsformen und -umfang von technischer Infrastruktur und Beratungsangeboten, damit verbundene Ressourcenzuordnungen sowie Organisations- und Entscheidungsstrukturen“ [HKL21, S. 240]. Dieser systemische Ansatz hat zu einer strukturellen Veränderung des Systems Hochschule (z. B. neue zentrale Ansprechpartner, Service-Center etc.) und Verbesserungen in der nachhaltigen Verankerung des gezielten Medieneinsatzes in der Hochschullehre beigetragen.

Um moderne und leistungsfähige IT-Infrastrukturen zu gewährleisten, hat sich die Modellhochschule mit anderen Hochschulen und wissenschaftlichen Einrichtungen am Standort bzw. in der Region zusammengeschlossen, und betreibt ein gemeinsames Rechenzentrum,

- mit dem es auch möglich wird, den Wissenschaftler:innen und Forscher:innen der beteiligten Einrichtungen Hoch- und Höchstleistungs-Rechner-Kapazitäten zur Verfügung zu stellen, die diese insbesondere für rechenintensive Simulationen und

Modellrechnungen aus Natur-, Ingenieur-, Geistes- und Gesellschaftswissenschaften benötigen.

Um diese Anwendungen und Anwendungsfälle umsetzen zu können, wurden der Datenschutz und die IT-Sicherheit an der Modellhochschule

- konsequent ausgebaut und ein System etabliert, das sich als Service- und Beratungsstelle für die Anwender:innen und Nutzer:innen versteht und sich dabei gleichzeitig an höchsten Standards orientiert, diese aber versucht, so pragmatisch und nutzenorientiert wie möglich umzusetzen, z. B. durch etablierte Security Awareness-Kampagnen.

Darüber hinaus begegnet man den Herausforderungen der stetig voranschreitenden Digitalisierung durch vielfältige hochschul- und institutionenübergreifende Kooperationen.

- So ist die Modellhochschule mit den anderen Hochschulen des Landes vernetzt und profitiert von einer gemeinsamen Bearbeitung von Themenbereichen, gemeinsamen Initiativen und hochschulübergreifenden Strukturen.⁵ Die Landeseinrichtung ist wiederum in einen übergreifenden Austausch mit weiteren Landesinitiativen oder internationalen Organisationen eingebunden.
- Parallel dazu werden fachspezifische Netzwerke, Vernetzungsstrukturen für besondere Studierendengruppen (z. B. Austauschstudierende, duales Studium) oder für spezielle Verfahren (z. B. Eignungsfeststellungsverfahren) gepflegt.

Die Realisierung der genannten Digitalisierungsprozesse ist der Modellhochschule gelungen, indem sie

- einerseits immer wieder Prioritäten in ihren Digitalisierungsbestrebungen gesetzt, diese dann konsequent bearbeitet und sowohl durch Nutzung staatlicher Förderprogramme als auch durch Umschichtung eigener Mittel mit entsprechenden Ressourcen ausgestattet hat,⁶
- andererseits ein System aus Beratungs- und Supportdiensten sowie Anreizen geschaffen hat, mit dem es gelungen ist, die vielfältigen dezentralen Ideen, Aktivitäten und Initiativen einzufangen, zu unterstützen und miteinander zu vernetzen, um damit nicht nur eine hohe Durchdringung der Modellhochschule mit digitalen Anwendungen und Formaten zu erreichen, sondern auch eine Atmosphäre von Innovation zu erzeugen, die neue Ideen entstehen lässt.

⁵ Als Beispiel sei hier die *Digitale Hochschule NRW* genannt, eine Kooperationsgemeinschaft von 42 Hochschulen aus Nordrhein-Westfalen. Zusammen mit dem Ministerium für Kultur und Wissenschaft arbeiten sie daran, die Digitalisierung der nordrhein-westfälischen Hochschulen zu koordinieren und zu fördern, vgl. <https://www.dh.nrw/akteure/ueber-uns> [30.06.2021].

⁶ Aspekte der Ressourcenplanung, der Finanzierung wie auch des Controllings sind für die Umsetzung einer umfassend digital unterstützten Hochschule von großer Bedeutung, können aufgrund der erheblichen Breite der adressierten Digitalisierungsaspekte in diesem Zusammenhang jedoch nicht ausführlich erörtert werden.

Eine wesentliche Rolle hat hierfür die Erarbeitung und Umsetzung eines umfassenden Personalentwicklungskonzepts gespielt,

- mit dem es gelungen ist, die Beschäftigten auf allen Ebenen nicht nur mit den jeweils benötigten digitalen Kompetenzen umfassend zu qualifizieren, sondern auch
- das Stellentableau insbesondere in den zentralen und dezentralen Verwaltungsbereichen so umzugestalten, dass durch Digitalisierung verzichtbar gewordene Aufgaben und Stellen entfallen und dafür weniger, aber höherwertige Stellen im Digitalisierungs-Kontext geschaffen wurden [RfIII19; Gi20].
- Die Beschäftigten wurden, soweit möglich, über entsprechende Qualifizierungsmaßnahmen in die Lage versetzt, auch neue Aufgaben zu übernehmen. Gleichzeitig hatten die Tarifparteien durch Anpassung der Bestimmungen des TV-L dafür gesorgt, dass Höhergruppierungen bei entsprechenden neuen (technischen) Stellenanforderungen sowie bei Weiterqualifizierungen möglich sind. Insbesondere den IT-Fachkräften sollten Zulagen gezahlt werden können, um damit die Hochschule in die Lage zu versetzen, trotz hohem Wettbewerb mit der privaten Wirtschaft ausreichend IT-Fachkräfte zur Bewältigung der Digitalisierung zu gewinnen.

3 Fazit

Die Modellhochschule ist somit in Bezug auf ihren Stand der Digitalisierung und deren Rahmenbedingungen weit fortgeschritten und hat die im Gefolge der Corona-Pandemie erforderlichen Anpassungen nicht nur genutzt, um die digitale Infrastruktur der Zukunft weiter auszubauen, sondern hat die zwischenzeitlich eingeübten neuen Möglichkeiten in Forschung, Studium und Lehre und der Verwaltung zugleich zu Anpassungen in der strategischen Ausrichtung der eigenen Hochschule genutzt.

Dabei ist sie sich bewusst, dass die Digitalisierung zwar dabei hilft, aktuelle und künftige Anforderungen zu bewältigen und in Wissenschaft, Forschung und Lehre voranzuschreiten, dabei jedoch keinen Selbstzweck darstellt. Insofern gilt es, immer wieder einzuhalten und gemeinsam mit den Hochschulmitgliedern ein Commitment herzustellen, wie die weitere Entwicklung ausgestaltet werden soll.

Abschließend sei erwähnt, dass es sich bei den hier aufgeführten Elementen überwiegend um real an Hochschulen existierende Beispiele handelt, die aber in dieser Gesamtheit bisher noch an keiner Hochschule aufzufinden sind. Ebenso gibt es keinen allgemeingültigen Weg, wie eine weitgehende Digitalisierung an Hochschulen erreicht werden kann. Insofern ist diese Modellhochschule beispielhaft – jede Hochschule muss für sich das jeweils am besten auf ihre spezifischen Zielsetzungen, Besonderheiten und Bedürfnisse angepasste Vorgehen definieren und umsetzen.

4 Literatur

- [ALK20] Arndt, C.; Ladwig, T.; Knutzen, S.: Zwischen Neugier und Verunsicherung. Interne Hochschulbefragungen von Studierenden und Lehrenden im virtuellen Sommersemester 2020. Technische Universität Hamburg-Harburg, Hamburg, 2020.
- [Au17] Auth, G.: Die Rolle von Campus-Management-Systemen für die Digitalisierung der Hochschulen. Campus-Management-Systeme. Prozessorientierte Anwendungssoftware für die Organisation von Studium und Lehre. Die Hochschule: Journal für Wissenschaft und Bildung 26 (2017) 1, S. 40-58.
- [BGW21] Book, A.; Gilch, H.; Wannemacher, K.: Kooperationen zur Digitalisierung in Lehre, Forschung und Verwaltung an den Hochschulen. Sekundärauswertung einer bundesweiten Erhebung. In: Hochschulforum Digitalisierung (Hrsg.): Digitalisierung in Studium und Lehre gemeinsam gestalten. Innovative Formate, Strategien und Netzwerke. Springer VS, Wiesbaden, 2021 (im Erscheinen).
- [BM18] Beyer, A.; Mehler-Bicher, A.: Diese Veränderung ist dauerhaft. Was bedeutet die Digitalisierung für die Hochschulen? In: DUZ Wissenschaft & Management 05/2018. <https://www.duz.de/beitrag!/id/481/diese-veraenderung-ist-dauerhaft>.
- [Bo20] Bosse, E.; Lübcke, M.; Book, A.; Würmseer, G.: Corona@Hochschule. Befragung von Hochschulleitungen zur (digitalen) Lehre. HIS-HE, Hannover, 2020 (HIS-HE:Medium7|2020).
- [De20] Deimann, M.; Friedrich, J.-D.; Neubert, P.; Stelter, A.: Das digitale Sommersemester 2020: Was sagt die Forschung? Hochschulforum Digitalisierung, Berlin, 2020.
- [DFG21] Deutsche Forschungsgemeinschaft: Erkenntnisgeleitete Forschung stärken, von Wissensspeichern profitieren. Impulse der Deutschen Forschungsgemeinschaft für die 20. Legislaturperiode des Deutschen Bundestags. Deutsche Forschungsgemeinschaft, Bonn, 2021. https://www.dfg.de/download/pdf/presse/20210505_pm_impulspapier_legislaturperiode.pdf
- [DINI18] Deutsche Initiative für Netzwerkinformation (2018): Thesen zur Informations- und Kommunikationsinfrastruktur der Zukunft. https://edoc.hu-berlin.de/bitstream/handle/18452/19876/DINI-Thesen_2018_2.pdf?sequence=1&isAllowed=y.
- [GHS19] Gesing, R.; Hoffmann, L.; Schneider, B.: Stand der Digitalisierung an deutschen Universitäten. Eine Auswertung der Studie Digitalisierung. Digitalisierung – Chancen und Herausforderungen für die Universitäten Deutschlands. DUZ Special. Beilage zur DUZ – Magazin für Wissenschaft & Gesellschaft. DUZ, Berlin, 2019, S. 10–13.
- [Gi19] Gilch, H.; Beise, A. S.; Krempkow, R.; Müller, M.; Stratmann, F.; Wannemacher, K.: Digitalisierung der Hochschulen. Ergebnisse einer Schwerpunktstudie für die Expertenkommission Forschung und Innovation. Expertenkommission Forschung und Innovation (EFI), Berlin, 2019 (Studien zum deutschen Innovationssystem, Nr. 14-2019).
- [Gi20] Gilch, H.; Beise, A. S.; Krempkow, R.; Müller, M.; Stratmann, F.; Wannemacher, K.: Digitale Kompetenzen in der Hochschulstrategie – Quo vadis? Ergebnisse einer bundesweiten Schwerpunktstudie zur Digitalisierung an Hochschulen. In (Fürst, R. A., Hrsg.):

Digitale Bildung und Künstliche Intelligenz in Deutschland. Springer, Wiesbaden, 2020 (AKAD University Edition), S. 443-456.

- [He18] von der Heyde, M.: Mehr oder weniger CIO – Überblick zu CIO-Formen an deutschen Hochschulen. Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn, 2018.
- [HFD16a] Hochschulforum Digitalisierung: Strategieoptionen für Hochschulen im Digitalen Zeitalter. Hochschulforum Digitalisierung, Berlin, 2016 (Arbeitspapier Nr. 29).
- [HFD16b] Hochschulforum Digitalisierung: The Digital Turn. Hochschulbildung im digitalen Zeitalter. Hochschulforum Digitalisierung, Berlin, 2016 (Arbeitspapier Nr. 27).
- [HKL21] Hafer J.; Kostädt P.; Lucke U.: Das Corona-Virus als Treiber der Digitalisierung? In (Dittler U.; Kreidl C., Hrsg.): Wie Corona die Hochschullehre verändert. Springer Gabler, Wiesbaden, 2021. https://doi.org/10.1007/978-3-658-32609-8_15
- [Ho18] Holon IQ (Hrsg.): Education in 2030. Five scenarios for the future of learning and talent. Holon IQ, New York, San Francisco, Beijing, Sydney, 2018.
- [HRK12] Hochschulrektorenkonferenz: Hochschule im Digitalen Zeitalter. Informationskompetenz neu begreifen – Prozesse anders steuern. Entschließung der Mitgliederversammlung, 20. Nov. 2012.
- [Ko20] Konrad, U.; Förstner, K.; Reetz, J.; Wannemacher, K.; Kett, J.; Mannseicher, F.: Positionspapier Digitale Dienste für die Wissenschaft. Hrsg. von der Arbeitsgruppe Forschungssoftware im Rahmen der Schwerpunktinitiative Digitale Information der Allianz der deutschen Wissenschaftsorganisationen. Helmholtz-Zentrum Potsdam Deutsches GeoForschungsZentrum – GFZ, Potsdam, 2020. <http://doi.org/10.5281/zenodo.4301924>
- [LG17] Licka, P.; Gautschi, P.: Befragung. Die digitale Zukunft der Hochschule – Wie sieht sie aus und wie lässt sie sich gestalten? Berinfor, Zürich, 2017. <https://www.berinfor.ch/assets/docs/befragung/2017-Bericht-Befragung-Berinfor-Die-digitale-Zukunft-der-Hochschule.pdf>.
- [LW14] Lang, U.; Wimmer, M. (Hrsg.): CIOs und IT-Governance an deutschen Hochschulen. Hochschule Heilbronn, ZKI-Geschäftsstelle (Zentren für Kommunikation und Informationsverarbeitung in Lehre und Forschung), Heilbronn, 2014.
- [Ma20] Mah, D.-K.; Rampelt, F.; Dufentester, C.; Bernd, M.; Gamst, C.; Weygandt, B.: Digitale Lernangebote zum Thema Künstliche Intelligenz. Überblicksstudie zu kostenlosen Online-Kursen auf deutschen Lernplattformen. KI-Campus, Berlin, 2020. DOI: 10.5281/zenodo.4293318
- [Or20] Orr, D.; Lübcke, M.; Schmidt, J. P.; Ebner, M.; Wannemacher, K.; Ebner, M.; Dohmen, D.: Higher Education Landscape 2030. A Trend Analysis Based on the AHEAD International Horizon Scanning. Springer Nature Switzerland, Cham, 2020 (SpringerBriefs in Education).
- [PH17] Pasternack, P.; Hechler, D.: Digitalisierungsstrategien und Digitalisierungspolicies an Hochschulen. In: Die Hochschule: Journal für Wissenschaft und Bildung, 26 (2017) 2, S. 84–105. https://www.hof.uni-halle.de/journal/texte/17_2/Hechler_Pasternack.pdf.

- [Po17] Pomerantz, J.: IT Leadership in Higher Education, 2016: The Chief Information Officer. Research report. Louisville, CO: ECAR, March 2017, 2017.
- [PWC21] PricewaterhouseCoopers (Hrsg.): Die Digitalisierung an den Universitäten steuern. Die Sicht der Rektorate/Präsiden der 31 größten deutschen Universitäten zur Governance der Digitalisierung. PricewaterhouseCoopers, Frankfurt a. M., 2021.
- [RfII16] RfII – Rat für Informationsinfrastrukturen: Die Entwicklung von Konzepten für Informationsinfrastrukturen in der Bundesrepublik Deutschland seit den 1960er Jahren. Kurzfassung eines Berichts des Redaktionsausschusses Konzepte an den RfII. Rat für Informationsinfrastrukturen, Göttingen, 2016.
- [RfII19] RfII – Rat für Informationsinfrastrukturen: Digitale Kompetenzen – dringend gesucht! Empfehlungen zu Berufs- und Ausbildungsperspektiven für den Arbeitsmarkt Wissenschaft, Göttingen. 2019.
- [Ru20] Ruschmeier, R.; Gilch, H.; Lessig, M.; Stratmann, F.; Wannemacher, K.: Herausforderungen bei der Umsetzung des Onlinezugangsgesetzes im Kontext der Digitalen Hochschulbildung. Hochschulforum Digitalisierung, Berlin, 2020 (Arbeitspapier Nr. 55).
- [Sc17] Schmid, U.; Goertz, L.; Radomski, S.; Thom, S.; Behrens, J.: Monitor Digitale Bildung. Die Hochschulen im digitalen Zeitalter. Bertelsmann Stiftung, Gütersloh, 2017.
- [SEW20] Seyfeli, F.; Elsner, L.; Wannemacher, K.: Vom Corona-Shutdown zur Blended University? ExpertInnenbefragung Digitales Sommersemester. Tectum, Baden-Baden, 2020.
- [St20] Stifterverband: Lage und Entwicklung der Hochschulen aus Sicht ihrer Leitungen, Ausgabe 2020. Hochschul-Barometer. Stimmungsbarometer, Bildung für Future Skills, Hochschulpolitik. Stifterverband für die Deutsche Wissenschaft e. V., Essen, 2020.
- [Ste19] Stelzig, S.: Maßnahmen zur Innovation in der Lehre im digitalen Zeitalter: Rahmenbedingungen und Anreizsysteme für Lehrende. In: Morisse, K. (Hrsg.): Hochschulmanagement im digitalen Zeitalter, Hochschule Osnabrück, 2019, S. 15–26.
- [Wa20] Wannemacher, K.: Welche Bildungswege wird eine umfassende digital unterstützte Hochschule künftig eröffnen? In (Henke, J; Pasternack, P., Hrsg.): Wie die Hochschulen durch das Zeitalter des Frühdigitalismus kommen. Springer VS, Wiesbaden, 2020, S. 163-165.
- [We20] Wertz, I.: Zukunftsorientierte Lernräume. Kompetenzorientierung im Fokus. HIS-HE, Hannover, 2020 (HIS-HE:Medium2|2020).
- [WB21] Wannemacher, K.; Bodmann, L.: Künstliche Intelligenz an den Hochschulen. Potenziale und Herausforderungen in Forschung, Studium und Lehre sowie Curriculumentwicklung. Hochschulforum Digitalisierung, Berlin, 2021 (Arbeitspapier Nr. 59).
- [WLS19] Wannemacher, K.; Lübcke, M.; Seyfeli, F.: Things to Come. Digitalisierung und Bildungsteilhabe. Eine Trendanalyse zur Hochschulbildung der Zukunft. In (Hafer, J.; Mauch, M.; Schumann, M., Hrsg.): Teilhabe in der digitalen Bildungswelt. Waxmann, Münster etc., 2019, S. 78-88.
- [WRP20] de Witt, C.; Rampelt, F.; Pinkwart, N. (Hrsg.): Künstliche Intelligenz in der Hochschulbildung. Whitepaper, Oktober 2020. KI-Campus, Berlin, 2020.

Workshop: Where is the Science in
Data Science Projects?
Online-Workshop über die
Wissenschaftlichkeit von
Vorgehensmodellen für
Data-Science-Projekte (WISDAP)

Where is the Science in Data Science Projects?

Online-Workshop über die Wissenschaftlichkeit von Vorgehensmodellen für Data-Science-Projekte (WISDAP)

Jens Kaufmann,¹ Stephan Kühnel,² René Theuerkauf,³ Emal M. Alekosai,⁴ Sayed Hoseini,⁵ Uwe Neuhaus,⁶ Michael Schulz⁷

Abstract: Als Ergebnis einer virtuellen Arbeitsgruppe aus Wissenschaftler:innen und Praktiker:innen entstand zwischen April 2019 und Februar 2020 das Data-Science-Vorgehensmodell DASC-PM, dessen Ziel es ist, vorhandenes Wissen über die Durchführung von Data-Science-Projekten für alle Interessensgruppen in geeigneter Form zu strukturieren. Unter Berücksichtigung dieses Modells, aber nicht exklusiv darauf gestützt, soll der Aspekt der Wissenschaftlichkeit bei der Umsetzung von Data-Science-Projekten im Rahmen des Workshops besonders betrachtet werden.

Keywords: Data Science, DASC-PM, Wissenschaftlichkeit, Rigor, Relevance

1 Einleitung

Daten und ihre Auswertung haben in der modernen Wirtschaftsgesellschaft einen so großen Platz eingenommen, dass sich nicht nur der Begriff der „Datenökonomie“ entwickelt hat, sondern Rankings über die Leistungsfähigkeit selbiger in den Ländern der Welt erstellt werden [BCS19]. Die Nachfrage nach datenzentrierten Leistungen hat entsprechend zugenommen und mir ihr gestiegen ist auch die Komplexität der Aufgaben in den Bereichen Datenaufbereitung und Datenanalyse. Diese Begriffe und Aufgaben können zusammen mit weiteren verbundenen Tätigkeiten als *Data Science* zusammengefasst werden [Sc20a].

¹ Hochschule Niederrhein, Fachbereich Wirtschaftswissenschaften, Webschulstr. 41-43, 41065 Mönchengladbach, Deutschland jens.kaufmann@hs-niederrhein.de

² Martin-Luther-Universität Halle-Wittenberg, Lehrstuhl für Wirtschaftsinformatik, insb. Betriebliches Informationsmanagement, Universitätsring 3, 06108 Halle (Saale), Deutschland, stephan.kuehnel@wiwi.uni-halle.de

³ Martin-Luther-Universität Halle-Wittenberg, Lehrstuhl für Wirtschaftsinformatik, insb. E-Business, Universitätsring 3, 06108 Halle (Saale), Deutschland, rene.theuerkauf@wiwi.uni-halle.de

⁴ Robert Bosch GmbH, Borsigstr. 4, 70469 Stuttgart, Deutschland, mohammademal.alekozai@de.bosch.com

⁵ Hochschule Niederrhein, Fachbereich Elektrotechnik und Informatik, Reinarzstr. 49, 47805 Krefeld, Deutschland, sayed.hoseini@hs-niederrhein.de

⁶ NORDAKADEMIE Hochschule der Wirtschaft, Fachbereich Informatik, Köllner Chaussee 11, 25337 Elmshorn, Deutschland, uwe.neuhaus@nordakademie.de

⁷ NORDAKADEMIE Hochschule der Wirtschaft, Fachbereich Informatik, Köllner Chaussee 11, 25337 Elmshorn, Deutschland, michael.schulz@nordakademie.de

1.1 Data Science als Anwendungs- und Forschungsfeld

Traditionelle Datenanalysemethoden stoßen bei den beschriebenen Anforderungen häufig an die Grenze ihrer Möglichkeiten. Ausschlaggebend dafür sind mehrere Faktoren. So sind auf der einen Seite große Mengen an Daten frei verfügbar, was die Menge der Eingaben in die Verfahren stark erhöht, zum anderen sind diese aber häufig von heterogener Gestalt und Qualität sowie von unstrukturierter Form [He14]. Insbesondere durch diese häufig als *Big Data* bezeichnete Datenmenge ergeben sich bei der Integration von Daten neue Herausforderungen. Neben numerischen Daten auch Text, Audio und Video verarbeiten zu können, erfordert bei der Analyse eine Kombination aus Herangehensweisen verschiedener Disziplinen wie Mathematik und Informatik zur Bewältigung der reinen technischen und analytischen Schwierigkeiten sowie Wissen über die Anwendungsdomäne. Der so entstehende Dreiklang wird bei [Co10] beschrieben und häufig als Ausgangsdefinition der Data Science verwendet.

Während die bisher beschriebenen Aspekte noch als komplexere Formen etablierter Datenanalyse und ihrer Anwendung im betrieblichen Kontext aufgefasst werden können, zeigen weitere an diese angrenzende und Einfluss nehmende Entwicklungen, dass sich mit Data Science ein eigenes Forschungsfeld etabliert, das sich von bestehenden Methoden und Beschreibungen z. B. des Data-Mining abgrenzen lässt. Als ein Beispiel können Rechnersysteme genannt werden, die als Agenten eingesetzt werden, Daten automatisch auswerten und algorithmisch Entscheidungen treffen. Die dabei entstehenden Fragestellungen umfassen neben rein technischen und wirtschaftlichen Aspekten auch Betrachtungen der Ethik und des Datenschutzes [ABH17], welche eines umfassenden wissenschaftlichen Diskurses bedürfen. Ein weiterer Aspekt findet sich dann, wenn Techniken eingesetzt werden, die eine mangelnde Erklärbarkeit aufweisen und für den Nutzer selbst bei Kenntnis über die eingesetzte Methodik (z. B. neuronale Netze) der eigentliche Entscheidungsprozess auf den zugrunde liegenden Daten nicht nachvollziehbar ist. Dies ist insofern problematisch, da die Erklärung einer getroffenen Entscheidung ein Kernbestandteil menschlicher Kommunikation und Lernprozesse ist [Go18].

Data Science hat daher nicht nur die Aufgabe, Erkenntnisgewinn aus Daten zu ermöglichen, die notwendigen technischen wie analytischen Grundlagen zu schaffen oder bereitzustellen und die Analyseergebnisse mit fachlichem Wissen zu handlungsbefähigenden Interpretationen umzusetzen. Vielmehr muss sie auch dem Kern ihrer wörtlichen Bedeutung, der *Wissenschaft*, Respekt gezollt werden. Gemeint ist dabei allerdings nicht, dass ausschließlich Wissenschaftler die entsprechenden Methoden einsetzen sollen oder dass nur besonders zertifizierte Methodiken zum Einsatz kommen dürfen. Sie ist allerdings verpflichtet, nicht nur die reine Ergebnisproduktion in den Vordergrund zu stellen, sondern den Ansprüchen an Objektivität, Reliabilität und Validität nachzukommen, die auch von Praktikern gefordert wird und werden muss, um zum einen die Ergebnisse sicher und nachvollziehbar nutzen und präsentieren zu können und zum anderen die verwendeten Methoden korrekt auszuwählen oder bei Bedarf neu zu erstellen. Dies ist umso entscheidender, wenn beispielsweise die

Eignung bestimmter Verfahren für einen gegebenen Kontext geprüft werden soll und damit der Prozess des Erkenntnisgewinns im Vordergrund steht oder im Allgemeinen Verfahren aus dem akademischen Kontext in der Praxis für die konkrete Anwendung adaptiert werden.

1.2 Wissenschaftlichkeit der Data Science

Dem Workshop ging eine Studie der Workshop-Organisatoren voraus, die unter anderem vorhandene und gewünschte Kompetenzen sowie die aktuelle Nutzung von Vorgehensmodellen im Data-Science-Kontext untersuchte. Diese Studie motivierte konkret dazu, die Wissenschaftlichkeit der Data Science weiter aufzuarbeiten. So ist ein methodisches Vorgehen eine häufig von den Studienteilnehmer:innen verlangte Kompetenz von Data Scientist:innen. Zudem waren Wissenschaftler:innen mit einem großen Anteil in der Studie vertreten und gaben gleichzeitig an, Projekte durchzuführen, sodass sich mindestens, aber, wie vorher beschrieben, nicht ausschließlich, für diese Anwendergruppe die Notwendigkeit ergibt, Projektergebnisse auch wissenschaftlich angemessen aufzubereiten und nachvollziehbar zu gestalten, was die Disziplin selbst beeinflusst. Abschließend sind Vorgehensmodelle ein tradiertes Element wissenschaftlicher Methodik (siehe Abschnitt 4 dieses Beitrags) und das hier verwendete das *Data Science Process Model* (DASC-PM) als Vorgehensmodell soll aufzeigen, wie sein Einsatz die Wissenschaftlichkeit der Data Science unterstützt.

Es ist in diesem Zusammenhang ein viel diskutiertes Thema, Wissenschaft gegen die Praxis oder wissenschaftliche Strenge (Rigor) gegen die praktische Relevanz der Ergebnisse zu stellen. Tatsächlich schließen sich beide Thematiken nicht aus und sind auch mit der praktischen Relevanz, die die Anwender von Data Science mutmaßlich vordergründig interessiert, problemlos zu vereinbaren [Di12]. Wie also die Wissenschaft, die „Science“ in Data Science zu interpretieren ist, ist Gegenstand des in diesem Beitrag beschriebenen Workshops.

Dabei sollen aus Sicht von Wissenschaftler:innen und Praktiker:innen drei Kernfragen adressiert werden:

1. Was kennzeichnet die Wissenschaftlichkeit in der Disziplin Data Science?
2. Wie kann ein wissenschaftliches Vorgehen in datengetriebenen Projekten adäquat gehandhabt und sichergestellt werden?
3. Welche Kernaspekte sind bei der Durchführung, Dokumentation und Evaluation von Data-Science-Projekten aus wissenschaftlicher Sicht zu berücksichtigen?

Im Rahmen des Workshops sollen diese Fragen nicht einfach nur beantwortet werden, vielmehr ist es ein Kernziel, aus den gewonnenen Erkenntnissen Guidelines abzuleiten, die Wissenschaftler:innen und Praktiker:innen bei der Adressierung wissenschaftlicher Aspekte im Rahmen von Data-Science-Projekten angemessen unterstützen.

1.3 Aufbau des Beitrags

Neben den allgemein gültigen Überlegungen und Vorkenntnissen der Teilnehmer:innen des Workshops bildet das Vorgehensmodell DASC-PM eine weitere Grundlage der Diskussion. Dieses wird im zweiten Abschnitt des Beitrags nochmal zusammenfassend vorgestellt. Der dritte Abschnitt beschreibt eine Studie der Workshops-Organisatoren zu vorhandenen und gewünschten Kompetenzen sowie zur aktuellen Nutzung von Vorgehensmodellen im Data-Science-Kontext. Die Ergebnisse der Studie werden als rahmengebendes Element in die Diskussion eingebunden und liefern den Teilnehmer:innen wertvolle Erkenntnisse über die Anforderungen und Einschätzungen von Wissenschaftler:innen und Praktiker:innen zur aktuellen Anwendung und Durchführung von Data-Science-Projekten, sodass die Fragestellung des Workshops nach der Wissenschaftlichkeit im Vorgehen keine Betrachtung zum Selbstzweck fördert, sondern für die Überführung in praktische Handlungsanweisungen verwendbare Ergebnisse hervorbringt. Abschließend wird in Abschnitt vier das Workshopkonzept vorgestellt. Dieses wird methodisch aus den zuvor aufgearbeiteten Ergebnissen und Anforderungen an wissenschaftliches Vorgehen abgeleitet. Abschnitt und Beitrag schließen mit einem Ausblick auf weitere Schritte.

2 DASC-PM – Ein Modell für Data-Science-Aktivitäten

Data Science wird – wie ein wesentlicher Teil der Aktivitäten in Unternehmen – häufig in Projektform betrieben. Entsprechend ist das Umfeld dieser Projekte ebenso mit einzubeziehen wie die spezifische Ausgestaltung der Aktivitäten selbst. Da Data Science mehr umfasst als reine Analyseverfahren, benötigt sie auch ein eigenes, geeignetes Vorgehensmodell [Sc20b]. Um dieses Ziel zu erreichen, haben 22 Personen aus Wissenschaft und Wirtschaft in einer virtuellen Arbeitsgruppe ein Prozessmodell für Data-Science-Projekte entwickelt. Dort wurden umfangreiche Interviews zu den Aspekten der Disziplin sowie zu Kompetenz- und Rollendefinitionen geführt. Die Ergebnisse wurden im Anschluss von allen Mitgliedern der Arbeitsgruppe evaluiert. Dieser Prozess aus Interviews und Aufarbeitungen wurde in einem iterativen Verfahren zehn Mal wiederholt, sodass die Ergebnisse jedes Durchlaufs in die darauf folgenden mit eingeflossen sind. In einem abschließenden Aufarbeitungsprozess wurden die Arbeitsergebnisse zusammengeführt und angereichert. Das Resultat ist das *Data Science Process Modell* (DASC-PM) - ein Vorgehensmodell, welches Data-Science-Projekte anhand von Schlüsselbereichen beschreibt und die Kompetenzen sowie Rollen der Beteiligten darstellt. Der so dokumentierte Prozess zu Data-Science-Projekten sowie die Kompetenzbereiche sind in Abbildung 1 dargestellt. Das Modell selbst inklusive der hier vorgestellten Auszüge wird in [Sc21] beschrieben.

Das Modell beschreibt in der Arbeitsgruppe ausgewählte Aspekte der Data Science in einer möglichst umfassenden, nicht aber abschließenden Form. Insbesondere da die Data Science ein sehr aktives Forschungsfeld darstellt, ist DASC-PM keineswegs ein statisches Modell, sondern ein dynamisches Konstrukt, das sich z. B. durch die Ergebnisse von

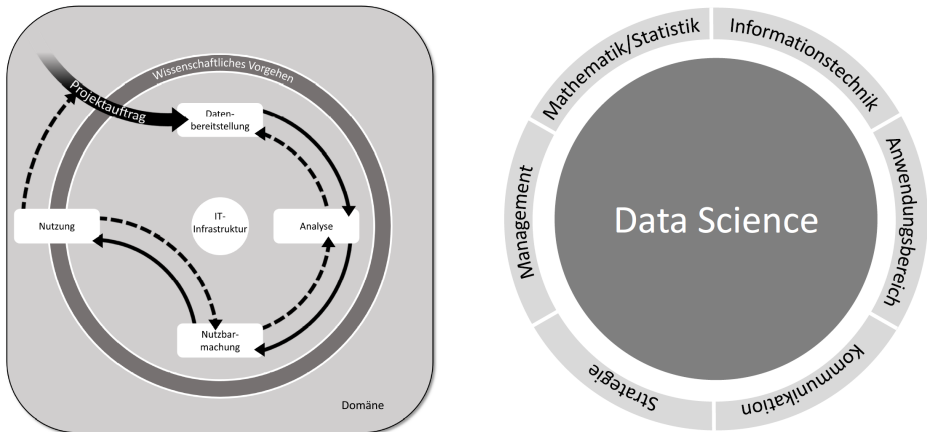


Abb. 1: (a) Data-Science-Vorgehensmodell DASC-PM (b) Nötige Kompetenzen in einem Data-Science-Projekt, beide aus [Sc21]

Workshops weiter verändern wird. Der in diesem Beitrag beschriebene Workshop soll seinen Anteil leisten, inhaltliche Lücken in Bezug auf den Aspekt der Wissenschaftlichkeit bei der Durchführung von Projekten zu schließen.

3 Studie

Im Zeitraum vom 30. November 2020 bis zum 21. Dezember 2020 wurde eine Umfrage zu Data-Science-Projekten durchgeführt. Adressiert wurden (unter Zuhilfenahme von E-Mail-Verteilerlisten und Social-Media-Kanälen) Personengruppen, die bereits Data-Science-Erfahrungen sammeln konnten. Durch dieses Vorgehen ist die genaue Zahl der Empfänger:innen des Aufrufes zur Teilnahme nicht zu benennen. Als Anreiz für die Beteiligung an der Umfrage wurde die Zurverfügungstellung der aufbereiteten Ergebnisse in Aussicht gestellt.

3.1 Informationen zu der Umfragedurchführung

Insgesamt kam es zu 107 Aufrufen der Umfrage. In 67 Fällen wurde die Bearbeitung unterbrochen. Ein Abbruchmuster war dabei nicht festzustellen. 40 Teilnehmer:innen haben die Umfrage beendet. Das arithmetische Mittel der Bearbeitungszeit lag bei 14,87 Minuten, der Median bei 10,45 Minuten. Sämtliche Ergebnisse wurden anonymisiert gespeichert. Keiner der Datensätze musste auf Grund üblicher Filterkriterien (zu kurze Bearbeitungszeit, auffälliges Bearbeitungsmuster, etc.) aus dem Datenset entfernt werden.

Abbildung 2 fasst einige Informationen zu den Teilnehmer:innen zusammen. 15 Personen arbeiten bei kleinen Unternehmen mit maximal 250 Mitarbeiter:innen, elf bei Unternehmen mit 251 bis 2000 Mitarbeiter:innen und 14 bei Unternehmen mit mehr als 2000 Mitarbeiter:innen. Mit einer Anzahl von 21 besaß etwas mehr als die Hälfte der Teilnehmer:innen keine Führungsposition, bei 19 Teilnehmer:innen war eine Führungsposition gegeben. Ebenfalls 19 Personen haben angegeben, große Erfahrungen im Data-Science-Bereich zu besitzen (Einschätzung 4 oder 5 auf einer Skala von "1 – n. v." bis "5 - Experte"), sieben eher wenig (Einschätzung 1 oder 2). 37 der 40 Teilnehmer:innen haben bereits mindestens ein Data-Science-Projekt mit durchgeführt, 23 bereits mehrere.

Von den Teilnehmer:innen schätzten sich acht am ehesten als Berater, elf als Wissenschaftler, 14 als Praktiker und sieben als Lernende ein. Letztere gehören jedoch vermutlich nicht zu einer Gruppe „Studierende“, sondern sind einer der zuvor genannten Gruppen zuzuordnen, betrachten sich selbst aber offenbar nicht als erfahren genug, um sich „Praktiker“ zu nennen. Diese Annahme wird auch dadurch gestützt, dass alle Teilnehmer:innen dieser Gruppe angeben, bereits Projekte durchgeführt zu haben, drei davon in Unternehmen.

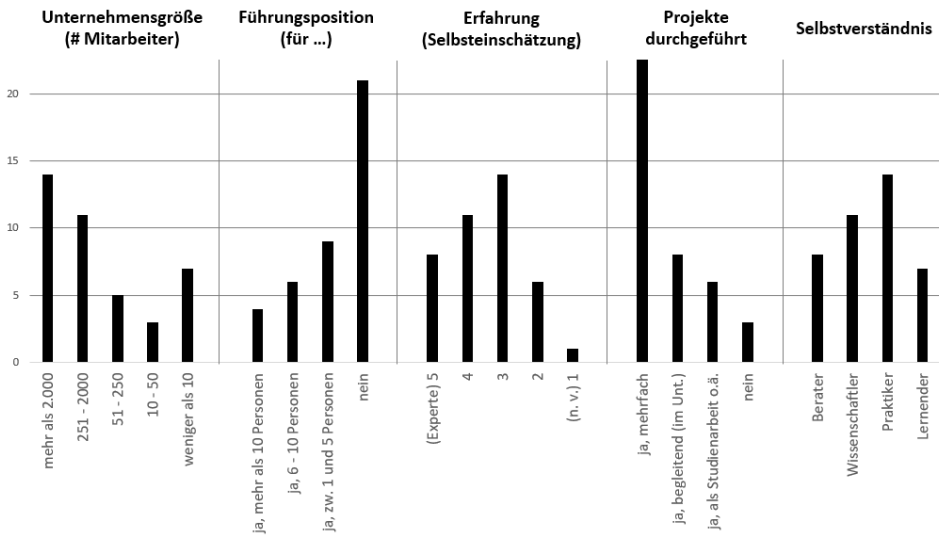


Abb. 2: Informationen zu den Teilnehmer:innen an der Umfrage

3.2 Status quo: Data-Science-Projekte und Kompetenzen

Der erste inhaltliche Themenblock der Umfrage beschäftigte sich damit, wer Data-Science Projekte anstößt (Management, Fachabteilung, Data-Science-Team, IT-Abteilung, Externe) und auf welche Geschäftsfelder (Hauptgeschäftsfeld, weitere bestehende Geschäftsfelder, neue Geschäftsfelder) die Use-Cases der Data-Science-Projekte vorwiegend einzahlen. Dabei wurde zwischen den aktuell vorhandenen Gegebenheiten in der Organisation der Umfrageteilnehmer:innen (Ist-Zustand) und den angestrebten Verhältnissen (Soll-Zustand) unterschieden.

Aktuell ist häufig ein dediziertes Data-Science-Team Initiator eines Data-Science-Projekts (34 %), gefolgt von den Fachabteilungen (27 %) und dem Management (21 %). Von geringerer Bedeutung sind Externe (14 %) und die IT-Abteilung (4 %). Für den Soll-Zustand wünschen sich die Umfrageteilnehmer:innen, dass Data-Science-Projekte noch häufiger von den Fachabteilungen (35 %) und dem Management (25 %) angestoßen werden, dem Data-Science-Team kommt in dieser Frage aber weiterhin eine wichtige Rolle zu (26 %). Der Einfluss externer Projekt-Initiatoren sinkt entsprechend (7 %), die IT-Abteilung ist weiterhin nur gelegentlich als Ideengeberin für Projekte vorgesehen (7 %).

Die Mehrheit der Use-Cases von Data-Science-Projekten zählt bislang auf das Hauptgeschäftsfeld des Unternehmens oder der Organisation ein (51 %), auf weitere bestehende Geschäftsfelder entfallen immerhin 29 %. Nur 20 % der Use-Cases zielen bislang auf neue oder neu zu entwickelnde Geschäftsfelder. Beim Soll-Zustand ändert sich dies, dort sollen Haupt- und neue Geschäftsfelder in etwa gleich stark durch Data-Science-Projekte unterstützt werden (36 % bzw. 34 %).

Als nächstes wurden die Umfrageteilnehmer:innen zu den Projektdimensionen *Teamgröße*, *Laufzeit* und *Budget* befragt. Die überwiegende Mehrheit aller Data-Science-Projektteams (77 %) umfasst zwischen 3 und 5 Personen, 17 % sind kleiner und 7 % größer. Die typische Projektdauer beträgt zwischen drei und sechs Monaten (52 %), 28 % der Projekte sind bereits nach ein bis zwei Monaten abgeschlossen, 17 % dauern länger als sechs Monate. Auch sehr kurze Projekte (Laufzeit von weniger als einem Monat) gibt es gelegentlich (3 %). Dabei könnte es sich um Proof-of-Concepts handeln. Das Budget von Data-Science-Projekten kann sehr unterschiedlich ausfallen. In 36 % der in der Umfrage genannten Fälle betrug das Budget weniger als 20.000 Euro, in 32 % zwischen 20.000 und 50.000 Euro und in 21 % zwischen 50.000 und 100.000 Euro. Nur wenige Projekte verfügten über größeres Budget (100.000 - 200.000 Euro: 4 %; mehr als 200.000 Euro: 7 %).

Stark ausgeprägt ist gemäß der Umfrageergebnisse der Zusammenhang von Projekt- und Unternehmensgröße. Kleinere Unternehmen (≤ 250 Mitarbeiter:innen) haben tendenziell kleinere Projektteams, kürzere Projektlaufzeiten und geringere Budgets. In größeren Unternehmen (> 250 Mitarbeiter:innen) findet sich hingegen die ganze Projektbandbreite, angefangen von kurzen und preiswerten Proof-of-Concepts bis zu langlaufenden Projekten mit großem Data-Science-Teams und entsprechendem Projektbudget.

Ein weiterer Block der Umfrage befasste sich mit den für ein Data-Science-Projekt benötigten Kompetenzen. Zuerst wurden die Teilnehmer:innen befragt, ob die im weiteren Projektverlauf benötigten Kompetenzen bereits zu Projektstart bekannt sind. Nur 16 % von ihnen haben diese Frage vollständig oder überwiegend bejaht. 62 % brachten hingegen zum Ausdruck, dass die benötigten Kompetenzen zu Projektstart noch nicht oder größtenteils nicht vorhanden sind. Dies spiegelt den häufig stark ausgeprägten Forschungscharakter von Data-Science-Projekten wieder, bei denen das konkrete Vorgehen und die eingesetzten Methoden vielfach erst während der Laufzeit festgelegt werden können.

Unabhängig von der Frage, ob die benötigten Kompetenzen bereits feststehen, ergibt sich bei Data-Science-Projekten häufig ein weiteres Problem: Sind die Mitarbeiter:innen mit den entsprechenden Kompetenzen im Unternehmen bzw. in der Organisation verfügbar? Nur 32 % der Umfrageteilnehmer:innen haben diese Frage vollständig oder überwiegend bejaht. 55 % sagten aus, dass entsprechend qualifiziertes Personal ganz oder überwiegend fehlt.

Um ein möglichst detailliertes Bild davon zu erhalten, an welchen konkreten Kompetenzen ein Mangel herrscht, wurden die Teilnehmer:innen zu einem vorher definierten Data-Science-Kompetenzspektrum befragt, das an die Ergebnisse von [Sc21] angelehnt ist und die folgenden elf Kompetenzen umfasst: *Mathematik, Methoden, Daten, Programmierung, Systemadministration, Anwendbarkeit, Domäne, Kommunikation, Strategie, Projektmanagement, Riskmanagement*. Die Teilnehmer:innen sollten dazu Auskunft geben, ob sie die jeweilige Kompetenz generell für relevant in Data-Science-Projekten halten, in welchem Maße die Kompetenz im eigenen Unternehmen / in der eigenen Organisation vorhanden ist und in welchem Maße die Kompetenz tatsächlich benötigt wird. Für die letzten beiden Fragen wurde ein Skala von 1 (Kompetenz überhaupt nicht vorhanden) bis 5 (Kompetenz vollständig vorhanden) verwendet.

Alle elf aufgeführten Kompetenzen wurden von den Teilnehmer:innen als relevant für Data-Science-Projekte eingestuft. Wie Abbildung 3 zeigt, gibt es jedoch erhebliche Diskrepanzen zwischen den benötigten und den tatsächlich vorhandenen Kompetenzen. In allen elf Bereichen liegt der Wert für die vorhandenen Kompetenzen unter dem der benötigten. Am stärksten ausgeprägt ist diese Differenz bei den Data-Science-Methoden, der Mathematik und der Programmierung. Die kleinsten Unterschiede ergaben sich bei der Befragung bei den Punkten Projektmanagement, Domäne und Systemadministration.

3.3 Vorgehensmodelle in der Data Science

Bei der Durchführung von Data-Science-Projekten werden häufig Vorgehensmodelle eingesetzt (vgl. Kapitel 2), die einen strukturierten Projektablauf mit vorgegebenen Phasen und definierten Aufgaben unterstützen. In der Umfrage wurde daher erfragt, welche dieser Vorgehensmodelle bekannt sind, ob sie in der Praxis tatsächlich eingesetzt werden und wie die Umfrageteilnehmer:innen die Nützlichkeit der Modelle generell einschätzen.

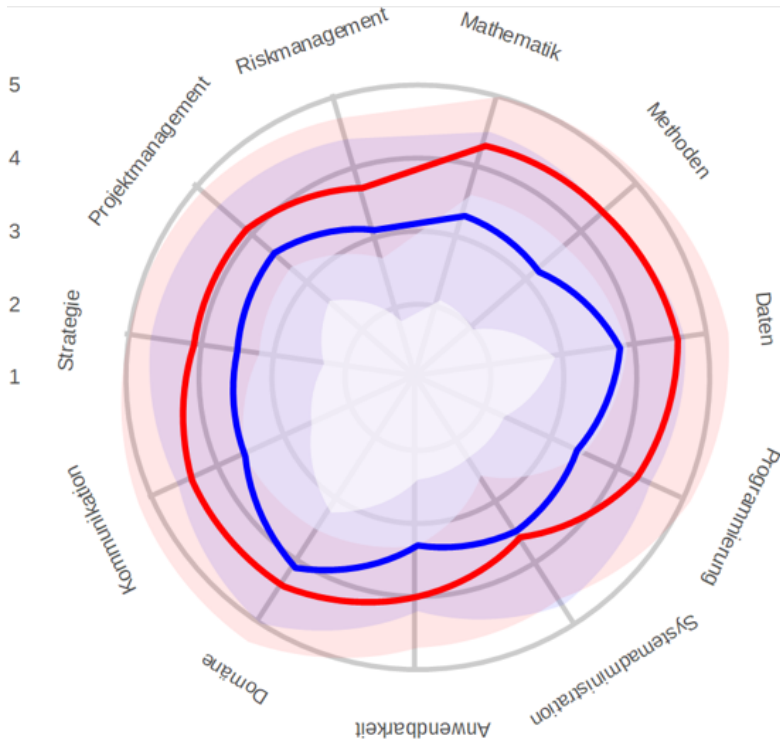


Abb. 3: Differenzen zwischen vorhandenen (blau) und benötigten (rot) Kompetenzen. (Skala: 1 = überhaupt nicht vorhanden (Kreismittelpunkt) bis 5 = vollständig vorhanden (Kreisrand))

Fast alle Umfrageteilnehmer:innen (85 %) kennen mindestens ein Vorgehensmodell für Data-Science-Projekte, viele sogar zwei (58 %), drei (38 %), vier (25 %) oder noch mehr (10 %). Die beiden am häufigsten angegebenen Modelle sind dabei der *Cross-Industry Standard Process for Data Mining (CRISP-DM)* und der *Knowledge Discovery in Data Bases Process (KDD)*. Sie sind 75 % bzw. 40 % der Teilnehmer:innen bekannt. Der Bekanntheitsgrad des *Team Data Science Process (TDSP)*, der *Analytics Solutions Unified Method for Data Mining/Predictive Analytics (ASUM-DM)* und von *Sample, Explore, Modify, Model, and Assess (SEMMA)* ist mit 18 %, 13 % und erneut 13 % hingegen geringer.

Die große Mehrheit (83 %) der Umfrageteilnehmer:innen sind vollständig oder überwiegend der Auffassung, dass der Einsatz eines Vorgehensmodells die Erfolgswahrscheinlichkeit von Data-Science-Projekten erhöht. Nur 8 % stimmten eher nicht zu. Der Aussage, dass der Einsatz von Vorgehensmodellen allgemein der Standard bei Data-Science-Projekten ist, haben allerdings nur 30 % der Umfrageteilnehmer:innen ganz oder überwiegend zugestimmt, 53 % gaben an, dass dies bislang nicht bzw. überwiegend nicht der Fall ist. Eine mögliche Erklärung dafür könnte sein, dass für Data-Science-Projekte andere,

allgemeinere Vorgehensmodelle genutzt werden. Dies wäre aufgrund der Erfahrungen der Teilnehmer:innen jedoch weniger empfehlenswert, denn 60 % sind der Auffassung, dass sich ein Vorgehensmodell für Data-Science-Projekte von einem allgemeinen Vorgehen für Projekte unterscheidet. Nur 15 % waren eher nicht dieser Meinung.

4 Konzept des WISDAP-Workshops und weitere Schritte

Wissenschaftlichkeit („Science“) und rigoroses wissenschaftliches Vorgehen nehmen in der Data Science bereits auf Basis der Gebietsbezeichnung eine entscheidende Rolle ein. Die Ergebnisse der Studie zeigen, dass die Teilnehmer:innen sowohl vor dem Hintergrund ihrer eigenen Tätigkeit als auch vor der geplanten Anwendung von Ergebnissen durchgeführter Data-Science-Aktivitäten einen Bedarf sehen, die Verlässlichkeit von Algorithmen und Analysen durch Nutzung stringenter Methoden und Vorgehensmodelle zu sichern.

4.1 Vorgehensmodelle in der Wissenschaft

Dass Vorgehensmodelle in der (Wirtschafts-)Informatik eine hohe Relevanz besitzen, zeigt sich bereits an der Existenz einer entsprechenden Fachgruppe *Vorgehensmodelle für die betriebliche Anwendungsentwicklung (WI-VM)* innerhalb der Gesellschaft für Informatik e.V. [Ge21]. Dabei werden Vorgehensmodelle üblicherweise als eine Form von Referenzmodellen beschrieben ([Br12; VF16]), die ihrerseits eine Standard-Methode zur Erstellung von IT-Artefakten im Rahmen der Forschung in der Wirtschaftsinformatik darstellt [Fr17]. Tatsächlich existieren anerkannte Vorgehensmodelle zur Entwicklung von Referenzmodellen [Fe14]. Die Verwendung von Vorgehensmodellen kann daher als etabliertes Vorgehen nicht nur im praktischen, sondern auch im wissenschaftlichen Kontext angesehen werden.

Die Frage nach der Wissenschaftlichkeit der Data Science wird aus diesem Grund mit Hilfe dreier Kernfragen (siehe Abschnitt 1) und anhand des Vorgehensmodells DASC-PM untersucht, sodass die gewonnenen Erkenntnisse auch am Modell in handlungsunterstützende Artefakte überführt werden können. Dass diese konkrete Form eines Ergebnisses auch im wissenschaftlichen Kontext angemessen sein kann, zeigt sich beispielsweise bei vom Brocke et al. (2015), die ihre Richtlinien zur Literatursuche um eine Tabelle im Checklisten-Format ergänzen [Br15].

4.2 Konzept des Workshops

Der Workshop verfolgt den Ansatz des User-centered Design (UCD), der Design-Prozesse beschreibt, bei denen potentielle Anwender:innen frühzeitig aktiv in den Design-Prozess einbezogen werden [AMP04]. Der Ansatz des UCD wurde bereits bei der initialen Konzeption

des DASC-PM verfolgt, wobei die Erfahrungen und das Wissen von Wissenschaftler:innen und Praktiker:innen, die als potentielle Anwender:innen des DASC-PM angesehen werden, erhoben und berücksichtigt wurden. Im Einklang mit diesem Vorgehen werden auch im WISDAP-Workshop beide Nutzergruppen einbezogen. Da bei der Untersuchung ein besonderer Fokus auf den unterschiedlichen Sichtweisen von Wissenschaft und Praxis liegt, erscheint eine initiale Unterteilung der Workshop-Teilnehmer:innen in Wissenschaftler:innen und Praktiker:innen als sinnvoll, um gegenseitige Beeinflussungen zu vermeiden. Innerhalb der beiden Gruppen - Wissenschaft und Praxis - werden schließlich moderierte Fokusgruppendifkussionen durchgeführt.

Die moderierte Fokusgruppe wird als eine bedeutende Methode des UCD angesehen [AMP04]. Dabei handelt es sich um eine qualitative Forschungsmethode, bei der ein:e Moderator:in eine Gruppendiskussion leitet, die sich auf die Interaktion zwischen den Teilnehmern:innen stützt, um neue Erkenntnisse zu generieren [Mo97]. Die erforderliche Anzahl moderierter Fokusgruppen, die für verlässliche Ergebnisse notwendig ist, ist in der Literatur stark umstritten. Basierend auf den empirischen Ergebnissen von Guest et al. [GNM17] sind zwei Fokusgruppen jedoch als ausreichend anzusehen.

Das Vorgehen in beiden moderierten Fokusgruppen besteht aus insgesamt vier Schritten:

- 1) Einführung in das Forschungsthema,
- 2) Vorstellung des Problemkontextes,
- 3) Diskussion über die Fragen 1-3 (vgl. Kapitel 1) und
- 4) Überführung der Diskussionsergebnisse in Guidelines.

4.3 Weiteres Vorgehen

Im Nachgang des Workshops werden die gewonnenen Ergebnisse zusammengeführt und der offenen und virtuellen DASC-PM Arbeitsgruppe zu Evaluationszwecken vorgestellt. Hierbei soll geprüft werden, ob die Guidelines verständlich, vollständig, nachvollziehbar, praktisch anwendbar und nützlich sind oder ob es weiterer Überarbeitungen bedarf.

Die Erkenntnisse fließen in die regelmäßige Weiterentwicklung des Vorgehensmodells ein. Soweit sie allgemeingültig scheinen, sollen sie der wissenschaftlichen Gemeinschaft auch durch geeignete weitere Publikationen zur Diskussion und Weiterentwicklung zur Verfügung gestellt werden.

Literatur

- [ABH17] van der Aalst, W.M.; Bichler, M.; Heinzl, A.: Responsible Data Science. *Business and Information Systems Engineering* 59/5, S. 311–313, 2017.
- [AMP04] Abras, C.; Maloney-Krichmar, D.; Preece, J.: User-Centered Design. In (Bainbridge, W. S., Hrsg.): *Berkshire Encyclopedia of Human-Computer Interaction*. Berkshire Publishing Group, S. 763–768, 2004.
- [BCS19] Bhalla, A.; Chakravorti, B.; Shankar Chaturvedi, R.: Which Countries Are Leading the Data Economy?, 2019, URL: <https://hbr.org/2019/01/which-countries-are-leading-the-data-economy>, Stand: 20.07.2021.
- [Br12] Breitner, M. H.: Vorgehensmodell - Enzyklopädie der Wirtschaftsinformatik, 2012, URL: <https://www.enzyklopaedie-der-wirtschaftsinformatik.de/wi-enzyklopaedie/lexikon/is-management/Systementwicklung/Vorgehensmodell>, Stand: 20.07.2021.
- [Br15] vom Brocke, J.; Simons, A.; Riemer, K.; Niehaves, B.; Plattfaut, R.; Cleven, A.: Standing on the shoulders of giants: Challenges and recommendations of literature search in information systems research. *Communications of the Association for Information Systems* 37/, S. 205–224, 2015.
- [Co10] Conway, D.: The Data Science Venn Diagram, 2010, URL: <http://www.dataists.com/2010/09/the-data-science-venn-diagram/>, Stand: 22.06.2021.
- [Di12] Dilger, A.: Rigor, wissenschaftliche und praktische Relevanz. In: Diskussionspapier des Instituts für Organisationsökonomik. Diskussionspapier des Instituts für Organisationsökonomik 03/2012, Institut für Organisationsökonomik, Westfälische Wilhelms-Universität Münster, 2012, URL: https://www.wiwi.uni-muenster.de/io/sites/io/files/forschen/downloads/dp-io_03_2012.pdf.
- [Fe14] Fettke, P.: Eine Methode zur induktiven Entwicklung von Referenzmodellen. In: Tagungsband Multikonferenz Wirtschaftsinformatik 2014. S. 1034–1047, 2014.
- [Fr17] Friedrich, T.; Schlauderer, S.; Weidinger, J.; Raab, M.: On the Research Paradigms and Research Methods Employed in the BISE Journal - A Ten-Year Update. In: *Proceedings of Wirtschaftsinformatik 2017*. S. 1111–1125, 2017.
- [Ge21] Gesellschaft für Informatik e.V.: Startseite FG WI-VM, 2021, URL: <https://fg-wi-vm.gi.de/>, Stand: 20.07.2021.
- [GNM17] Guest, G.; Namey, E.; McKenna, K.: How Many Focus Groups Are Enough? Building an Evidence Base for Nonprobability Sample Sizes. *Field Methods* 29/1, S. 3–22, 2017.
- [Go18] Goebel, R.; Chander, A.; Holzinger, K.; Lecue, F.; Akata, Z.; Stumpf, S.; Kieseberg, P.; Holzinger, A.: Explainable AI: The new 42? 2018.

-
- [He14] Hendler, J.: Data integration for heterogenous datasets. *Big Data* 2/4, S. 205–215, 2014.
- [Mo97] Morgan, D. L.: *Focus groups as qualitative research*. Sage Publ, Thousand Oaks, Calif., 1997.
- [Sc20a] Schulz, M.: Data-Science-Projekte und ihre Besonderheiten. *Wirtschaftsinformatik & Management* 12/5, S. 376–381, 2020.
- [Sc20b] Schulz, M.; Neuhaus, U.; Kaufmann, J.; Badura, D.; Kuehnel, S.; Badewitz, W.; Dann, D.; Kloker, S.; Alekozai, E. M.; Lanquillon, C.: Introducing DASC-PM: A Data Science Process Model. In: *ACIS 2020 Proceedings*. 45. 2020.
- [Sc21] Schulz, M.; Neuhaus, U.; Kaufmann, J.; Badura, D.; Kerzel, U.; Welter, F.; Prothmann, M.; Kühnel, S.; Passlick, J.; Rissler, R.; Badewitz, W.; Dann, D.; Gröschel, A.; Kloker, S.; Alekozai, E. M.; Felderer, M.; Lanquillon, C.; Brauner, D.; Gölzer, P.; Binder, H.; Rhode, H.; Gehrke, N.: *DASC-PM v1.0 - Ein Vorgehensmodell für Data-Science-Projekte*. 2021.
- [VF16] Vom Brocke, J.; Fettke, P.: Referenzmodell - Enzyklopädie der Wirtschaftsinformatik, 2016, URL: <https://www.enzyklopaedie-der-wirtschaftsinformatik.de/lexikon/is-management/Systementwicklung/Softwarearchitektur/Wiederverwendung-von-Softwarebausteinen/Referenzmodell/index.html>, Stand: 20.07.2021.

Standardization Workshop on
Industrial Automation and Control
Systems (6thIACSWS'21)

Presentation of the Programme of the 6th GI/ACM Standardization Workshop on Industrial Automation and Control Systems

The 6th IACS WS'21 September 28, 2021 Virtual

Jan de Meer¹, Karl Waedt², Axel Rennoch³, Hans-Joachim Hof⁴

Abstract: One of the important basic objectives of the 6th IACS WS'21 is to contribute to Sustainability achieved by standardization that is based on new I4.0 technologies such as Smart Manufacturing, Digital Twin, AI-based Robotics, Industrial Internet of Things and more.

Keywords: Industrial Automation and Control Systems, I4.0, Standardization, Smart Manufacturing, Production Life Cycle, Security Levels, Functional Safety.

1 The 6th IACS WS'21 Specification

The 6th IACS Standardization WS'21 in series is again aligned with the yearly assembly of the 'GI Jahrestagung 2021'. Hence the Workshop schedule is planned in accordance with the conference schedule

of the [51st GI Jahrestagung from September 27 to October 01, 2021](#)

organized by GI-Berlin-Virtual:



All the communications to prepare the workshop including submissions uploads and reviewing, communicating with participants, organizers, PC members and last not least,

¹ Smartspace@lab.eu GmbH, Berlin Germany, demeer@ACM.ORG

² Framatome GmbH, Erlangen Germany, karl.Waedt@framatome.com

³ Fraunhofer-FOKUS, Berlin Germany, axel.rennoch@fokus.fraunhofer.de

⁴ Technische Hochschule Ingolstadt Germany, Hans-Joachim.Hof@thi.de

authors has been achieved via the *easychair tool* by registering and logging-in to the **6th IACS WS'21!** —> <https://easychair.org/account/signup>

The use and deployment of *EasyChair* enabled the organizers to track all the constraints and conditions on the numerous submissions of this event. On the *easychair* platform the members of the co-chair board (CB) and the programme committee (PC) have performed diverging communications EU-wide and quite heterogeneously comprising industrial, scientific, governmental start-up SMEs and many other interested parties and organizations integrated to a ‘virtual round table’.

The workshop’s joint CB/PC Board appreciates the continuous support from Fraunhofer FOKUS in providing logistics and scientific experiences for preparing the workshop but also for printing posters, organizing reviewing decision-making meetings etc!

One of the **general objective of the 6th IACS WS'21** and of the conf rence ,*GI Jahrestagung 2021*‘ is on how to achieve sustainability by standardization of I4.0 Automation and Control Systems (IACS) according to the current industrial system and component requirements of one of the most important industrial standards series IEC 62443!

Work of Standards Developing Organizations (SDOs) and international Technical Standardization Committees compared to the concepts and approach of IACS is also part (but not limited to) of IEC TC65 ,Smart Manufacturing‘ WG23 (system) and WG24 (AAS), ISO/JTC1 SC27 on Security Technologies, SC38 on Cloud Computing, SC41 on Industrial Internets of Things, SC42 on Artificial Intelligence Technology in Smart Factories, Smart Cities, Smart Grids etc.

The 6th GI/ACM IACS WS'21 nominated Board of Co-Chairs:

Jan B. de Meer(General),

Karl Waedt(Co),

Axel Rennoch(Co),

Hans-Joachim Hof(Co)

The 6th GI/ACM IACS WS'21 Programme Committee Members:

Axel Rennoch

Fraunhofer Institute FOKUS Berlin

Hans-Joachim Hof

Technische Hochschule Ingolstadt INSI

Jan-Bernhard De Meer

German Chapter ACM

Karl Waedt	Framatome GmbH Erlangen
Olga Meyer	Fraunhofer Institute IPA Stuttgart
Peer Reymann	ITQS GmbH Germany
Rainer Falk	Siemens AG München
Sabine Kruspig	Kanzlei Schwarz & Kollegen München
Scott Cadzow	C3L UK
Steffen Fries	Siemens AG München
Ulrich Seldeslachts	Leuven Belgium

The 6th IACS WS'21 Time Plan anticipated

1. Early Registration of Abstracts of Intentional Submissions: April 11;
2. LNI-ready papers for the 6th IACS WS'21 LNI Proceedings: June 30; (End of Reviewing)
3. Final LNI Proceedings Preparation: July 31
4. The 6th IACS WS'21, GI Berlin-virtual: September 28 (accepted presentations due).

2 The 6th IACS WS'21 Supporters

Starting with the GI Conference *Informatik2016* held in Klagenfurt the IACS Workshop Series appreciates cooperation with national and EU industrial and research standardization supporting organizations:



3 The 6th IACS WS'21 Targets

One of the important basic objectives of the 6th IACS WS'21 is to contribute to Sustainability achieved by standardization that is based on new I4.0 technologies such as Smart Manufacturing, Digital Twin, AI-based Robotics, Industrial Internet of Things and more.

According to the multistandard IEC 62443 prescribing requirements on security, safety, privacy, quality of work, asset management etc. the 6th IACS WS'21 adopts the following objectives derived from IEC 62443:

- IACS Modeling, Vocabularies and Concepts
- System Security Conformance Metrics
- Production Lifecycle and Use Cases of New Technologies
- IACS Risk Assessment and Security Levels
- Administration Shell of Repositories for I4.0 Objects and IoT Devices
- Knowledge Derivation from Big Data Lakes
- Ontology Language and Unique Object Identification
- Security Algebra and Verification Techniques
- Human-Machine Interoperability
- Semantic Interoperability in SM/IIoTs
- New Artificial Intelligence Techniques and Approaches
- Functional Safety and Trustworthiness
- ICT Application Regulations and Ethics.

4 The 6th IACS WS'21 Narrative

The industrial multipart standard IEC 62443 IACS addressed by the GI/ACM I4.0 standardization workshop series has been and still is developed in accordance with other standardization organizations aiming at requirements, methods and techniques of CRITIS, ETSI TC Cyber, IEC TC65 Smart Factoring, ISO JTC1 SC27/WG4 IT Security, Artificial Intelligence (SC42) and Industrial IoT (SC41) and more.

From early beginning of the IEC 62443 series development in 2013 and since the multipart standard IEC 62443 is continuously growing. Almost every 2 to 3 years a new standard

part is published and comprises to-day 13 parts ranging from system to component specifications. The roadmap towards a complete view on I4.0 Systems comprises a system view of four groups of standard parts, i.e. General Concepts (1), Policies and Procedures (2), System Aspects (3), Component Aspects (4).

In 2018 a 5th group of 'Industry Profiles' has been defined and became started and is now nearly to be finished. Industry Profiles are based on available normative parts of IEC 62443-2.4 for solutions suppliers and IEC 62443-4.1 for product developers.

This is now the 6th IACS Workshop in series joined with 'GI Jahrestagung' 2021 in Berlin-Virtual. All former I4.0 standardization workshops organized under the auspices of 'GI Jahrestagung' happened in Karlsruhe (the first WS organized virtually), Kassel (4th), Berlin(3rd), Chemnitz (2nd) and Klagenfurt (1st).

The members of the PC and its Chairing Board (CB) appreciated and acknowledging the manifold support from the associations of 'German Chapter ACM' and 'Gesellschaft für Informatik (GI)' of the D-A-CH countries, the EU 'ECSO' and, the national 'SCI4.0' and 'Plattform I4.0' organizations and finally the support of Fraunhofer FOKUS Berlin.

The 'European Cyber Organization (ECSO)' supports the workshop's main issue of addressing industrial development and harmonization by and with standards as it is key to the IEC 62443 IACS series. ECSO is structured into 6 working groups comprising but not limited to Standardization of Supply Chain Management (1), International Collaborations (2), Sectorial Demanding I4.0 (3), Coordination with Regions (4) and Awareness of Cyber Ranges (5).

July 31, 2021, Berlin, for the joint board of Co-Chairs and PC members:

Jan-Bernhard deMeer (General Chair)

Axel Rennoch, Hans-Joachim Hof, Karl Waedt (Co-Chairs).

5 The 6th IACS WS'21 Programme

On the following poster you'll find the list of accepted authors and their speeches:

**Preliminary Programme of the 6th GI/ACM I4.0 Standardization Workshop
on Industrial Automation and Control Systems (6th IACS WS'21),
Berlin-Virtual, 2021 September 28, 9h-16h (3 sessions)**



- **Opening Keynote Prof. Hannes Federrath, President of GI:** Science, Standardization and Industry4.0 – How do they together make innovations
- **Joseph Schindler FA University Erlangen et al:** Secure OPC UA Server Configuration for Smart Charging Stations
- **Robert Altschaffel OVG University Magdeburg et al:** Supporting Security in Industrial Automation and Control Systems using Domain-specific Modelling
- **Asmaa Tellabi University of Siegen et al:** ABAC and RBAC for IACS for Industry4.0 Access Control Management
- **Martin Szemkus HS Magdeburg-Stendal et al:** Primary and Supporting Assets for IACS Risk Management
- **Keynote Sebastian Fritsch secuvera GmbH:** Evaluation Concepts for Security of IACS-Systems



- **Invited Keynote N.N. German-Chinese Cooperation:** Cyber Security Testing Systems for Sino-German Intelligent Manufacturing I4.0
- **Christele Larissa Moussi-Djeukoua FA University Erlangen:** Secure Unidirectional Security Gateways for Industry 4.0
- **Raman Barakat Fraunhofer FOKUS Berlin et al:** Towards a Certification Scheme for IoT Security Evaluation
- **Keynote Detlef Tenhagen JTC1/SC41 Convenor et al:** Facets of the Digital Twin
- **Anja Simon Labs NW Industrie 4.0 e.V:** Neutral Interoperability Testbeds
- **Keynote Olga Meyer Fraunhofer IPA Stuttgart:** German Standardization Roadmap 4.0 – From national recommendations for action to global harmonization








6 The 6th IACS WS'21 Speakers

Anja Simon (f) received her diploma degree in engineering for technical cybernetics and automation at the HTWK Leipzig in 1989 and a degree in economics of engineering from an association of the European Union and German IHK in 1996. In addition, she graduated as master of consulting excellence at Siemens Munich in 2006. She joined Siemens AG in 1992 and has worked in multiple different positions since then. Her professional career has led Anja Simon from software developer, product and sales engineer as well senior business consultant for major operational projects for industrial customers to central management tasks for global multifunctional shared service centers and the responsibility of international IT service outsourcing deals. Recently, as a program manager for cross-

company digitization projects, she was responsible for the R&D functions and the operational roll-out of joint solutions. Anja Simon will take over the CTO role for LNI 4.0 (Lab Networks for Industrie 4.0) on July 1st, 2021.

Axel Rennoch is computer scientists at the Fraunhofer Institute for Open Communication Systems in Berlin. As a member of the System Quality Competence Center, he is involved / responsible for validation and testing projects on next generation networks and software technologies. Axel Rennoch has been working in the field of Formal Methods, Testing Methodologies and Quality of Service considerations in various scientific and industrial projects since 1985. His experiences address the application of Formal Description Techniques, the development and execution of protocol tests, and software testing. During this work he contributed to several national and international standardization groups (e.g. DIN, ISO, ATM-Forum, OMG, ETSI) and published research papers continuously.

Joseph Schindler is currently a member of the Graduate Program at Framatome GmbH Erlangen. After finishing his Bachelor of Mechanical Engineering, Josef Schindler decided to reorientate and to study Electrical Power Engineering at Friedrich-Alexander-University Erlangen-Nuremberg. His key motivation for this change for the master studies was the "Energy Transition/Energiewende". Luckily, he found an ideal master thesis for this concern: "Modelling of a Hybrid Energy Storage (Vanadium Redox Flow Battery & Flywheel Storage) with a Neural Network-based Control". In 2018, he started a PhD graduate program at Framatome, where he researches the impact of cross-commodity sharing at neighbourhood-level. Aims are the reduction of peak loads and better integration of Renewable Energy Sources. There is a high potential for cybersecurity attacks in the network interconnection of the neighbours. Hence, the research focuses increasingly on cybersecurity-related topics.

Markus Rentschler completed his studies in 1993 and holds degrees in Communications Engineering from the University of Applied Sciences in Konstanz/Germany and Digital Systems Engineering from the Heriot-Watt University in Edinburgh/Scotland. Since then, he could gain over 25 years of experience in development and quality assurance of embedded communication systems and is holder of several patents. In his current position he is with his team responsible for the standardization of system interfaces for the *Balluff* product families and an active member or lead in several external standardization activities. Besides that, he is a part-time lecturer for "Software Engineering" at the Cooperative State University in Stuttgart and is regularly publishing on national and international technical conferences.

Olga Meyer works as a research associate at the Fraunhofer Institute for Manufacturing Engineering and Automation IPA. She leads several public and contract research projects in the field of manufacturing IT and develops innovative solutions for cloud manufacturing as well as IT architectures for cyber-physical production systems and communication technologies. Olga Meyer is an active member of several standardization working groups at national and international level working on the development of standards in the areas of Smart Manufacturing, Digital Twin, Industrial Internet of Things and other related technologies. Within the Horizon 2020 project QU4LITY, she contributed to the "German Standardization Roadmap for Industry 4.0", which is one of the central communication media for Industry 4.0 in Germany.

Secure OPC UA Server configuration for smart charging stations

Josef Schindler,¹ Erkin Kirdan,² Karl Waedt³

Abstract: Direct-To-Consumer (D2C) marketing recently gains popularity in society by bypassing unnecessary mediators and thus avoiding cost. In Information and Communication Technology (ICT) terms, it is comparable with Machine-To-Machine (M2M) communication overcoming additional mediators, i.e. remote servers. In this paper, we research M2M communication for battery boosted charging station. Therefore, we consider a setup comprising an OPC Unified Architecture (OPC UA) Client and OPC UA Server. The server represents the smart charging station, where the client can reserve time slots for charging sessions.

In this paper, we answer how to facilitate this using the different services provided by OPC UA. Additionally, we apply an OPC UA Metasploit module on two differently - according to the function manuals - configured OPC UA Servers. Afterwards, we validate the results.

Keywords: Machine-To-Machine; OPC Unified Architecture; cybersecurity; penetration testing; charging station

1 Introduction

In previous work [SWW19], we demonstrated the standardisation of OPC UA Server interfaces with OPC UA Namespaces. Therefore, we exemplarily considered a boosted charging station. A battery helped to overcome insufficient electricity grid connection for high power charging there. Between two charging sessions, the power grid recharges the battery. That refreshes the battery so it can boost the following charging process. In figure 1, yellow arrows depict the power flows by pointing from one component to another. As can be seen, the grid connection is weak (thin line), while the power flow into the Electric Vehicle (EV) is greater. For reservation (customer), organisation (with nearby charging entities) and remote maintenance purposes, there is an OPC UA Server on-site, listening to OPC UA Clients (see figure 1).

According to Erba et al. [EMT21], a secure configuration is crucial, and a remarkable share of vendors do not provide explicit recommendation in their manuals for it. We want to

¹ Friedrich-Alexander-University (FAU) Erlangen-Nuremberg, Chair of Electrical Energy Systems (EES), Cauerstraße 4, 91058, Erlangen, Germany josef.s.schindler@fau.de

² Technical University Munich erkin.kirdan@tum.de

³ Framatome GmbH, ICETA-G Department, Paul-Gossen-Straße 100, 91052, Erlangen, Germany karl.waedt@framatome.com

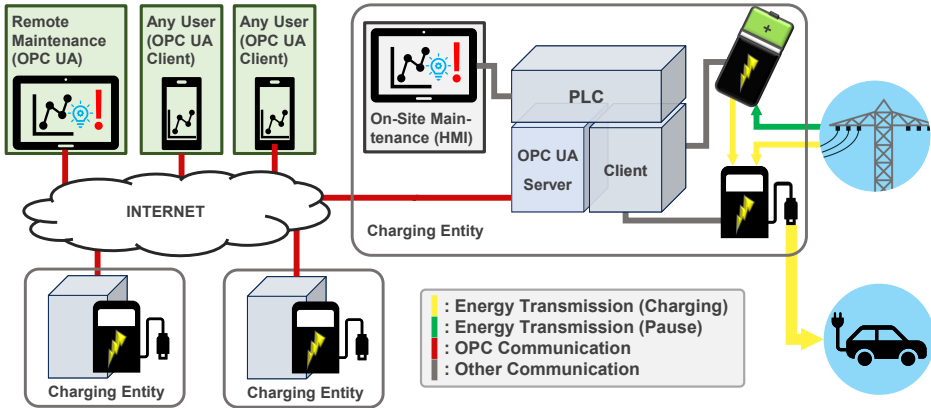


Fig. 1: system communication and power flow overview

look a little bit deeper into the security of Siemens' PLC platform, which is used widely in automation. Our contributions are:

1. Cybersecurity consideration of the boosted EV charger with reservation system (according to [SWW19])
2. A comparison of the manuals concerning OPC UA provided by Siemens [Si18, Si21] and their alignment with a secure configuration
3. Penetration test of the OPC UA Server hosted by the S7-1518 with Metasploit OPC UA-extension [Ro20]

In the following sections, we do a literature review (section 2). Next, in section 3, we depict the hardware configuration (subsection 3.1) and the penetration testing environment and setup (subsection 3.2). Sections 4 & 5 depict the results & conclusion, respectively.

2 Related Work

The security models of the four most commonly used open-source implementations are analyzed in [Mu20a]. According to the results, they comply with the standard's security model with a few vulnerabilities in 2 of the implementations. Furthermore, their scalability is compared for an increasing number of OPC UA Clients and nodes. Scalabilities vary considerably based on the programming language of the OPC UA stack. The same implementations are also analyzed in [Mu20b] with respect to their features and interoperability. According to the results, they can work together seamlessly being deployed as client and servers.

Erba et al. study practical challenges to configure OPC UA securely [EMT21]. Their results show considerable security deficiencies in the implementations that inevitably lead to insecure configurations and deployments.

Roepert et al. assess the security of OPC UA deployments in [Ro20]. As part of their work, they develop a Metasploit module for assisting network-based security assessments of OPC UA deployments. This module is also used to test our setup.

3 System Overview

3.1 Hardware Configuration

The OPC UA Server at the charging entity is reachable from the internet (see section 1 and figure 1). Therefore, a secure configuration is crucial, according to Erba et al. [EMT21]. Additionally, the authors attest Siemens to fail in guiding users to a secure PLC configuration with the OPC UA manual [Si18]. Indeed, Siemens does not give explicit recommendation for a secure configuration there. As users might not be punctual or have time issues, this could lead to weak configurations. Accordingly, one tested configuration is the default one:

- Security Policy: Basic256, Basic256Sha256 (both 'sign' or 'sign and encrypt'), but also None!
- Accept any client certificate!
- Guest authentication!

Recently, Siemens updated manuals dedicated to secure communication. One chapter dedicates to the OPC UA Server configuration and applies to the PLC S7 1518 [Si21]. Accordingly, we employed a second configuration:

- allow only signed and encrypted communication, algorithm: Basic256Sha256
- do not allow guest-authentication
- do not allow access to PLC tags and Data Block (DB) components
- use a specific set or list of trusted clients (certificates) on the PLC

3.2 Metasploit for OPC UA

We wanted to test different cybersecurity configurations for the OPC UA Server described in the previous subsection 3.1. A recently published module for Metasploit dedicates to OPC UA [Ra21, Ro20]. Roepert et al. outline four steps for the assessment of security, which are [Ro20]:

1. Discovery: searching for potential targets, i.e. OPC UA Servers
2. Authentication: testing for different weak authentication methods, such as anonymous, easy-to-know credentials or default credentials. Check whether the server accepts self-signed client certificates.
3. Configuration Check: Deriving OPC UA Server information & configuration
4. Vulnerability Check: Checking for potential exploits such as known Common Vulnerabilities and Exposures (CVE) or Denial-of-Service (DoS).

During our tests, we only performed phases 1-3 (see figure 2). The vulnerability check is out of scope. For phase 4, Roepert et al. refer to existing Metasploit exploits, too [Ro20].

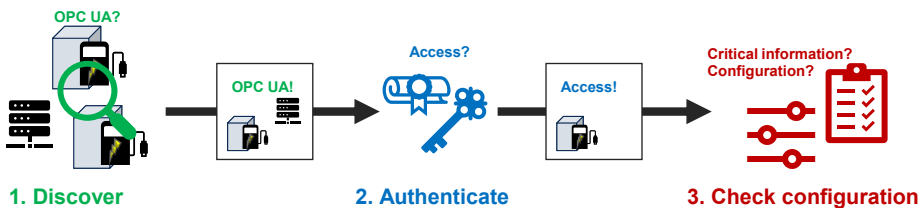


Fig. 2: Security assessment scheme for OPC UA Servers with three steps, according [Ro20]

4 Results

Table 1 depicts the test results for the different phases and configurations. As can be seen, the default configuration revealed information to potential attackers. A login was initially possible without providing credentials. Later on, during the configuration check, one could gather sensitive information. That includes read- and write-able nodes.

With the strict configuration according to the secure communication's manual [Si21], no login was possible. Consequently, no sensitive data could be derived from the OPC UA Server in this test.

5 Conclusion

In this research, we showed that a cyber-secure configuration for respective hardware is crucial. Differently than Erba et al. [EMT21] stated, there is secure guidance for Siemens PLC S7-1500 OPC UA Servers, but one has to look for and find it. That highlights the importance of dedicated policies and training for engineers configuring such hardware. The recently released Metasploit module for OPC UA provides a good option for configuration and penetration testing. Future research could focus on other implementations and vendors than the here considered one, too.

Tab. 1: Security assessment results for PLC S7-1518, with two different configurations

Phase	Configuration	
	Default	According to security manuals [Si21]
1. Discovery	Found server	Found server
2. Authentication	Login with empty credentials was possible	No login possible (no fitting credentials found and not possible with self-signed certificate)
3. Configuration	Revealed unencrypted communication (MessageSecurityMode: MessageSecurityMode.None_) Revealed all nodes, even writeable ones and those reserved for privileged users!	Login was not possible, hence no configuration could be derived

References

- [EMT21] Erba, Alessandro; Müller, Anne; Tippenhauer, Nils Ole: Practical Pitfalls for Security in OPC UA. arXiv preprint arXiv:2104.06051, 4 2021.
- [Mu20a] Muehlbauer, Nikolas; Kirdan, Erkin; Pahl, Marc-Oliver; Carle, Georg: Open-Source OPC UA Security and Scalability. In: 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA). pp. 262–269, 2020.
- [Mu20b] Muehlbauer, Nikolas; Kirdan, Erkin; Pahl, Marc-Oliver; Waedt, Karl: Feature-based Comparison of Open Source OPC-UA Implementations. GI-Jahrestagung, pp. 367–377, 2020.
- [Ra21] Metasploit-framework. <https://github.com/rapid7/metasploit-framework>, Accessed: 2021-06-12.
- [Ro20] Roepert, Linus; Dahlmanns, Markus; Fink, Ina Berenice; Pennekamp, Jan; Henze, Martin: Assessing the Security of OPC UA Deployments. In: Proceedings of the 1st ITG Workshop on IT Security (ITSec). 2020.
- [Si18] Siemens AG: OPC UA .NET Client for the SIMATIC S7-1500 OPC UA Server. Function manual, Siemens AG, 02 2018.
- [Si21] Siemens AG: S7-1500, ET 200MP, ET 200SP, ET 200AL, ET 200pro, ET 200eco PN Communication. Function manual, Siemens AG, 05 2021.
- [SWW19] Schindler, Josef; Watson, Venesa; Waedt, Karl: Interoperability of fast charging station with battery booster. GI-Jahrestagung (Workshops), pp. 295–307, 2019.

Supporting Security in Industrial Automation and Control Systems using Domain-Specific Modelling

Robert Altschaffel¹, Ivo Hempel¹, Oliver Keil¹, Josef Schindler², Martin Szemkus³, Jana Dittmann¹, Matthias Lange³, Karl Waedt² and Yongjian Ding³

Abstract: This paper explores how domain specific modelling can be used to support the identification of potential vulnerabilities and risks in Industrial Automation and Control Systems (IACS) to enhance security by enabling a mitigation of these vulnerabilities. This approach can be used to support already deployed IACS or to include Security-by-Design and Security Defence-in-Depth principles in the planning of future facilities. This paper explores the requirements for such a modelling approach including domain and security specific aspects. Three interlinked aspects of IACS which require different modelling approaches are identified leading to three distinct types of models: Infrastructure, cyber-process, and physical process. These three types are relevant for different attack vectors and to judge the potential impact of any attack. This paper shows examples for these three models and how these models can be used to identify vulnerabilities with the aim to close them.

Keywords: Industrial Automation and Control Systems, Security, Modelling, Standards

1 Introduction

Security is of increasing importance in the domain of Industrial Automation and Control Systems (IACS) as shown by the increasing number of recent attacks [AI20, AS19]. IACS are cyber-physical systems - computer systems which can directly affect the physical world by attached actuators. Hence, a threat to the security of an IACS often carries negative implications for the safety of the physical process or environments associated with the IACS. Hence, an increase of security in IACS is needed.

This paper aims to improve the security of already deployed or currently planned IACS by using domain specific modelling. This domain specific modelling allows to describe the IACS in question in a way that supports the identification of potential vulnerabilities

¹ Otto-von-Guericke-University, Department of Computer Science, Universitätsplatz 2, Magdeburg, 39106, robert.altschaffel@iti.cs.uni-magdeburg.de, ivo.hempel@ovgu.de, oliver.keil@ovgu.de, jana.dittmann@iti.cs.uni-magdeburg.de

² Framatome GmbH, ICETA-G Department, Paul-Gossen-Straße 100, 91052, Erlangen, Germany, josef.schindler@covalion.net, karl.waedt@framatome.de

³ Hochschule Magdeburg-Stendal, Magdeburg, 39114, martin.szemkus@h2.de, mathias.lange@h2.de, yongjian.ding@h2.de

and risks.

To achieve this, the domain specific modelling must consider all the factors which might affect the overall security of the IACS in question. Section 2 provides necessary background information as well as the requirement definition in Section 2.3. Based on these requirements, an approach for security-aware domain specific modelling is presented in Section 3. Section 4 closes with an outlook on further research.

2 Requirement Definition and Related Background

This chapter describes relevant terms and technologies in the scope of modelling IACS. Based on this knowledge, we define requirements that must be met by our modelling approaches proposed later in this paper.

2.1 Components of IACS

According to [Al20] an IACS is defined as "A communication network of Actors, Sensors and Processing units geared towards controlling a physical process". Therefore, these components are essential for IACS and relevant for modelling in this scope:

- **Sensor:** "Collects information about the environment [...]" [Al20]
- **Actuator:** "Manipulates the environment [...]" [Al20]
- **Processing Unit** (short **PU**; here **Programmable Logic Controller**): "Evaluates the data gathered by sensors and/or gives control signals to actors." [Al20]
- **Communication Wiring:** "The physical and logical carrier that facilitates communication between sensors, actors, and processing units." [Al20]

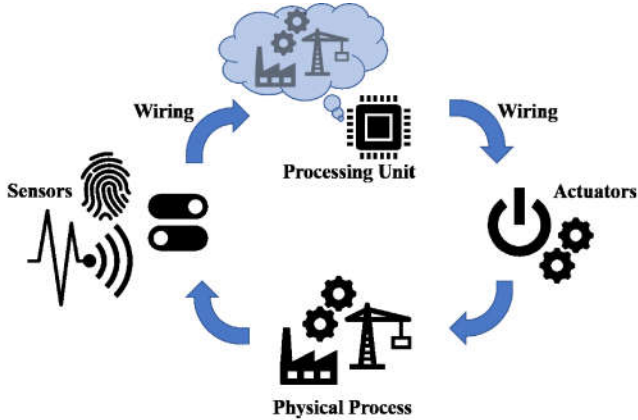


Fig. 1: Control loop with sensors, processing units and actuators in a physical process

Sensors measure a physical process while processing units compute how the physical process should be affected by the actuator implementing a control loop (see Fig. 1). This necessitates communication between sensors, processing units and actuators. This communication is handled by direct cable connections or more complex digital communication buses. Multiple control loops may be used to control complex physical processes. In addition to these control loops, other systems are usually included to provide an overview on the state of the entire physical process to an operator. Such a system is referred to as Supervisory Control And Data Acquisition (SCADA).

According to [Al20], the structure of these components follows a control hierarchy dictating the flow of communication within the network. A common and generalized way to describe such a control hierarchy is the Purdue Enterprise Reference Architecture (PERA; see [Wil92]). PERA consists of multiple levels of hierarchy. A brief overview on these levels can be found in [Ro11]:

- **Level 0:** Process sensors and actuators involved in the basic manufacturing process
- **Level 1:** Basic controllers (typically a PU) that direct and manipulate the manufacturing process
- **Level 2:** Area supervisory control applications and functions associated with the cell/area zone runtime supervision and operation
- **Level 3:** Site level plant-wide ICS functions

The higher levels provide business functions performed in an attached IT system. Although this summary references a manufacturing process, it is applicable to any other physical process as well. Within this structure, Level 0 components communicate with Level 1 components which in turn communicate with Level 2 components and so on. Hence, they form a control hierarchy which also acts as a communication hierarchy since

it dictates the flow of communication.

The hierarchy (and therefore the communication) might also be specified by further domain-specific restrictions. In this paper, IACS in the domain of Nuclear Power Plants (NPP) are taken as an example having high complexity and requirements for security. The IAEA provides an exemplary implementation for a Defence-in-Depth-Architecture (also Graded Approach, see [IA11]) based on different Security Levels (SL) and Security Zones which restrict the communication flow. The SLs are based on the impact a potential failure or compromise could have on the physical process. A summary of these SLs is provided in [Hi20]:

- **SL1:** systems vital to the facility (e.g. physical emergency protection)
- **SL2:** operational control systems which require high security
- **SL3:** supervision systems not required for operations
- **SL4:** technical data management systems (e.g. used for maintenance)
- **SL5:** business systems

While lower security levels should be able to send information to higher security levels, the information flow in the other direction should be highly restricted. Therefore, the levels are connected via access systems like firewalls or data diodes (see [Hi20]).

2.2 Attacks on IACS

Three different principal targets of typical attacks on IACS were identified in [St14]:

- communication stack of the deployed devices
- lack of authentication requirements within the hardware
- security problems in the respective software implementations

Various attacks on IACS were reviewed in [Al20] to identify forensic traces caused by these attacks. These traces can be present in Non-Volatile Memory (Mass storage), Volatile Memory (Main memory) or Communication (Network communication). All reviewed attacks identified potential traces in the communication. This is because communication between various devices is a necessary component of all remotely executed attacks on IACS. Potential exceptions are supply chain attacks in which case a component is manipulated by an attacker before it is installed within an IACS. If the component in question can directly affect the physical process, communication with different Security Levels or Zones is not necessary.

2.3 Requirement Definition

This section discusses how the components and their communication within an IACS translate to the requirements for a modelling approach in order to increase security.

The general aim of the modelling is to describe both the current and the target state of the system. This enables system planning to be carried out more efficiently. It facilitates the creation of test cases and discovery of potential attack vectors. Also, it simplifies the risk assessment of changes and weak spots of the system.

Based on the previous considerations, we define the following requirements for modelling:

- **Technical requirements** (general modelling techniques and tools):
 - Any present physical components, logical entities and possible communication channels must be depicted.
 - The visualization must enable fast and easy comprehension of the infrastructure with its static components and dynamic behaviour.
 - Model elements must be both standardizable and customizable.
 - The resulting model files must be of reasonable size in data.
- **Subject-specific requirements** (dependent on the domain and system):
 - The PERA [Wil92] levels must be depicted.
 - The security zones and levels from the “Graded Approach” [IA11] must be depicted.

3 Modelling Aspects and Approaches

Covering security and its implications in an industrial facility requires various aspects to be explored and modelled. This chapter discusses these aspects and explores means to model them.

3.1 Modelling Aspects

To provide a complete view of the industrial facility, we define the following three aspects which must be covered by the models:

- **Static infrastructure:** These models visualize the physical components contained in the infrastructure together with their connections between each other. They include the control technology for the physical process (sensors, actuators and control devices) and network devices (e.g. switches and firewalls).

- **Physical process:** These models visualize the sequence of the physical process implemented by the control technology. They show the individual steps that happen within the process which also includes switching between physical states. This modelling aspect describes the dynamic behaviour of the system.
- **Control loop and its technology:** These models visualize the components which supervise and control the physical process. This includes sensors, actuators and their controllers as well as the communication between these devices. Hence, these models describe the IACS. This modelling aspect emphasizes the signals and information that is being exchanged between the devices as well as their consequences. It enables a more detailed reflection of attack vectors for the industrial facility. Thus, it also describes the dynamic behaviour.

Together, these three aspects allow modelling of the infrastructure and behaviour for the entire industrial facility. The distinction between physical process and control circuit is made due to the following reasons:

The physical process focuses on the sequences of physical states that are based on logical conditions and resulting actions.

The modelling of control circuits and technology then describes the implementation of these logical conditions and resulting actions with actual devices and their actual communication between each other.

This distinction allows a more fine-granular view on the infrastructure with its components and processes depending on the needs of the administrator or researcher.

However, all these models are closely interlinked with each other. The model of the control circuit includes communication with other systems. Hence, remote attacks would have to move through the modelled IACS to have an impact. In this case, the model of the physical process is necessary to understand the potential consequences of a vulnerability within the control loop. This connection allows for the identification of a safety impact due to a security problem. The modelling of the infrastructure covers supply chain attacks. Again, the connection to the control loop model is necessary to understand how such an attack could propagate within the network. The connection to the model of the physical process is again necessary to consider potential impacts on the physical process.

3.2 Modelling Approaches

This section describes two possible approaches: Manual and automated modelling.

Manual Modelling

For the process of manual modelling, we chose the tool “draw.io” (available at [Dr21]). This tool is a JavaScript based open-source software for creating diagrams. It is easy to

use, the created files are usually smaller than one megabyte and the creation of custom symbols is possible. The models are highly reusable and adaptable. Also, linking multiple models is possible. This makes the tool a practical solution for the modelling process.

The exemplary models in this paper are based on the Integrated Nuclear Evaluation System 7 (INES-7) of the Research Group of Multimedia and Security at OvGU. It is a demonstrator for the physical and control processes of power plants. In this case, the reactor is represented by a radio-controlled water heater.

One aspect of modelling is the static infrastructure. The modelling of a plant takes place in separated plant areas. That means different factory buildings are depicted in different sections of the model. Actuators and sensors control the physical process. Hence, they belong to PERA level 0 in the model. All physical and logical controllers (e.g. PLCs) are denoted as entities. Each entity gets assigned to a zone, a PERA level and a security level. Each entity has symbols assigned to it as well as communication channels drawn for it. Also, they are classified as bidirectional, unidirectional receiving or unidirectional sending. The modelling of the static infrastructure is separated into two zones: Water heater and phone (see Fig. 2). The heating control circuit consists of a sensor for the water temperature, a sensor for the water level and an actuator for the heating element. The sensors use an ellipse as the symbol, the actuator uses an adapted valve symbol.

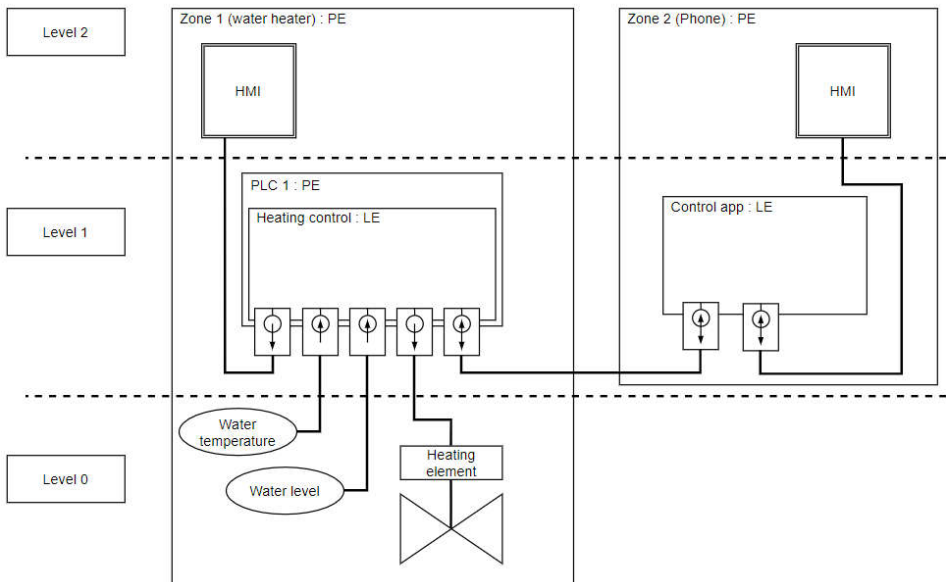


Fig. 2: Modelling of the static infrastructure of a Wi-Fi water heater

For the modelling of the physical process, a flow chart or a more complex Unified Modelling Language (UML) diagram for the process steps is sufficient. A simple flow

chart (see Fig. 3) is sufficient to show the process of heating up water.

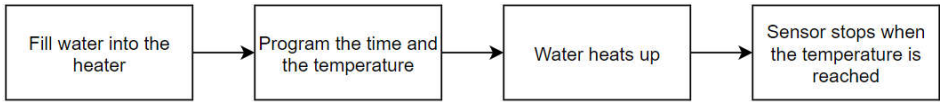


Fig. 3: Modelling of the physical process of heating water with a flow chart

In the modelling of the control loop and its technology, we differentiate between static and dynamic behaviour. In the static case, we depict each PLC individually. The model includes these elements:

- **Sensors:** Gathers information about the environment.
- **Calculation:** Transforms the sensor input into usable data representation.
- **Distribution:** The result of the calculation is distributed on the wiring.
- **Aggregation:** The data is aggregated and processed using a defined logic.
- **Logic:** Definition for the behaviour of the actuators.
- **Actuators:** Affects the environment based on the logic.

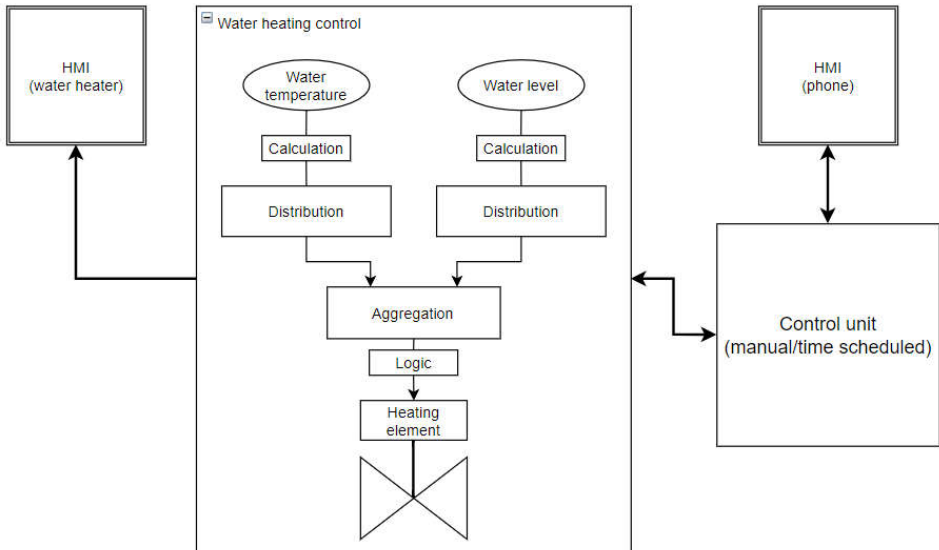


Fig. 4: Generic modelling of the control loop and its technology of a Wi-Fi water heater

The sensors and actuators are connected to the PLC with an indicator for the unidirectional communication. The physical entity of the PLC consists of the logical entity of the heating control. The PLC communicates with the second zone in a

bidirectional way. A logical entity in form of an app controls the process parameters and displays sensor values. A bidirectional channel exists for the Human Machine Interface (HMI) on the phone. Each entity is aligned according to its PERA level. The security levels could be colour-coded with different background of the entities.

Fig. 4 shows the static behaviour of the control loop and its technology. Sensors and actuators are put into a control loop with these steps: Calculation, distribution, aggregation and logic. A control unit is used as the logical entity for defining the process parameters if necessary. Additionally, the HMIs are depicted. With this simple, but meaningful example the weak spots for potential threats of a system can be assessed fast and they can be easily visualized.

Automated Modelling

Besides manual modelling, we also want to point out the possible use of automation in modelling. Creating models automatically could come with several advantages over manual modelling:

- **Less time cost for human resources:** Manual modelling of infrastructures requires an employee to invest time to create the model. Automatic modelling would only require the employee to start the process. Then, the automated modelling tool creates models based on its configuration. A final manual check might be required but would consume far less time than going through the entire process of manual modelling.
- **Gather additional information:** An automated modelling process may be able to gather new information that would have been undetected when using manual modelling with reasonable time investment. For instance, an administrator knows that he uses a specific network protocol for any communication within all applications on a device. To verify that this is the only protocol on the wire to and from this device, he looks at the network traffic for a certain amount of time with analysis tools like Wireshark [Wi21]. However, the operating system may communicate with other devices via more protocols which the administrator does not yet know about. In this case, an automated modelling tool could gather more information as it is a continuous process of gathering and modelling information.
- **Overcoming human failure factors:** Modelling a communication setup could include scanning and assessing all ports in use on a specific device. When having long lists of information (here ports), manual assessment may lead to lapse (defined in [An18]) due to human error. This can result in security issues along the modelling, development & engineering pipeline. Automation provides higher reliability and overcomes these potential failures.

4 Outlook

This paper discussed how domain specific modelling can support security in IACS by enabling the identification of vulnerabilities. To achieve this, the modelling has to fulfil domain specific requirements including the ability to model domain specific security measures. Various possible attack vectors and the complexity of IACS and the physical processes they control lead to the establishment of three types of models which are inter-linked. Some basic methods to conduct such a modelling are proposed - including the establishment of specific sets of modelling symbols to describe domain specific elements.

The Research Group of Multimedia and Security at the Otto-von-Guericke-University (OvGU) of Magdeburg is currently working on two publications based on this paper. First, an automated approach on modelling will be created and evaluated. In section 3.2, the Research Group Multimedia and Security at OvGU Magdeburg has already introduced the possibility for automation. Second, the Research Group of Multimedia and Security at OvGU introduced the interlinked approach with its aspects in section 3.1 in this paper and plans to create an interlinked concept for models which combines the three aspects with regards to popular forensic models like data streams, data types and incident ontologies (see [BS11]).

Acknowledgments

The work on this paper was partly funded by the German Federal Ministry for Economic Affairs and Energy within the project SMARTTEST2 (Grant No.: 1501600B).

Bibliography

- [AG21] AGCS: Managing the impact of increasing interconnectivity: Trends in cyber risk. Report, Allianz Global Corporate & Specialty, 03 2021.
- [Al20] Altschaffel, R.: Computer forensics in cyber-physical systems: applying existing forensic knowledge and procedures from classical IT to automation and automotive, <https://opendata.uni-halle.de/handle/1981185920/35574>, accessed: 17/06/2021, 2020.
- [An18] Anu, V.; Hu, W.; Carver, J.; Walia, G.; Bradshaw, G.: Development of a human error taxonomy for software requirements: A systematic literature review. *Information and Software Technology*, 103:112–124, 2018.
- [AS19] A., Saravanan; S., Bama S.: A Review on Cyber Security and the Fifth Generation Cyberattacks. *Oriental Journal of Computer Science and Technology*, 12:50–56, 2019
- [BS11] BSI: Leitfaden „IT-Forensik“. Technical guideline, Bundesamt für Sicherheit in der Informationstechnik (BSI), Bonn (DE), 03 2011.

- [Dr21] draw.io. Online: <https://github.com/jgraph/drawio> (Last access: 2021/06/18)
- [Hi20] Hildebrandt, M.; Altschaffel, R.; Lamshoft, K.; Lange, M.; Szemkus, M.; Neubert, T.; Vielhauer, C.; Ding, Y.; Dittmann, J.: Threat Analysis of Steganographic and Covert Communication in Nuclear I&C Systems. In: In Third International Conference on Nuclear Security: Sustaining and Strengthening Efforts (ICONS 2020). 2020.
- [IA11] IAEA: Computer Security at Nuclear Facilities. Technical guidance: Reference manual, International Atomic Energy Agency (IAEA), Vienna (AT), 12 2011
- [Ro11] Rockwell Automation: Converged Plantwide Ethernet (CPwE) Design and Implementation Guide. Design and implementation guide, Rockwell Automation, 09 2011.
- [St14] Stirland, J.; Jones, K.; Janicke, H.; Wu, T.: Developing Cyber Forensics for SCADA Industrial Control Systems. In: The International Conference on Information Security and Cyber Forensics (InfoSec). pp. 98–111, 2014.
- [Wi21] Wireshark. <https://www.wireshark.org/about.html>, Online; Accessed: 2021-06-18.
- [Wi92] Williams, T. J.: The Purdue enterprise reference architecture: a technical guide for CIM planning and implementation. 1992

Security aspects of FPGA and virtualization case studies

Asmaa Tellabi¹, Abdelbast Sabri², Christoph Ruland³, Karl Waedt⁴

Abstract: Virtualization technology is a technology that has been integrated a lot lately, thanks to its interoperability and enhanced performances. An essential element in virtualization is the hypervisor that is responsible of managing virtual machines; it allows various guest operating systems (OS) to run on one hardware entity simultaneously. Many researchers recognized the challenge of allowing multiple critical applications to share one hardware platform without interfering, and ensuring security of transactions a major safety and security challenge in virtualization. Usually, security by design is also comprises the protection of intellectual property (IP), possible safety related damages and extend financial losses. By the extension of the usage of programmable logic outside commercial markets to avionics and military applications, security by design integrates extra aspects to ensure safety and cybersecurity. Solutions for protecting application data during transmission and storage exist, but solutions for protecting Field Programmable Gate Array (FPGA) configuration data are not well-known. In the context of Industry 4.0, the FPGA hardware solutions provide the needed level of flexibility and performance. The flexibility, portability and even, to a wide extent, technology independence are due to the software based definition of the hardware via Hardware Description Languages (HDL).

In this paper, threats and vulnerabilities surrounding FPGAs will be addressed. An example of a type 1 hypervisor called XtratuM will be ported on top of a Xilinx Soc 7000 zc706 will be integrated, as well as a prototype of embedded system using PetaLinux will be provided.

Keywords: Virtualization, hypervisor XtratuM, embedded system using PetaLinux, FPGA Xilinx Soc 7000 zc706.

1 Introduction

Currently, the global marketplace has presented multiple novel opportunities but also emerging threats. Corporations and governments must cope with these threats that vary from counterfeiting till espionage. Consequences of such threats can go far beyond financial losses as they can also harm personal safety [JP09]. Enterprises deploy various layers of protection to ensure security; they contain firewalls, and authentication/encryption, security protocols and intrusion detection/intrusion prevention systems [IC19]. Cybersecurity has been a critical subject for many companies for over 25

¹ University of Siegen, Faculty of Science and Engineering, Chair for Data Communication Systems, Hölderlinstraße 3, Siegen, 57068, asmaa.tellabi@student.uni-siegen.de

² University Friedrich-Alexander-University, Martensstraße 5a, 91058 , Erlangen, abdelbast.sabri@fau.de

³ University of Siegen, Faculty of Science and Engineering, Chair for Data Communication Systems, Hölderlinstraße 3, Siegen, 57068, christoph.ruland@student.uni-siegen.de

⁴ Framatome GmbH, Henri-Dunant-Str. 50, 91058, Erlangen, karl.waedt@framatome.com

years; although it is only recently that it caught engineers' attention working on embedded systems [IC19].

Easily reconfigurable architectures such as FPGAs are offering multiple remarkable features that can be integrated in embedded systems in case security is a vital matter [Ge12]. FPGAs from 2007's came out with multiple functional features than the ones created in the mid-1980s. These new FPGAs integrate embedded processors, Giga-bit serial transceivers, clock managers, analogue to digital converters, digital signal processing blocks, Ethernet controllers, large memory capacity. Such evolution in the capacity and application levels of FPGAs has two main security consequences. To start with, current FPGA designs display a major development investment that must be protected. In addition, FPGAs are progressively more being deployed in applications that necessitate security features that are does not existing yet, or that must be sufficiently inspected. Together, they have shed the light on security attributes related to FPGAs in the military, automotive, consumer industries, and also within the research community, each with its specific requirements and security perceptions [ST12].

2 Assessment of Threats and Vulnerabilities surrounding FPGA

FPGAs are based on programmability and a set of identical logic blocks in order to produce a flexible computing entity, which is can lower design costs, decrease system's complexity, and reduce the time to market, by means of parallelism and hardware acceleration to reach better performances. FPGAs rising popularity has pushed integrators to integrate security controls in the design's level, but the way resources in embedded systems are allocated forms challenges in offering a secure entity [EE19]. FPGAs are programmed using a bitstream, which is a binary data that is passed to the FPGA by means of precise I/O ports on the device. They specify the way internal resources are going to be used for executing logic operations [ST12].

2.1 The Threats

There is a wide range of threats surrounding the design and each one has its particular implications. Some threats aim to gain financial benefits of a company, while other might aim to harm individuals or even national security [ST12]. Industries are slowly gravitating towards FPGAs thanks to their programmability's benefits and proof against obsolescence. FPGAs are less susceptible to reverse engineering attacks but to another different kind of threats [Zh18].

- **Reverse Engineering**

Using an already existing product in the market, attackers analyze the design by observing the layout, the used devices, downloading the firmware, and examining transactions

between devices. By exploiting this information, attackers might rebuild the design, so they can create new products or help their future products development [ST12].

- **Cloning**

For this type of threats, attackers do not try to fully comprehend and analyze the design. Their goal is to produce replicas of a current product, and realize a better profit by selling it since they have not went through the same the time and expense of product development and marketing processes [SD08]. As attackers are less likely to use a better quality modules and quality assurance components, the developed replicas can harm the company's reputation and also its finances [ST12]. Serious impacts happen when cloning is used to create replicas products dedicated to high availability applications, e.g., in aviation which can negatively impact flight's safety [Zh18].

- **Overbuilding**

One of the easiest and simplest types of design's theft is overbuilding. With the growing trend of outsourcing, usually original equipment manufacturers (OEM) count on off-shore suppliers to create its products [SD08]. As a result, a dishonest subcontractor might create some additional components more than those ordered by the OEM. Even though this can be seen as a type of counterfeiting, these extra components that are produced are similar to the originals, which make the detection challenging [Zh18].

- **Tampering**

This refers to an outside attacker trying to gain illegal access to an electronic device. Tampering might be either a part of a reverse engineering attack, or it can aim to realize malicious or criminal acts [ST12]. For example, an attacker can attempt to mine operating data or firmware, or can attempt to alter the firmware in the system so they can put the system in jeopardy [SD08].

- **JTAG Threats**

A different kind of threats exploits the JTAG port. Even if the port is used for board connectivity and to test system's functionalities, these ports can be exploited to conclude the FPGAs configuration [ST12]. JTAG is always set up in most FPGAs, it can be used to apply the design methodically so it makes reverse engineer attacks easier, thus stealing the design's information [Zh18]. On the other hand, similar to timing analysis on an FPGA, this type of attacks consumes a lot of time and requires specific skills but it does not require dedicated equipment [SD08].

2.2 Weaknesses

- **Complacency**

This occurs in the design teams and companies' side. Some companies forget about integrating security in the design because it requires more time or to the certainty that legal

protection is enough [Ge12]. Therefore, integration of security in the design is not advised, which make companies consider only few steps to protect its valuable intellectual property [ST12]. This type of weakness can be mitigated by proper training, education and the acknowledgment that even if a legal protection exists it does not ensure entire protection [Zh18].

- **Inadequate Security Controls**

This represents another category of vulnerability at the FPGA level [EE19]. For instance, a company can integrate an anti-tamper detection device in a system including an FPGA to warn end users of tampering attacks, but in case these security controls do not include encryption of FPGA bitstreams the system will still be exposed to reverse engineering attacks [Zh18]. Furthermore, security controls must not be included just at the device or FPGA level, but also at the board and system levels, taking into consideration possible threats at each level. Obviously, designs need to be carefully reviewed to guarantee that all security characteristics are examined [ST12].

- **Back Doors**

Some extra configurations applied in a design to help with the debug might create security holes that can be exploited by attackers. From analogous to software systems, some back doors are left to facilitate the access to system administrators, but the hardware design can have some security holes as well [Ge12][ST12]. For instance, in case device debugs modes/cores they can be exploited in order to go beyond security controls and anti-tampering measures. Although these back doors can be helpful in the period of the design, they have to be cut out from the design before final assembly.

- **Design Flaws**

These flaws can create exploitable security holes, e.g., untested cases that are present in the design can make it vulnerable [XW19]. Consequently, careful attention must be given to testing before delivering the final product [ST12].

- **Device Flaws**

Same as design flaws, a device can have manufacturing flaws that make it susceptible to an attack. Choosing vendors with in-depth testing structures and unconventional quality assurance techniques can significantly reduce this risk [XA19].

- **Single-Event Upsets**

Single-event upsets (SEU) arise when device memory structures have their states modified because of the impact of high-energy neutrons. An SEU is able of modifying the functionality of a device and compromising security controls included inside the device [ST12]. Applying SEU mitigation methods not only can neutralize this risk but it can enhance system reliability.

Virtualization use case study

Xtratum is a bare-metal or type 1 hypervisor, it was conceived especially to adhere to temporal and spatial requirements for safety critical applications [FE18]. Xtratum offers virtualization extensions to partitions; it runs in supervisor processor mode and virtualizes the CPU, memory, interrupts, and some defined peripherals [A115]. The first version was intended for x86 Architectures (version 2.0), then it has been completely transformed to be compatible with SPARC v8 architecture, to be exact for LEON2 processor. Present versions mainly include the necessary functionalities required to create safety Critical systems based on ARINC 653, AUTOSTAR and other standards [As18].

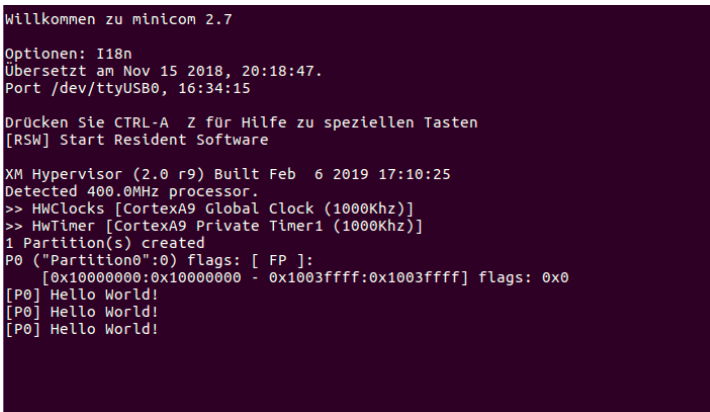


Fig. 1: Execution Results of the 'Hello World' Partition.

```

<?xml version="1.0"?>
- <SystemDescription name="hello_world" version="1.0.0" xmlns="http://www.xtratum.org/xm-arm-2.x">
  - <HwDescription>
    - <MemoryLayout>
      <Region size="1MB" start="0x0" type="rom"/>
      <Region size="1023MB" start="0x00100000" type="sdram"/>
    </MemoryLayout>
    - <ProcessorTable>
      - <Processor frequency="400Mhz" id="0">
        - <CyclicPlanTable>
          - <Plan id="0" majorFrame="200ms">
            <Slot start="0ms" id="0" partitionId="0" duration="200ms"/>
          </Plan>
        </CyclicPlanTable>
      </Processor>
    </ProcessorTable>
    - <Devices>
      <Uart name="Uart" id="1" baudRate="115200"/>
    </Devices>
  </HwDescription>
  - <XMHypervisor console="Uart">
    <PhysicalMemoryArea size="1MB"/>
  </XMHypervisor>
  - <PartitionTable>
    - <Partition name="Partition0" id="0" console="Uart" flags="boot fp">
      - <PhysicalMemoryAreas>
        <Area size="256KB" start="0x10000000"/>
      </PhysicalMemoryAreas>
    </Partition>
  </PartitionTable>
</SystemDescription>
    
```

Fig. 2: XML file of the 'Hello World' Partition.

In this paper, the version used of Xtratum is the one compatible with Arm processors [ARM18]. The ZC 706 board includes 2 Cortex A9 processors, in this example only one processor will be used. Xtratum has a default configuration made for this board specifically [As18a]. In this implementation, a hello world application is implemented. To be able to run application on Xtratum, an XML file has to be first configured, this file include all physical resources, and a scheduling plan for partitions [FE18]. To compile applications on the board, the command make is used. Figure 1 shows the results of the running application on the board.

2.3 PetaLinux

PetaLinux Tools offer the necessary tools used to customize, build and deploy Embedded Linux solutions on Xilinx processing systems. Designed to speed up design productivity, these tools are used together with the Xilinx hardware design tools to facilitate the development of Linux systems for Zynq-7000, MicroBlaze and PowerPC [XW19]. PetaLinux can be used by developers to configure, build and deploy the required open source and systems software to Xilinx, including [XA19]:

First Stage Boot Loader (FSBL) : is responsible of loading partitions (software programs, the bitstream) in the image, that is stored in Non-Volatile Memory (NVM) , to the destination partitions. Usually, the destination can be Double Data Rate (DDR), On-Chip Memory (OCM), or Advanced eXtensible Interface battery-backed RAM (AXI BBRAM). The destination of the bitstream is the PL configuration memory

U-BOOT: which is an open source software that is executed on Zynq devices. It is generally used to load Linux. Other U-Boot functions contain reading DDR memory, erasing, reading and writing into NVM.

Linux kernel: is the core of the Linux OS, it is responsible of controlling operations in the system. Kernel.org together with Xilinx additions (BSP and drivers) provides the Linux kernel for Xilinx Zynq platforms. Linux kernel is the main component of the OS that connects the hardware to applications.

The device tree: is basically a data structure in byte code that contains helpful information used by the kernel when booting up. The boot loader copies data into a fixed address in the RAM before jumping to the kernel's entry point.

Root filesystem (Rootfs): comprises various specific configuration files related to systems. Some examples contain a kernel that is created for a system, a specific hostname, etc. Their contents have to be suitable to boot, restore, recover, and/or repair the system. To boot a system, a certain amount of software and data must be included on the root partition to mount other filesystems. This consists of utilities, configuration, boot loader information, and other necessary start-up data.

First, the FSBL is executed followed by the execution of U-Boot, which loads the device tree and Linux kernel into the memory and then mounting rootfs are executed. After the hardware bit stream and software images have been built, the new PetaLinux platform with the Zynq kernel can be booted via an SD card. Figure 1 shows the design flow on Vivado and PetaLinux.

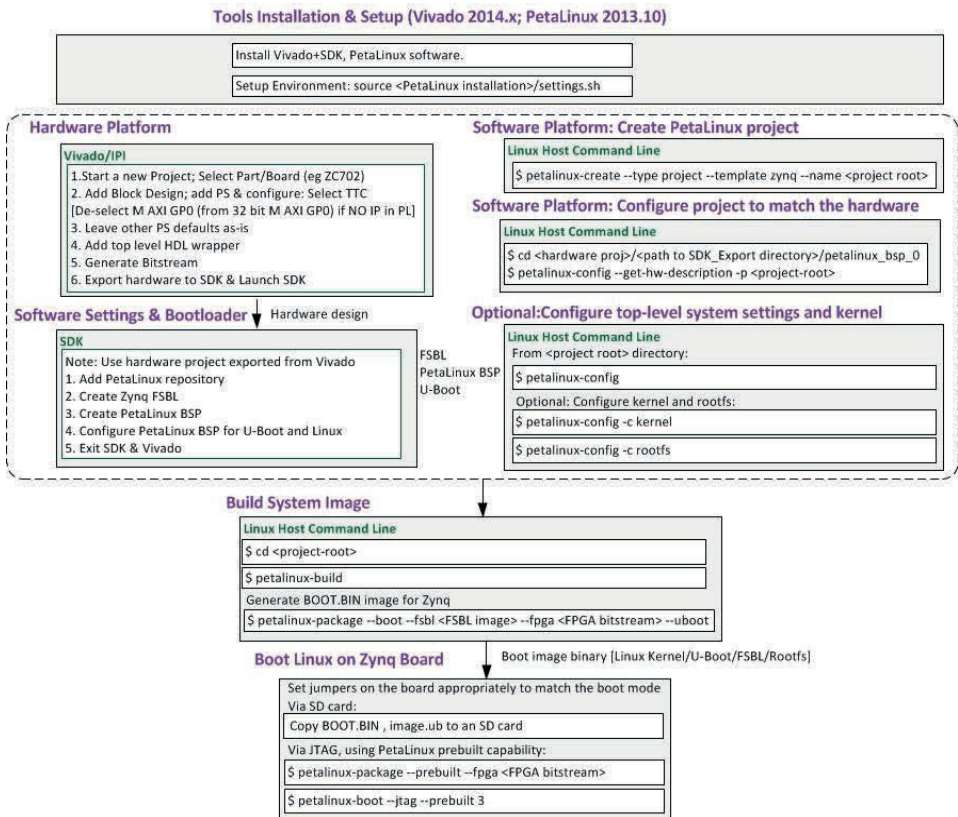


Fig. 3: Design flow using Vivado (v2014.x) and PetaLinux (v2013.10)[Ja19].

The Xilinx Vivado Software Development Kit (SDK), PetaLinux tools and Board Support Package (BSP) is a set of customized drivers to the selected hardware; they must be installed before any design development. PetaLinux and Vivado SDK must be of the same version, e.g. if Vivado and SDK 2018.2 are installed, the PetaLinux version must be 2018.2. In this example booting Linux on Zynq can be done using JTAG or SD card [XA19][XW19].

The `petalinux-create` command creates partitions that are part of a PetaLinux project, a partition can be either a new project, a component or an application in the same project [PC19]. Components or applications can then be either integrated or excluded from the final system using the `petalinux-config -c rootfs` command. Dropbear is a similar tool to OpenSSH. To use `openssh-server`, dropbear must be set up in the PetaLinux Root File System before applying the `petalinux-build` using the following command: `petalinux-config -c rootfs`. This will configure the project to use SSH between the board and other external systems [XW19].

In this paper, 2 applications were developed and tested in PetaLinux.

The first one is a Hello World application that was created using Petalinux 2018.02. PetaLinux provides a tool to create user application templates for either C or C++ [PC19]. These templates include applications' source code and Makefiles to facilitate the configuration and compilation of applications in the target, and then installing them into the root file system. The new application can be found in the `<project-root>/components/apps/myapp` directory.

A terminal emulator program should be configured to connect to the USB serial port at 115,200 baud. The emulator `minicom` is used in this example to connect to the serial port. The second one is related to OpenSSH. With SSH into the board can files be edited and compiled on the desktop, and then executables copied over Ethernet. When the `ifconfig` is executed in the emulator, first the `inet addr` shows which represents the ip-address of the board, in this case 192.168.0.10. SSH can be activated by executing the following command: `ssh root@192.168.1.69` in a new terminal.

3 Conclusion

Virtualization technologies introduce new threats, risks, challenges, new assets and components to the digital environment. Consequently, novel security controls and measures are necessary to offer a trustworthy system for such environments. Multiple techniques and applications for different types of virtualization exist; but they all have a same objective which is to erase the limitations present in the physical hardware by building secure partitions in the hardware, which make it operates like various virtual environments. This paper listed most of the common vulnerabilities and weaknesses found in FPGAs. An example of a Type 1 hypervisor running on top of the board is presented as

well. An illustration of embedded systems procreation using PetaLinux on a Xilinx ZC 706 board is shown. In the future, more hypervisors such as Xen are going to be tested on the board and other boards with different architectures like x86.

Bibliography

- [JP09] Juwayriyah, H; Paul, Q.: Protecting the FPGA Design from Common Threats; 2009.
- [ST12] Steven, M: Solving Today's Design Security Concerns; 2012.
- [Ge12] Gedare, B.; Eugen, L.: Bhagirath, N.: Rahul, S.: Hardware and Security: Vulnerabilities and Solutions. In (Elsevier Inc.): Handbook on Securing Cyber-Physical Critical Infrastructure, New York 1999. Noah & Sons, San Francisco, pp. 305-331, 2012.
- [IC19] Icon Labs: Security Requirements for Embedded Devices – What is Really Needed?, www.iconlabs.com, accessed: 21/02/2019.
- [EE19] EETimes: New Advances in FPGA Security, https://www.eetimes.com/document.asp?doc_id=1327477#, accessed: 04/05/2019.
- [Zh18] Zhiming, Z.; Laurent, N.: Charles, A.K.: Qiaoyan, Y.: Thwarting Security Threats From Malicious FPGA Tools With Novel FPGA-Oriented Moving Target Defense, 2018.
- [SD08] Saar, D.: Solving Today's Design Security Concerns; 2008.
- [XW19] Xilinx Wiki, <https://xilinx-wiki.atlassian.net/wiki/spaces/A/pages/18842250/PetaLinux>, accessed: 30/04/2019.
- [XA19] Lester Sanders: XAPP1175, Secure Boot of Zynq-7000 All ProgrammableSoC, 2015, https://www.xilinx.com/support/documentation/application_notes/xapp1175_zynq_secure_boot.pdf, accessed: 30/04/2019.
- [FE18] Fentiss: Fent Innovative Software Solutions, www.fentiss.com/xtratum, accessed: 20/08/2018.
- [Al15] A. Alonso; et.al.; Safety Concept for a Mixed Criticality On-Board Software System .IFAC-PapersOnLine, Volume 48, Issue 10, Pages 240-245, ISSN 2405-8963, 2015.
- [As18] Asmaa, T.; Ines, B.Z: Christoph, R.: Karl, W.: Virtualization on Secure Platforms for Industrial Applications Current use cases and future perspectives, 2018.
- [ARM18] ARM: ARM Security Technology - Building a Secure System using TrustZone Technology. Tech. Rep., 2009.
- [As18a] Asmaa, T.; Ludger, P.: Christoph, R.: Karl, W.: Security Aspects of Hardware Virtualization Technologies for Industrial Automation and Control Systems, 2018.
- [Ja19] Janani Subraveti: Implementation of UDP communication on a ZYNQ platform for processing clustered data based on multiple Gigabit Ethernet ports for phenoPET, 2017, https://user.fz-juelich.de/record/844070/files/J%C3%BCI_4406.pdf, accessed: 30/04/2019.

- [PC19] PetaLinux Command Line Reference UG1157, https://www.xilinx.com/support/documentation/sw_manuals/xilinx2018_3/ug1157-petalinux-tools-command-line-guide.pdf, accessed: 30/04/2019.

Secure Unidirectional Security Gateways for Industrie 4.0

Christele Larissa Moussi Djeukoua,¹ Timothee Guiraud,² Edita Bajramovic,³ Josef Schindler,⁴ Karl Waedt⁵

Abstract: Secure data exchange between different geographical sites (e.g. industrial manufacturing facilities, power plants, research labs, and manufacturer test facilities) is an important part of cybersecurity. It is for example addressed by section 13 “Communications Security” of ISO/IEC 27002:2013. ISO/IEC 27033-4:2014 gives guidance for securing communications between networks using security gateways (firewall, application firewall, Intrusion Protection System, etc.). While the newest part of the multipart ISO/IEC 27033 standards series, ISO/IEC WD 27033-7 “Information technology Network security Part 7: Guidelines for network virtualization security” is more on the abstraction of physical components involved in communication, this paper aims to emphasize the fiber optical network communication-based security between very specific Cyber-Physical Systems (CPS). In this case, secure means a strictly reduced feedback channel. According to ISO/IEC 27033-4, essentially, the approaches described in this paper would be classified as physically unidirectional security gateways. In this paper, requirements and concepts will be presented which are beneficial for a scalable use in Industry 4.0 applications for highest (hardware-enforced) unidirectional communication and which can coexist with Firewall and Demilitarized Zone (DMZ) approaches that are also needed for complex IACS architectures.

Keywords: Security Gateway; Fiber Optical Network; Secure Communication; Forensic Readiness; Industrial Automation; Control System

1 Introduction

Mentioned for the first time in 2011 at the Hanover Fair, the term Industrie 4.0 (fourth industrial revolution) was created by the German government to promote connected production and digital convergence between industry, businesses, and other processes. It is a collective term for the technologies and concepts of a value chain organization that creates together Cyber-Physical Systems (CPS) and the Internet of Things (IoT) [HPO16, IE20a].

¹ Friedrich-Alexander-University (FAU) Erlangen-Nuremberg, IT Security Infrastructures Lab, Martensstrasse 3, 91058, Erlangen, Germany larissadjeukoua@yahoo.fr

² Framatome GmbH, ICETA-G Department, Paul-Gossen-Straße 100, 91052, Erlangen Germany timothee.guiraud@protonmail.com

³ University of Erlangen-Nürnberg, IT Security Infrastructures Lab, Martensstrasse 3, 91058 Erlangen, Germany edita.bajramovic@fau.de

⁴ Friedrich-Alexander-University (FAU) Erlangen-Nuremberg, Chair of Electrical Energy Systems (EES), Cauerstraße 4, 91058, Erlangen, Germany josef.s.schindler@fau.de

⁵ Framatome GmbH, ICETA-G Department, Paul-Gossen-Straße 100, 91052, Erlangen, Germany karl.waedt@framatome.com

IoT belongs to one of the most dynamic and exciting areas of ICT (Information and Communication Technologies). Aiming to connect physical entities to computer systems through networks, IoT uses electronic components that interact with the physical world as building blocks. It bases on sensors and actuators of any type, such as thermometers, accelerometers, video cameras, microphones, relays, or industrial equipment for manufacturing or process-controlling. While actuators will act on physical entities, sensors will collect information about the physical world. For controlling physical entities, provision of contextual, real-time, and predictive information is necessary that impacts both physical and virtual entities. Therefore, mobile technology, cloud computing, big data, and deep analytics (predictive, cognitive, real-time, and contextual) play an important role in collecting and processing data to achieve business objectives [IE18, IE20b].

2 Problem statement

The digitalization of manufacturing introduces danger to the integration of IT and critical infrastructures, which could affect the industrial manufacturing process (malware, spyware, loss of data integrity, or problems with the availability of information). With increased interconnectivity and potential impact on functional or nuclear safety, Cyber-physical Systems (CPS) and related Industrial IoT (IIoT) gradually become more sensitive targets of cybersecurity attacks [TSS17]. To effectively detect such attacks with minimum delay, forensic readiness and means for fast evaluation of related events are mandatory. Forensic readiness intends to maximize an environment's ability to collect credible digital evidence on the one side and to minimize the cost of forensics in incident response on the other [BSJ10].

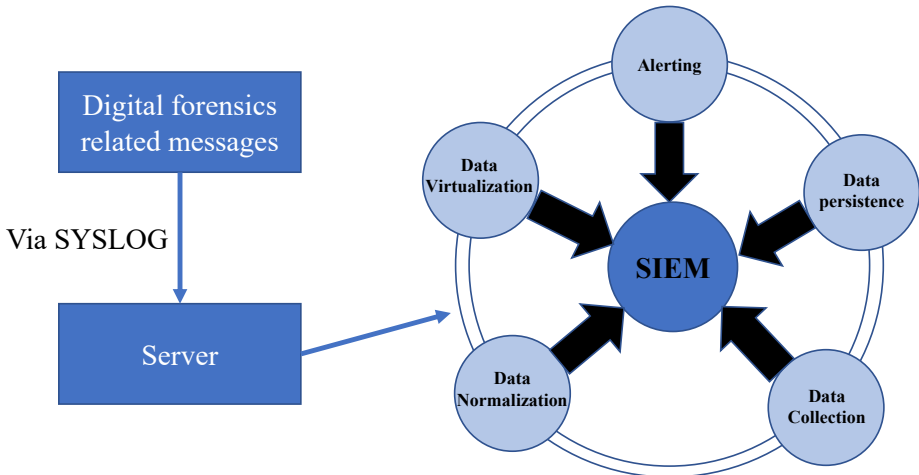


Fig. 1: Current IT-solutions for detecting cybersecurity attacks

In the global economy, there is a new kind of software criminality that can cause data availability and reliability problems for companies. Current commercial solutions are not suited for jointly meeting the specific safety and cybersecurity requirements [ET20]. As a common practice in the IT domain, digital forensics readiness-related messages can be sent via SYSLOG towards a server that processes them into a Security Information and Event Management (SIEM) system. This idea is presented in figure 1 [FS07]. This paper discusses how to design and deploy secure unidirectional gateways to minimize potential threats and unexpected situations. In this context, the following Research Questions (RQ) arise:

- What is important from the security side according to the requirements of the implementation of Industry 4.0 by using unidirectional secure gateways?
- What high-reliability real-time behavior can be assured despite unidirectional communication?

This paper aims to present a solution to these research questions.

3 Secure unidirectional gateways: data diode

The International Electrotechnical Commission proposes new standards for power plants and critical infrastructure [IE19b, IE19a, IE20a]. As requested by those standards with joint requirements on safety and security, there shall be no retroaction from a SIEM system (or security monitoring system in general) to the safety automation systems, IIoT, and thus to the cyber-physical systems (CPS) [IE19b, IE19a]. The gateways replace firewalls in some industrial network environments, providing absolute protection to control systems and Operational Technology (OT) networks for attacks originating in external networks. Accordingly, the data related to forensic readiness (e.g. from gateways and service units) has to be sent via data diodes towards a SIEM system, like described in figure 2.

3.1 Definition

An Optical Data Diode is an analogy to the electronic diode that permits current flow in one direction only. It is a method to execute a unidirectional data transfer between two computers. Usually, a sender of a unidirectional connection will transmit data only if the transmission is acknowledged at the physical layer (ISO/OSI Layer 1) and the transport layer (ISO/OSI Layer 4). The data diode assures that data is transmitted despite not having any acknowledgment mechanism. In the case of an optical fiber solution, there is no fiber optic link between the transmit port of the transceiver on the receiving side of the data diode and the receiving port of the transceiver on the transmitting side of the data diode. Visual inspection of the transceivers should easily show that they have not been tampered or replaced with a component that can both transmit and receive [Ba16].

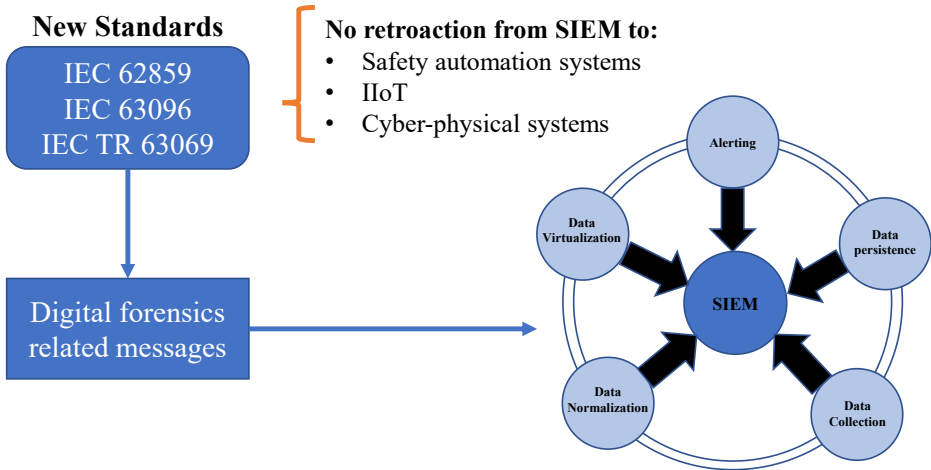


Fig. 2: Solutions for power plants and critical infrastructure with joint requirements on safety and security

3.2 Advantages of a data diode

Data diodes play an important role in the comprehensive protection of networked control systems. High-security, one-way data communication at sensitive interfaces enables the transfer of information to a maintenance service without the system being exposed to cyber risks. This section uses protection principles to illustrate the benefits of using data diodes for Industrie 4.0 [Ba16].

- **Integrity:** The entire data flows to the receiver together with strong cryptographic hash values. That assures the verification of data integrity by the receiver.
- **Availability and Reliability:** The transferred data is almost simultaneously accessible on both systems. The transmission works even if individual telegrams fail.
- **Confidentiality:** Optional data is encrypted if sensitive. Unauthorized individuals have no access to confidential data. Additionally, the latter has protection against any malicious attacks (e.g. intrusion or hacking). The hardware structure is not complex. This factor is only relevant if the data is not 'short-lived'.
- **High volume and real-time:** High volume data transmission is possible in real-time with configurable cycles.
- **Flexibility:** Different numbers of transmitters and receivers.
- **Portability:** Optical Data Diodes can be applied with different operating systems, such as Windows or Linux.

- **Transmission:** On ISO/OSI levels 4 (UDP) and 2 (LLC)
- **End-to-end logging:** On sender-side for forensic readiness and for offline evaluation of logs to detect manipulations
- **Forensics:** Ability to smartly compare logs from redundant receivers (one receiver does not need to know who the other receivers are)

3.3 Application fields of data diodes

To guarantee information security or protection of critical digital systems, such as Industrial Automation and Control Systems (IACS), from cyber attacks, unidirectional network devices are typically used [IE19b]. In many industrial application unidirectional data flow is sufficient, e.g. for monitoring the health of equipment that is susceptible to ageing, e.g. the shafts of steam turbines, or the concentration of chemical solutions used in different industrial applications, especially with slow gradients. For security applications secure unidirectional communication is also a major benefit, e.g. for the transfer of logging information towards a Security Information and Event Management (SIEM) system. Other fields where data diodes are in using are: IT bridge - Sending/receiving emails from open to critical/confidential networks - Secure cloud connectivity of critical OT networks - Database replication - Data mining - Trusted back-end and hybrid cloud hosted solutions (private / public) - Secure data exchange for data marketplaces - Secure credential/ certificate provisioning - Secure cross-data base sharing - Secure printing from a less secure network to a high secure network (reducing print costs) - Transferring application and operating system updates from a less secure network to a high secure network - Time synchronization in highly secure networks - File transfer - Streaming video - Sending/receiving alerts or alarms from open to critical/confidential networks - Government - Commercial companies [Wik21].

For these applications there are some commercial solutions available on the market. However, these are targeting selective unidirectional transmission lines via a set of protocols that are in principle designed to be bidirectional, e.g. file sharing via TCP/IP. In order to be transparent towards the user applications these solutions mimic the acknowledgements at the network communication layer, while making use of dedicated hardware and servers (for protocol conversion/acknowledgement mimicking) at both sides of the unidirectional communication. This dedicated hardware typically certified according to a high Evaluation Assurance Level (EAL), with additionally increases the equipment cost.

A transport protocol is necessary so that the data can be transferred from one computer to another. A distinction is made between TCP, UDP or LLC. The UDP standard allows the fast transport of data packets and to track the order of transmission of information. In contrast, TCP establishes a connection to the end host, where the data transmission excludes the mixing of information, your loss, delay [TW09]. LLC is a bit like TCP and UDP, except that it runs on simple Ethernet and is limited to the maximum datagram length at the ISO/OSI layer 2. This type LLC1 has no confirmation and basically sends unconfirmed data to a

receiver, making it very similar to UDP. LLC2 is a variant where connections are made and properly received packet data is acknowledged. So this is more like TCP - except for many fine-grained protocols for dynamically changing networks, so there are no flags or windows or other complexities that make TCP suitable for long-distance traffic. Since data must be sent unidirectionally, the UDP protocol is suitable. The User Datagram Protocol (UDP) is a connectionless high-level communication protocol. It offers better real-time characteristics than TCP, because in contrast to TCP it employs no transparent retransmission or error correction mechanisms which checks for every packet whether it has arrived including proper sequence of packets. Nevertheless it can be guaranteed, that no packets are lost in local networks, provided point-to-point topology is in use [EFS17].

Currently data diodes for different application scenarios are prototypically implemented with a focus on reliability for different use cases and forensic readiness. By the integration as a thin layer on existing automation hardware for example, use cases with multiple receivers can be considered: data is transferred unidirectionally from one sender to various receivers that are located in diverse zones. It is also possible to regulate which data has to flow towards which receiver (see figure 3). There are also use cases with given unreliability assumptions regarding the transmission lines, real-time requirements, provisions against spoofing and user interface considerations (e.g. regarding SIEM supported log evaluations and incident handling). To ensure the effectiveness of these application scenarios, practical results shall be provided based on different prototype implementations, including the consideration of high data volumes and transmission performance.

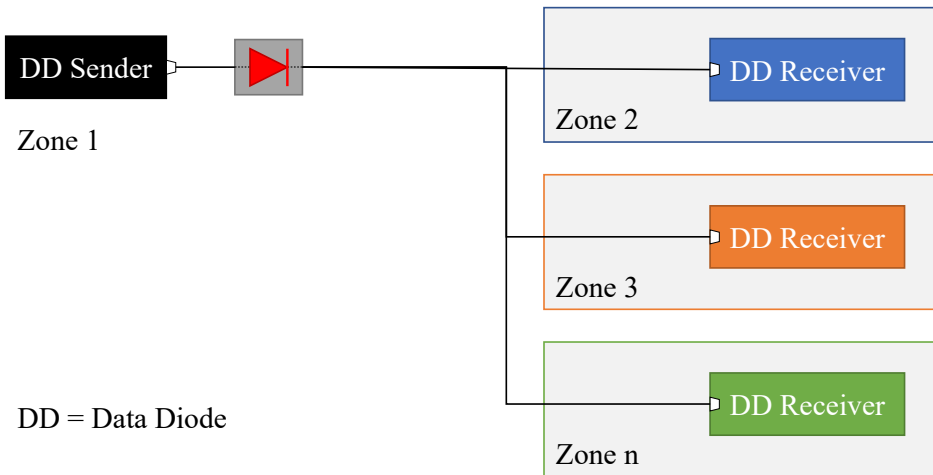


Fig. 3: Data transfer from one sender to various receivers

4 Conclusion

The aim of the paper was to conduct research in the importance of secure unidirectional Security Gateways for Industrie 4.0. The section 2 of this paper showed that the explosion of interconnectivity creates a new potential danger, what can have an impact on all the processes of a company. The fact of the need for increased data volumes and availability in real time requires new infrastructures and adaptations to the handling of information. It can be expected that current commercial solutions are not able to guarantee both specific security and cyber security requirements or to reduce the risks of cyber-attacks, such as the loss of data integrity, data availability etc. According to the on the figure 2 related standards for power plants and critical infrastructure, computer systems should be typically airgapped with data diodes to assure that no information can physically flow back . The results of in this paper presented analysis show that the usage of unidirectional gateways (data diodes) can reduce common security risk factors in the manufacturing area.

Optical Data Diodes are applicable in many different ways, depending on how data need to be transferred. There can be either applied with two computers only or with numerous ones at the same time. As messages cannot be acknowledged, each message has to be sent multiple times (with appropriate identifiers, hash values and optionally encryption). It shall be assumed that any single message can be lost (when sending via UDP or LLC) but at least one of the duplicates is not lost with sufficiently high reliability.



References

- [Ba16] Bartels et al., K.: Handreichung zum 'Stand der Technik im Sinne des IT-Sicherheitsgesetzes (ITSiG)'. Technical report, TeleTrusT – Bundesverband IT-Sicherheit e.V., Berlin, DE, 2016.
- [BSJ10] Barske, D.; Stander, A.; Jordaan, J.: A Digital Forensic Readiness framework for South African SME's. In: 2010 Information Security for South Africa. pp. 1–6, 2010.
- [EFS17] Eggert, L.; Fairhurst, G.; Shepherd, G.: UDP Usage Guidelines (RFC8085). Standard, Internet Engineering Task Force (IETF), 03 2017.
- [ET20] ETSI: Cyber Security for Consumer Internet of Things: Baseline Requirements (ETSI EN 303 645:2020). Standard, European Telecommunications Standards Institute (ETSI), Valbonne, FR, 06 2020.
- [FS07] Freiling, Felix C.; Schwittay, Bastian: A Common Process Model for Incident Response and Computer Forensics. In: 2010 Information Security for South Africa. pp. 19–40, 2007.
- [HPO16] Hermann, Mario; Pentek, Tobias; Otto, Boris: Design Principles for Industrie 4.0 Scenarios. In: 2016 49th Hawaii International Conference on System Sciences (HICSS). pp. 3928–3937, 2016.
- [IE18] IEC: Internet of things (IoT) – Reference Architecture (ISO/IEC 30141:2018). Standard, International Electrotechnical Commission (IEC), Geneva, CH, 08 2018.

- [IE19a] IEC: Industrial-process measurement, control and automation - Framework for functional safety and security (IEC TR 63069:2019). Standard, International Electrotechnical Commission (IEC), Geneva, CH, 05 2019.
- [IE19b] IEC: Instrumentation and control systems - Requirements for coordinating safety and cybersecurity (IEC 62859:2016+AMD1:2019 NPPs). Standard, International Electrotechnical Commission (IEC), Geneva, CH, 10 2019.
- [IE20a] IEC: I&C and EPS – Security controls - Version 4 (IEC 63096:2020 NPPs). Standard, International Electrotechnical Commission (IEC), Geneva, CH, 10 2020.
- [IE20b] IEC: Internet of things (IoT) – Industrial IoT (PD ISO/IEC TR 30166:2020). Standard, International Electrotechnical Commission (IEC), Geneva, CH, 05 2020.
- [TSS17] Tupa, Jiri; Simota, Jan; Steiner, Frantisek: Aspects of Risk Management Implementation for Industry 4.0. *Procedia Manufacturing*, 11:1223–1230, 2017. 27th International Conference on Flexible Automation and Intelligent Manufacturing, FAIM2017, 27-30 June 2017, Modena, Italy.
- [TW09] Trick, U.; Weber, F.: IP, TCP/IP und Telekommunikationsnetze: Anforderungen - Protokolle - Architekturen. Oldenbourg Verlag, Munich, DE, 2009.
- [Wik21] Unidirectional network. https://en.wikipedia.org/wiki/Unidirectional_network, Online; Accessed: 2021-06-04.

Towards a certification scheme for IoT security evaluation

Workshop paper on Industrial Automation and Control Systems (IACS)

Raman Barakat¹, Faruk Catal², Sascha Hackel³, Axel Rennoch ⁴, Martin A. Schneider ⁵

Abstract: Many European and international standardization bodies and industrial organizations do provide more or less detailed specification catalogues addressing IoT product security requirements, test cases and evaluation methods. In this contribution, a dedicated set of relevant standards, guides and recommendations which recently have been recognized by the European Union Agency for Cybersecurity (ENISA) will be introduced. Special attention is given to their contribution for the security evaluation process and the product quality itself, including the level of details regarding their suitability for test definition and execution.

Keywords: IoT, security, evaluation, testing, certification.

1 Introduction


Caused by the massive usage of IoT products and lots of attacks on such products, e.g. by the worm Mirai [Ant17], a strong request for the evaluation and certification of significant security aspects is being discussed. Various organizations have published standards that provide requirements for the evaluation and testing of IoT products. It is important to understand the different approaches and possibilities of these standards.


The remainder of this paper is organized as follows: Section 2 provides a brief description of standards and what they are intended for. Section 3 proposes an initial classification of these standards and assigns them to the pairs product (requirements) and (assessment of the) process but also product design and product quality.

¹ Fraunhofer FOKUS, Kaiserin-Augusta-Allee 31, 10589 Berlin, ramon.barakat@fokus.fraunhofer.de

² Fraunhofer FOKUS, Kaiserin-Augusta-Allee 31, 10589 Berlin, faruk.catal@fokus.fraunhofer.de

³ Fraunhofer FOKUS, Kaiserin-Augusta-Allee 31, 10589 Berlin, sascha.hackel@fokus.fraunhofer.de

⁴ Fraunhofer FOKUS, Kaiserin-Augusta-Allee 31, 10589 Berlin, axel.rennoch@fokus.fraunhofer.de, 
<https://orcid.org/0000-0003-3419-298X>

⁵ Fraunhofer FOKUS, Kaiserin-Augusta-Allee 31, 10589 Berlin, martin.schneider@fokus.fraunhofer.de, 
<https://orcid.org/0000-0002-8864-6492>

2 Overview of considered standards

The following list is a selection of standards for further discussion. The documents have been identified by ENISA's Stakeholder Cybersecurity Certification Group (SCCG) [Eni21] as relevant for the development of a cybersecurity certification scheme for IoT products.

2.1 IEC 62443 Industrial communication networks - IT security for networks and systems

IEC 62443:2018 is a series of standards that describes security aspects of industrial products. Part 4 of this standard series addresses the cybersecurity for industrial automation and control systems (IACS). IEC 62443-4 defines secure development lifecycle (SDL) requirements related to cybersecurity for products intended for use in the IACS and provides guidance on how to meet these requirements. These requirements can be applied to new or existing processes and cover the entire lifecycle, including developing, maintaining and retiring hardware, software or firmware.

2.1.1 IEC 62443-4-1 Secure product development lifecycle requirements

Part 4-1 of IEC 62443 specifies generic SDL process for IoT products. The covered lifecycle phases include security requirements definition, secure design, secure implementation, including coding guidelines, verification and validation, defect management, patch management and product end-of-life. Each requirement is supported with a rationale and supplemental guidance that supports the implementation of the requirement. An annex provides an example of possible metrics to measure the effectiveness of the implementation of the security requirements in a specific process.

2.1.2 IEC 62443-4-2 Technical security requirements for IACS components

In contrast to Part 4-1 of 62443, Part 4-2 specifies technical security requirements for the product itself. The standard provides detailed component requirements (CRs) associated with the seven foundational requirements (FRs) that are described in Part 1-1 of the 62443 standard series and include defining the requirements for control system capability security levels and their components. These seven foundational requirements (FRs) are

- identification and authentication control (IAC),
- use control (UC),
- system integrity (SI),
- data confidentiality (DC),
- restricted data flow (RDF),
- timely response to events (TRE), and

- resource availability (RA).

These FRs are the basis for defining four security capability levels where the lowest level would imply minimal security capabilities, preventing systems from unauthorized disclosure of information via eavesdropping and casual exposure, and the highest level means that components are resistant against information disclosure in case sophisticated methods are actively applied by an adversary with specific IACS knowledge and sufficient resources. In addition to the FRs the document also specifies specific requirements for software applications (SAR), embedded devices (EDR), host devices (HDR) and network devices (NDR). Defining security capability levels for the control system component is the goal and objective of this document as opposed to targeted and achieved security levels.

2.2 ISO/IEC 11889 Trusted Platform Module

ISO/IEC 11889:2015 defines the Trusted Platform Module (TPM), a device that enables trust in computing platforms in general. TPM meeting the requirements may enable establishing trust in platform scenarios involving security and privacy. TPMs require hardware protections to provide three roots of trust: storage, measurement, and reporting. Basing TPM roots of trust in hardware is an improvement over software-based solutions whose protections are vulnerable to malicious software. The architecture defines a TPM that is a component that receives commands and returns responses. By sending commands to a TPM and processing the responses, security benefits accrue for the platform as a whole. The root of trust for storage consists primarily of creating, managing and protecting cryptographic keys and other data values. Artefacts protected by or associated with encryption keys, like passwords, certificates or other credentials, can be used for authentication and many other security scenarios.

ISO/IEC 11889 consists of the following four parts:

- Part 1: Architecture
- Part 2: Structures
- Part 3: Commands
- Part 4: Supporting routines

Part 1 introduces the relevant modules while parts 2, 3 and 4 do provide some sources for the implementation of the modules using C programming language.

2.3 ISO/IEC 27402 (committee draft)

The ISO/IEC committee draft (CD) 27402 Cybersecurity — IoT security and privacy — Device baseline requirements standard is intended to specify a security baseline or platform for 'IoT devices' [things] supporting information security and privacy controls. Examples of baseline [information security] requirements cover the following topics:

- Unique device identifier that should be immutable and verifiable
- Factory reset functionality
- Delete all user data information' functionality
- Protection of data
- Patching/updating capability for firmware and software)

It is anticipated that additional security controls will be required and may be defined in further standards for specific applications (e.g., medical things). Currently, the standard is at 1st committee draft stage.

2.4 GP Security Evaluation Standard for IoT Platforms (SESIP)

The Security Evaluation Standard for IoT Platforms (SESIP) provided by GlobalPlatform Inc. (GP) is designed specifically for the IoT platforms and platform parts on which IoT products are based. SESIP provides a common and optimized approach for evaluating the security of connected products that meets the specific compliance, security, privacy and scalability challenges of the evolving IoT ecosystem. SESIP follows all mandatory aspects of ISO 15408 Common Criteria standard. Based on “Mapping” documents, the SESIP Certificate is reusable in another certification schemes by composition of certified parts, and reuse of certification across different evaluations.

2.5 ETSI TC CYBER series

Within ETSI TC CYBER, experts are working on “Cyber Security for consumer Internet of Things”. Basic requirements have been published in ETSI EN 303 645. Starting from this baseline, the experts are currently defining related test cases and assessment criteria.

2.5.1 ETSI EN 303 645 Cyber Security for Consumer Internet of Things: Baseline Requirements

The ETSI EN 303 645 standard specifies high-level security and data protection provisions for consumer IoT devices that are connected to network infrastructure (such as the Internet or home network) and their interactions with associated services.

In addition to the requirements, it provides basic guidance through examples and explanatory text on how to implement these requirements.

2.5.2 ETSI TS 103 701 (draft)

The ETSI TS 103 701 document specifies test scenarios for assessing consumer IoT products against the provisions of EN 303 645. It is to set out mandatory and recommended assessments, guidance and examples to support their implementation. The document

targeting testing labs and certification bodies that provide assurance on the security of relevant products. It is additionally targeting manufacturers that wish to carry out a self-assessment. The assurance schemes that this document is used in and their outcomes are out of scope. The proposed document also does not set out detailed testing protocols. However, the proposed document is intended as input to a future EU common cybersecurity certification scheme as proposed in the Cybersecurity Act.

3 Classification

3.1 Common Criteria

In order to find a classification and relevant approach for the comparison of standards, we first provide a brief introduction of the international concept of Common Criteria (CC) for security evaluation [Iso17]. The Common Criteria standard provides a generic catalogue for both the security functional requirements (SFR) of the product under consideration, i.e., the target of evaluation (TOE), and a systematic approach for the evaluation facility to address the relevant steps and documentation for the assurance of the desired evaluation level. The SFRs have been ordered in eleven main classes, e.g. for the product audit functions, cryptographic support, identification and authentication facilities or the protection of the access to the TOE. The CC also introduces the concept of Protection Profiles (PP), addressing a category of products (e.g., firewalls) that cover the specific SFRs of such product groups. Evaluators of the security product or protection profile are requested to check security assurance requirements (SAR) that in most cases cover six assurance classes, including lifecycle, development aspects (e.g., product architecture and specification), testing and vulnerability analysis. The Common Criteria Quick Reference Card [RD18] provides a comprehensive overview of the CC systematic.

Due to the broad scope and generic systematic of the Common Criteria we use them for classifying the documents introduced in Section 2, i.e., to distinguish requirements related to the product itself and the assurance process.

3.2 Classification of considered documents

In this section, the scope of the selected documents has been considered w.r.t. to the main two requirement groups proposed by the CC, i.e., if the contents include SFRs for the product and SARs for the assessment of the evaluation process. Furthermore, it has been checked if the documents support the design and contribute to the testing of the product quality.

IEC 62443-4-1 focus on the design aspects for the target industrial security product. It provides development guidance to ensure an advanced development process for some secure target product. The content is on a general level and can be described as a best

practice guide without much detail on functionality and evaluation aspects. As part of the product lifecycle, the document contains details on security verification and validation testing. However, they have been provided on a high level and do not contain concrete test scenarios. W.r.t. certification schemes, the contents of the document can contribute to the aspects of the assessment process (ADV, ALC, ATE).

IEC 62443-4-2 provides a very detailed list of industrial security product requirements. It can be stated that this document considers all SFR classes from the CC. However, the ordering and naming of the requirements in this document differ from the generic requirements from the CC. It should be noted that the product requirements have been related to security levels 0 to 4. In contrast, the CC introduce evaluation assurance levels (EAL) related to the depth of the evaluation process that range from 1 to 7.

ISO/IEC 11889 provides a detailed specification for a trusted platform module, including many recommended specific technical details for the target product. Parts 2 to 4 represent a reference implementation for the required functionality. In this sense, the standard also pre-defines some design and implementation aspects that can be regarded as prerequisites for the product evaluation in addition to true SFRs. The details within the multi-part standard provide many design recommendations but no scenarios for quality testing.

The ISO/IEC 27402 (committee draft) includes both recommendations for the device manufacturer regarding some steps in the development process (e.g., risk assessment) and product declarations (e.g., security features and vulnerability disclosure). Additionally, it provides concrete requirements for the security product itself (e.g., secure storage, reset function). This understanding from the scope of the CC provides a relatively small selection towards ten SFRs (FIA, FMT, FDP, FPR, FTA, FAU) and three SARs (ADV, ASE/AVA).

GP SESIP focusses on IoT Platforms and is an adaptation of the CC for IoT platforms, and therefore addresses both SFRs and SARs. Furthermore, the document introduces five assurance levels defined by the scope of SARs. In this understanding, the document shares similarities to a product-specific protection profile as known from the CC. Due to this generic approach, the document does not provide concrete design decisions or test objectives.

ETSI EN 303 645 addresses requirements for the security product itself (storage of security parameters, secure communication, removal of user data) but also assurance requirements (software update process). It is possible to find relationships both to SFRs, e.g., password rules to FMT_MSA.2 (Secure security attributes), and to SARs, e.g., ALC_CMC (software integrity).

ETSI TS 103 701 (draft) focusses on the assessment of requirements listed in ETSI EN 303 645. In this sense, it repeats the SFRs and SARs. However, it also addresses the definition of concrete tests using an informal description of test purposes, test actions and conditions for the assignment of verdicts. Therefore, the document provides also a

refinement of the original requirements, e.g., Provision 5.3-7 on cryptography methods has been refined w.r.t. the details “not to be known to be vulnerable” (e.g., deprecated or inappropriate for the intended product lifetime).

Table 1 provides a summary of the contents of the selected documents. Some documents address several aspects, “X” stands for *covered*, “(X)” for *partially addressed*. The main focus of the documents has been identified using I (industry), C (consumer), S (system), D (device) and is emphasised using the grey colour of the corresponding field. The column “level” indicates if the document includes some grouping of product requirements (security level, SL), evaluation assurance levels (EAL) or categorizes selected requirements as *optional* (o). The number of requirements or tests has been given in the last column to get an idea of the volume of the documents (ISO 11889 does not include dedicated requirements but a specification of a reference implementation).

	TOE	Product	Assess Process	Design	Quality Test	Level	#req
IEC 62443-4-1	I S		(X)	(X)	(X)	-	48
IEC 62443-4-2	I S	X		(X)		SL	88
ISO/IEC 11889	S	X		X		-	N/A
ISO/IEC 27402	D	(X)	(X)			-	13
GP SESIP	S	X	(X)	X		EAL	53
ETSI EN 303645	C D	X	X			o	67
ETSI TS 103701	C D	(X)	X		X		109

Tab. 1: Content classification of selected standards

4 Conclusion

Currently, multiple aspects of IoT certification approaches are under discussion in various working groups of standardization bodies and industrial associations. The technical viewpoints differ due to the various stakeholders in the organisations and interest groups. It appears very important to understand the differences, e.g., if “security levels” are proposed, since they may refer to product requirements or are related to the evaluation process. However, there is still a need for harmonization and common strategies towards a security certification scheme for IoT products.

In summary, it can be noted that most documents address technical requirements for the security product itself, and many are also considering assurance requirements. Concrete test case specifications to be used for quality evaluation are rare. Therefore, we like to mention that further documents by ETSI on the Test specification for foundational Security IoT-Profile [Ets21] have recently been released that include formalized test purpose specifications for IoT.

Acknowledgements


This contribution has been partly supported by the European commission H2020-EU.2.1.1, Grant agreement ID: 952684: <https://cordis.europa.eu/project/id/952684>.

Bibliography

- [Iso17] ISO15408 Common Criteria for Information Technology Security Evaluation, Version 3.1 Revision 5. April 2017.
- [Eni21] Enisa Stakeholder Cybersecurity certification group. Report on Draft URWP Consultation, January 2021.
- [Ets21] ETSI TS 103 701, draft V0.0.7, CYBER; Cyber Security for Consumer Internet of Things: Conformance Assessment of Baseline Requirements, March 2021.
- [Ets20] ETSI EN 303 645. V2.1.1, CYBER; Cyber Security for Consumer Internet of Things: Baseline Requirements, June 2020
- [Ets21] ETSI TS 103 646 V1.1.1 (2021-01) Methods for Testing and Specification (MTS); Test Specification for foundational Security IoT-Profile
- [GP20] GlobalPlatform Technology, Security Evaluation Standard for IoT Platforms (SESIP) – Public Release v1.0. March 2020
- [Iec19a] IEC 62443-4-1. Security for industrial automation and control systems – Part 4-1: Secure product development lifecycle requirements, February 2019.
- [Iec19b] IEC 62443-4-2. Security for industrial automation and control systems – Part 4-2: Technical Security Requirements for IACS Components, February 2019.
- [Iso20a] ISO/IEC 11889-1. Information technology - Trusted Platform Module Library - Part 1: Architecture
- [Iso20a] ISO/IEC 1st CD 27402. Cybersecurity – IoT Security and Privacy- Device baseline requirements, November 2020
- [Ant17] Antonakakis, M., April, T., Bailey, M., Bernhard, M., Bursztein, E., Cochran, J., ... & Zhou, Y. (2017). Understanding the mirai botnet. In 26th {USENIX} security symposium ({USENIX} Security 17) (pp. 1093-1110).
- [RD18] Rennoch, A., deMeer, J.: Common Criteria Quick Reference Card, <http://www.school-of-technology.de/resources/ccQRC.pdf>.

Neutral Interoperability Testbeds

Industrie 4.0 neutral cross-vendor interoperability testbeds together with medium enterprises

Dr. Dominik Rohrmus , Anja Simon¹, Dr. Alexander Willner², Markus Rentschler³


Abstract: The open, neutral, pre-competitive, ministerial financed five testbeds of the non-profit German industry association Labs Network Industrie 4.0 are based on more than 120 active Industrie 4.0 use cases. Almost all address the topic of cross-vendor interoperability. The testbeds support future standard development by a neutral validation process. This process is fully aligned with the standardization bodies like IEEE or IEC and covered by liaison contracts. The testbeds cover the areas of the Asset Administration Shell, OPC UA companion specifications, TSN and Edge Management. The BMWi Industry 4.0 Competence Centers like Augsburg are the host of the testbed and provides a factory hall as well as all the technical equipment for the document work and the implementation.

Keywords: Industrie 4.0, neutral testbeds, standardization, cross-vendor interoperability, TSN, OPC UA, Companion Specifications, Edge Management

1 Introduction

The pre-competitive and non-profit German association Labs Network Industrie 4.0 e.V. (LNI 4.0)⁴ established five neutral testbeds on cross-vendor interoperability since 2017. The association was founded in late 2015 together with the German Plattform Industrie 4.0 and the Standardization Council Industrie 4.0 (SCI 4.0)⁵, which is carried out jointly by DIN⁶ and DKE⁷ to support Industrie 4.0 standardization activities. LNI 4.0 was established by an alliance of major companies, industry associations, policymakers and researchers. The goal of LNI 4.0 is to support the pioneering work of small and medium enterprises (SME) in the area of digitalization of industrial production. LNI 4.0 acts as a competence and experimentation platform for these stakeholders.

The target of the open, neutral, pre-competitive, ministerial financed testbeds is to prepare and to validate standardization documents in their draft status based on liaison

¹ Labs Network Industrie 4.0 e.V., Ernst-Reuter-Platz 7, 10587 Berlin, info@lmi40.de,  <https://orcid.org/>

² Fraunhofer FOKUS, Kaiserin-Augusta-Allee 31, 10589 Berlin, alexander.willner@fokus.fraunhofer.de

³ Balluff GmbH, Schurwaldstrasse 9, 73765 Neuhausen, Markus.Rentschler@balluff.de

⁴ <https://lmi40.de/?lang=en>

⁵ <https://www.sci40.com/>

⁶ <https://www.din.de/en>

⁷ <https://www.dke.de/en>

contacts. The testbeds focus on the manufacturing industry with respect to the emerging Industrie 4.0 technologies.

The outline of the paper is as follows: Section 2 gives an overview on four interoperability testbeds and emphasizes on an overview. Section 3 presents the fifth, the Edge Management testbed, in depth. Finally, Section 4 concludes this contribution.

2 Neutral and open interoperability testbeds

2.1 Time-Sensitive Networking (TSN)

The testbed for the validation of the extensions to the IEEE 802.1 standard family for time-sensitive networking (TSN)⁸ based on SME use cases includes the IEC/IEEE 60802⁹ TSN Profiles for Industrial Automation activities. The IEEE 802.1 extensions enable to run heterogeneous real-time applications in a single TSN network. Among the 40 partners, there are 14 SME partners. Liaison contracts with the IEEE 802.1 Working Group and the OPC Foundation¹⁰ exist. With the liaisons it is possible to comment the draft versions of the standardization bodies and by that share the validation results bi-directionally in a neutral way.

The TSN testbed is conceived as a continuous plug-festivals which always starts with SME requirements and the LNI 4.0 Industrie 4.0 use cases. These use cases provide the basis for the architecture, electronics and mechanics used and are reflected in the testbed demonstrator. Plug-festival means that all 40 testbed partners continuously try out their (pre-)products with one another. The Federal Ministry Industrie 4.0 Competence Center in Augsburg¹¹ hosts the testbed and provides a factory building and all industrial technical equipment to enable the validation and the demonstrator.

The IEEE 802.1 standards form the basis of the testbed. TSN is validated as deterministic, real-time Ethernet communication. The testbed also deliberately works with standard templates, so as to be able to validate the different use cases against the IEEE standards and standard projects. These are used to define the technology, the architecture and the network. The key technologies concerned are switches and end-devices, time synchronization, decentralized network configuration, the forwarding of real-time data with bounded latency and the setting up of time-critical data streams. Robots and control components from different manufacturers are networked. The testbed works closely together with other international testbeds like Fraunhofer FOKUS¹² and the IIC¹³ testbeds.

⁸ <https://1.ieee802.org/>

⁹ <https://1.ieee802.org/tsn/iec-ieee-60802/>

¹⁰ <https://opcfoundation.org/about/opc-technologies/opc-ua/>

¹¹ <https://kompetenzzentrum-augsburg-digital.de/>

¹² https://www.fokus.fraunhofer.de/en/fokus_testbeds/tsn-iop-lab

¹³ <https://www.iiconsortium.org/time-sensitive-networks.htm>

2.2 Asset Administration Shell (AAS)

The AAS testbed validates the Asset Administration Shell standards initiated by the German Plattform Industrie 4.0¹⁴. The AAS is the implementation of the „Digital Twin“ for Industrie 4.0¹⁵. It establishes cross-company interoperability and is available for non-intelligent and intelligent products. The AAS covers the complete life cycle of products, devices, machines and facilities and enables integrated value chains. The AAS is the digital basis for autonomous systems and AI.

The testbed is hosted by the Deutsche Messe Technology Academy¹⁶ and the Kompetenzzentrum 4.0 Hannover¹⁷. A total of 35 industry partners with a focus on SME are collaborating to validate the normative activities. The testing covers the validation of the AASX open-source software packages¹⁸, the support of the IEC TC65 WG 23 & 24 normative projects. Beside the cross-vendor interoperability the connectivity with 5G is of importance for the testbed. This is enabled by the 5G campus network of the Deutsche Messe Technology Academy. All is tested with a demonstrator that is developed in a virtual, hybrid and physically modularized way.

The testbed has close international cooperations with CESMII, USA¹⁹ and with South Korea²⁰. Very important is the cooperation with the AAS user association, the Industrial Digital Twin Association e.V. (IDTA)²¹.

2.3 Companion Specifications with OPC UA

The Companion Specifications testbed validates the VDMA²² OPC UA companion specifications. The draft versions of the different VDMA professional bodies are at different stages, up to the release candidate level. Nine SMEs of a total of 20 partners are involved in the joint testing. The OPC UA VDMA companion specifications can partially be represented as sub-models of the asset administration shell.

The University of Applied Sciences Ravensburg-Weingarten, Faculty Mechanical Engineering²³, hosts the testbed and provides a factory hall and all the technical facilities needed to implement the use cases and the manufacturing demonstrator. Companion specifications in robotics, image processing, the packaging industry and in future of MES,

¹⁴ <https://www.plattform-i40.de/PI40/Redaktion/EN/Standardartikel/specification-administrationshell.html>

¹⁵ https://www.plattform-i40.de/PI40/Redaktion/EN/Downloads/Publikation/VWSiD%20V2.0.pdf?__blob=publicationFile&v=2

¹⁶ <https://www.technology-academy.group/en/homepage/>

¹⁷ <https://mitunsdigital.de/>

¹⁸ <https://github.com/admin-shell-io/aasx-package-explorer/releases>

¹⁹ <https://www.cesmii.org/technology-sm-profiles/>

²⁰ https://www.youtube.com/watch?v=z1_Bba2S9Lc

²¹ <https://idtwins.org/en/>

²² <https://opcua.vdma.org/>

²³ <https://www.rwu.de/hochschule/fakultaeten/maschinenbau>

ERP and cloud systems are all validated in the testbed. The aim is to test them and thereby support the international standardization. For this, solutions from different manufacturers are tested with OPC UA client/server and in future pub sub services.

2.4 Industrial Cloud to Cloud

The Industrial Cloud to Cloud testbed validates cloud transport protocols. Parts of its work are transferred to the Technical Committee of the GAIA-X European Association for Data and Cloud AISBL²⁴ for further global-scaled elaboration.

3 Neutral and open Edge Management Testbed

The target of the Edge Management testbed is to prepare the standardization in the context of the manufacturing industry with respect to the emerging edge computing technology. The testbed focuses on the Management of Edge Technologies and is described in [1]. Currently, a suitable released standard for this focus does not exist. The testbed will develop proposals for this aspect in the form of functional primitives including parameter sets (“functional view”), which afterwards must be implemented (“implementation view”). For this purpose, concepts are developed, practically implemented and validated. The results and experiences will be made available to the standardization activities to feed them into the further or new development of standards.

3.1 Related Work

The relevance of Edge Management becomes visible by the wide range of emerged organizations that are dealing with the definition and implementation of edge infrastructure solutions, usually initiated by product vendors or solution providers with a focus on the implementation level. Nevertheless, high-level concepts are in most cases developed firstly and subsequently detailed in a top-down approach. Examples of these organizations are the association Open Industry 4.0 Alliance (OI40)²⁵ with a focus on the European industrial automation market, the Chinese-based Edge Computing Consortium (ECC)²⁶, the European Edge Computing Consortium (EECC)²⁷, the US-based Open Edge Computing Initiative (OECI)²⁸, the Industrial Internet Consortium (IIC)²⁹ or the Japanese based Edgex Consortium³⁰, as well as the Korean Electronics Technology Institute

²⁴ <https://www.gaia-x.eu/>

²⁵ <https://openindustry4.com/>

²⁶ <https://ecconsortium.eu>

²⁷ <http://en.ecconsortium.org/>

²⁸ <https://www.openedgecomputing.org/>

²⁹ <https://www.iiconsortium.org/fog-and-edge-white-papers.htm>

³⁰ <https://www.edgexcross.org/en/>

(KETI)³¹ IoT Edge Platform. Even an Automotive Edge Computing Consortium³² has been founded, dealing with the problem of big data from connected cars. In [2] the authors have even identified 75 edge computing related activities.

The main focus of the completely neutral LNI 4.0 testbed compared to these testbeds is the ambition towards normative actions and full cross-vendor interoperability regarding edge management, which is seen as crucial for making the overall approach more feasible for broad offerings and applications.

3.2 Stakeholder Views

The Industrial Internet Reference Architecture [3] defines a reference model of stakeholder concerns related to technical system, called views. This methodology is also applied by the LNI 4.0 testbeds. A view is framing the description and analysis of specific system concerns as shown in Figure 1.

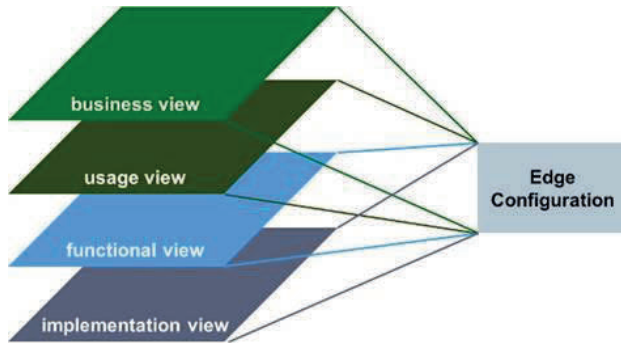


Fig. 1: Methodological classification according to [3]

Based on this, various concepts (in Figure 1 called a functional view) and solution approaches (in Figure 1 called an implementation view) are discussed and developed within the LNI 4.0 testbed Edge Configuration working group.

Specific requirements were provided by the testbed partner companies with a focus SME, describing the business concerns of various companies and market stakeholders like users and operators [4]. Most of these are businesses, which are highly relevant for the manufacturing automation domain.

³¹ https://www.keti.re.kr/eng/tech/tech_list04.php

³² <https://aecc.org/>

3.3 Architectural Concept

From an architectural viewpoint, the work of the LNI 4.0 testbed Edge Management is based on a layered architecture as shown in Figure 2.

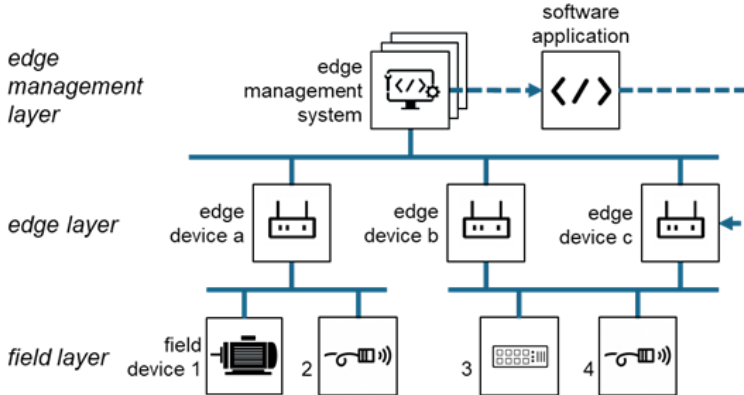


Fig. 2: Architectural context [4]

The following basic system entities have been identified:

1. Field device

Field devices are physical computing resources with often deterministic communication capabilities. Field devices communicate with edge devices, can be configured via parameters and the firmware of a field device can be updated. Field devices do not support the deployment of applications.

2. Edge device

Edge devices are physical computing resources with capabilities for communication and edge runtimes to be deployed on the edge device. Edge devices also can be configured by parameters and the firmware of edge devices can be updated. Edge devices can be connected to field devices and for each connected field device there is a data endpoint representing the communication capabilities between field and edge device. This data endpoint can be configured by an edge management system.

3. Edge management system

An edge management system is a software program deployed on an IT infrastructure. An edge management system can provide configuration capabilities for field devices and edge devices, an application store for providing software applications, edge runtimes and firmware, and configuration and deployment capabilities, which are provided via the application store of the edge management system.

4. Software application

Software applications are executable software-programs, which can be deployed, executed and configured on an edge runtime or an IT infrastructure. The software-programs are provided via the application store of an edge management system. These software applications are usually utilizing the information from connected field devices, i.e. for data analysis. Basic applications for generic usage may be for example Asset Management [5] and System Monitoring [6].

3.4 View Concept

The framing in views as shown in Figure 1 in Section 3.2 is briefly outlined.

1. Business View

The business view covers the elements to establish an edge system in a business and regulatory context and identifies the requirements how the edge system should support the business objectives through fundamental system capabilities. The following business stakeholders and their concerns have been identified: Component Supplier, Data-based Service Provider, Machine Supplier, System Integrator, Machine User, Edge Infrastructure Provider and Edge Management Provider. The details are being presented in [4].

2. Usage View

The usage view addresses the expected system usage. It is typically represented as sequences of activities involving human or logical, e.g. system or system components users that deliver its intended functionality in ultimately achieving its fundamental system capabilities. The stakeholders of these concerns are typically consisting for example of system engineers, product managers or developer as well as the logical user IT infrastructure who are involved in the usage of the considered edge system. The details are being presented in [7].

3. Functional View

The functional view describes the functional components in an edge system. These are the structure and interrelation, the interfaces and interactions between them, and the relation and interactions of the system with external elements in the environment, to support the usages and activities of the overall system. These elements are of particular interest to system and component architects, developers and integrators. The activities are analysed and functional requirements derived from them. The goal is that the functional requirements finally define the interfaces of the involved system entities.

4. Implementation View

The implementation viewpoint deals with the technologies needed to implement the functional components, their communication schemes and their lifecycle procedures. These elements are of particular interest to system and component architects, developers, integrators and system operators. The implementation specifics of the functional view and the implementation view are not only in the focus of this LNI 4.0 testbed, but mainly targeted by the already in Section 3.1 mentioned organizations.

4 Conclusion and Outlook

The key success factors for the acceptance of cross-vendor industry interoperability are the standardization and the easy integration into the existing infrastructure and systems (brown field).

The pre-competitive and non-profit German association LNI 4.0 published a series of documents providing key definitions regarding the terminology, architecture, roles and interactions for testbeds with a specific application to the topic of the management of edge technologies [4], [7]. The documents were developed in the testbeds.

For the upcoming work of the international standardization, the SCI 4.0 will give support and will ensure the international synchronization. For the necessary detailing for the implementations other organizations like the ones mentioned in Section 3.1 will refer to these documents.

Interoperability creates the essential basis (technical foundation) for the next steps in digitization. The next steps for LNI 4.0 are the evaluation of the business requirements, the initiation and testing of further modules in the AAS like energy efficiency, decarbonization or the circular economy.

In addition to this, additional functionalities such as Knowledge Management (Knowledge Graph), Human-machine interface (service & commissioning), Simulation, etc. are taken into consideration.

In the future, it will also be possible to look at the entire lifecycle management and the resulting new business models.

5 References

- [1] M. Rentschler, et al.: Standardization of Edge Configuration. IEEE Proceedings of the 25th Conference on Emerging Technologies and Factory Automation, Vol. 1, 1147-1150, Vienna, 2020.

- [2] Hamm, A., Willner, A., Schieferdecker, I. (2020), „Edge Computing: A Comprehensive Survey of Current Initiatives and a Roadmap for a Sustainable Edge Computing Development.” 15th International Conference on Wirtschaftsinformatik, Potsdam, 2020.
- [3] The Industrial Internet Reference Architecture Technical Report, accessible at https://www.iiconsortium.org/IIC_PUB_G1_V1.80_2017-01-31.pdf.
- [4] LNI 4.0 testbed “Edge Management” – Business View document, accessible at https://lni40.de/lni40-content/uploads/2021/03/BusinessView-LNI_Testbed-Edge-Management_V2.0.pdf
- [5] M. Wang, J. Tan and Y. Li, "Design and implementation of enterprise asset management system based on IOT technology," 2015 IEEE International Conference on Communication Software and Networks (ICCSN), Chengdu, 2015, pp. 384-388.
- [6] M. D. Unde and P. S. Kurhe, "Web based control and data acquisition system for industrial application monitoring," 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), Chennai, 2017, pp. 246-249.
- [7] LNI 4.0 testbed “Edge Management” – Usage View document, accessible at https://www.plattform-i40.de/PI40/Redaktion/DE/Downloads/Publikation/LNI4.0-Testbed-Edge-Configuration_UsageView.html

Workshop; Herausforderungen beim
Testen moderner Softwaresysteme
(TAV-2021)

TAV-2021: Herausforderungen beim Testen moderner Softwaresysteme

Dehla Sokenou¹, Mario Friske², Baris Güldali³, David Faragó⁴, Mario Winter⁵, Wilhelm Ruckdeschel⁶

Abstract: Die letzten Jahre zeigten einige neue Trends in der Entwicklung moderner Softwaresysteme. Dazu zählen massiv verteilte Systeme (Smart Metering, Internet of Things), der Trend zu containerbasierten Microservices auf Clustern, die vermehrte Entwicklung von KI-basierten Systemen bis hin zum Quantencomputing. Diese zeichnen sich oft durch eine Komplexität aus, die ganz neue Herausforderungen für den Test bereithält. Wir wollen im Workshop diese Herausforderungen beleuchten und Lösungsansätze für das effiziente Testen moderner Softwaresysteme diskutieren.

Keywords: Softwaretest, Testautomatisierung, Testdatenmanagement, KI-Systeme, Quantencomputing

1 Einleitung

Moderne Softwaresysteme erfordern häufig ein Umdenken in Bezug auf etablierte Entwicklungsmethoden. In diesem von der GI-Fachgruppe *Test, Analyse und Verifikation von Software* (TAV) initiierten Workshop wollen wir insbesondere die Herausforderungen beleuchten, mit denen sich die Softwareentwicklung bei der Qualitätssicherung dieser Systeme konfrontiert sieht.

Um die Herausforderungen besser zu verstehen, betrachten wir zunächst die Neuerungen, die moderne Softwaresysteme mit sich bringen.

2 Herausforderungen durch moderne Softwaresysteme

Bei massiv verteilten Systemen wie Smart Metering oder dem Internet of Things verschwimmen die Systemgrenzen. Oft werden Systeme integriert, die von unterschiedlichen Herstellern stammen und nie zusammen getestet wurden. Gesamtsysteme, wie bspw. ein Smart Home, bestehen oft aus sehr vielen Einzelkomponenten, die interoperie-

¹ WPS – Workplace Solutions

² Freelance Software Test Consultant

³ S&N CQM

⁴ EclipseSource & QPR Technologies

⁵ TH Köln

⁶ DHBW Ravensburg

ren müssen, obwohl bei ihrer Entwicklung die anderen Komponenten noch gar nicht bekannt waren. Hier helfen standardisierte Schnittstellen und Protokolle, deren korrekte Umsetzung im Einzelfall, aber auch die Eignung der Schnittstellen und Protokolle als solches für den Aufbau eines Gesamtsystems geprüft werden muss.

Eine Microservice-Landschaft, die auf einem Cluster betrieben wird, lässt das Deployment einzelner Systembestandteile unabhängig voneinander zu. Dabei können mehrere Versionen des gleichen Microservice zu einer Zeit existieren. Der Test des einzelnen Microservice ist gut verstanden und oft – aufgrund der Kompaktheit – sehr einfach mit hoher Abdeckung umzusetzen. Die Komplexität entsteht durch das Zusammenspiel der Microservices, bei denen zudem oft Aufgaben wie Logging und Security durch den Cluster übernommen werden. Hier sind effiziente Teststrategien gefragt, die die Lücke zwischen dem Test des einzelnen Microservice und dem des Gesamtsystems schließen.

KI-basierte Systeme lernen selbstständig, aber das Ergebnis des Lernprozesses ist häufig weder offensichtlich noch deterministisch. Selbst auf ersten Augenschein korrekt arbeitende Systeme können in zuvor nicht bedachten Situationen zu überraschenden Ergebnissen kommen, wie Experimente eindrucksvoll zeigen⁷. Testen darf sich nicht nur auf die Normalfälle konzentrieren, sondern muss auch immer Ausnahmen und Fehlerfälle berücksichtigen, doch wie findet man die relevanten Fälle im KI-Umfeld?

Eine weitere Klasse von Systemen, die auf einer vollständig anderen Basis als die bisherigen Systeme aufgebaut sind, sind Quantencomputer. Hier wird die Eigenschaft von Quanten genutzt, zu einer Zeit mehr als einen Zustand annehmen zu können – eine Zustandsüberlagerung oder Superposition. Quantencomputer arbeiten mit Wahrscheinlichkeiten und die Messung des Zustands reduziert die Superposition auf einen einzigen Zustand. Nicht nur die Algorithmen für Quantencomputer sehen deshalb anders aus als für klassische Computer, es muss auch neue Verfahren des Testens geben. So würde bspw. die Beobachtung eines Testobjekts während der Testausführung im Fall eines Quantencomputers das Ergebnis der Berechnung beeinflussen.

Betrachtet man moderne Softwaresysteme, so darf das Augenmerk nicht nur auf das Produkt gelegt werden, sondern muss auch den Prozess betrachten. Deshalb widmet sich der Workshop nicht nur den Testverfahren, sondern auch den Testprozessen.

3 Themenbereiche

Aus einer Reihe von Einreichungen wurden neun Papiere im Peer-Review-Verfahren zur Aufnahme in den Workshop ausgewählt. Diese umfassen sowohl Diskussionsbeiträge zum aktuellen Stand der Technik, die Weiterentwicklung von Testverfahren für moderne Softwaresysteme bis hin zu konkreten Ansätzen zum Testen der einzelnen Systemklassen.

⁷ vgl. Arne Grävemeyer: Autos sehen Gespenster – Pixelmuster irritieren die KI autonomer Fahrzeuge. ct-Magazin, 17/2020.

3.1 Testautomatisierung

Heute ist Testautomatisierung als Stand der Technik zu betrachten. Ob dies jedoch immer und überall anwendbar ist, fragen Matthias Hamburg und Stephan Weißleder in *Der Glaubenskrieg um Testautomatisierung*. Ein konkretes Verfahren zur Testautomatisierung stellen David Georg Reichelt, Stefan Kühne und Wilhelm Hasselbring mit *Testselektion für Performanzregressionsbenchmarks in CI-Prozessen* vor.

3.2 Methoden und Strategien

Konkrete Testmethoden sind notwendig, um den Test moderner Softwaresysteme effizient zu gestalten. Jonas Klamroth, Max Scheerer und Oliver Denninger stellen ein Verfahren zur *Holistischen Verifikation von Hybriden Quantenprogrammen durch Software Bounded Model Checking* vor. Ebenfalls im Bereich der formalen Verfahren bewegt sich der Beitrag von Dierk Ehmke zum *Testdatenmanagement mit formalen Testdatenspezifikationen in einem großen Migrationsprojekt*.

Die strategischen Herausforderungen betrachtet Ralf Somplatzki in *QA-Strategie oder Flickenteppich?*

3.3 Test in der Praxis

Wie ist der aktuelle Stand beim Testen moderner Softwaresysteme? Fließen neue Erkenntnisse in die Softwareentwicklung ein und wird bereits allgemein anerkannt, dass das Testen sich an neue Herausforderungen anpassen muss, um auch zukünftig die Qualität von Software sicherstellen zu können?

Eine Antwort auf diese Fragen gibt die von Mario Winter, Karin Vosseberg und Frank Simon durchgeführte *Softwaretest-Umfrage 2020 - (Wie) passt die Forschung zum Bedarf der Praxis?* Neben den Stand der Technik wurde auch der zukünftige Forschungsbedarf erfragt.

Simon André Scherr, Frank Elberzhager und Adeline Schäfer geben eine praktischen Einblick in die *Schnelllebige Qualitätssicherung im digitalen Ökosystem für mobile Dienste*.

3.4 Innovative Ansätze

Künstliche Intelligenz ist nicht nur die Basis für neuartige Anwendungssysteme, sondern bereichert auch die Softwareentwicklung selbst, bspw. durch innovative Testverfahren. So zeigen Hans-Werner Wiesbrock, Siddique Reza Khan, Mozhdeh Massah, Jens Bielefeldt, Kai Uwe Basener, Stefan Scharoba und Michael Hübner mit ihrem Beitrag, wie *Deep Learning für den Test Eingebetteter Systeme* eingesetzt werden kann. Matthias

Hamburg präsentiert *Metamorphes Testen* als Technik, um automatisiert Testorakel abzuleiten.

4 Zusammenfassung und Ausblick

Der Workshop, ausgerichtet von der GI-Fachgruppe *Test, Analyse und Verifikation von Software* (TAV)⁸ zusammen mit seinen Arbeitskreisen *Innovative Testmethoden* (TOOP)⁹ und *Testen und KI* (TKI)¹⁰, gibt einen spannenden Einblick in die Herausforderungen beim Testen moderner Softwaresysteme.

Die eingereichten Beiträge bilden die Grundlage für eine breite Diskussion und sollen Impulse für zukünftige Forschungsarbeiten geben. Die Arbeit werden wir in zukünftigen Workshops und Fachgruppentreffen fortführen. In diesem Workshop fokussieren wir speziell auf den Test. Zukünftige Workshops werden auch die anderen Themenschwerpunkte der TAV, die Analyse und die Verifikation, betrachten und im Hinblick auf moderne Softwaresysteme untersuchen.

⁸ <https://fg-tav.gi.de>

⁹ <https://ak-toop.gi.de>

¹⁰ <https://ak-tki.gi.de>

Der Glaubenskrieg über Testautomatisierung

Matthias Hamburg¹, Stephan Weißleder²

Abstract: Mit fortschreitender Technologie mehren sich die Erfolgsberichte über Testautomatisierung, z.B. im Rahmen von DevOps und Continuous Delivery. Protagonisten des explorativen Testens setzen entgegen, dass nur die kreative menschliche Arbeit das Wort „Testen“ verdient, und jede Art von Automatisierung einen kreativen Prozess unmöglich macht. Die Autoren untersuchen deshalb die Kosten und den Nutzen der Automatisierung der Aktivitäten des dynamischen Testens. Sie gehen auch darauf ein, welche Orientierung Lehrpläne und Glossar des ISTQB® bei dieser Frage bieten. Daraus leiten sie ihre Antwort auf die Streitfrage ab, ob die Testautomatisierung eine neue Wunderlösung oder ein Irrglaube ist.

Keywords: Testautomatisierung, modellbasiertes Testen, Continuous Delivery, Verifizierung, Validierung.

1 Einleitung

In der aktuellen Fachliteratur werden grundsätzlich gegensätzliche Meinungen zur Automatisierung des Testens vertreten. Einerseits lautet das bekannte Mantra von DevOps „automatisiere alles, was du mehr als zweimal ausführst“. Insbesondere die Befürworter der Continuous Delivery setzen als qualitätssichernde Maßnahme auf automatisierte Tests, die vor jedem Deployment durchgeführt werden. Andererseits setzen Kritiker entgegen, dass jede Art von Automatisierung eine standardisierte Wiederholung voraussetzt, die einen kreativen Prozess unmöglich macht. Entsprechend würde das bedeuten, dass „das einzig wahre Testen“ kreativ sein muss und deshalb unbedingt manuell durchzuführen sei. Alles andere sei eher eine technische Nachprüfung („checking“).

Automatisierung lohnt sich grundsätzlich für wohldefinierte, klar abgegrenzte Aktivitäten, wenn sie häufig wiederholt werden, und nicht durch manuelle Tätigkeiten unterbrochen werden. Außerdem muss der Nutzen die Investition und Wartung der Automatisierung überwiegen. Ausgangspunkt für die Entwicklung eines validen Standpunktes in dieser Kontroverse soll deshalb eine genauere Betrachtung der zu automatisierenden Testaktivitäten sein. Dabei ist die Kontroverse auf die Kernaktivitäten des dynamischen Testens fokussiert. Testmanagementaktivitäten wie Planung und Steuerung des Testens, Einrichtung und Verwaltung der Testumgebung oder Fehlermanagement sind weniger umstritten und werden an dieser Stelle ausgeklammert.

¹ German Testing Board e.V., Koldestrasse 8b, 91052 Erlangen, m.hamburg@gtb.de

² German Testing Board e.V., Koldestrasse 8b, 91052 Erlangen, s.weissleder@gtb.de

2 Hauptaktivitäten des dynamischen Testens

Unser Ausgangspunkt ist das Domänenmodell der Testentwicklung in Abb. 1, das von der Arbeitsgruppe Glossar des ISTQB® entwickelt wurde [Ha19].

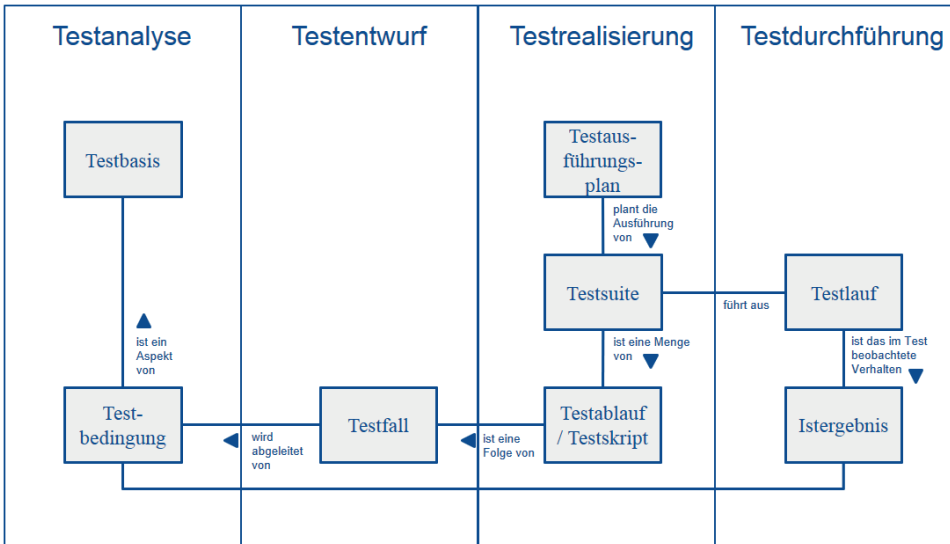


Abb. 1: Dynamische Testaktivitäten und ihre Ergebnisse

Dieses Modell in Abb. 1 zeigt die Beziehung verschiedener Artefakte in den vier Hauptaktivitäten des dynamischen Testens nach ISTQB® [CT19]: Testanalyse, Testentwurf, Testrealisierung und Testdurchführung. Im Folgenden betrachten wir die Automatisierung in den drei Aktivitäten Testentwurf, Testrealisierung und Testdurchführung. Wieweit können wir die Erstellung dieser Arbeitsergebnisse automatisieren, welche Kosten sind damit grundsätzlich verbunden, und was ist der Nutzen? Während der Nutzen mit den Arbeitsergebnissen in Abb. 1 verbunden ist, hilft beim Einschätzen der Kosten die generische Testautomatisierungsarchitektur des ISTQB® [CT16].

Wir beginnen mit der Automatisierung bei der letzten Aktivität der Kette und schreiten in Richtung früherer Aktivitäten voran, um unnötige Effizienzverluste zu vermeiden. Wenn man zum Beispiel Testfälle durch automatischen Testentwurf aus Testbedingungen generieren kann, aber danach die Testrealisierung (z.B. die Bereitstellung von Testdaten, der Entwurf von Testskripten) manuell erfolgt, kann es zu hohen Wartungsaufwänden kommen. Nach einer kleinen Änderung im Verhaltensmodell können die generierten abstrakten Testfälle radikal anders aussehen als vorher, und die manuelle Realisierung müsste fast komplett wiederholt werden.

3 Nur die Testdurchführung automatisieren

Die Automatisierung der Testdurchführung ist laut ISTQB-Glossar „Die Verwendung einer Software, um die Ausführung von Tests zu steuern, tatsächliche mit erwarteten Ergebnissen zu vergleichen, die definierten Vorbedingungen herzustellen sowie weitere Testüberwachungs- und Berichtsfunktionen durchzuführen“ [CT21a]. Sie ist oftmals der erste und intuitive Ansatz zur Automatisierung von wiederholten Aktivitäten des Testens.

Kosten

Die Architektur der automatischen Testdurchführung enthält eine Testausführungsschicht, welche die Steuerung der Ausführung der Testskripte, inkl. der Behandlung von Abweichungen, die Protokollierung der Ausführung, und die Generierung von Testberichten umfasst [CT16]. Darüber hinaus enthält die Architektur eine Testadaptierungsschicht, die die relativ abstrakten Testskripte in die Zielsprache der verschiedenen Komponenten, Konfigurationen oder Schnittstellen des Testobjekts übersetzt. Die Kosten dieser Schichten hängen wesentlich davon ab, wie verbreitet die verwendeten Technologien sind.

Eine weitere Herausforderung sind Unterbrechungen durch die manuelle Untersuchung der Abweichungen des Istergebnisses vom erwarteten Ergebnis. Neuerdings wird an dieser Stelle künstliche Intelligenz zur Automatisierung eingesetzt [Za21], deren Kosten ebenfalls berücksichtigt werden müssen.

Nutzen

Automatisiert man die Durchführung von Testsuiten, so entsteht eine Kostenersparnis bei der wiederholten Durchführung ein und derselben Testsuite. Solche Wiederholungen schaffen ein Vertrauen in die Lauffähigkeit des Testobjekts. Zum Finden von Fehlerwirkungen durch neue oder geänderte Teile des Testobjekts sind sie allerdings weniger geeignet.

Effizienz

Die Automatisierung der Testdurchführung allein lohnt sich nur unter günstigen Umständen. Dazu gehören eine gängige oder standardisierte Testumgebung auf der Kostenseite sowie eine stark inkrementelle Entwicklung mit zahlreichen Regressionstests auf der Nutzenseite.

Missionare des explorativen Testens kritisieren also die hohen Kosten für Architektur und manuelles Skripten einerseits und den beschränkten Nutzen andererseits nicht ganz zu Unrecht. Auf der anderen Seite muss man auch den Missionaren der Automatisierung Recht geben, dass menschliches Ad-Hoc Testen das Vertrauen in die Lauffähigkeit immer größerer Softwaresysteme nicht begründen kann.

Um durch die Testautomatisierung mehr Nutzen zu erzielen, müssen die ausgeführten Testskripte flexibler werden. Dies führt zum Bedarf nach der Automatisierung der Testrealisierung.

4 Zusätzlich die Testrealisierung automatisieren

Die Testrealisierung bereitet auf Basis von Testanalyse und -entwurf die Testmittel für die Testdurchführung vor [CT21a]. Ausgangspunkt sind die Testbedingungen und Testfälle.

Kosten

Um diese Aktivität zu automatisieren, sollten Testdaten aus Datenmodellen und den Inhalten der Testfälle (Eingaben und erwartete Ergebnisse) generiert werden. Testskripte können aus einem Verhaltensmodell des Testobjekts und aus den Testfallinhalten generiert werden. Entsprechend braucht eine generische Testautomatisierungsarchitektur auch eine Testdefinitionsschicht, die die Testrealisierung durch Definition von Testsuiten und/oder Testfällen unterstützt, z.B. durch Anbieten von Vorlagen bzw. Richtlinien.

Nutzen

Diese Automatisierungsstufe entlastet die Tester von der manuellen Pflege der Testskripte und Testsuiten, ohne auf die Dokumentation dieser Arbeitsergebnisse verzichten zu müssen. Darüber hinaus können Testdaten für den gleichen abstrakten Testfall leicht variiert werden. Dies ist insbesondere bei variantenreichen Anwendungen nützlich, die entweder auf zahlreichen Plattformen oder mit zahlreichen Umgebungseinstellungen laufen. Die auszuführenden abstrakten Testfälle in den Testsuiten können dynamisch ausgewählt werden.

Effizienz

Die Mehrkosten für die Entwicklung und Wartung der Testdefinitionsschicht sind im Allgemeinen gering. Der Zusatznutzen aus Skriptgenerierung und mehr Flexibilität der Testfälle ist es allerdings auch. Die mehrfache Realisierung und Ausführung des gleichen abstrakten Testfalls hat relativ wenig Chancen, neue Fehlerzustände aufzudecken. Nur in bestimmten Situationen – z.B. wenn die Dokumentation und Verfolgbarkeit der Testskripte vorgeschrieben ist, oder bei variantenreichen oder komplexen datenorientierten Anwendungen mit komplexen Testskripten – bringt dieser Automatisierungsschritt einen wesentlichen Effizienzgewinn.

Kritiker der Testautomatisierung greifen oft das Problem der Ineffizienz des „geskripteten Testens“ auf. Wenn man aber die (abstrakten) Testfälle von den Testskripten unterscheidet (Abb. 1), wird dieser Einwand durch die Automatisierung der Testrealisierung entkräftet. Auf der anderen Seite wird das Ziel einer effektiven Fehlerfindung durch diese Ausbaustufe der Automatisierung noch nicht ausreichend unterstützt. Dies legt die zusätzliche Automatisierung des Testentwurfs nahe.

5 Zusätzlich den Testentwurf automatisieren

Beim Automatisieren des Testentwurfs spricht man oft von modellbasiertem Testen (MBT). Das ISTQB widmet den Grundlagen des MBT einen eigenen Lehrplan [CT15]. Im Grunde ist aber modellbasiertes Testen ein sehr allgemeiner Begriff, definiert als Testen, das auf Modellen basiert oder diese involviert [CT21a].

Kosten

Diese vergleichsweise umfassende Automatisierung des Testens erfordert es, alle Informationen in Modellen zu pflegen, aus denen alle anderen Arbeitsergebnisse des dynamischen Testens generiert werden können. Das ergibt eine gleichfalls komplexe und mächtige Grundlage. Die für den Testentwurf passende Schicht einer generischen Testautomatisierungsarchitektur ist die Testgenerierungsschicht, die den manuellen oder automatisierten Entwurf von Testsuiten und/oder Testfällen unterstützt.

Wollen wir z.B. Black-Box-Tests generieren, so brauchen wir ein Verhaltensmodell des Testobjekts als Basis. Auch die Testbedingungen (wie z.B. 2-Switch Zustandsübergänge) müssen dem Modell angefügt werden. Um konsistente abstrakte Testfälle generieren zu können, kann auch ein Modell der Testdaten oder der Struktur des Testobjekts hilfreich sein (z.B. ein UML-Klassendiagramm). Für konkrete Testfälle können weitere mögliche UML-Objektmodelle hinzukommen.

Nutzen

Modellbasiertes Testen ist nicht neu und birgt viel Potenzial [Wi09]. Durch die erreichte volle Flexibilität bei der Generierung abstrakter Testfälle erreichen wir eine effektive Überdeckung und Fehlerfindung. Es erleichtert auch den Einsatz von Testverfahren wie paarweises Testen oder n-Switch Überdeckung [CT21b], die höhere Fehlerfindungsraten bei einer relativ geringen Anzahl von Testfällen versprechen, also geringe Kosten für die Testrealisierung und –durchführung, die jedoch ohne werkzeuggestützte Testfallgenerierung einen unpraktikabel hohen Aufwand im Testentwurf bedeuten. Darüber hinaus ist die Ableitung der Testmodelle aus der Testbasis eine frühe Qualitätssicherungsmaßnahme [CT15].

Effizienz

Die Automatisierung des Testentwurfs bringt einen hohen Zusatznutzen in der Erreichung der Testziele. Die damit verbundene frühe Prüfung der Testbasis auf Testbarkeit erhöht zusätzlich die Effizienz des Testens. Die Zusatzkosten sind allerdings auch hoch. Der wichtigste Engpass ist in der Praxis die Qualifikation der Tester. Bei guter Qualifikation (insbesondere Zertifizierungen nach [CT15] und [CT16]) kann hiermit in der Zukunft die höchste Effizienzstufe erreicht werden. Für die Protagonisten der Continuous Delivery wäre damit das eigentliche Ziel erreicht: so weit wie möglich automatisieren. Auf Kritiker der Testautomatisierung kann man soweit eingehen, dass die strukturierten (Black-Box oder White-Box) Testverfahren durch erfahrungsbasiertes Testen ergänzt werden sollten.

6 Schlussfolgerung

Unsere Antwort an die Befürworter der vollständigen Testautomatisierung ist: Testentwurf, -realisierung und -durchführung lassen sich im Rahmen der Verifizierung vollständig automatisieren. Die heutige Technologie ermöglicht das alles. Anwender sollten dabei Kosten- und Nutzenfaktoren der Automatisierung der einzelnen Testaktivitäten im Auge behalten. Unter passenden Rahmenbedingungen kann automatisches dynamisches Testen der effektive und effiziente Kern der Qualitätssicherung sein. Zu den günstigen Rahmenbedingungen gehören idealerweise eine verbreitete Standardtechnologie mit umfangreichen verfügbaren Testbibliotheken für eine vollumfängliche Testschnittstelle des Testobjekts, eine langlebige Anwendung mit häufigen und schnellen Änderungszyklen, gegebenenfalls Variantenreichtum, und nicht zuletzt ein hochqualifiziertes Testteam, mit Testanalysten die in abstrakten Modellen denken können, sowie mit Automatisierungsingenieuren, die eine komplexe Automatisierungsarchitektur entwickeln und pflegen.

Unsere Antwort an die Skeptiker der Testautomatisierung lautet: Softwaretester müssen klar zwischen Verifizierung und Validierung unterscheiden können. Beide sind unverzichtbar. Das Ziel der Verifizierung ist der Nachweis, dass festgelegte Anforderungen erfüllt worden sind. Hingegen ist das Ziel der Validierung, die Eignung für den beabsichtigten Gebrauch zu bestätigen [CT21a]. Die Verifizierung kann man unter günstigen Umständen sehr weitgehend automatisieren, und seine Effektivität durch Automatisierung sogar verstärken. Das wesentlich informellere Validieren hingegen nicht, hierbei ist Kreativität gefordert. Deswegen ist aber nicht das eine das „wahre Testen“ und das andere „etwas für Roboter“. Beide sind sehr wichtig und müssen sich sinnvoll ergänzen.

Literaturverzeichnis

- [CT15] ISTQB®, Certified Tester Model-Based Tester Syllabus, 2015.
- [CT16] ISTQB®, Certified Tester Test Automation Engineer Syllabus, 2016.
- [CT19] ISTQB®, Certified Tester Foundation Level Syllabus v3.1, 2019.
- [CT21a] ISTQB®/GTB Standardglossar der Testbegriffe, <https://glossary.istqb.org/de>, Stand: 13.06.2021.
- [CT21b] ISTQB®, Certified Tester Advanced Test Analyst Syllabus v3.1, 2021.
- [Ha19] Hamburg, M.: Ein Domänenmodell des Testens. German Testing Magazin 01/19, S. 10-11, 2019.
- [Wi09] Winter, M.: Modellbasiertes Testen – Alter Wein in neuen Schläuchen? Proc. 10. SQC-Kongress, Düsseldorf, 28. Mai 2009.
- [Za21] Zax, M.: Self-Healing Tests – Der heilige Gral der Testautomation. Proc. German Testing Day, 2021

Testselektion für Performanzregressionsbenchmarks in CI-Prozessen

David Georg Reichelt,¹ Stefan Kühne,² Wilhelm Hasselbring³

Abstract: Um Performanzregressionen zu finden, werden in Softwareprojekten Performanzregressionsbenchmarks (PRB) eingesetzt. Die Ausführung der PRBs ist zeitaufwändig und wird deshalb oft unregelmäßig, bspw. nach jedem Release, ausgeführt. Dadurch können Regressionen übersehen werden. Darüber hinaus ist die Ursachenanalyse für länger vergangene Regressionen schwer. Durch Quelltextänderungen verursachte Performanzänderungen können schneller gefunden werden, indem nur die PRBs ausgeführt werden, die geänderten Quelltext aufrufen. Wir stellen eine Erweiterung des Testwerkzeugs Peass vor, mit dem die Regressionstestselektion für PRBs möglich wird. Hierbei werden PRBs unterstützt, die mit dem Benchmarkingframework JM⁴ implementiert sind. Wir evaluieren unsere Erweiterung anhand der PRBs des Anwendungsservers jetty⁵ und zeigen, dass Peass die Benchmarkausführungsdauer um 97,9 % reduziert und alle von den PRBs abgedeckten Regressionen findet.

1 Einleitung

Eine Herausforderung beim Testen moderner Softwaresysteme ist die Sicherstellung der Einhaltung von Performanzanforderungen. Ineffiziente Implementierungen führen zu erhöhter Rechenzeit, zu einem schlechteren Nutzererlebnis und zu einem erhöhten Energieverbrauch. Durch Lasttests und Benchmarks kann die Performanz einer Anwendung geprüft und mit der vorherigen Version verglichen werden. Diese Ansätze gewinnen in modernen Softwaresystemen an Bedeutung, um bspw. die Performanz von Softwarekomponenten in automatisch deployten containerbasierten Microservices oder akkubetriebenen Endgeräten zu überprüfen. Durch Performanzregressionsbenchmarks (PRB), die bspw. in JM implementiert werden, wird es möglich, Regressionen in regelmäßig stattfindenden Tests zu identifizieren. Als Regressionen werden dabei Verschlechterung der Performanz desselben Anwendungsfalls gegenüber der Vorgängerversion betrachtet.

Moderne Software wird oft auf virtuellen Maschinen, wie der Java Virtual Maschine, ausgeführt. Die Performanz in virtuellen Maschinen wird von nichtdeterministischen Effekten wie Just-in-Time-Compilation, Garbage Collection und Speicherfragmentierung beeinflusst [GBE07]. Um Performanzänderungen zuverlässig zu bestimmen sind daher

¹ Universität Leipzig, dg.reichelt@uni-leipzig.de

² Universität Leipzig, kuehne@uni-leipzig.de

³ Universität Kiel, hasselbring@email.uni-kiel.de

⁴ Java Measurement Harness, <https://openjdk.java.net/projects/code-tools/jmh/>

⁵ <https://github.com/eclipse/jetty.project>

Wiederholungen der gemessenen Workloads nötig. Aufgrund der nötigen Wiederholungen benötigt der Messprozess für PRBs in 38 % aller Projekte über vier Stunden [St17]. Messungen finden daher oft nur unregelmäßig, bspw. nächtlich [WEH15] oder zu Releases, statt. Meist existieren dabei mehrere Benchmarks und nicht alle Benchmarks rufen den gesamten Quelltext auf. Um Regressionen zu erkennen, ist es ausreichend, die Benchmarks auszuführen, deren aufgerufener Quelltext geändert wurde.

Das bestehende Werkzeug Peass (**P**erformanz**a**nalyse von **S**oftwaresystemen)⁶ automatisiert die Regressionstestsselektion, die Messung und die Ursachenanalyse für Performanztests, die bisher aus transformierten JUnit-Tests erzeugt werden. In dieser Arbeit stellen wir eine Erweiterung von Peass vor, mit der es möglich wird, die Messung von in JMH implementierten PRBs zu beschleunigen.

Dieser Beitrag ist folgendermaßen gegliedert: Eingangs wird eine Einführung in das bestehende Werkzeug Peass gegeben. Anschließend wird dargestellt, welche Erweiterungen für die Unterstützung der Ausführung von JMH-Benchmarks durch Peass benötigt werden. Darauf basierend wird durch eine Fallstudie evaluiert, wie effizient Peass die Regressionstestsselektion für PRBs durchführen kann. Danach wird der hier verfolgte Ansatz zur Regressionstestsselektion mit verwandten Arbeiten verglichen. Abschließend wird eine Zusammenfassung gegeben.

2 Peass

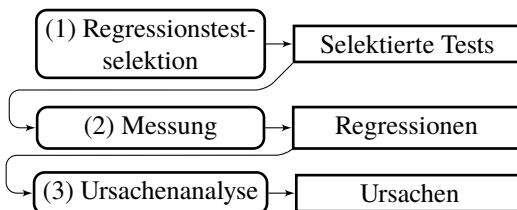


Abb. 1: Ansatz von PeASS

Um Performanzänderungen zu erkennen, ist es notwendig, gleiche Workloads in verschiedenen Versionen zu messen. Da die Implementierung und Wartung von Benchmarks und Lasttests aufwändig ist, sind diese nur selten in Quelltextrepositorien vorhanden [St17]. Um dennoch die Performanz von Software in verschiedenen Versionen zu vergleichen, misst Peass die Performanz von Unittests [RKH19].

Unittests sind nicht immer für die Performanzmessung nutzbar, da Testwerkzeuge wie Mocking das Ergebnis verfälschen können oder das Testen von Grenzfällen nicht repräsentativ für die Gesamtanwendung sein kann. Da Unittests in vielen Projekten vorhanden sind und meist eine hohe Abdeckung des Quelltextes aufweisen, sind sie dennoch oft die einzige Quelle für messbare Workloads und werden daher in Peass genutzt.

Peass besteht aus drei Schritten, die in Abbildung 1 dargestellt sind: (1) In neuen Commits werden oft nur kleine Teile der Software geändert. Tests, die den geänderten Teil der Software

⁶ Verfügbar unter <https://github.com/DaGeRe/peass>, Jenkins-Plugin unter <https://github.com/DaGeRe/peass-ci>

nicht untersuchen, müssen daher nicht getestet werden. Peass führt eine **Regressionstestselektion** aus, die bestimmt, welche Tests in der aktuellen Version eine geänderte Performanz haben können. Hierfür werden (ggf. vor dem neuen Commit) die Ausführungstraces aller Tests bestimmt. Wird ein neuer Commit erstellt, kann durch statische Quelltextanalyse auf Basis der geänderten Methoden und Klassen ermittelt werden, welche Tests Performanzänderungen aufweisen können. Durch erneute Ermittlung der Traces und Vergleich der aufgerufenen Methoden kann die Anzahl der Tests mit potentiellen Performanzänderungen weiter reduziert werden. (2) Anschließend wird durch die **Messung** bestimmt, welche Tests eine Performanzänderung besitzen. Hierfür werden nacheinander für jede Version eine zu definierende Anzahl VMs gestartet. In jeder VM wird zuerst eine zu definierende Anzahl von Aufwärmiterationen ausgeführt. Anschließend werden Messiterationen ausgeführt. Die Mittelwerte der Messausführungen bildet anschließend die Basis für den Vergleich der Messwerte. (3) Abschließend wird eine **Ursachenanalyse** durchgeführt, bei der für jeden Test durch Messung aller Knoten im Aufrufbaum ermittelt wird, welche Methode(n) die Performanzänderung verursachen.

3 Unterstützung von JMH-Benchmarks

Performanzregressionen können im Produktivbetrieb durch den wiederholten Aufruf ähnlicher Workloads entstehen. Diese Workloads können ggf. nicht durch transformierte Unittests abgedeckt werden, bspw. wenn Unittests einen höheren Anteil von Grenzfällen enthalten als die Workloads im Produktivbetrieb. Daher implementieren Entwickler bzw. Tester Benchmarks und Lasttests, anhand derer Regressionen vor der Überführung in den Produktivbetrieb festgestellt werden können. Um Performanzregressionen in Peass auch durch Benchmarks erkennen zu können, wurde Peass für die Nutzung anderer Workloaddefinitionen erweitert. Um andere Workloaddefinitionen wie Benchmarks und Lasttests messen zu können, wurden generische Schnittstellen extrahiert. Die bestehende Implementierung der Messung von JUnit-Tests in Maven- und Gradleprojekten wurde an die Schnittstellen angepasst. Darüber hinaus wurden die Schnittstellen prototypisch für JMH-Benchmarks in Maven-Projekten implementiert.

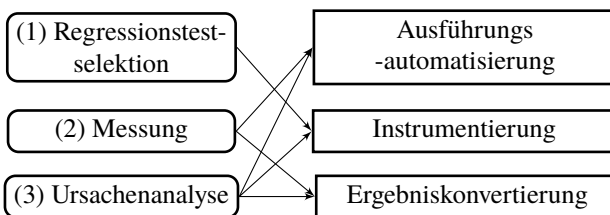


Abb. 2: Erweiterungen von Peass

Für die Unterstützung verschiedener Typen von Workloaddefinition wurden folgende Anforderungen an Peass festgestellt: (1) **Automatisierung** der Ausführung: Für die Messung und Ursachenanalyse ist es nötig, die Workloads

zu ermitteln, zu prüfen, ob die aktuelle Version lauffähig ist und die Workloads selbst mit der definierten Konfigurationen, bspw. Anzahl von VM-Ausführung, Iterationen und Warmup

auszuführen. (2) **Instrumentierung** der Ausführung: Um die Regressionstestsselektion und die Ursachenanalyse auszuführen, ist es nötig, Ausführungstraces zu erzeugen. Dies geschieht für Unittests durch Instrumentierung mit Kieker-Probes [Hv20]. Die Instrumentierung kann durch direkte Änderung des Quelltextes oder eine Weaving-Technologie wie AspectJ geschehen. Diese Instrumentierung muss für den jeweiligen Workloaddefinitionstyp und dessen Quelltextpfade angepasst werden. (3) **Transformation der Messergebnisse**: Peass verwaltet Ergebnisse einzelner VM-Ausführungen in einem XML-Datenformat und speichert die Kieker-Ergebnisdaten in einer standardisierten Ordnerstruktur. Die Messergebnisse von anderen Workloadausführungen müssen in das Peass-Format überführt werden.

Um die Regressionstestsselektion, die Testausführung und die Ursachenanalyse von JMH-Benchmarks mit Peass zu automatisieren, wurden die Schnittstellen folgendermaßen implementiert: (1) **Automatisierung** der JMH-Benchmark-Ausführung: Für die Ermittlung der Workloads wurde das Parsing so angepasst, dass `org.openjdk.jmh.annotations.Benchmark` implementierende Klassen gefunden werden. Um JMH-Benchmarks auszuführen, ist es notwendig, diese vorab zu kompilieren und anschließend eine generierte jar mit den definierten Konfigurationsparametern aufzurufen. Für die Ermittlung der Lauffähigkeit und die Generierung der jar konnte die bestehende Automatisierung der Maven-Ausführung weitergenutzt werden. Darüber hinaus wurde die Ausführung der Benchmark-jar mit der übergebenen Konfiguration implementiert. (2) **Instrumentierung der Messausführung**: Die Instrumentierung von Maven-Projekten konnte weitestgehend weiterverwendet werden. (3) **Transformation der Messergebnisse**: JMH erzeugt Ausgabedaten in einem eigenen JSON-Datenformat. Die Messergebnisse werden dementsprechend aus dem JMH-JSON-Format in das Peass-XML-Format transformiert.

4 Fallstudie

Um zu überprüfen, ob Peass-JMH in der Lage ist, alle Performanzänderungen zu finden, wurden die Benchmarks der aktuellsten 100 Versionen des Anwendungsservers jetty analysiert. Hierfür wurden die *herkömmliche Ausführung ohne Peass* und die *Ausführung mit Peass* untersucht. Alle Messungen wurden auf OpenJDK 11.0.10, laufend auf CentOS 7.9.2009 mit Intel E5-2620 v3 2.40GHz ausgeführt. Die Ausführung aller Benchmarks erfolgte mit 30 VMs [GBE07] und unter Nutzung der in der Jenkinsfile⁷ definierten Parameter. Die Benchmarkausführungen wurden unter Nutzung eines slurm-Clusters verteilt. Die Ausführung der Benchmarks dauert durchschnittlich 15,6 Stunden.

Für die *herkömmliche Ausführung ohne Peass* wurden jede der 100 Versionen ausgecheckt. Anschließend wurden alle 52 Benchmarks ausgeführt. Darauf basierend wurde unter Nutzung des zweiseitigen T-Tests ermittelt, wo Performanzänderungen vorhanden sind.

⁷ https://github.com/eclipse/jetty.project/blob/jetty-10.0.x/Jmh_Jenkinsfile

Die Ausführung der 5.200 Benchmarks (100 Versionen á 52 Benchmarks) würde ohne Parallelisierung 65 Tage benötigen.

Für die *Ausführung mit Peass* wurde Peass mit der Regressionstestselektion für jede Version ausgeführt. Durch die Regressionstestselektion wurden 98 Benchmark-Versionspaare selektiert. Diese ließen sich in 29,4 Stunden ausführen. Die Regressionstestselektion für alle 100 Versionen dauerte 2,9 h. Die gesamte Ausführungsdauer konnte also insgesamt um 97,9 % reduziert werden. Alle durch die vollständigen PRBs aufgedeckten Regressionen wurden auch durch die selektierten PRBs aufgedeckt.

5 Verwandte Arbeiten

Die nötigen auszuführenden Regressionstests können durch Regressionstestselektion, -priorisierung und -reduktion eingegrenzt werden [YH12]. Dabei bestimmt die Regressionstestselektion die auszuführenden Tests direkt, die Priorisierung ordnet Tests danach, ob eine Regression wahrscheinlicher ist und die Reduktion wählt Tests aus, deren Ausführung nicht nötig ist. In der Praxis haben sich bislang vor allem Methoden für die Regressionstestselektion funktionaler Tests etabliert, bspw. das Maven-Plugin Ekstazi⁸ und das eclipse-Plugin Infinitest⁹. Diese Werkzeuge sind für die Regressionstestselektion für funktionale Tests optimiert und haben daher einen geringen Overhead. Allerdings selektieren sie mehr Tests als die Regressionstestselektion von Peass [RK18] und würden damit zu einer höheren Ausführungsdauer führen.

Verschiedene Arbeiten entwickeln Regressionstestselektionen bzw. -priorisierungen für Performanztests. In der Regel wird ein Kostenmodell entwickelt, das einzelnen Methodenaufrufen oder Quelltextkonstrukten Kosten zuordnet, anhand derer dann (wahrscheinliche) Performanzänderungen selektiert werden. Alcocer et al. [SABV16] entwerfen eine PRB-selektion für Pharo-Benchmarks. Das Kostenmodell basiert auf Analyse von Ausführungstraces. Benchmarks werden anhand der Häufigkeit des Vorkommens von Typen von Quelltextänderungen (bspw. Hinzufügen von Methodenaufrufen oder Schleifen) selektiert. Durch die Ausführung von 17 % aller Versionen können 83 % aller Regressionen identifiziert werden. Mostafa et al. [MWX17] entwickeln ebenfalls ein Kostenmodell, in dem die Ausführungsdauer einzelner Methoden erfasst wird. Anschließend werden Änderungen priorisiert, die Ausführungen von Schleifen über Collections hinzufügen oder ändern. Huang et al. [Hu14] entwickeln ebenfalls ein Kostenmodell, in dem Änderungen, die oft aufgerufene und teure Methodenaufrufe betreffen, als risikoreich eingeschätzt werden. Durch die Ausführung von maximal 22 % der Testfälle können in den untersuchten Fallstudien mindestens 87 % aller Performanzänderungen bestimmt. Eine Methode zur Regressionstestselektion für PRBs mit JM existiert hingegen bislang nicht.

⁸ <http://ekstazi.org/>

⁹ <https://infinitest.github.io/>

6 Zusammenfassung und Ausblick

Peass ist ein Werkzeug, um Performanzmessungen im CI-Prozess durchzuführen. Peass enthält eine Regressionstestselektion zur Beschleunigung der Messung und eine Ursachenanalyse zur Untersuchung von Änderungen. In dieser Arbeit wird eine Erweiterung von Peass vorgestellt, die das beschleunigte Erkennen von Performanzänderungen in JMH-Benchmarks ermöglicht. Es wurde gezeigt, dass die Peass auf den letzten 100 Commits des Anwendungsservers jetty alle Testfälle selektiert, die potentielle Regressionen aufdecken können. Diese Arbeit ist ein Schritt hin zu **verlässlicher und effizienter Performanzmessung** im CI-Prozess äquivalent zum aktuellen Vorgehen für Unittests. Wir planen (1) durch weitere Fallstudien die Effizienz von Peass für JMH-Benchmarks vertieft zu betrachten, (2) weitere Spezifikationen von Workloads für Performanzmessung, wie JMeter-Benchmarks, zu unterstützen, und (3) durch den praktischen Einsatz von Peass zu evaluieren, inwiefern die Aufdeckung realer Performanzprobleme durch die Unittestbasierte Messung möglich ist.

Danksagung Diese Arbeit wurde durch das Bundesministerium für Bildung und Forschung im Rahmen des Projekts „Performance Überwachung Effizient Integriert“ (*PermanEnt*, BMBF 01IS20032D) finanziert. Berechnungen für diese Arbeit wurden mit Ressourcen des Universitätsrechenzentrums Leipzig durchgeführt.

Literaturverzeichnis

- [GBE07] Georges, Andy; Buytaert, Dries; Eeckhout, Lieven: Statistically rigorous java performance evaluation. *ACM SIGPLAN Notices*, 42(10), 2007.
- [Hu14] Huang, Peng; Ma, Xiao; Shen, Dongcai; Zhou, Yuanyuan: Performance Regression Testing Target Prioritization via Performance Risk Analysis. In: *ICSE*. ACM, 2014.
- [Hv20] Hasselbring, Wilhelm; van Hoorn, André; Kieker: A monitoring framework for software engineering research. *Software Impacts*, 5, 2020.
- [MWX17] Mostafa, Shaikh; Wang, Xiaoyin; Xie, Tao: PerfRanker: prioritization of performance regression tests for collection-intensive software. In: *ISSTA*. ACM, 2017.
- [RK18] Reichelt, David Georg; Kühne, Stefan: Better Early Than Never: Performance Test Acceleration by Regression Test Selection. In: *Companion of ICPE*. 2018.
- [RKH19] Reichelt, D. G.; Kühne, S.; Hasselbring, W.: PeASS: A Tool for Identifying Performance Changes at Code Level. In: *Proceedings of the 33rd ACM/IEEE ASE*. ACM, 2019.
- [SABV16] Sandoval Alcocer, Juan Pablo; Bergel, Alexandre; Valente, Marco Tulio: Learning from Source Code History to Identify Performance Failures. In: *ICPE*. ACM, 2016.
- [St17] Stefan, P.; Horky, V.; Bulej, L.; Tuma, P.: Unit Testing Performance in Java Projects: Are We There Yet? In: *Proceedings of ACM/SPEC ICPE 2017*. ACM, S. 401–412, 2017.
- [WEH15] Waller, J.; Ehmke, N. C.; Hasselbring, W.: Including Performance Benchmarks into Continuous Integration to Enable DevOps. *SE Notes*, 40(2), March 2015.
- [YH12] Yoo, S.; Harman, M.: Regression testing minimization, selection and prioritization: a survey. *Software Testing, Verification and Reliability*, 22(2):67–120, 2012.

Holistische Verifikation von Hybriden Quantenprogrammen durch Software Bounded Model Checking

Jonas Klamroth,¹ Max Scheerer,¹ Oliver Denninger¹

Abstract:

Quantencomputer erschließen uns durch ihren überpolynomiellen Speedup neue Anwendungsfelder für schwer-berechenbare Probleme. Der Entwurf von Quantenalgorithmen ist bisher allerdings komplex und fehleranfällig. Daher ist zu erwarten, dass vorerst nur einzelne Subroutinen eines Programms auf Quantencomputern umgesetzt werden. Um die Korrektheit solcher Programme garantieren zu können, sind neue Ansätze erforderlich. In dieser Arbeit stellen wir einen Ansatz zum vollautomatischen Nachweis der Korrektheit von Programmen mit eingebetteten Quantenalgorithmen vor. Dazu bauen wir auf Bounded-Model-Checking-Verfahren auf, welche die Fehlerfreiheit hinsichtlich einer gegebenen Spezifikation beweisen können. Als Spezifikationssprache verwenden wir JML. Dabei werden die Quantenalgorithmen als Quantenschaltkreis beschrieben und in Java eingebettet. Wir zeigen die Umsetzbarkeit unseres Ansatzes an zwei etablierten Quantenalgorithmen.

Keywords: Bounded Model Checking; Quantencomputing; Verifikation; hybride Quantenprogramme

1 Einleitung

Das Potential von Quantencomputern ist nicht mehr nur theoretischer Natur, sondern es sind bereits verschiedene Quantenalgorithmen für praktische Problemstellungen bekannt: z.B. der *Grover*-Algorithmus für die unstrukturierte Suche in Datenbanken oder der *Shor*-Algorithmus zur Faktorisierung von natürlichen Zahlen. Der Shor-Algorithmus erzielt einen überpolynomiellen Speedup und ist somit deutlich effizienter als jeder bekannte klassische Algorithmus. Während der Shor-Algorithmus sich stark auf Kryptographie und somit auf IT-Sicherheit auswirkt, adressieren andere Quantenalgorithmen andere Anwendungsgebiete.

Obwohl Quantencomputer einen immensen Speedup erzielen können, kommen Quantenalgorithmen in der Praxis bisher kaum zum Einsatz. Dies ist maßgeblich auf hardwareseitige Einschränkungen zurückzuführen, die sogenannte *Noisy Intermediate-Scale Quantum* (NISQ) Computer mit sich bringen, z.B. geringe Dekohärenzzeit und hohe Gatterfehlerraten. Aus diesem Grund haben sich sogenannte *Variational Quantum Algorithms* (VQA) etabliert [Ce20], die zuverlässiger auf NISQ-Computern ausgeführt werden können. Ein VQA besteht aus einem parametrisierten Quantenschaltkreis, dessen Parameter auf einem klassischen Rechner iterativ optimiert werden, bis eine akzeptable Lösung, der durch den

¹ FZI Forschungszentrum Informatik, Haid-und-Neu-Straße 10-14, 76131 Karlsruhe, Deutschland
klamroth/scheerer/denninger@fzi.de

Quantenschaltkreis codierten Funktion, erreicht ist. Wie alle Quantenalgorithmen können VQA nur in Kombination mit klassischen Programmteilen verwendet werden, da immer zumindest eine Vor- sowie Nachverarbeitung nötig ist [LB20]. Ein System, bestehend aus klassischem sowie Quantenprogrammteil, wird als *hybrides System* bzw. *hybrides Quantenprogramm* bezeichnet.

Ein weiterer Hinderungsgrund für den Einsatz von (hybriden) Quantenalgorithmen bildet ihr inhärent hoher Grad an Komplexität. Quantencomputer nutzen quantenmechanische Effekte, wie z.B. *Quantenverschränkung*, *Superposition* oder *Messungen* von *Quantenzuständen*. Somit kann das Implementieren von Quantenprogrammen nicht mit konventionellen Programmen verglichen werden, da Quantencomputer ein völlig neues Paradigma mit sich bringen. Entsprechend schwer ist auch das Debuggen und Testen, so dass formale Verifikation eine noch bedeutendere Rolle bei der Qualitätssicherung einnimmt [Ba20].

In dieser Arbeit stellen wir einen Ansatz zum vollautomatischen Nachweis der Fehlerfreiheit von Programmen mit eingebetteten Quantenalgorithmen hinsichtlich einer gegebenen Spezifikation vor, indem die Quantenalgorithmen in Java-Code übersetzt werden. Unser Ansatz ist auf das Schaltkreismodell für Quantencomputing anwendbar. Dabei werden Algorithmen durch Schaltkreise beschrieben, die wiederum aus einer Anordnung von Gattern bestehen. Jedes Gatter manipuliert dabei ein oder mehrere Qbits. Nachfolgend stellen wir die notwendigen mathematischen Grundlagen vor. Für eine ausführlichere Einführung in Quantencomputing verweisen wir auf [NC10].

2 Grundlagen Quantencomputing

Ein Qbit kann, in Gegensatz zu einem klassischen Bit, nicht nur in den Zuständen 1 oder 0 sein, sondern auch in einer beliebigen Superposition davon. Um dies auszudrücken, werden Qbits als Elemente in \mathbb{C}^2 beschrieben. Qbits werden typischerweise in der Dirac-Notation wie folgt dargestellt: $\alpha |0\rangle + \beta |1\rangle$, mit $|0\rangle = (1\ 0)^T$, $|1\rangle = (0\ 1)^T$ und $\alpha, \beta \in \mathbb{C}$. Dabei gilt immer $|\alpha|^2 + |\beta|^2 = 1$. Der Zustand eines Systems mit mehreren Qbits kann als Tensorprodukt der einzelnen Zustände dargestellt werden. Dies ist im Allgemeinen umgekehrt nicht der Fall. Als Beispiel betrachte man folgenden Zustand: $\sqrt{2}^{-1} |00\rangle + \sqrt{2}^{-1} |11\rangle$. Hierfür lässt sich keine äquivalente Zerlegung in Zustände einzelner Qbits finden. Solche Zustände nennt man "verschränkt", d.h. die Manipulation eines einzelnen Qbits verändert den Zustand des gesamten Systems. Die Manipulation eines Qbits geschieht durch unitäre Transformationen. Da wir immer nur endlich viele Qbits betrachten, kann eine solche Transformation als Matrix ausgedrückt werden. Dabei gilt: Eine Matrix M ist unitär, wenn: $MM^\dagger = M^\dagger M = I$ wobei I die Identität ist und M^\dagger die komplex konjugierte transponierte Matrix von M .

Um den Zustand eines Qbits beobachten zu können, muss auf einem Quantencomputer eine Messung des Qbits durchgeführt werden. Eine solche Messung hat zwei Effekte: Sie liefert als Ergebnis ein klassisches Bit (abhängig vom Zustand des Qbits) und das Qbit kollabiert. Konkret bedeutet das für ein Qbit in Zustand $|\phi\rangle = \alpha |0\rangle + \beta |1\rangle$: Die Wahrscheinlichkeit

$|0\rangle$ zu messen beträgt $|\alpha|^2$ bzw. $|\beta|^2$ für $|1\rangle$. In jeden Fall kollabiert $|\phi\rangle$ in den gemessenen Zustand. Das bedeutet insbesondere, dass eine Messung den Zustand des Systems verändert.

3 Ansatz zur Verifikation hybrider Quantenprogramme

Die Grundidee unseres Ansatzes ist die Einbettung von Quantenschaltkreisen in eine klassische Programmiersprache (in unserem Fall Java) basierend auf einer entsprechenden Übersetzung. Ein Java-Programm mit Quantensubroutinen kann somit in ein reines Java-Programm übersetzt werden und dadurch wie ein klassisches Programm spezifiziert, getestet und verifiziert werden. Der Vorteile dieses Ansatzes ist, dass durch die Einbettung des Quantenteils der Software in den klassischen Teil die Komplexität für Entwickler drastisch reduziert wird. Sowohl Spezifikation als auch Verifikation erfolgen analog zu klassischer Software. Auch die Möglichkeit Tests für den entsprechenden Quantenalgorithmus zu schreiben und Debuggingmöglichkeiten ergeben sich hierdurch. Darüber hinaus funktioniert unser Ansatz vollautomatisch, so dass kein quantenmechanisches Detailwissen nötig ist. Die Übersetzung basiert auf der Darstellung von Quantenzuständen als Arrays von Fließkommazahlen. Diese werden im Laufe des Programms entsprechend des gegebenen Quantenschaltkreises manipuliert. Zur Umsetzung müssen zwei grundlegende Operationen ermöglicht werden: unitäre Transformationen und Messungen. Transformationen lassen sich als Matrixmultiplikation abbilden und können somit direkt als Summe bzw. Produkt von Fließkommazahlen ausgedrückt werden. Es sei kurz erwähnt, dass die Anwendung eines 1-Qbit-Gatters (bspw. Hadamard-Gatter) auf einen Zustand mit mehr als einem Qbit erreicht werden kann, indem mit Hilfe des Tensorprodukts und der Identitätsmatrix eine äquivalente Operation auf alle Qbits angewendet wird.

Der spannendere Aspekt ist die Umsetzung von Messungen (in unserem Fall ausschließlich in der Standard-Basis). Messungen lassen ein Qbit in den gemessenen Zustand kollabieren. Für einen Zustand mit mehr als einem Qbit bedeutet dies, dass die Wahrscheinlichkeit für alle Zustände, in denen das Qbit nicht den gemessenen Zustand annimmt, auf 0 fällt. Betrachten wir als Beispiel den 2-QBit-Zustand: $\alpha |00\rangle + \beta |01\rangle + \gamma |10\rangle + \delta |11\rangle$. Messen wir in diesem Zustand eine 1 für das erste Qbit, so ergibt sich der Zustand: $\frac{1}{\sqrt{\gamma^2 + \delta^2}}(\gamma |10\rangle + \delta |11\rangle)$. Der vorgestellte Faktor normalisiert den Zustand, sodass der resultierende Vektor wieder Länge 1 hat. Wir unterstützen zwei Varianten für Messungen: klassische Messung und Maximumsmessungen. Beide sind praktisch motiviert: Während die klassische Messung alle möglichen Messergebnisse für die Verifikation betrachtet, wird bei der Maximumsmessung ausschließlich das wahrscheinlichste Ergebnis der Messung weiter verfolgt. Dies hat den Grund, dass Quantenalgorithmen oftmals das gewünschte Ergebnis nicht mit Sicherheit, allerdings mit hoher Wahrscheinlichkeit liefern. Mit dieser Art der Modellierung können wir Aussagen für das wahrscheinlichste Ergebnis des Algorithmus beweisen. Umgesetzt werden beide Messvarianten durch eine Fallunterscheidung. Im klassischen Fall wird als Unterscheidungskriterium ein nichtdeterministisches Bit genutzt und somit beide Fälle betrachtet. Bei der Maximumvariante wird hingegen die Summe

der beiden Wahrscheinlichkeiten für das entsprechende Bit als Unterscheidungskriterium herangezogen.

4 Fallstudien

Um unseren Ansatz zu evaluieren haben wir zwei typische Beispiele für Quantenalgorithmen mit der vorgestellten Technik spezifiziert und verifiziert. Die Übersetzung von der Gatterdarstellung des Quantenalgorithmus in den entsprechenden Java-Code wurde mit Hilfe eines von uns entwickelten Werkzeugs automatisch ausgeführt. Als Verifikationswerkzeug wurde JJBMC [Be20] genutzt, welches (mit JBMC [Co18] als Backend) für die Verifikation von mit JML [Le08] spezifiziertem Java-Code entwickelt wurde.

Die erste Fallstudie ist der Deutsch-Algorithmus. Gegeben eine Funktion $f : \{0, 1\} \rightarrow \{0, 1\}$, entscheidet der Deutsch-Algorithmus, ob diese Funktion konstant ist oder nicht. Hierfür benötigt der Algorithmus nur eine einzige Evaluation der Funktion. Mit unserem Ansatz waren wir in der Lage vollautomatisch in weniger als einer Sekunde zu zeigen, dass die Implementierung des Algorithmus für jede möglich Funktion das erwartete Ergebnis berechnet. Als zweite Fallstudie wurde das Quantenteleportationsprotokoll betrachtet. Hierbei wird gezeigt, dass es möglich ist, einen Quantenzustand, alleine mit der Übertragung zwei klassischer Bits zwischen zwei Parteien, zu teleportieren. Auch in diesem Fall konnte mit unserem Ansatz gezeigt werden, dass durch das Protokoll der Quantenzustand tatsächlich "teleportiert" wird. Übersetzung und Verifikation erfolgten mit geringem manuellen Aufwand.

Die Fallstudien zeigten allerdings auch zwei kritische Punkte für das vorgestellte Verfahren auf. Als erstes ist anzumerken, dass wir durch den Einsatz von Fließkommazahlen anstatt reeller Zahlen eine natürlich Fehlerquelle durch Rundungsfehler in Kauf nehmen. In den vorgestellten Fallstudien haben diese Rundungsfehler keine entscheidenden Probleme verursacht. Es ist jedoch anzunehmen, dass diese sich bei größeren Quantenprogrammen potenzieren würden und somit gegebenenfalls die Aussagen der Verifikation in Frage stellen könnten. Dies führt auch zum zweiten Nachteil unseres Ansatzes: Skalierbarkeit. Die vorgestellten Fallstudien kamen mit maximal drei Qbits aus und bereits hier konnten wir exponentielle Aufwände für sowohl Übersetzung als auch Verifikation beobachten. Die Übersetzung ist aktuell als Proof-of-Concept umgesetzt und enthält keine Optimierungen. Entsprechend kann die Skalierbarkeit unseres Ansatzes aktuell noch nicht abschließend bewertet werden.

5 Verwandte Arbeiten

In der Literatur finden sich mehrere Ansätze, um beweisbare Aussagen über Quantenprogramme zu treffen. Wir unterscheiden hierbei zwischen Ansätzen, die rein theoretischer

Natur sind und solchen, die Werkzeuge und somit Automatisierungsmöglichkeiten vorstellen. Auf theoretischer Seite finden sich mehrere Logiken wie beispielsweise LQP [BS06] (eine dynamische Logik für Quantenprogramme), EEQPL [CMS06] (angelehnt an die klassische Hoare-Logik) oder eine vollständige Floyd-Hoare-Logik für eine Quanten-While-Sprache [Yi12]. Im Why3-Framework können mit Hilfe von parametrisierten Pfadsummen teilautomatisierte Beweise über Quantenschaltkreise geführt werden [Ch21]. Die Schaltkreise werden dabei in einer DSL (QBricks) beschrieben und mit Hilfe einer weiteren DSL (QBricks-spec) spezifiziert. Bemerkenswert ist hierbei, dass Schaltkreise mit variabler Größe und unbeschränkt vielen Qbits betrachtet werden können. Ein weiterer Ansatz wurde für den Theorembeweiser Coq vorgestellt [PRZ17]. Auch hier wird das Schaltkreismodell als Grundlage genutzt und die Sprache für die Beschreibung der Programme heißt Qwire. Qwire nimmt außerdem eine abstrakte Hostsprache an, mit der zusammen das Quantenprogramm ausgeführt wird. Dieser Ansatz ist allerdings deutlich weniger automatisiert als der zuvor beschriebene. Auch einen Model Checker für Quantenprogramme gibt es bereits [GNP08]. Spezifikationen werden hierbei in QCTL geschrieben. Allerdings wird hier die Integration in eine Hostsprache nicht betrachtet.

6 Zusammenfassung und weitere Arbeiten

Wir haben gezeigt, dass hybride Software spezifiziert und verifiziert werden kann, indem die Quantenalgorithmen in den klassischen Softwareteil eingebettet werden. Die Reduktion auf die Verifikation klassischer Software hat hierbei den Vorteil, dass gewohnte Sprachen und Praktiken unverändert übernommen werden können. Als weitere Arbeiten sehen wir vor allem die formale Beschreibung der Übersetzung und damit verbunden einen formalen Korrektheitsbeweis. Außerdem sollen in Zukunft andere Hostsprachen (Übersetzungen in andere Sprachen als Java) sowie andere Werkzeuge zur Verifikation untersucht werden. Zusätzlich sollen weitere Fallstudien durchgeführt werden, um den Einsatz auch für komplexere Algorithmen zu testen.

Die Arbeit erfolgte im Projekt SEQUOIA, gefördert durch das Wirtschaftsministerium Baden-Württemberg.

Literatur

- [Ba20] Barbosa, L. S.: Software Engineering for 'Quantum Advantage'. In: Proceedings of the IEEE/ACM 42nd International Conference on Software Engineering Workshops (ICSEW'20). ACM, S. 427–429, 2020.
- [Be20] Beckert, B.; Kirsten, M.; Klamroth, J.; Ulbrich, M.: Modular Verification of JML Contracts Using Bounded Model Checking. In: Leveraging Applications of Formal Methods, Verification and Validation: Verification Principles, International Symposium on Leveraging Applications of Formal Methods (ISoLA 2020). Springer, S. 60–80, 2020.

- [BS06] Baltag, A.; Smets, S.: LQP: The Dynamic Logic of Quantum Information. *Mathematical Structures in Computer Science* 16/3, S. 491–525, 2006.
- [Ce20] Cerezo, M.; Arrasmith, A.; Babbush, R.; Benjamin, S. C.; Endo, S.; Fujii, K.; McClean, J. R.; Mitarai, K.; Yuan, X.; Cincio, L.; Coles, P. J.: Variational Quantum Algorithms, 2020, arXiv: 2012.09265 [quant-ph].
- [Ch21] Chareton, C.; Bardin, S.; Bobot, F.; Perrelle, V.; Valiron, B.: An Automated Deductive Verification Framework for Circuit-building Quantum Programs. In: *Programming Languages and Systems, European Symposium on Programming (ESOP 2021)*. Springer, S. 148–177, 2021.
- [CMS06] Chadha, R.; Mateus, P.; Sernadas, A.: Reasoning About Imperative Quantum Programs. *Electronic Notes in Theoretical Computer Science, Proceedings of the 22nd Annual Conference on Mathematical Foundations of Programming Semantics (MFPS XXII)* 158/, S. 19–39, 2006.
- [Co18] Cordeiro, L.; Kesseli, P.; Kroening, D.; Schrammel, P.; Trtik, M.: JBMC: A bounded model checking tool for verifying Java bytecode. In: *Computer Aided Verification, International Conference on Computer Aided Verification (CAV 2018)*. Springer, S. 183–190, 2018.
- [GNP08] Gay, S. J.; Nagarajan, R.; Papanikolaou, N.: QMC: A Model Checker for Quantum Systems. In: *Computer Aided Verification, International Conference on Computer Aided Verification (CAV 2008)*. Springer, S. 543–547, 2008.
- [LB20] Leymann, F.; Barzen, J.: The bitter truth about gate-based quantum algorithms in the NISQ era. *Quantum Science and Technology* 5/4, S. 044007, 2020.
- [Le08] Leavens, G. T.; Poll, E.; Clifton, C.; Cheon, Y.; Ruby, C.; Cok, D.; Müller, P.; Kiniry, J.; Chalin, P.; Zimmerman, D. M. et al.: *JML reference manual*, 2008.
- [NC10] Nielsen, M. A.; Chuang, I. L.: *Quantum Computation and Quantum Information*. Cambridge University Press, 2010.
- [PRZ17] Paykin, J.; Rand, R.; Zdancewic, S.: QWIRE: A Core Language for Quantum Circuits. In: *Proceedings of the 44th ACM SIGPLAN Symposium on Principles of Programming Languages (POPL 2017)*. ACM, S. 846–858, 2017.
- [Yi12] Ying, M.: Floyd–Hoare Logic for Quantum Programs. *ACM Transactions on Programming Languages and Systems* 33/6, 19:1–19:49, 2012.

Testdatenmanagement mit formalen Testdatenspezifikationen in einem großen Migrationsprojekt

Dierk Ehmke¹

Abstract: Für ein großes Migrationsprojekt in der Finanzindustrie wurde ein spezifikationsbasierter Testdatenmanagementansatz entwickelt. Die Testdaten werden automatisch im Ticket-system Jira bereitgestellt. Beschrieben werden Ansatz, Resultate und Voraussetzungen für eine erfolgreiche Einführung.

Keywords: formal specification; test data; test data management; migration project test

1 Einleitung

Aus einer jahrzehntealten Mainframe-Anwendung werden Millionen Verträge in ein bestehendes, modernes Zielsystem migriert. Dort existiert ein Bestand, zu dem die migrierten Daten hinzukommen. Nur im Altsystem vorhandene Geschäftsprozesse müssen implementiert werden. Wichtigstes Testziel ist die Korrektheit der migrierten Verträge sowie der neuen und der bestehenden Geschäftsprozesse.

Mehrere Testteams testen gemeinsam oder exklusiv auf den Testumgebungen. Das Testmanagement wird von erfahrenen internen und externen Mitarbeitern durchgeführt. Testfälle werden von besonders erfahrenen Sachbearbeitern erstellt. Sachbearbeiter sind es auch, die die Tests ausführen. Dem Testdatenmanagement steht eine neunmonatige Vorbereitungsphase zur Verfügung.

Als Ticketsystem wird Jira mit dem Testmanagementwerkzeug XRay eingesetzt. Für Regressionstests werden Testfälle automatisiert. Die Datenbanken des Zielsystems sind auf den Testumgebungen mit SQL erreichbar. Der Hersteller des Zielsystems unterstützt bei der Umsetzung fachlicher Spezifikationen in Datenbankabfragen.

Das Mengengerüst umfasst Millionen Verträge, hunderte Geschäftsprozesse, hunderte Projektbeteiligte (teilweise temporär, geschätzt ein Drittel davon im Test) und tausende Testfälle.

Die Aufgabenstellung besteht darin, zu jedem Testfall für die Testdurchführung nach fachlichen Kriterien passende Verträge auszuwählen. Wegen etwaiger Testwiederholungen müssen es mehrere sein. Da die Testdurchführungen oft die Daten so verändern, dass sie

¹ Freiberuflicher Softwaretestberater mit Schwerpunkt Testdatenmanagement, Herdweg 76, 64285 Darmstadt, mail@d-ehmke.de, <https://www.d-ehmke.de>

für folgende Testdurchführungen unbrauchbar werden, müssen die Verträge jeweils für genau eine Durchführung eines Testfalls reserviert werden.

Testdaten sind die migrierten anonymisierten Produktionsdaten. In einem Migrationsprojekt sind die migrierten Daten selbst Gegenstand des Tests, die Verwendung synthetischer Daten ist hier nicht zielführend, denn dann würde man durch die Migration verursachte Datenfehler nicht entdecken.

Dem Stand der Technik entsprechende Lösungsalternativen wurden in Betracht gezogen [Tr18]. A: die Tester suchen sich zu Beginn der Testdurchführung passende Verträge. B: in die Testfallbeschreibung wird die Nummer eines passenden Vertrages aufgenommen. C: Ein Testdatenmanager führt manuell Listen von passenden Verträgen zu jedem Testfall. D: Anschaffung eines geeigneten Testdatenmanagementsystems. E: Implementierung einer automatisierten Individuallösung.

A, B und C wurden verworfen. Wenn sich die Datenbasis ändert, weil zum Beispiel der Vertrag im Bestand verändert wurde und man (neue) Verträge suchen muss, ist nicht sichergestellt, dass man die Spezifikationen der erforderlichen Testdateneigenschaften hat. Für automatische Regressionstests ist das ein großes Problem, dass dadurch verschärft wird, dass man für die Testwiederholungen viele Verträge je Testfall benötigt. Alternative D – die Anschaffung eines passenden Testdatenmanagementsystems wurde verworfen, da dies zeitlich und organisatorisch nicht machbar war und nicht sichergestellt war, dass man am Ende ein im Budgetrahmen befindliches Produkt findet, dass alle Anforderungen erfüllt. Eine mit kleinen Schritten einführbare Individuallösung (E) wurde entwickelt und hat sich bewährt.

2 Lösung

Kern der Lösung sind formale Testdatenspezifikationen, die automatisch in Datenbanksuchabfragen für die Vertragssuche konvertiert werden. Die Notation ist eine einfach gehaltene Domain Specific Language und orientiert sich an der Fachterminologie in der Organisation. Beispiel:

```
TESTDATENSPEZIFIKATION
  benötigt wird ein Vertrag mit den Merkmalen
    Tarif = XYZ_2012
    Zahlart = 57
    Beginn >= 01.01.1970
    Beitrag >= 100
ENDE TESTDATENSPEZIFIKATION
```

Jeder Testfall beinhaltet eine Testdatenspezifikation, für die in Jira ein Feld implementiert wurde. Werden ausnahmsweise keine Testdaten benötigt, wird mit einem speziellen Text

ausgedrückt, dass dies intendiert ist und kein Versehen vorliegt. Aktuell werden etwa 75 Attribute und 1000 Schlüssel unterstützt. Vereinfacht verläuft der Prozess von der Testdatenspezifikation bis zur Testdatenbereitstellung wie folgt:

Erstellung: Die Testfallersteller müssen für jeden Testfall Testdatenspezifikationen erstellen, die in Jira eingestellt werden, siehe Abb. 1. Da die Erstellung einer korrekten formalen Spezifikation schwierig ist, wird sie mit einem Erfassungsformular unterstützt, in dem formale Fehler abgefangen werden. Realisiert wurde es mit einem marktgängigen Tabellenkalkulationsprogramm.

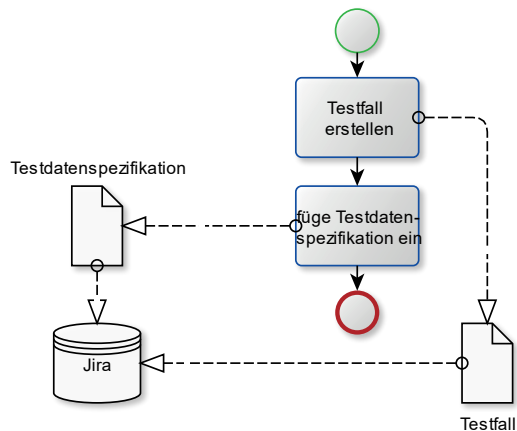


Abb. 1: Prozess der Erstellung einer Testdatenspezifikation

Reservierung: Vor Beginn einer Testkampagne stellen die Testmanager die durchzuführenden Testfälle zusammen und exportieren deren Testdatenspezifikationen aus Jira für die Reservierung durch das Testdatenmanagement. Ein Compiler übersetzt diese in ein SQL-Programm, das die Suche und Reservierung im Datenbestand der zugewiesenen Testumgebung durchführt.

Ein Compiler übersetzt diese in ein SQL-Programm, das die Suche und Reservierung im Datenbestand der zugewiesenen Testumgebung durchführt. Folgender SQL-Code setzt die Testdatenspezifikation weiter oben um:

```

select V.NR from VERTRAG as V
  where V.TARIF='XYZ_2012' AND V.BEGINN >= '01.01.1970'
        AND V.ZAHLART=57 AND V.BEITRAG = 100
        FETCH FIRST 5 ROWS ONLY
  
```

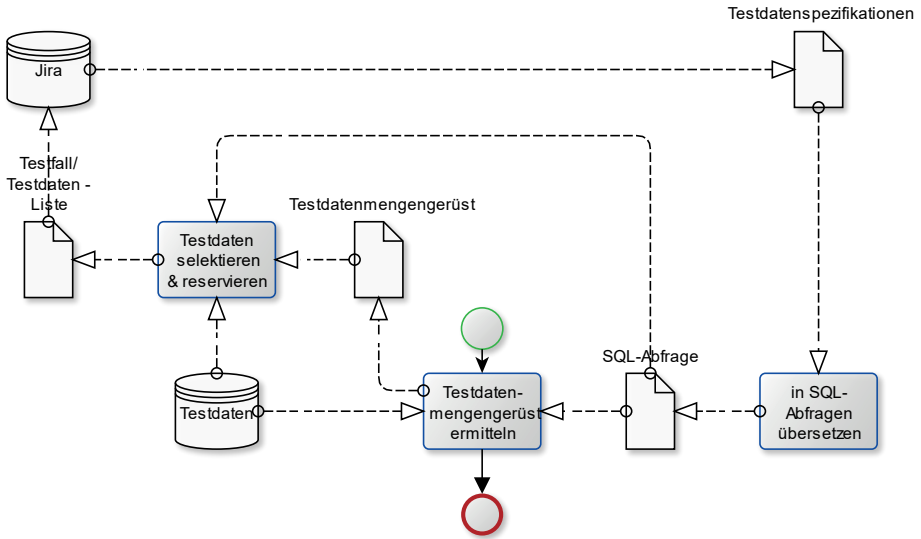



Abb. 2: Prozess der Testdatenbereitstellung

Bereitstellung, siehe Abb. 2: Ergebnis ist eine Liste von Testfallschlüsseln mit zugeordneten Verträgen, die so aufgebaut ist:

Testfall	Vertrag
MIG-2799	V52643921
MIG-2799	V20363262
MIG-9156	V98879715
MIG-9156	V59287826

Sie wird in Jira abgelegt.

Zugriff, siehe Abb. 3: Vor Teststart werden sogenannte Testruns für die Durchführung der Testfälle angelegt, das ist der Prozess in XRay. Automatisch wird jedem Testrun ein hinterlegter Vertrag zugewiesen, indem im Kommentarfeld ein Text eingefügt wird: "Dieser Testausführung ist Vertrag V52643921 zugewiesen".

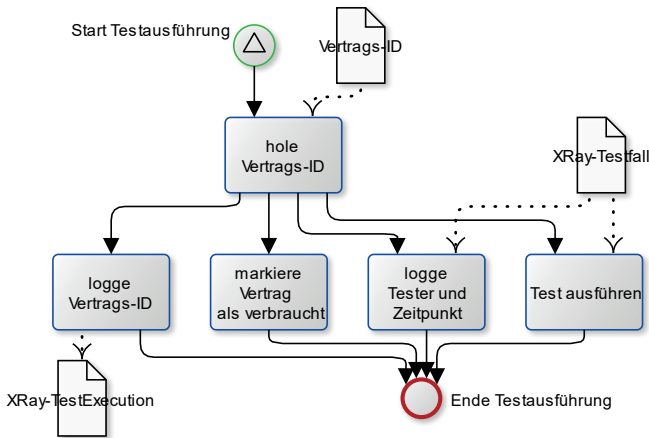


Abb. 3: Zugriff auf Testdaten während der Testausführung

In der Systemübersicht in Abb. 4 sieht man die beteiligten Systeme und den Datenfluss.

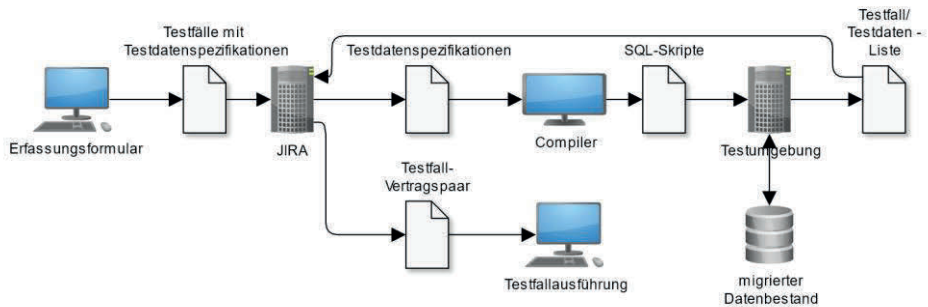


Abb. 4: Systemübersicht

Bei Inbetriebnahme des Verfahrens stellte sich heraus, dass für viele Testdatenspezifikationen keine passenden Testdaten gefunden wurden, weil die Testfallersteller nicht darauf geachtet hatten, dass es zu den Testdatenspezifikationen Verträge im Bestand gibt. Da die zugehörigen Testfälle ohne Testdaten nicht ausgeführt werden, konnte die geplante Testabdeckung nicht erzielt werden. Mit folgenden Maßnahmen wird gegengesteuert:

Schulung der Testfallersteller, damit sie nur die für den Testfall relevanten Attribut/Schlüssel-Kombinationen definieren.

Automatische Vereinfachung der Testdatenspezifikationen während der Testdatensuche: es gibt zum Beispiel Vertragseigenschaften, die durch das Zielsystem hergestellt werden

können. Die zugehörigen Suchkriterien dürfen eliminiert werden, wenn sie Treffer verhindern. Der Vorgang wird protokolliert und ist transparent für die Testmanager. Beispiel: der Vertrag soll im Zustand „beendet“ sein. Dieses Kriterium kann entfernt werden und der Tester führt zu Beginn des Tests den Geschäftsvorfall „Vertrag beenden“ durch, damit der Vertrag für den Testfall geeignet ist.

Erweiterung des Erfassungsformulars um eine Zählfunktionalität, mit der die Anzahl der Treffer für eine Testdatenspezifikation ermittelt wird.

Seltene Datenkonstellationen müssen für die Testfälle reserviert werden, die sie benötigen. Es hat sich bewährt, die Zuteilung bei den Spezifikationen mit den wenigsten Treffern zu beginnen.

Teilen sich mehrere Testfälle eine Testdatenspezifikation und können sie nicht alle die benötigte Anzahl Testdaten mit den geforderten Eigenschaften bekommen, werden die verfügbaren Verträge gleichmäßig an sie verteilt.

Nach Teststart benötigen Testmanager oder Tester aus verschiedenen Gründen spontan Testdaten mit bestimmten Eigenschaften. Dazu gibt es im Erstellungsformular eine Reservierungsfunktionalität, mit der sie passende, konsistente Verträge reservieren.

3 Resultate

Folgende Erfahrungen gibt es; auf die grundlegenden Funktionen, wie Reservierung, Bereitstellung passender Verträge usw. wird nicht näher eingegangen.

Ohne Akzeptanz gibt es keine Spezifikationen, das musste manches modell- oder spezifikationsbasierte Verfahren in der Praxis erfahren (wenn er nicht grundsätzlich an der Komplexität der Aufgabe gescheitert ist). Obendrein ist eine aus Sicht eines Informatikers einfache Notation dies noch lange nicht für die Mitarbeiter der Fachabteilung. Der vorgestellte Ansatz wird in fast allen Testteams konsequent umgesetzt. Testteams, die anfangs ein anderes Verfahren wählten, wechselten im Laufe des Projekts zum vorgestellten Ansatz.

Ein Datenbereitstellungsverfahren in einem solch großen Projekt muss skalieren. Mit steigender Anzahl Testfälle nimmt lediglich die Laufzeit der Skripte linear zu. Zusätzliche Testteams werden bedient, ohne dass das Testdatenmanagement personell aufgestockt werden muss.

Das Verfahren liefert für die automatisierten Regressionstestfälle beliebig - sofern es der Bestand hergibt - viele passende Testdaten, mit denen Regressionstestläufe ausgeführt werden und schafft so die Voraussetzung für langfristig stabil laufende automatische Tests.

Zwei wichtige Testdokumentationsaspekte werden durch das Verfahren sichergestellt. Zum einen werden die Testdateneigenschaften im Testfall exakt mit den formalen Testdatenspezifikationen dokumentiert und zum anderen wird lückenlos automatisch dokumentiert, welche Verträge im Testlauf verwendet werden.

Die Datensparsamkeit und damit die Verwendung kleiner dimensionierter Testumgebungen werden unterstützt. Bei großen Datenbeständen kann es vorteilhaft werden, auf den Testumgebungen mit Teilbeladungen zu arbeiten. Die Ladeskripte laufen wesentlich schneller und Ausgaben für Lizenzen und Hardware für die Testumgebungen sind kleiner als bei der Verwendung des gesamten Produktionsdatenbestands. Mit dem Verfahren lässt sich vorab der exakte Testdatenbedarf ermitteln, was die Voraussetzung für diese Vorgehensweise schafft.

Da Testdaten automatisch genau einer Durchführung eines Testfalls zugeteilt werden, können viele Testteams eine Testumgebung ohne Konflikte nutzen.

Dateninkonsistenzen werden effizient abgefangen. Aufgrund von Migrationsproblemen in mehreren Teilsystemen waren bis zu ein Drittel aller Daten betroffen. Beispielsweise sind dann versicherte Personen nicht in der ERP-Software hinterlegt und dadurch gibt es bei der Testausführung einen Fehler. Die Kriterien bzw. die Schlüssel der betroffenen Verträge werden in den Skripten hinterlegt und diese bei der Datensuche ausgeschlossen.

Der Code-Umfang für die Implementierung beträgt über 12.000 Lines of Code in Groovy, SQL, Shell-Skript und VBA.

4 Erfolgsfaktoren

Neben der erfolgreichen handwerklichen Umsetzung bedarf es der Akzeptanz in der Organisation, damit ein solcher Ansatz wirksam wird. Folgende Faktoren haben dazu beigetragen:

- die Einbeziehung der betroffenen Mitarbeiter während der Konzeption
- die Wahl einer Domain Specific Language für die Notation
- das komfortable Erstellungsformular mit Zählfunktion für schnelles Feedback für die Ersteller und Reservierungsfunktion für die schnelle Lösung von Testdatenproblemen durch das Testmanagement

- die automatische Zuweisung der Verträge beim Anlegen der Testausführungen
- die Unterstützung der Testautomation

Es braucht Erfahrung, um gute Testdatenspezifikationen zu erstellen, die einerseits zum Testfall passen und die Testabdeckung erhöhen und andererseits passende Testdaten liefern. Daher müssen Schulungen erfolgen und eine Lernkurve einkalkuliert werden, bis die Trefferquote der Testdatenspezifikationen akzeptabel wird.

5 Zusammenfassung

Die in das spezifikationsbasierte Testdatenmanagement gesetzten Erwartungen werden erfüllt. Neben technischen Aspekten tragen auch „weiche“ Faktoren für die Herstellung der Akzeptanz bei den Beteiligten zum Erfolg bei. Der vorgestellte Ansatz orientiert sich dabei eng an den vorgegebenen Prozessen und der vorhandenen technischen Infrastruktur. Die exakten Testdatenspezifikationen erlauben die Pflege der Testfälle und das Beschaffen von Testdaten auch noch dann, wenn die Ersteller der Testfälle nicht mehr zur Verfügung stehen. Das schafft nachhaltige automatisierte Regressionstests über die Laufzeit des Migrationsprojektes hinaus.

6 Danksagungen

Besonders bedanke ich mich bei Mario Friske für den Anstoß zu diesem Beitrag, Peter Weber für anregende Diskussionen in der Konzeptionsphase, den Kollegen im GI-AK TOOP für inspirierende Diskussionen, sowie Tanja Tremmel, Eckehard Kruse, Klaus Franz und Michael Kurz aus der Arbeitsgruppe ASQF Testdatenmanagement für viele Diskussionen und Anregungen in den vergangenen Jahren.

7 Literatur

[Tr18] Franz, K.; Tremmel, T.; Kruse, E.: Basiswissen Testdatenmanagement. dpunkt.verlag, Heidelberg, S. 114-120, 2018.

QA-Strategie oder Flickenteppich?

Herausforderungen an Teststrategien für moderne Softwaresysteme

Ralf Somplatzki¹

Abstract: Moderne Software-Projektstrukturen stellen veränderte und neue Anforderungen an das Testvorgehen. Aufgaben aus den Bereichen Entwicklung und Softwaretest nähern sich symbiotisch an. Die klassischen Rollenbilder verschwimmen. Was ermöglicht agilen Teams, dem Spagat der Verantwortung für das funktionale Ergebnis, wie auch der erwarteten Qualität gerecht zu werden? Wie kann Qualität in hybriden Systemlandschaften gewährleistet werden? Der Beitrag sucht Antworten auf diese und weitere Fragen, mit denen sich heutige Software-Dienstleister auseinandersetzen. Es wird ein möglicher Ansatz aufgezeigt, um mit der Vielgestalt der Projekte, geprägt durch Umfang, Teamstrukturen und Kundenkulturen, umzugehen.

Keywords: Qualitätssicherung, Softwaretest, Teststrategie, Testmanagement, Vorgehensmodell, Projektstruktur, Agile Softwareentwicklung, SCRUM.

1 Motivation

Die Qualitätssicherung (QA) moderner Softwareprojekte ist über die Evolutionsstufen der Softwareentwicklung in den vergangenen Jahren erheblich anspruchsvoller geworden. Verteilte Systeme, agile Methoden und kultureller Wandel fordern neue Konzepte zur Qualitätssicherung, die über reinen Softwaretest hinaus gehen. Benötigt wird eine Strategie, die der Mannigfaltigkeit moderner Softwareprojekte Rechnung trägt und dabei flexibel und leichtgewichtig bleibt.

2 Ausgangslage

Software-Dienstleister stehen vor der Herausforderung, unterschiedliche Projekt-Setups ebenso wie Kultur und Prioritäten der Kunden, in ihrer Vielschichtigkeit im eigenen Unternehmen zu vereinen. Gleichzeitig ist es aus Gründen der Effizienz und der Pflege einer eigenen Unternehmenskultur und -identität erforderlich, einheitliche Vorgehensweisen zu etablieren.

¹ Ralf Somplatzki/GEBIT Solutions GmbH, Oststraße 10, D-40211 Düsseldorf, ralf.somplatzki@gebit.de

3 Zielsetzung

Eine moderne QA-Strategie für Software-Dienstleister muss unterschiedliche Projektausprägungen unterstützen. Das Spektrum reicht von überschaubaren Dienstleistungsaufträgen mit einem kleinen Team über große, langlaufende Projekte mit verteilten Teams bis hin zu komplexen Lösungsszenarien in Kollaboration zahlreicher Dienstleister und Lieferanten.

An erster Stelle steht nicht, einzelne Tätigkeiten zu optimieren oder Tools einzuführen, sondern geeignete Methoden und Strategien zu finden, die für heutige Projektsituationen geeignet sind.

4 Herausforderungen

Dieser Beitrag adressiert Problemfelder im Kontext der Qualitätssicherung, die im Zusammenhang mit Entwicklungsmethoden und modernen Softwaresystemen auftreten.

Als Lösungsansatz für die heutige Vielschichtigkeit wird mehr als die eine, „richtige“ Teststrategie benötigt. Es ist ein adaptierbarer Ansatz erforderlich, der es ermöglicht, unter einem „Unternehmensdach“ vielgestaltig und dennoch kompatibel zu arbeiten. Dabei sind insbesondere die folgenden Aspekte zu berücksichtigen:

4.1 Unterstützung der Eigenständigkeit der Teams

Agile Konzepte setzen darauf, dass Teams in Eigenverantwortung für ein Ergebnis „von Wert“ (für die Anforderer) stehen. In diesem Kontext ist die Einbeziehung des Teams in die Prozessgestaltung ein unverzichtbares Konzept, auch in Bezug auf die Qualitätssicherung.

4.2 Beachtung der Ziele agiler Strategien, insbesondere „Time-To-Market“

Unter anderem bedingt durch hohen Wettbewerbsdruck der beauftragenden Unternehmen erfreuen sich agile Ansätze einer großen Beliebtheit. Mit „agil“ wird dabei assoziiert, gewünschte Lieferungen zu einem bestmöglichen, nahen Termin erwarten zu dürfen.

Um dies in der Praxis tatsächlich zu ermöglichen, ist zu beachten, dass eine zeitnahe Lieferung nur dann sinnvoll ist, wenn das Produkt „von Wert“ ist. Also die gestellten fachlichen Anforderungen erfüllt. [SS20]

Hier setzt die Herausforderung für die Qualitätssicherung deutlich früher an, als zum Zeitpunkt der Softwareentwicklung. Allzu häufig stellt sich mit der Lieferung heraus, dass Erwartungshaltungen und Lieferqualität oder -umfang voneinander abweichen.

4.3 Wandel von Rollenbildern, Skills und Verantwortung

Gegenüber vertikalen Verfahren, wie dem Wasserfallmodell, werden zahlreiche Disziplinen im Rahmen eines Softwareprojektes in einem Team vereint, das aus gleichberechtigten Teilnehmern zusammengesetzt ist. Dies führt dazu, dass auch das Thema QA im Verantwortungsbereich des Teams liegt. Die Aufgabenfelder Softwareentwicklung und Softwaretest wachsen zusammen und stellen neue Anforderungen an die Teammitglieder.

5 Lösungsansatz und Ergebnisse

Für eine vielgestaltige Projektlandschaft werden modulare Methodeneinheiten bereitgestellt. Zwei Schlüssel-Elemente versetzen Teams und Projektverantwortliche in die Lage, dem jeweiligen Projektsetup gerecht zu werden:

- QA-Module (Methoden-Beschreibungen, Vorlagen und Trainingseinheiten)
- Projekt-Schablonen

Die QA-Module beschreiben, die für die Projekte eines Unternehmens in Frage kommenden Vorgehensweisen und Techniken. Projekt-Schablonen stellen beispielhafte Projektmodelle dar, unter Einbeziehung der vorhandenen QA-Module. Im Idealfall kann ein Projektteam eine Projekt-Schablone auf ein eigenes Projekt anwenden und dabei individuelle Anpassungen an den eingesetzten QA-Modulen vornehmen.

Darüber hinaus werden messbare Kriterien benötigt, um Qualität bewerten zu können [Li09]. Sie lassen sich zum Beispiel aus den Testmanagement-Aspekten des ISTQB herleiten. Entscheidend ist hier, dass alle Projektbeteiligten die Messkriterien verstehen, um eine gemeinsame Einigung herstellen und tragen zu können.

Nach diesem Konzept werden einzelnen Etappen und Meilensteinen des Entwicklungsprozesses Qualitäts-Messpunkte zugeordnet. Abb. 1 zeigt ein minimalistisches Beispiel:

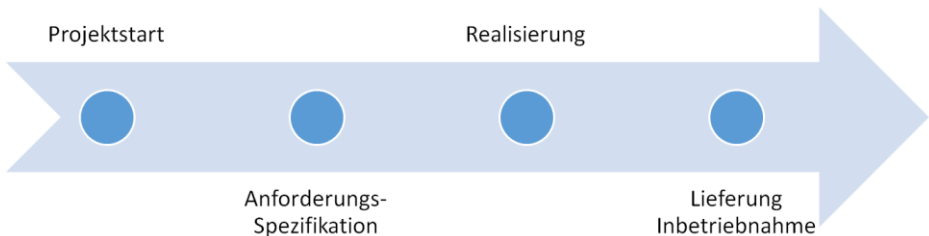


Abb. 1: Qualitäts-Messpunkte im Projekt-Lebenszyklus

Zur Veranschaulichung eine Aufstellung möglicher Messpunkte und Messkriterien. Diese können 1-zu-1 adaptiert oder je nach Kundenanforderungen und weiteren Rahmenbedingungen angepasst werden.

Messpunkt	Messparameter
Projektstart	Bestimmung der Wichtung der Qualitätsparameter: Zeit, Funktionalität, Qualität, Kosten
Anforderungen (DoR)	Erforderlicher Detailgrad, Verständlichkeit, Abnahme durch welche Rollen?
Tagesabschluss (Build)	Fehlertoleranz (z.B. Wertung von Warnungen)
Sprintergebnis (DoD)	Umfang erforderlicher Artefakte, wie kommentierter Sourcecode, Spezifikation, Testfallbeschreibungen, Review-Abnahme, Testabdeckung
Produktionskandidat	Funktionale Testergebnisse, Testabdeckung nach Teststufen, kommentierte Testberichte

Tab. 1: Beispiele für Messpunkte und Messparameter

Für alle Messpunkte werden Vorlagen zur Verfügung gestellt, die an die projektspezifischen Bedürfnisse angepasst werden. Auf diese Weise werden Erwartungshaltungen schon vor Produktionsbeginn abgestimmt und können in weiteren Evolutionsstufen, zum Beispiel bei jeder Retrospektive überprüft und angepasst werden [Be01]



Abb. 2: Kontinuierlicher Abgleich der Qualitätskriterien im Projekt-Lebenszyklus

Die folgenden Absätze zeigen auf, wie das Lösungskonzept den Herausforderungen, aus Kapitel 4, Rechnung trägt:

5.1 Unterstützung der Eigenständigkeit der Teams

QA-Module unterstützen Teams dabei, im Hinblick auf Umfang und Nutzen von QA-relevanten Aspekten eigenverantwortlich zu agieren. Als Beispiele seien „DoR“ („Definition of Ready“) und „DoD“ („Definition of Done“) genannt. Während die „DoR“ Qualitätskriterien für Anforderungen festlegt, bestimmt die „DoD“, wann eine Aufgabe (ein Teilprodukt) abgeschlossen/fertiggestellt ist. Darüber hinaus bieten QA-Module

Vorlagen für Testverfahren, Techniken und Werkzeuge, über deren Einsatz die Teammitglieder autark entscheiden können.

5.2 Beachtung der Ziele agiler Strategien, insbesondere „Time-To-Market“

Im Rahmen einer tragfähigen QA-Strategie einigen sich die Projekt-Beteiligten über die Kriterien zur Definition „von Wert“. Zum Beispiel ist für einen Kunden, der ein neues Softwareprodukt entwickeln lässt, „von Wert“, dass sein Produkt zum Termin einer Messe vorführbar ist ohne vollständig zu sein. Für einen Kunden, der eine Softwarelösung für den produktiven Einsatz benötigt, ist das Produkt „von Wert“, wenn es stabil und zuverlässig die erforderlichen Standard-Geschäftsvorfälle unterstützt.

QA-Module ermöglichen Teams, im Hinblick auf die abgestimmte Erwartungshaltung, geeignete Elemente zur Qualitätssicherung auszuwählen. In Bezug auf das Beispiel „Messemmodell“ kann ein minimalistisches Automatisierungsverfahren und für die Vorführszenarien ein explorativer Testansatz gewählt werden.

5.3 Wandel von Rollenbildern, Skills und Verantwortung

Um die Eigenständigkeit der Teams zu unterstützen hat sich die Integration von Teammitgliedern als sinnvoll erwiesen, die im Bereich Softwaretest ausgebildet und erfahren sind. Dies führt dazu, dass in einem ausgewogenen Team-Setup nicht nur Wissen über konkretes Testvorgehen, sondern auch das Verständnis („Mindset“) in Bezug auf Qualität, und damit die Kultur, entwickelt wird.

Die QA Module helfen den Teams, die Vielfältigkeit in der SW QA zu überblicken. QA-erfahrene Teammitglieder sind nicht nur auf SW-Test fokussiert, sondern unterstützen die Teammitglieder in ihrer QA-Entwicklung.

6 Lösungsbeispiel

Zur Veranschaulichung des Lösungsansatzes werden drei Projektvarianten angenommen²: Ein präsentationsfähiges Pilotsystems, das lediglich Funktionen bereitstellt, die für eine weitergehende Entscheidung relevant sind, ein Softwareprojekt, das eine kundenspezifische Lösung implementiert und ein Standardprodukt, das Drittanbietern („Consumer“) den Zugriff auf bereitgestellte Funktionen ermöglicht.

Zu Beginn des Projektes werden die Qualitätsaspekte priorisiert (Abb.3).

² Für das Beispiel wird ein Prozess in Anlehnung an das SCRUM-Modell angenommen. Das zugrunde liegende Konzept der Adaptierbarkeit lässt sich aber auch auf andere Vorgehensweisen anwenden.

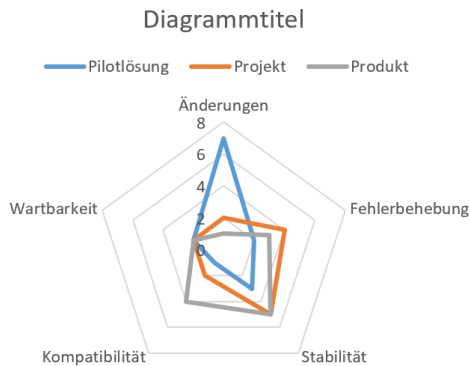


Abb. 3: Gewichtung von Qualitätskriterien in Relation zur Projektart

Anhand der jeweiligen strategischen Ausrichtung werden Projektschablone und QA-Module gewählt und an den Projektbedarf angepasst. Teams sind somit in der Lage, sowohl Unternehmens-Standards einzusetzen und dennoch das eigene Projektsetup zu gestalten. Zum Beispiel für die Pilotversion auf eine offene „DoR“, für das Kundenprojekt auf die Mindestanforderungen an Spezifikationsartefakte und für die Service-Komponente auf eine weitere Teststufe „Contract“ [CI14] Wert zu legen.

7 Fazit und Ausblick

Die Lösung versteht sich als Basis, als „Methoden-Bibliothek“, die beliebig erweitert oder zugeschnitten werden kann. In diesem Sinne ermöglicht die Methodik allen Teammitgliedern, an der Gestaltung und Nutzung der Bibliotheks-Elemente mitzuwirken.

In Anlehnung an den Titel dieses Beitrages liegt es in der Hand der Beteiligten, kein Flickwerk, sondern eine abgestimmte QA Strategie zu gestalten, die auf das jeweilige Projekt strategisch zugeschnitten ist.

8 Literaturverzeichnis

- [Be01] Beck, K. et.al.: Principles behind the Agile Manifesto; Principle #12; <https://agilemanifesto.org/principles.html>; Stand: 16.06.2021
- [CI14] Toby Clemson, Testing Strategies in a Microservice Architecture, <https://martinfowler.com/articles/microservice-testing/>. Stand: 15.06.2021
- [Li09] Liggesmeyer P. (2009) Software-Messung. In: Software-Qualität. Spektrum Akademischer Verlag. https://doi.org/10.1007/978-3-8274-2203-3_7. Stand: 13.06.2021
- [SS20] Schwaber, K.; Sutherland, J.: The Scrum Guide, S. 3&11, 2020

Softwaretest-Umfrage 2020 – (Wie) passt die Forschung zum Bedarf der Praxis?

Mario Winter ¹, Karin Vosseberg², Frank Simon³

Abstract: Unter wissenschaftlicher Leitung der TH Köln und der Hochschule Bremerhaven hat das German Testing Board (GTB e.V.) zusammen mit dem Austrian Testing Board (ATB) 2020 zum dritten Mal seit 2011 und 2015/16 die Umfrage „Softwaretest in Praxis und Forschung“ im deutschsprachigen Raum durchgeführt. Über 1.250 Teilnehmende beantworteten Fragen zur Qualitätssicherung aus Sicht des Managements, der operativen Umsetzung und der Forschung. Der Beitrag vergleicht zunächst einige aktuelle Ergebnisse zur Softwaretest-Forschung mit denen aus der Umfrage 2015/16. Hauptsächlich untersucht er dann, inwieweit die aktuellen Antworten zur Forschung dem von der Praxis geforderten Forschungsbedarf entsprechen. Den Fragebogen zur Softwaretest-Forschung haben über 250 Personen aufgerufen. Da jedoch nur 46 Teilnehmer*innen alle Fragen beantwortet haben, konnten wir die Umfrage zur Softwaretest-Forschung nur deskriptiv analysieren. Dennoch haben wir einige interessante Beobachtungen gemacht.


- **Keywords:** Softwaretest; Qualitätssicherung; Stand der Technik; Forschung; Umfrage.

Einleitung

Im Sommer 2020 wurde bereits zum dritten Mal nach 2011 [WSV+12] und 2015/16⁴ [WSV17] (s.a. [StU21]) unter wiss. Leitung der TH Köln und der Hochschule Bremerhaven eine der größten Umfragen zu Test und Qualitätssicherung im deutschsprachigen Raum durchgeführt. Ideeller, personeller (und finanzieller) Förderer ist das German Testing Board (GTB e.V.) zusammen mit dem Austrian Testing Board (ATB).

Die drei Fragebögen aus Sicht des Managements, der operativen Umsetzung und der Forschung waren vom 5. Mai bis zum 30. Juni 2020 online. Der Fragenkatalog wurde 2020 bewusst ähnlich zu den Softwaretest-Umfragen 2011 und 2016 gehalten, damit unsere aktuellen Analysen die damaligen Einblicke in aktuelle und zukünftige Trends und Herausforderungen rund um die Qualitätssicherung von Software und Systemen fortführen.

Die Ergebnisse sollen auf der einen Seite einem möglichst breiten Interessentenkreis erlauben, Impulse für die praxis- und forschungsorientierte Ausrichtung der Aus- und Weiterbildung sowie Qualifizierung abzuleiten, sowie auf der anderen Seite den Unternehmen Grundlagen für ein Benchmarking zur Verfügung stellen. Eine zusammenfassende Broschüre sowie Diagramme mit den Antwortverteilungen zu allen

¹ TH Köln, Institut für Informatik, Steinmüllerallee 1, 51643 Gummersbach, mario.winter@th-koeln.de, 

² Hochschule Bremerhaven, Fachbereich Management und Informationssysteme, An der Karlstadt 8, 27568 Bremerhaven, kvosseberg@hs-bremerhaven.de

³ German Testing Board e.V., Leitung Business Development, Koldestraße 8 b, 91052 Erlangen, f.simon@gtb.de

⁴ Die Umfrage 2015/16 wird im weiteren Text als Umfrage 2016 aufgeführt.

Fragen sind im Internet auf [StU21] veröffentlicht. Der technische Report zur Umfrage [WVS+21] enthält detaillierte Analysen und die vollständigen Daten aller Antworten.

Insgesamt zählten wir über 1.250 Teilnehmende. Den Fragebogen zur Softwaretest-Forschung haben über 250 Personen aufgerufen. Da im Mittel zu den Fragen nur ca. 50 Antworten vorliegen und nur 46 Teilnehmer*innen alle Fragen beantwortet haben, konnten wir diese nur deskriptiv analysieren. Wir vermuten, dass viele Forscher*innen die Umfrage aufgrund des Aufwands für die COVID-19 bedingte Online-Lehre und des Ausfalls einschlägiger Fachtagungen nicht wahrgenommen haben. Dennoch haben wir einige interessante Beobachtungen gemacht.

In Kapitel 0 vergleichen wir zunächst einige aktuelle Ergebnisse zur Softwaretest-Forschung mit denen aus der Umfrage 2016. Danach untersuchen wir in Kapitel 2, inwieweit die aktuellen Antworten aus der Forschung dem von der Praxis geforderten Forschungsbedarf entsprechen. Dazu stellen wir dem Forschungsbedarf anhand der Antworten der operativ Umsetzenden die entsprechenden Antworten der Forscher*innen gegenüber und analysieren, inwieweit sich aktuelle Themen, Fragestellungen und Probleme in Praxis und Forschung überschneiden. Ein kurzes Fazit schließt den Beitrag ab.

Forschung 2020 vs. Forschung 2016

Zur Einstimmung in die aktuellen Entwicklungen und Forschungsthemen zeigt die Begriffswolke in Abb. 1 häufige Begriffe in den Titeln von Konferenzbeiträgen und Zeitschriftenartikeln aus den Jahren 2020-21 zum Thema.

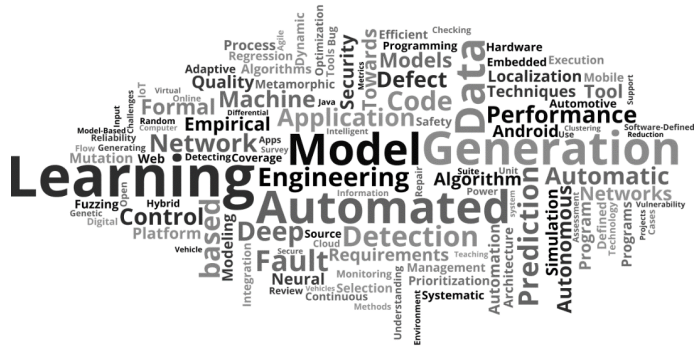


Abb. 1 Begriffswolke IEEE Explore Recherche 2020-2021

Erzeugt wurde die Wortwolke aus den ersten 2000 Treffern für den Suchvektor (software AND (testing OR validation OR verification)) in der digitalen Bibliothek der IEEE. Die Größe entspricht dabei der Häufigkeit des Begriffs, Ausrichtung und Anordnung sind zufällig. Die Suchbegriffe wurden entfernt. Im Vergleich zur Wortwolke von 2016 ([WSV17]) sind neue prominente Begriffe wie z. B. *Learning*, *Generation*, *Data* und *Automated* auffällig. Davon war 2016 nur *Data* prominent und *Generation* mittelgroß vertreten, am prominentesten waren 2016 jedoch *Embedded*, *Reliability*, *Engineering*, und *Process*.

1.1 Forschung: Finanzierung, Instituts- und Projektgröße

Bei 43% (2016: 44%) der Befragten werden die Forschungsprojekte überwiegend (>80%) aus öffentlichen Mitteln finanziert. Eine Mischfinanzierung geben 25% (2016: 30%) an und nur bei 13% (2016: 11%) überwiegen (>80%) privat geförderte Projekte. 30% (2016: 34%) sehen die Ausrichtung ihrer Forschung als ausgewogen zwischen Auftrags- und Grundlagenforschung an. 22% (2016: 31%) konzentrieren sich auf Auftragsforschung und 27% (2016: 28%) auf Grundlagenforschung. Im Bereich der Fördergeber und der Forschungsausrichtung gab es somit von 2016 bis 2020 nur unwesentliche Unterschiede, wobei anzumerken ist, dass 18% (2016: 15%) keine Angabe bei der bei der Finanzierung und 22% (2016: 8%) keine Angabe bei der Ausrichtung der Forschung gemacht haben.

24% (2016: 36%) der Forschungsinstitute haben 1-9 Mitarbeiter*innen, bei 22% (2016: 32%) sind es 10-49. Bei nur 7% (2016: 6%) sind es 50-99 Mitarbeiter*innen, und 22% (2016: 14%) der Forschungsinstitute haben mehr als 100 Mitarbeiter*innen. 24% (2016: 12%) machten keine Angabe zur Anzahl der Mitarbeiter*innen. Zu der Projektgröße bzw. der Anzahl der Projektmitarbeiter*innen zeigt Abb. 2 die Ergebnisse. Interessant ist auch hier wieder der Blick auf die Kategorie „keine Angabe“: Bei der Mitarbeit*innen-Anzahl sind es 24% (2016: 12%), bei der Anzahl Personen im größten Forschungsprojekt sind es 30% (2016: 15%) und bei denen im kleinsten Projekt sind es 24% (2016: 12%). Hat die Arbeit im (COVID-19-bedingten?) Homeoffice den Blick auf die Mitarbeiter*innen getrübt?

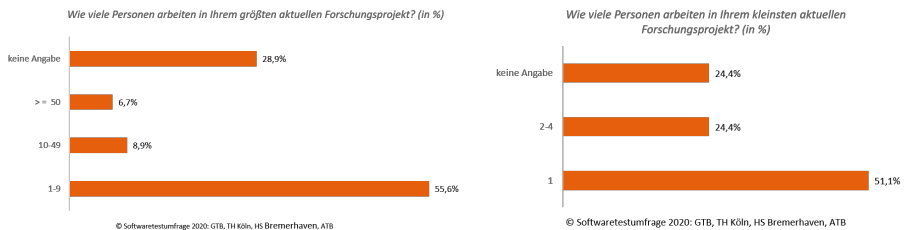


Abb. 2 Anzahl Personen in den Projekten

Zusammenfassend existiert auch 2020 eine ausgewogene überwiegend öffentlich geförderte Forschung im Bereich der Qualitätssicherung, die weitgehend in Ein-Personen-Projekten durchgeführt wird. Unsere Vermutung lautet: Softwaretest-Forschung findet in Form von Promotionsprojekten statt, große Leuchtturmprojekte fehlen weiterhin!

1.2 Forschung: Qualitätsmanagement

Die Frage nach dem Bedarf an Forschung zu Qualitäts-, Vorgehens- und Verbesserungsmodellen wurde 2020 erweitert, so dass für den Vergleich zu 2016 einige Kategorien geeignet zusammengefasst werden müssen. In den Antworten zum **Fehler! Verweisquelle konnte nicht gefunden werden.** fällt im Vergleich zu 2016 auf, dass der Forschungsbedarf zu herkömmlichen organisatorischen Aspekten des Testens rückläufig ist. Die Forschung in diesem Bereich sehen 2020 z.B. zu Crowd Testing nur noch 32%

der Antwortenden als zu gering an (2016: 53%), verbunden allerdings mit einem Anstieg der Kategorie „keine Angabe“ von 37% (2016: 30%). Klare Bedarfs-Gewinner sind hier die Forschung zu organisatorischen Aspekten bei DevOps, die von 54% als zu gering angesehen wird (2016: nicht gefragt) und bei agilen Vorgehensmodellen (für Teams) mit 48% (2016: 57%). Auch die Forschung zu organisatorischen Aspekten agiler Vorgehensmodelle für Multiteams (SAFe, LeSS, DAD, Nexus, ...) wird von 32% (2016: nicht gefragt) und zu speziellen agilen Prozessbewertungs- und Prozessverbesserungsmodellen (AAM, Adept, AAIM, ...) von 27% (2016: nicht gefragt) der Teilnehmer*innen als zu gering angesehen.

In Bezug auf bestimmte Qualitätsmaßnahmen (Aufgaben, Aktivitäten bzw. Nachweise) wird die Forschung bzgl. dem Nachweis der Wirksamkeit und Effizienz der Qualitätsmaßnahmen von 57% (2016: 68%) als zu gering angesehen. 48% sehen die Forschung zur Vergleichbarkeit der unterschiedlichen Qualitätsmaßnahmen (Benchmarking) als zu gering an (2016: 56%), ebenso sehen 48% dies in der Forschung zum Einsatz von KI zur Unterstützung des Qualitätsmanagements (2016: nicht gefragt).

2 Forschung 2020 vs. Forschungsbedarf der Praxis

2.1 Forschungstransfer

Zufrieden mit den Forschungsaufträgen und Ideen aus der Praxis sind 21% (2016: 22%) der Forscher*innen, 27% (2016: 31%) sehen diese als angemessen an. Der umgekehrte Weg, also die Annahme und Umsetzung ihrer Forschungsergebnisse durch die Praxis, wird nur noch von 26% (2016: 30%) der Forscher*innen als angemessen angesehen, und nur noch 17% (2016: 22%) sind damit zufrieden. Unzufrieden mit der Annahme und Umsetzung ihrer Forschungsergebnisse durch die Praxis sind 2016 mit 21% (2016: 27%) etwas weniger Forscher*innen. Hier liegt tatsächlich weiterhin Verbesserungspotenzial, denn die Unzufriedenheit von 25% (2016: 25%) der Forscher*innen mit den Aufträgen aus der Praxis ist immer noch recht hoch.

Ein Grund dafür kann die weiterhin recht geringe Attraktivität (oder Bekanntheit?) der Kursunterlagen und Forschungsberichte von Hochschulen für Tester*innen und Entwickler*innen sein, von denen lediglich 6% angeben, diese zur Informationsbeschaffung meistens oder sogar immer zu nutzen (2016: 6%). Als meistgenutzte Quelle geben Tester*innen und Entwickler*innen wie auch bereits 2016 das Internet (Foren, Blogs etc.) mit 81% an (2016: 87%). Nur noch 36% (2016: 48%) nutzen Bücher oder Zeitschriften, wohingegen der Anteil von 21% (2016: 21%), die zu ihrer Informationsbeschaffung Konferenzen nutzen, konstant geblieben ist.

2.2 Forschung zu speziellen Testarten und Testverfahren

Wir greifen nun einige Antworten auf Fragen zu Forschungsaktivitäten in Bezug auf Testarten und Testverfahren bei speziellen Testansätzen bzw. -ausrichtungen heraus. Als

zu gering bezeichnen 60% der Forscher*innen die Anstrengungen im Bereich des Testens von KI-Systemen, gefolgt vom Security Testing, bei dem dies 56% so sehen. Performance Testing sehen 38% als zu gering erforscht an, und Usability Testing wird von 33% als zu gering erforscht betrachtet.

Die Antworten der Praktiker*innen hinsichtlich der Berücksichtigung von Testarten und Testverfahren passen recht gut zu dem Bild der Forscher*innen. So antworten in fast allen Branchen ca. 55%, dass sich Security Testing selten oder nie durchführen – hier erscheint der mit 56% hoch eingeschätzte Forschungsbedarf in Hinblick auf praktikable Verfahren angemessen. Performance Testing und Usability Testing wird von jeweils ca. 45% selten oder nie durchgeführt. Auch hier sehen die Forscher*innen einen erhöhten Bedarf an Forschungsaktivitäten. Für den Test von KI-Systemen sehen sich 44% der Praktiker als schlecht oder sehr schlecht gewappnet, so dass hier die Forscher*innen den sehr hohen Bedarf an Forschung angemessen einschätzen.

2.3 Forschung zu agilen Praktiken

43% der Forscher*innen sehen zu geringe Forschungsaktivitäten im Bereich der Integration des Testens in agilen Vorgehensweisen, zum CI/CD als agile Praktik sowie zu entsprechenden methodischen Ansätzen für den agilen Test. Mit 37% etwas geringer ist der Anteil derer, die das auch für die organisatorischen Aspekte von DevOps so sehen. Die Antworten aus der Praxis lassen Forschungsbedarf im Bereich CI erkennen, da 51% der Tester*innen und Entwickler*innen angeben, immer oder meist einen CI-Server für den Integrationstest einzusetzen.

Ein tieferer Blick auf die Forschung im Bereich agiler Praktiken ergibt zwiespältige Ergebnisse. So sehen zwar 38% der Forscher*innen die Forschung zu TDD (test driven development) als zu gering an, 31% sehen dies auch beim ATTD (acceptance-test driven development) und 27% bzw. 29% auch bei BDD (behavior driven development) und bei „specification by example“. Für das explorative Testen sehen nur knapp 23% der Forscher*innen die Forschung als zu gering an. Auf der anderen Seite machen jeweils deutlich über 30% der Forscher*innen keine Angabe zu diesen Themen. Ausnahme ist hier lediglich TDD, zu dem nur 23% keine Angabe machten.

In der Praxis ist der anwendungsfallbasierte Test mit 71% immer noch das am häufigsten eingesetzte Testverfahren. Mit über 82% sehen die meisten Tester*innen und Entwickler*innen die Testautomatisierung als die bedeutendste Praktik zur Qualitätssicherung an, dicht gefolgt von Code Reviews (79%) und CI (75%). Exploratives Testen ohne Testcharta wird von 33% immer oder meist eingesetzt, jedoch nur 17% setzen immer oder meist das explorative Testen ohne Testcharta ein. Gerade im Bereich exploratives Testen scheinen die Forscher*innen den Forschungsbedarf in der Praxis deutlich zu unterschätzen.

2.4 Zukünftiger Forschungsbedarf

Erheblich erhöhten zukünftigen Forschungsbedarf zur Qualitätssicherung sehen 52% der Forscher*innen im Bereich autonomer KI-Systeme, gefolgt von 47% zu IoT und Industrie 4.0 und 45% zur QS von kontinuierlich vor und während des Einsatzes lernender Systeme sowie 43% im datenschutzrelevanter Systeme/Security. Es folgen eingebettete Systeme, sicherheitskritische Systeme und Big Data/Smart Data, bei denen jeweils über 33% erhöhten zukünftigen Forschungsbedarf sehen. Den Forschungsbedarf zur Qualitätssicherung selbstlernender Algorithmen und Systeme, die nur vor dem Einsatz lernen, sehen lediglich 29% der Forscher*innen als zukünftig erhöht an.

Wie sieht sich die Praxis für diese zukünftigen Herausforderungen gewappnet? Lediglich 14% geben an, dass sie sich für Herausforderungen im Bereich autonomer Systeme gut oder sehr gut gewappnet sehen, 18% sehen das so für KI, 21% für IoT, 23% für Industrie 4.0, 24% für Big Data, und 26% für eingebettete Systeme. Hier scheint die Forschung die richtigen Themen anzugehen. Nur der von den Forscher*innen gesehene erhöhte Forschungsbedarf für Security (43%, 2016 Safety/Security: 51%) passt nicht ganz zur Meinung der Tester*innen und Entwickler*innen, die sich im Bereich Safety/Security gut (35%, 2016: 36%) bzw. angemessen (21%, 2016: 23%) gewappnet sehen.

Resümé

Die (relativ) geringe Anzahl der Teilnehmer*innen von knapp 50 beim Forschungs-Fragebogen lässt natürlich nur eingeschränkte Aussagen zu. Die Forscher*innen sehen noch erheblichen Bedarf an fundierten Erkenntnissen im Bereich der Qualitätssicherung. Dies betrifft zum einen die Auswahl, die Wirksamkeit und die gegenseitigen Abhängigkeiten von einzelnen Maßnahmen zur Qualitätssicherung allgemein, zum anderen die agilen Praktiken und Vorgehensweisen.

Im Bereich Qualitätssicherung scheinen Promotionen die Treiber der Forschung zu sein. Die großen Leuchtturm-Projekte fehlen. Mit Ein-Personen-Projekten, die die Mehrzahl der Forschungsprojekte im Gebiet der Qualitätssicherung derzeit ausmachen, wird der Durchbruch nur bedingt, in kleinen Schritten oder auch gar nicht gelingen. In einigen Bereichen wie z.B. Testautomatisierung und exploratives Testen scheinen die Forscher*innen den Forschungsbedarf in der Praxis deutlich zu unterschätzen. Autonome Systeme, das Internet der Dinge/Industrie 4.0 sowie künstliche Intelligenz sind die großen Herausforderungen in den kommenden Jahren, gefolgt von Safety und Security. Da auf den Gebieten auch erheblicher Bedarf an Grundlagenforschung besteht, sind größere Forschungsprojekte erforderlich.

Literaturverzeichnis

[WSV+12] Winter, M.; Spillner, A.; Vosseberg, K.; Haberl, P.: Softwaretest-Umfrage 2011 - Erkenntnisziele, Durchführung und Ergebnisse. In: Proc. Software Engineering (SE 2012), Berlin, 2012, GI-LNI, Vol. P-198, S. 157-168.

[WSV17] Winter, M.; Spillner, A.; Vosseberg, K.: Softwaretest in der Forschung – Ergebnisse der Umfrage 2015/16. 40. TAV-Workshop der GI-FG „Test, Analyse und Verifikation von Software“ (TAV), In: Softwaretechnik-Trends, Vol. 37, Nr. 1, Februar 2017.

[WVS20] Winter, M.; Vosseberg, K.; Simon, F.: Ergebnisse der Softwaretest-Umfrage 2020. 45. TAV-Workshop der GI-FG „Test, Analyse und Verifikation von Software“ (TAV), In: Softwaretechnik-Trends, Vol. 41, Nr. 1, Februar 2021.

[WVS+21] Winter, M.; Vosseberg, K.; Simon, F.; Sakarya, M.; Simon, A.: Technischer Report zur Umfrage 2020 “Softwaretest in Praxis und Forschung”. TH Köln und GTB e.V., 2021, (abrufbar ab Q3/2021 unter [StU21]).

[StU21] Softwaretest-Umfrage, www.softwaretest-umfrage.de/, Stand: 30.06.2021.

Schnellebige Qualitätssicherung im digitalen Ökosystem für mobile Dienste - vom ganz großen Schmerz, viralen Posts und dem Aufenthalt in der Bahn

Simon André Scherr¹, Frank Elberzhager², Adeline Silva Schäfer³

Zusammenfassung: Der digitale Wandel führt immer mehr zu komplexen digitalen Ökosystemen. Mobile Dienste sind vielfach Teil davon. Dadurch, dass heute fast jeder ein Smartphone besitzt, können solche Systeme für eine Vielzahl potenzieller Nutzer entwickelt werden, um Mehrwerte zu schaffen. Daher stellt sich die Frage, wie man die Qualität dieser sicherstellen kann. Dieser Artikel gibt einen Überblick über verschiedene Möglichkeiten der Qualitätssicherung im Ökosystem und geht auf unsere persönlichen Erfahrungen ein.

Schlüsselworte: Digitales Ökosystem, Qualitätssicherung, Mobile Apps, Nutzerfeedback.

1. Einleitung und Motivation

Heutige Softwaresysteme sind vielfach miteinander vernetzt. Insbesondere Apps auf Mobilgeräten, Smart Home und Internet of Things bestimmen einschneidend die Softwarelandschaft. Digitale Ökosysteme gehören zu den jüngsten Entwicklungen, welche die Komplexität heutiger Softwaresysteme noch einmal deutlich erhöhen, u.a. weil viele Systeme miteinander interagieren.

Die Frage, wie man trotz Budget- und Zeitvorgaben ein Produkt in der gewünschten Qualität bekommt, wird daher komplexer denn je. In diesem Beitrag wollen wir unsere Erfahrungen bei der Softwareentwicklung in digitalen Ökosystem aufzeigen und einen Fokus auf mobile Dienste legen. Dabei zeigen wir anhand von zwei Projekten, eines im ländlichen und ein anderes im urbanen Raum, wie wir einen Dienst entwickelt haben und wie dabei die Qualitätssicherung (QS) ausgesehen hat.

Bei der Produktentwicklung im Ökosystem müssen weitere Dinge beachtet werden, um erfolgreich zu sein. Die Charakteristiken der regulären Entwicklung bleiben natürlich bestehen, werden aber um die Dimension des Ökosystems erweitert. Da eine Anwendung nicht mehr der „Herrscher“ über das System ist, sondern verschiedene Systeme, die typischerweise von getrennten Teams entwickelt werden, nimmt die Fehlerwahrscheinlichkeit zu. Releasepläne müssen zwar aufeinander abgestimmt werden, dennoch arbeitet jedes Team unterschiedlich schnell. Dies stellt verglichen mit der

¹ Fraunhofer IESE, Fraunhofer-Platz 1, 67663 Kaiserslautern, simon.scherr@iese.fraunhofer.de

² Fraunhofer IESE, Fraunhofer-Platz 1, 67663 Kaiserslautern, frank.elberzhager@iese.fraunhofer.de

³ Fraunhofer IESE, Fraunhofer-Platz 1, 67663 Kaiserslautern, adeline.schaefer@iese.fraunhofer.de

klassischen Softwareentwicklung neue Anforderungen im Bereich des Systemtests und an die Kompatibilität der Teilelementen des Ökosystems. Durch die Heterogenität von Ökosystemen wird die Ursachenforschung für Fehlverhalten komplexer.

Aktuell forschen wir an der Frage, wie die Attraktivität des ländlichen Raums in Deutschland durch digitale Dienste gesteigert werden kann. Dazu haben wir, u.a. im Projekt Digitale Dörfer, in den letzten Jahren eine Plattform sowie unterschiedliche Dienste entwickelt, die zusammen ein Ökosystem für den ländlichen Raum bilden. Die entstandenen Lösungen werden in zahlreichen Regionen in Deutschland, Österreich und Belgien genutzt. Darüber hinaus sind wir dabei, ein digitales Ökosystem für ein klimaneutrales Stadtquartier zu entwickeln, in dem ebenfalls eine Plattform sowie digitale Dienste angeboten werden sollen („Pfaffprojekt“).

Seit Ende 2017 arbeiten wir am DorfFunk (siehe Abbildung 1) und hatten den ersten öffentlichen Release Anfang 2018. Seitdem hat DorfFunk im Schnitt ein Update im Monat erhalten. Im Pfaffprojekt wird der DorfFunk als PfaffFunk an die Bedürfnisse des Stadtquartiers angepasst.



Abb. 1: Startseite des DorfFunkes

2. Und es hat „Boom“ gemacht

Die Plattform DorfFunk mit 14.000 registrierten Nutzern Ende 2019, wurde mit Beginn der Covid-19-Pandemie in Deutschland den Gemeinden dreier Bundesländer kostenlos zur Verfügung gestellt. Infolgedessen verdreifachte sich die Anzahl der registrierten DorfFunk-Benutzer innerhalb von nur drei Monaten von 15.500 Ende Februar 2020 auf mehr als 49.000 Ende Mai.

Diese Benutzerexplosion hat sich besonders auf den Kundensupport ausgewirkt, sondern erforderte auch Anpassungen der App (z.B. die bundesweite Anzeige von Beiträge, die zuvor nur lokal in einem Radius von 20 km angezeigt wurden). Unsere klassischen QS-Ansätze kamen weiterhin zum Einsatz und wurden durch innovative Ansätze ergänzt; beides skizzieren wir nachfolgend, um ein Gesamtbild unserer Qualitätssicherung aufzuzeigen.

3. Klassische Qualitätssicherungsansätze

Zu Beginn der Entwicklung verließen wir uns bei der Qualitätssicherung stark auf unsere Systemtests, da das Team schnell neue Funktionen entwickeln musste. Mit der Zeit ersetzen wir den reaktiven durch einen proaktiven Ansatz, um Fehler zu verhindern, statt sie zu korrigieren.

Wir haben verschiedene Werkzeuge zur statischen Code-Analyse eingeführt, um die Codequalität zu verbessern. Da viele der Analysen sich direkt in den Compiler integrieren, werden die gefundenen Elemente sofort den Entwicklern in ihrer Entwicklungsumgebung angezeigt. Dadurch werden einige Fehler noch vor der eigentlichen QS Phase erkannt. Für verschiedene Probleme haben wir dazu auch eigene Analyseregeln entwickelt. Mit der Zeit konnten wir so Codequalität verbessern. Die compilerbasierten Analysen erlaubten uns auch einstige Warnungen auf das Niveau eines Fehlers anzuheben.

Im Rahmen unserer Bemühungen, die Qualität unserer App zu verbessern, fordert unsere „Definition of Done“ Unit-Tests und je nach Aufgabe auch UI Tests. Eine Aufgabe ist erst dann erledigt, wenn die entsprechenden Tests erstellt wurden. Wenn ein Fehler gefunden wird, müssen zusätzlich zum Fix neue Testfälle hinzugefügt werden. Mit unseren automatisierten UI-Tests können wir das System einschließlich der Interaktionen mit dem Backend vollständig testen. Allerdings sind die Erstellung und insbesondere Wartung von UI-Tests sehr teuer. Sie können jedoch sehr hilfreich sein, um zu verhindern, dass neue Funktionen vorhandene Funktionen beschädigen. Sie dienen daher nicht nur als Ende-zu-Ende-Tests, sondern auch als Regressionstests.

Neben Systemtests zur Sicherstellung der Qualität der App, haben wir weitere Aktivitäten gestartet. In jedem Sprint führen wir regelmäßig Systemtests durch. Diese Tests finden parallel zur Entwicklung statt. In der Regel wird der in Sprint n entwickelte Code in Sprint n+1 vollständig getestet. DorfFunk ist eine App, die mit zwei anderen Produkten in unserem Ökosystem interagiert. Deshalb testen wir sie gemeinsam. Wann immer Code, der diese Systeme verbindet, geändert wird – sei es in DorfFunk oder in den anderen Apps –, muss ein Integrationstest durchgeführt werden. Diese Tests ähneln unseren Systemtests, aber statt uns nur auf eine App zu konzentrieren, überprüfen wir mehrere Apps.

Abgesehen von Verbesserungen der Tests interner Versionen, haben wir ein Beta-Programm für alle DorfFunk-Releases gestartet. Dazu wird zunächst eine neue Produktivversion der Anwendung erstellt und in die zugehörigen App Stores geladen.

Anschließend wird die Version in den internen Test-Track (Android) bzw. in ein internes Test-Flight(iOS) gegeben. Hierbei erfolgen die Prüfung der Kompatibilität zur Backend-Version und die Durchspielung der Migration von vorhergehenden öffentlichen Releases auf die kommende Version. Daraufhin wird ein Android Beta-Programm mit geschlossenen Tracks gestartet bzw. ein externes Test-Flight Beta-Programm für iOS. Hier dürfen ausgewählte Nutzer einige Tage vor der breiten Masse die neue DorfFunk-Version testen. Dadurch identifizieren wir in erster Linie Kompatibilitätsprobleme in der praktischen Anwendung.

4. Innovative Methoden

Die Werkzeuge der klassischen Qualitätssicherung liefern einen großen Mehrwert für die App-Entwicklung, dieser kann jedoch durch das zusätzliche Beschreiten modernerer Wege wie Device Clouds, Szenario-Tests, Test Mobs und unserer „Eat your own dog food“-Qualitätssicherung gesteigert werden und zum schnellen Erreichen der gewünschten Qualität verhelfen. Einige dieser Methoden versetzen das Team in die verschiedensten Szenarien, um das eigene Produkt auf Herz und Nieren testen zu können. All diese Methoden werden immer wieder mit unseren Erfahrungen bezüglich der Gebiete, auf denen sie sich für Apps besonders gut bzw. weniger gut eignen, angereichert.

Test Mobs greifen eines der Trendthemen aus der agilen Softwareentwicklung, nämlich „Mob Programming“, auf und versuchen, es auf die Welt des Testens zu übertragen. Mob Programming erweitert Pair Programming um mehr als zwei Personen: Statt zu zweit vor einem Rechner zu sitzen, nimmt beim Mob Programming ein ganzes Team an der Lösung von Entwicklungsproblemen teil [KL16]. Das Team versammelt sich dabei um einen ausreichend großen Bildschirm und versucht, gemeinschaftlich mit nur einem Keyboard zu entwickeln. Das Team, bzw. der Mob, ist idealerweise bunt gemischt. Test Mobs setzen auf die gleichen Prinzipien. Wir nutzen Test Mobs seit Januar 2019 im Rahmen unserer regulären Qualitätssicherung und haben in [SES20] genauer darüber berichtet. Außerdem geben wir auf Konferenzen Tutorials zu Test Mobs und teilen unsere Erfahrungen. Bei Betrachtung der gefundenen Auffälligkeiten nach einem Test lässt sich über mehrere Monate als konstantes Teilnehmerfeedback festhalten, dass einzelne Personen nicht in der Kürze der Zeit zu einem derartigen Ergebnis gelangt wären. Durch die Test Mobs, die gruppenintern viel Freude verbreiten, werden gemeinschaftlich Szenarien erarbeitet, die einzelne Tester nicht abdecken können, zudem werden unterschiedliche Bereiche berücksichtigt. Test Mobs als Testverfahren funktioniert auch sehr gut in Zeiten von Home Office im Rahmen eines Gruppenvideoanrufs. Die Ziel- und Scopedefinition wird in einem geteilten Dokument gemacht, worin auch der Protokollant seine Notizen festhält. Der Driver gibt seinen Bildschirm für die Entdeckerphase frei.

Die meisten Testansätze werden in wesentlich kürzerer Zeit ausgeführt, als Endbenutzer zur Ausführung derselben Aufgabe benötigen. Dies ist einerseits gut, weil wir innerhalb von Minuten herausfinden können, welche dieser Funktionen nicht richtig funktionieren,

andererseits können wir keine Probleme feststellen, die nach Stunden oder Tagen des Gebrauchs auftreten. Um dieses Problem zu lösen, haben wir eine sogenannte „Langzeitqualitätssicherung“ entwickelt, bei der wir Tests durchführen, die nicht Teil eines Sprints sind, sondern eine kontinuierliche Nutzung der Anwendung darstellen, wie bei unseren Endbenutzern.

Hierzu erweiternd ist unsere „eat your own dog food“ QS. Wir haben eine Umgebung konfiguriert, die es uns erlaubt, unsere Apps als echte Endnutzer zu testen. Dafür haben wir nicht nur DorfFunk, sondern auch die anderen beiden Apps, mit denen DorfFunk interagiert, eingerichtet. Ziel hierbei ist es explizit, dass man keine künstlichen Testdaten erzeugt, sondern die Anwendung wirklich benutzt. Damit wird DorfFunk in die Kommunikation des Teams untereinander integriert. Der Fokus ist hierbei das bewusste Erleben des eigenen Produktes. Damit wird beim Team ein besseres Verständnis für die Nutzer gefördert. Dies ermöglicht es uns insbesondere, Fehler zu identifizieren, die nur in der Interaktion zwischen den Apps auftreten können. Bei diesem Testverfahren kann es auch sein, dass wir einzelnen Mitarbeitern ein Gerät für mehrere Tage ausleihen, um die Anwendung unter realen Umständen zu testen. Dadurch, dass alle Personen im Projekt eingeladen sind, die App zu testen, können alle die Apps besser kennenlernen, auch wenn sie nicht an der Entwicklung beteiligt waren. Von der hierdurch erreichten schnellen Problemidentifizierung profitiert das gesamte Team.

Eine immer wieder spannende Gelegenheit zum Testen ist das Reisen in der Bahn. Hierbei zeigt sich insbesondere bei der Durchquerung von Mittelgebirgen mit unterschiedlichen Netzstärken – sowohl innerhalb des WLANs an Bord als auch bei der Nutzung des Mobilgeräts –, wie gut die Anwendung mit ständig wechselnden Verbindungsszenarien klarkommt. Eine permanente Mischung aus keinem Empfang bis hin zu LTE bietet beim Reisen mit deutlich mehr als 100 km/h ein sehr vielfältiges Testszenario in kurzer Zeit.

Die meisten unserer Nutzer verwenden Android. Hierbei gibt es eine Bandbreite von High-End- zu Low-End-Telefonen oder Geräte mit alten Android-Versionen. Es ist schwer zu garantieren, dass das, was mit Handy A funktioniert, auch mit Handy B gut funktioniert. Um die zu Ressourceneffizienz zu testen haben wir einige günstige Geräte gekauft, darunter auch sehr billige unter 40 Euro. Neuerdings liegen Geräte mit Android Go sogar bei unter 30 Euro [BE19]. Wir möchten sicherstellen, dass wir ein größeres Spektrum an Geräten abdecken. Wenn die App auf solch einem Gerät funktioniert und die Leistung der App akzeptabel ist, funktioniert sie sehr wahrscheinlich auch auf besseren Handys.

Bei unserer Entwicklung haben wir einen szenariobasierten Ansatz eingeführt, um die Abdeckung unserer Tests zu verbessern. Hier testen wir Kombinationen verschiedener Funktionalitäten und Interaktionen. Um diese Szenarien realistischer zu gestalten, haben wir Personas entwickelt, die verschiedenen Arten von Nutzern simulieren, und verwenden diese, um die Aktionen zu beschreiben, die Personas im wirklichen Leben ausführen würden. Szenariobasierte Tests werden von uns bei Bedarf durchgeführt. Anders als bei Test Mobs ist mehr Aufwand erforderlich, um die Szenarien vorzubereiten.

5. Fazit

Auch wenn die altbewährten Methoden essenziell für die Qualitätssicherung von mobilen Diensten in einem digitalen Ökosystem sind, hat die Qualitätssicherung abseits der Norm bei uns eindeutig für eine weitere Verbesserungen gesorgt. Eine Vielzahl von vorher nicht bekannten Problemen wurde durch das gemeinsame Testen gefunden. Auch das Verständnis des eigenen Produktes wurde hierdurch verbessert und auch zwischen den Teammitgliedern angeglichen. Wichtig ist, dass der Prozess von Zeit zu Zeit reflektiert und an neue Bedürfnisse angepasst wird. Die klassische QS hat uns geholfen die Qualität trotz Nutzerexplosion adäquat sicher zu stellen, und neuere QS-Konzepte die Qualität weiter zu optimieren.

Besonders Test Mobs und Langzeittests haben eine starke Sensibilisierung für die Anwendung aus Nutzersicht geführt, die so von automatisierten Testansätzen nicht erreicht werden kann. Auch die bewusste Entscheidung günstige Geräte aktiv zu verwenden und auch über längere Zeiträume unserem Team zur Verfügung zu stellen hat insbesondere die Sensibilisierung für UX und auch Performance erhöht. Wo früher gerne mal mit leistungsstarken Geräten entwickelt wurde, greifen heute Entwickler auch selbstständig zu langsameren Geräten, um gleich zu prüfen, wie sich die App auf unterschiedlichen Geräten verhält.

Wir gehen davon aus, dass wir weitere Anpassungen nach Bedarf durchführen werden. Ein Beispiel hierfür könnte eine bessere Dokumentation und ggf. Klassifikation der gefundenen Probleme sein, so dass diese künftig noch besser ausgewertet werden können, was zu einer höheren Automatisierung von Tests führen könnte.

Danksagung

Dieser Beitrag wurde im Kontext des Projekts EnStadt: Pfaff (Förderkennzeichen: 03SBE112D und 03SBE112G) erstellt, gefördert vom Bundesministerium für Wirtschaft und Energie (BMWi) sowie vom Bundesministerium für Bildung und Forschung (BMBF).

Referenzen

- [KL16] Klose, A.-C.: Vom Pair Programming zum Mob Programming – Programmieretechniken im Wandel, <https://entwickler.de/online/agile/pair-programming-mob-programming-249211.html>, accessed: 18/07/2019.
- [SES20] Scherr, S. A.; Elberzhager, F.; Silva Schäfer, A.: Test Mobs – Unsere Erfahrungen mit gemeinschaftlichem Testen von mobile Apps, Softwaretechnik-Trends, Bd. 40, Nr. 2, S. 58-60, 2020.
- [BE19] Bergert, D.: Google kündigt Android 10 Go Edition an, <https://www.heise.de/newsticker/meldung/Google-kuendigt-Android-10-Go-Edition-an-4542404.html>, accessed: 28/11/2019.

Einsatz von Deep Learning für den Test Eingebetteter Systeme

Jens Bielefeldt, Kai-Uwe Basener,¹ Siddique Reza Khan, Mozhdeh Massah, Hans-Werner Wiesbrock,² Stefan Scharoba, Michael Hübner³

Abstract: Während der Durchführung von Funktions-, Stress- oder Laufzeittests eingebetteter Systeme fallen große Mengen von Testdaten an. Die Testläufe sind in der Regel, wenn auch nicht umfassend, von Testern und/oder Kunden ausgewertet worden, und können somit als Referenz für weitere Testdaten dienen.

In dem Forschungsprojekt **DeepTest** wird versucht, mit Hilfe von Techniken des Deep Learning diese riesigen Datenmengen für das Training neuronaler Netze zu nutzen, um weitere, neue Tests zu generieren und Testmuster in ihnen zu identifizieren. In dem vorliegenden Artikel wird über den derzeitigen Stand der Forschung berichtet.

Keywords: Test Eingebetteter Systeme; Testdaten Generierung; Deep Learning

1 Einleitung

Aufgrund der zunehmenden Komplexität heutiger Eingebetteter Systeme können nicht alle möglichen Wechselwirkungen systematisch und mit vertretbarem Aufwand über manuell spezifizierte Testfälle überprüft werden. Ergänzend werden deshalb neue Testfälle generiert, aus Modellen, anhand von Codestrukturen, s. [Re21, MB21] oder auch zufallsgeneriert gemäß von Profilen (z.B. Markov Automaten, s. [Wi15]).

Ein notorisches Problem dabei ist die Spezifikation von Orakeln: Was wird erwartet? Werden Testfälle automatisch aus dem Code abgeleitet, so auch zirkulär die Erwartungen. Und im Fall von Zufallsdaten kann nur die Einhaltung bekannter Regeln oder Muster überprüft werden, [Wi15].

In dem Forschungsprojekt **DeepTest** wird untersucht, wie man neue Testdaten mit Hilfe generativer neuronaler Netze erhalten kann. Dabei werden die Netze unüberwacht mit Signaldaten trainiert, undifferenziert in Eingangs- und Ausgangsdaten des SUT, s. Abb.1 Erst in der Anwendung als Testdaten werden sie in entsprechende Ein- und erwartete Ausgaben geteilt!

¹ Kostal Industrie Elektrik GmbH, Hagen, Germany {j.bielefeldt,k.basener}@kostal.com

² ITPower Solutions GmbH, ,Berlin, Germany {Siddique.Khan,Mozhdeh.Massah,hans-werner.wiesbrock}@itpower.de

³ BTU Cottbus–Senftenberg, Fachgebiet Technische Informatik, Cottbus, Germany {scharst1,huebner}@b-tu.de

Ein- und Ausgabedaten für Eingebettete Systeme sind nach der Prozessorzykluszeit getaktete Datenströme, im folgenden auch Signale genannt. Steuerungsdaten, in Abb.1 $A(t)$, und Feedback der Umgebung $B(t)$, gehen zur Zeit t ein und werden durch die ECU zu Ausgaben C verarbeitet.

Ein erfahrener Tester überprüft Testläufe zumeist in ihren grafischen Darstellungen, mit einem Oszilloskop oder anderem Werkzeug. Dazu wählt er ein Zeitfenster und lässt sich den Werteverlauf verschiedener Signale über diesen Zeitraum anzeigen. In den so erstellten Bildern der Signale überprüft er die Einhaltung bestimmter Signalmuster: Signal A geht nach oben, geht dann verzögert Signal B nach unten? In dem Bild selber sind Ein- oder Ausgabesignale nicht unterschieden.

Es seien deshalb allgemein Signaltbilder als Zusammenfassung von Werteverläufen verschiedener Signale über ein Zeitfenster definiert.

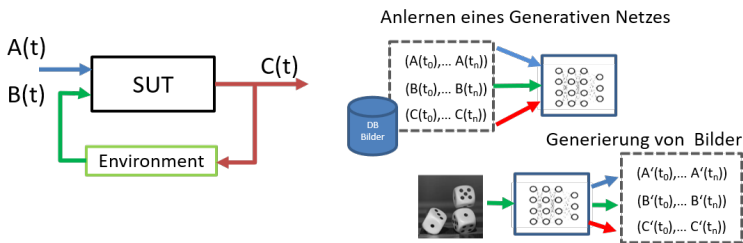


Abb. 1: Ein- und Ausgabedaten eines Testlaufes werden ununterschieden als Trainingsdaten eines Netzes verwendet. Erst in der Testanwendung werden die generierten Daten wieder in In- und Output unterteilt

Unser Ziel ist die Generierung neuer Testdaten. Dazu wählen wir Signaltbilder mit **allen** Eingangs- und Ausgangssignalen, in Abb.1 (A,B,C).

Generative Netze, Abb.2, ermöglichen die Erzeugung neuer Signaltbilder! Ein so erstelltes Signaltbild $\{(A(t), B(t), C(t)) | t \in [t_0, t_1]\}$ kann wieder in Eingangs- und Ausgangsströme unterteilt werden. Für den Test wird dann das SUT mit den Eingaben $\{(A(t), B(t)) | t \in [t_0, t_1]\}$ getrieben und ihre Ausgabe mit $\{C(t+1) | t \in [t_0, t_1 - 1]\}$ verglichen.

Testdaten sind Zeitreihen von Signalwerten und so bieten sich zur Modellierung auch Long-Short-Term-Memory (LSTM) Netze an, [HS97]. Da wir eine Vielzahl von Signalen betrachten, sind die Zeitreihen in unserem Fall vektorwertig. Unsere Experimente mit beiden Ansätzen zeigten jedoch, dass die Modellierung mit LSTM Netzen nicht so gut funktionierte, was vermutlich auch an ihren Schwierigkeiten liegt, Abhängigkeiten, Muster in den Signaltbildern zu fassen. Im Forschungsprojekt **DeepTest** konzentrieren wir uns deshalb vorrangig auf die Generierung von Signaltbildern durch Faltungsnetze, [Le98].

2 Ansatz

In dem grundlegenden Artikel [Go14] wurden Generative Adversariale Netze (GAN) eingeführt und gezeigt, wie sie zur Erzeugung neuer Bilder genutzt werden können. GANs wurden seitdem auch zur synthetischen Erzeugung von Zeitreihen [BWW19, AZM19] in verschiedenen Domänen eingesetzt.

Vereinfacht ausgedrückt, bestehen GANs aus zwei Netzen, die in einem 2-Personen-Nullsummenspiel miteinander verschränkt sind, Abb.2. Das Generator-Netz versucht, immer besser zu fälschen, das Diskriminator-Netz hingegen, immer genauer zwischen Fälschung und Original zu unterscheiden, (MinMax Prinzip).

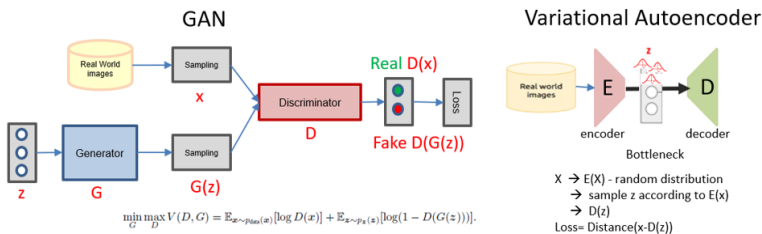


Abb. 2: (GAN) Das Generator-Netz wird in einem 2-Personen Nullsummenspiel mit einem Diskriminator-Netz gekoppelt. Ziel ist es, möglichst gute Fälschungen zu erstellen.

(VAE) Die Bilder werden durch ein Nadelöhr geschickt und wieder entpackt, genauer, ein Bild wird durch eine Zufallsverteilung kodiert, eine Stichprobe z entnommen und dekodiert $G(z)$. Angelernt wird über die Forderung, dass Original und dekodiertes Bild sich gleichen.

In dem Forschungsprojekt **DeepTest** werden GANs auf ihre Eignung zur Generierung von Signalbildern untersucht, d.h., Eingabebild x und dekodiertes Zufallsbild $G(z)$ in Abb.2 sind in diesem Fall reale bzw. künstlich erzeugte Signalbilder $\{(A(t), B(t), C(t)) | t \in [t_0, t_1]\}$.

Üblicherweise nutzen GANs für ihre Optimierungsfunktion die Kullback-Leibler Divergenz von Distributionen. Metriken oder andere Strukturen auf den Wahrscheinlichkeitsräumen spielen keine Rolle. Eine Weiterentwicklung der GANs sind Wasserstein GANs, WGAN, [ACB17, Gu17], welche für den Vergleich von Distributionen die Earth-Mover-Distanz oder auch Wasserstein Metrik genannt, nutzen. Diese basiert auf einer Metrik auf dem Grundraum!

Im Falle von Testsignalen gilt nun, dass kleine Zeitverschiebungen aufgrund der Trägheiten von Sensorik und Aktuator eine geringe Rolle darstellen. Dies spiegelt sich wider in der Auswahl geeigneter Metriken für ihren Vergleich. Bekannte Kriterien, die z.B. auf dem Dynamic Time Warping beruhen, sind jedoch nicht positiv definit, und müssen deshalb geeignet für den Einsatz in WGANs modifiziert werden. Wir haben diese Kriterien über Strafterme für Zeitverschiebungen so angepasst, dass sie eine Distanz definieren und damit für WGANs nutzbar sind.

Für die Wasserstein Metrik ist ein Integral über Lipschitz beschränkte Funktionen zu maximieren. In [ACB17] wurde die Lipschitz Bedingung durch starkes Beschränken der Parameter erzwungen, (Clipping), in [Gu17] durch einen Strafterm in der Optimierungsfunktion, welcher mit einem Gradienten gebildet wird. In beiden Fällen unterliegt den Konstrukten die Euklidische Metrik auf dem Bildraum. Wir haben indes eine neue Variante eingeführt, die die Lipschitz Bedingung effektiv auch für allgemeinere Metriken des Bildraums erzwingt. Ferner filtern wir bereits in der ersten Schicht über sämtliche Signale, (vertikale Bildachse), da wir alle möglichen Signalmuster in ihnen erfassen wollen.

Eine weitere in unserem Forschungsprojekt untersuchte Alternative ist die Konstruktion eines (Convolutional) Variational Autoencoders (VAE), [KW13]. In einem Autoencoder [GBC16] wird eine Eingabe komprimiert und dekodiert. Im Falle eines *Variational* Autoencoders wird die Eingabe durch eine Zufallsvariable kodiert, z.B. durch die Angabe von Mittelwert und Varianz einer multi-varianten Gaußverteilung. Anschließend wird eine Stichprobe aus dieser Verteilung genommen und diese dekodiert. Ziel des unüberwachten Anlernens ist es, die Gleichheit zwischen Original und Rekonstruktion zu erreichen, Abb.2. Der Decoder eines angelernten VAE kann dann zur Generierung neuer Bilder eingesetzt werden. Wir haben auch diese Konstrukte auf allgemeinere Metriken für den Vergleich von Original- und rekonstruierten Signalbilder erweitert.

3 Stand der Forschung

In [Pa20] verfolgen die Autoren einen sehr ähnlichen Ansatz. Ihr Ziel ist die Generierung kontinuierlicher Eingabestimuli für Eingebettete Systeme ohne damit verbundene Erwartungswerte. Sie lernen dazu einen Convolutional VAE an, bei dem die Metrik auf dem Raum der Testdaten durch die Diskriminatorauswertung eines trainierten GANs ersetzt wird, eine sogenannte VAE/GAN-Architektur, siehe [La15]. Die Qualität der generierten Signalbilder wird anschließend mit signal-spezifischen Ähnlichkeitsmaßen bewertet, z. B. Dynamic Time Warping. Diese Maße werden jedoch im Gegensatz zu unserem Ansatz nicht während des eigentlichen Trainings bereits benutzt.

DeepTest verfolgt hingegen das Ziel, zusätzlich auch Ausgaben zu generieren. Da diese den jeweiligen Eingaben folgen haben wir es stets mit Mustern zutun, die es zu erkennen gilt: Das Eingabesignal A impliziert das Ausgabesignal B. Allgemeiner interessiert uns das Erkennen impliziter Muster in den Signalen. In einem ersten Schritt lernen wir die Netze mit verschiedenen Architekturen durch Signalverläufe an, die bekannten Mustern folgen, und prüfen, inwieweit die generierten Signalverläufe eben diesen Mustern folgen. Dazu entwickelten wir speziell einen Testdatengenerator, der Signalbilder mit vorgegebenen realistischen Mustern und zeitlichen Verzögerungen erzeugen kann.

Im Beispiel Abb.3 kann man erkennen, wie einfache *Implies* und *Synch* Muster mit zeitlichen Verzögerungen in den generierten Signalen sich wiederfinden. Zugleich analysieren wir die Lernkurven der WGAN resp. Convolutional VAE Netze. Sie zeigen bereits ein gutes

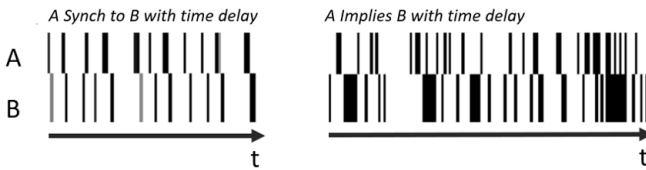


Abb. 3: Ausschnitt von WGAN generierten (binären) Signalbildern

In beiden Beispielen erkennt man deutlich, dass die Aktivierung des Signals A zeitlich verzögert eine Aktivierung von B auslöst, im rechten, dass B auch unabhängig von A ausgelöst werden kann.

Konvergenzverhalten. Allerdings sind diese Ergebnisse noch deutlich zu schärfen, ehe wir diese Techniken auf ein reales Testobjekt anwenden können, Abb.3.

4 Ausblick

Wie bereits im vorherigen Kapitel 3 beschrieben, sind die bisherigen Ergebnisse zur Testdatengenerierung über WGANs oder CVAE noch nicht befriedigend. Zur Zeit experimentieren wir deshalb weiter mit verschiedenen Architekturen, Hyperparametern und auch Straftermen in den Optimierungsfunktionen, um diese Resultate zu verfeinern.

Diese Technologie kann auch zur Überprüfung von Testläufen eingesetzt werden. Dazu lernen wir im einfachsten Fall einen VAE an, Abb.2. Als Autoencoder komprimiert und dekodiert der VAE ein vorgegebenes Signalbild und erzeugt so angelernt ein ähnliches Signalbild. Als statistisch generiertes Bild läßt es sich als ein *normalisiertes Bild*, d.h., angepasst an ein ähnliches durchschnittliches Signalbild lesen. Geben wir nun ein unbekanntes Signalbild

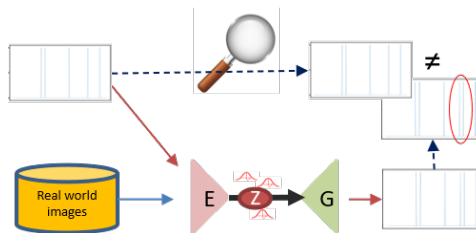


Abb. 4: Ein neuer Testlauf wird als Signalbild in das VAE eingegeben. Über den Vergleich mit dem dazu generierten Signalbild lassen sich Abweichungen erkennen und lokalisieren.

in den VAE, so wird dazu ein neues Signalbild generiert, in diesem Kontext auch gelesen als normales, weil durchschnittliches Testlaufbild, welches wir mit dem eingegebenen vergleichen können. Sind die Abweichungen in einem geeignet zu wählenden Sinne zu groß, so deutet dies auf eine Anomalie hin, auf einen Testlauf mit abweichendem Verhalten. Über einen Vergleich beider Bilder können wir auch die kritischen Zeit- und Wertstellen der Differenzen identifizieren. Wir hoffen, auf diese Weise automatisiert in großen Mengen von Testläufen Auffälligkeiten identifizieren und vorfiltern zu können, die anschließend von

einem Tester zu bewerten sind. Es gibt verschiedene Ansätze hierzu, siehe [Di19], die wir auf ihre Eignung zur Fehlerentdeckung in Testläufen untersuchen.

Danksagung

Diese Arbeit wurde durch das Bundesministerium für Bildung und Forschung (KMU-innovativ IKT project no.: 01IS19035) gefördert.

Literaturverzeichnis

- [ACB17] Arjovsky, Martin; Chintala, Soumith; Bottou, Léon: , Wasserstein GAN, 2017.
- [AZM19] Arnelid, Henrik; Zec, Edvin Listo; Mohammadiha, Nasser: Recurrent Conditional Generative Adversarial Networks for Autonomous Driving Sensor Modelling. In: 2019 IEEE Intelligent Transportation Systems Conference (ITSC). IEEE, oct 2019.
- [BWW19] Brophy, Eoin; Wang, Zhengwei; Ward, Tomas E.: Quick and Easy Time Series Generation with Established Image-based GANs. 2019.
- [Di19] Di Mattia, Federico; Galeone, Paolo; Simoni, Michele De; Ghelfi, Emanuele: , A Survey on GANs for Anomaly Detection, 27.06.2019.
- [GBC16] Goodfellow, Ian; Bengio, Yoshua; Courville, Aaron: Deep learning. MIT Press, Cambridge, Massachusetts and London, England, 2016.
- [Go14] Goodfellow, Ian J.; Pouget-Abadie, Jean; Mirza, Mehdi; Xu, Bing; Warde-Farley, David; Ozair, Sherjil; Courville, Aaron; Bengio, Yoshua: , Generative Adversarial Networks, 2014.
- [Gu17] Gulrajani, Ishaan; Ahmed, Faruk; Arjovsky, Martin; Dumoulin, Vincent; Courville, Aaron: , Improved Training of Wasserstein GANs, 2017.
- [HS97] Hochreiter, S.; Schmidhuber, J.: Long short-term memory. Neural computation, 9(8):1735–1780, 1997.
- [KW13] Kingma, Diederik P.; Welling, Max: , Auto-Encoding Variational Bayes, 2013.
- [La15] Larsen, Anders Boesen Lindbo; Sønderby, Søren Kaae; Larochelle, Hugo; Winther, Ole: , Autoencoding beyond pixels using a learned similarity metric, 2015.
- [Le98] Lecun, Y.; Bottou, L.; Bengio, Y.; Haffner, P.: Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11):2278–2324, 1998.
- [MB21] MBTsuite: sepp.med gmbh,<https://www.seppmed.de/de/portfolio/mbtsuite>, 11.05.2021.
- [Pa20] Parthasarathy, Dhasarathy; Bäckström, Karl; Henriksson, Jens; Einarsson, Sólrun: , Controlled time series generation for automotive software-in-the-loop testing using GANs, 2020.
- [Re21] Reactis, Testing Tools for Simulink and C Code,<https://reactive-systems.com/products.msp>, 11.05.2021.
- [Wi15] Wiesbrock, H.-W.: Eine Stochastische Testmethodik für den Robustheitstest Automobil Steuergeräte. Jgg. 246 in GI-Edition Lecture Notes in Informatics Proceedings, Ges. für Informatik, Bonn, 2015.

Metamorphes Testen

Matthias Hamburg¹

Abstract:

Metamorphes Testen ist ein relativ neues Testverfahren, das sich besonders gut eignet, wenn kein ausreichendes Testorakel verfügbar ist. Metamorphes Testen betrachtet mehrere Ausführungen des Testobjekts und prüft, ob die Eingaben und Ausgaben bei diesen Ausführungen zueinander konsistent sind. Die Konsistenzbedingungen werden meist aus der Spezifikation gewonnen. Erfüllt das Testobjekt bei einer Gruppe von Testfällen eine Konsistenzbedingung nicht wie erwartet, so wurde eine Fehlerwirkung gefunden. Der Bei-trag stellt das Verfahren vor, ordnet es in die Testmethodik ein, gibt einen Überblick über die bisherigen Ergebnisse und bewertet die Anwendbarkeit.

Keywords: Metamorphes Testen, Testorakel, Testverfahren, künstliche Intelligenz.

1 Einleitung

Beim dynamischen Testen werden Testfälle ausgeführt und bewertet, ob sie bestehen oder fehlschlagen. Dazu braucht man ein Testorakel: eine Informationsquelle zur Ermittlung des erwarteten Ergebnisses, um es mit dem tatsächlichen Ergebnis des Testobjekts zu vergleichen [CT21]. In der Praxis ist allerdings ein Testorakel nicht immer einfach bereitzustellen und man spricht vom Orakelproblem. Zum Beispiel ist das bei Systemen der Fall, die Ergebnisse über komplexe Algorithmen erzeugen und hierbei ganz besonders bei künstlicher Intelligenz.

Metamorphes Testen ist ein in der Praxis anwendbares Verfahren, um mit dem Orakelproblem umzugehen. Dabei wird das Ergebnis der Testfälle nicht individuell ausgewertet, sondern geprüft, ob verschiedene Ausführungen des Testobjekts bestimmte Konsistenzbedingungen, die sog. metamorphen Relationen erfüllen. Seit der Einführung dieses Testverfahrens durch Chen et al. [Ch98] sind zahlreiche wissenschaftliche Publikationen über das Verfahren erschienen. Insbesondere wegen der Relevanz des Orakelproblems bei Systemen mit künstlicher Intelligenz wurde das Verfahren auch in internationalen Standards aufgenommen [IS20]. Das Ziel dieses Beitrags ist, das Verfahren methodisch zu beschreiben und die Bekanntheit des Verfahrens auch in der deutschen Gemeinde der Testpraktiker zu fördern.

¹ German Testing Board e.V., Koldestrasse 8b, 91052 Erlangen, m.hamburg@gtb.de

2 Das Verfahren des metamorphen Testens

Im Unterschied zum eigenschaftsbasierten Testen (Englisch: „property-based testing“), das das Orakelproblem durch partielle Konsistenzbeziehungen zwischen Ein- und Ausgaben desselben Testfalls behandelt, geht metamorphes Testen davon aus, dass es oft einfacher ist, bei verwandten Eingaben die Konsistenz der Ausgaben des Testobjekts zueinander zu prüfen [Se20]. Der Name basiert darauf, dass Eingaben des Testobjekts umgeformt werden, um zu prüfen, ob sich die Ausgaben entsprechend ändern („Metamorphose“). Zum Beispiel kann bei einer Implementierung der Sinusfunktion getestet werden, ob Symmetrierelationen wie $\sin(-\alpha) = -\sin(\alpha)$ oder $\sin(\pi-\alpha) = \sin(\alpha)$ erfüllt sind. Darüber hinaus können auch Relationen zwischen mehreren Eingabe-/Ausgabepaaren verwendet werden, wie $\sin(\alpha+\beta) = \sin(\alpha)\sin(\pi/2-\beta) + \sin(\pi/2-\alpha)\sin(\beta)$. Formal gesehen ist eine metamorphe Relation für ein Testobjekt, das eine Funktion f ausführt, eine Relation zwischen einer Reihe von Eingaben x_0, x_1, \dots, x_n und den entsprechenden Ausgaben $f(x_0), f(x_1), \dots, f(x_n)$ [Se16]. Die Paare $(x_i, f(x_i))$ von Ein- und Ausgaben entsprechen den Testfällen. Der erste Testfall $(x_0, f(x_0))$ wird dabei Ausgangstestfall (Englisch: „source test case“) genannt; die weiteren Testfälle werden Folgetestfälle („follow-up test cases“) genannt. Beim metamorphen Testen wird zuerst ein Ausgangstestfall entworfen, um anschließend die Folgetestfälle anhand der metamorphen Relation herzuleiten.

Als ein weniger mathematisches Beispiel soll die Glossaranwendung des ISTQB® dienen, mit welcher nach Glossarbegriffen des ISTQB® gesucht werden kann [CT21]. Als Testbasis dient die Benutzeranleitung der Anwendung. Eine metamorphe Relation ist in diesem Beispiel: Wenn man die Textsuche im Glossar durch zusätzliche Filterkriterien einschränkt, dann muss das Suchergebnis eine Teilmenge des Ergebnisses der ursprünglichen Suche sein. Ein konkreter Ausgangstestfall könnte die ungefilterte Suche nach dem Text „Test“ bei der Spracheinstellung Deutsch sein und der Folgetestfall die gleiche Suche eingeschränkt auf den „Foundation Level“ Lehrplan. Beim aktuellen Stand der Glossaranwendung liefert der Ausgangstestfall eine Ergebnismenge mit 224 Begriffen und der Folgetestfall eine Menge mit 115 Begriffen. Enthält das Suchergebnis des Folgetestfalls einen Begriff, der im Suchergebnis des Ausgangstestfalls nicht vorkommt, so hat die Glossarsuche die metamorphe Beziehung nicht erfüllt: eine Fehlerwirkung. Das Beispiel zeigt die typische Situation, dass man beim metamorphen Testen Werkzeugunterstützung wie einen Komparator zum Abgleichen der Ergebnismengen gut gebrauchen kann. Ein zusätzlicher Filter nach Schlüsselbegriffen bzw. Nicht-Schlüsselbegriffen müsste eine disjunkte Partition ergeben – eine ternäre metamorphe Beziehung. Hierbei kann der Ausgangstestfall der vorherige Folgetestfall sein. Die Durchführung dieses Tests durch den Autor ergab 64 Schlüsselbegriffe und 45 Nicht-Schlüsselbegriffe: eine Fehlerwirkung war gefunden. Die Fehleranalyse ergab, dass das Attribut „IstSchlüsselbegriff“ bei den fehlenden 6 Begriffen mit einem falschen Wert belegt war. An diesem Beispiel sehen wir auch einen weiteren spezifischen Aspekt des metamorphen Testens: dass Testabläufe oft dadurch entstehen, dass Folgetestfälle einer metamorphen Relation als Ausgangstestfälle der nächsten Relation dienen können.

3 Erfahrungen aus Forschung und Praxis

Seit der Einführung dieses Testverfahrens durch Chen et al. [Ch98] sind zahlreiche Erfahrungsberichte aus verschiedenen Anwendungsgebieten erschienen. Ein aktueller Überblicksartikel [Se20] zählt 84 veröffentlichte Fallstudien zum Thema, die zwischen Januar 1998 und Mai 2018 erschienen sind. Diese berichten übereinstimmend über erfolgreiche Fehlerfindung. Einige davon bewerten sie empirisch durch Fehlereinpflanzung, d.h. durch absichtliches Einfügen von Fehlerzuständen in den Code. Andere generieren größere Mengen an konkreten Testfällen und bewerten deren Fehlerfindungsrate. Die Erfahrungsberichte stammen aus verschiedenen Anwendungsgebieten wie Webdiensten und -applikationen, Computergrafik, eingebetteten Systemen, Simulation und Modellierung, Bioinformatik, numerischer Programmierung, Compilerbau und künstlicher Intelligenz. An dieser Stelle wird nur auf einige von ihnen beispielhaft eingegangen.

Metamorphes Testen wurde beim Abnahmetest des NASA Data Access Toolkits erfolgreich angewendet [Li15]. Das System verwaltet Telemetriedaten der verschiedenen NASA Raumfahrtprogramme mit einer leistungsfähigen Abfrageoberfläche. Das riesige Datenvolumen stellte die Tester vor das Orakelproblem beim Feststellen, ob die Suchergebnisse korrekt sind. Metamorphes Testen wurde hier angewendet, indem die gleiche Abfrage auf verschiedene äquivalente Arten formuliert wurde, bei denen die Ausgaben identisch sein sollten. Mehrere Fehlerwirkungen wurden mit diesem Ansatz gefunden.

Websuchmaschinen sind aufgrund der komplexen Suchalgorithmen und der immensen Datenmenge im heutigen Internet gute Testobjekte für metamorphes Testen. Als Testbasis kann die Benutzerhilfe der Websuchmaschine dienen, welche über die Funktionen der Filterung und Parametrisierung informiert. Als Ausgangstestfall können Suchtexte ohne weitere Filterkriterien dienen. In den Folgetestfällen werden entsprechend Filter oder andere Parameter gesetzt, wie Einschränkung auf den Titel des ersten Suchergebnisses des Ausgangstestfalls, äquivalente Umformungen der Booleschen Suchbedingung oder Spracheinstellungen. Ein Forschungsteam hat die Qualität der vier weltweit populärsten Suchmaschinen Google, Bing, Chinese Bing und Baidu auf diese Weise getestet [Zh16]. Zum Beispiel wurde die Einschränkung auf den ersten Titel mit stündlich je 1000 zufällig generierten Begriffen in Englisch und in Chinesisch über 38 Stunden getestet und die Rate an Anomalien statistisch ausgewertet. Jede der Suchmaschinen erzeugte mindestens eine Anomalie pro Stunde. Die geringste Anomalierate hatte Google in Kombination mit Englisch, dicht gefolgt von Baidu mit Chinesisch.

Ein Forschungsbericht aus jüngster Zeit beschreibt einen Robustheitstest der Zitationszahlen wissenschaftlicher Publikationen bei den führenden Zitationsdatenbanken „Web of Science“ und „Scopus“ [Zh21]. Dabei wurde empirisch untersucht, wie sich typische Schreibfehler in den Referenzangaben auf die Erkennung der Zitation durch die Suchmaschinen auswirken. Ein überraschendes Ergebnis war, dass Bindestriche im Titel einen deutlich negativen Einfluss auf die Erkennung durch die Suchmaschinen haben. Auch wenn das Vorgehen eher eine empirische Analyse war, zeigt diese Studie dass metamorphes Testen auch bei dynamischen Robustheitstests sinnvoll anwendbar ist.

4 Metamorphes Testen von KI-basierten Systemen

Bei Systemen mit künstlicher Intelligenz liegt das Orakelproblem auf der Hand, weil das erwartete Verhalten nicht exakt vorhersagbar ist und über die Zeit nicht konstant bleibt. Metamorphes Testen ist hier ein empfohlenes Verfahren [IS20]. Zum Beispiel wird beim überwachten Lernen das System mit Ein- und Ausgabedaten trainiert. Sein anschließendes Verhalten hängt vom Trainingsalgorithmus und von den Trainingsdaten ab. Für metamorphes Testen können die Trainingsdaten als Ausgangstestfälle dienen und Folgetestfälle anhand metamorpher Relationen als Testdaten generiert werden. Da hierbei die metamorphen Relationen aus den Lernalgorithmen und Trainingsdaten abgeleitet werden, ist das ein eher seltenes Beispiel einer White-Box-Anwendung des metamorphen Testens.

Ein Forschungsteam hat ein Framework für das automatisierte Testen von autonomen Drohnen entwickelt [Li17]. Ausgangstestfall ist das Drohnenverhalten bei einem Flugszenario; Folgetestfälle sind verschiedene Variationen des Szenarios, bei denen das Ergebnis des Drohnenfluges ähnlich sein sollte. Zum Beispiel können die Lichtverhältnisse oder die Himmelsrichtung des Ziels geändert werden, wobei Abstände und relative Positionen von Hindernissen gleichbleiben. Mit diesem Ansatz konnten die Forscher eine Reihe von Anomalien im Verhalten der Drohne im Simulator aufdecken, einschließlich einiger Abstürze. Eine ähnliche Idee wurde von einer anderen Gruppe von Forschern verwendet, um das Werkzeug DeepTest zu entwickeln [Ti18]. Dieses testet selbstfahrende Autos im Fahrsimulator, die von tiefen neuronalen Netzen gesteuert werden. DeepTest generiert aus einem Ausgangstestfall automatisch Folgetestfälle für den Fahrsimulator mit unterschiedlichen Fahrbedingungen wie Regen oder Nebel, bei denen sich das Auto ähnlich verhalten soll. Unter Anderem fand DeepTest tausende von fehlerhaften Aktionen der Testobjekte bei verschiedenen realistischen Fahrbedingungen. Einige dieser Fehlerwirkungen würden im Straßenverkehr bei drei der leistungsstärksten Systeme zu tödlichen Unfällen führen.

5 Testmethodische Einordnung

Aus Sicht des ISTQB® kann man metamorphes Testen als ein Testverfahren betrachten, weil es Testbedingungen definiert, Testfälle entwirft und Testdaten spezifiziert [CT21]. Das Verfahren lässt sich leicht entlang der Hauptaktivitäten des dynamischen Testens erklären [CT19]. In der Testanalyse werden metamorphe Relationen gesucht, die als Testbedingungen dienen. Im Testentwurf werden zu jeder metamorphen Relation Ausgangstestfälle und Folgetestfälle entworfen, die die jeweilige metamorphe Relation erfüllen. In der Realisierung werden geeignete Testdaten für die Ein- und Ausgaben erzeugt und Testabläufe mit Ausgangs- und Folgetestfällen der verschiedenen Relationen aufgebaut. Anschließend werden die Testfälle ausgeführt und die Erfüllung der metamorphen Relation geprüft. Das Verfahren gewährleistet somit auch die bidirektionale Verfolgbarkeit zwischen Testbedingungen und den Testergebnissen.

Allerdings gibt es auch einige Unterschiede im Detail. Traditionell ist das erwartete Ergebnis ein Pflichtattribut des Testfalls [CT21], wofür ein Testorakel erforderlich ist. Eine Lösung des Orakelproblems macht deshalb eine etwas allgemeinere Auffassung erforderlich, ohne allerdings auf Erwartungen an die Ausgaben ganz zu verzichten. Der Ansatz des metamorphen Testens ist dabei, dass das Testergebnis nicht aus den Ausgaben einzelner Testfälle entsteht, sondern aus einer Gruppe von Testfällen. Ein Fehlschlag bedeutet, dass die Ausgaben von Ausgangstestfall und Folgetestfällen die metamorphe Relation nicht erfüllen. Dies ist ein eindeutiger Hinweis auf einen Fehlerzustand. Andererseits ergibt das Bestehen des Tests auf eine metamorphe Relation eine schwächere Aussage als beim traditionellen Testen. Somit ist metamorphes Testen eher für das Finden von Fehlerzuständen geeignet als für eine Begründung des Vertrauens in die Qualität des Testobjekts.

Metamorphes Testen wird meistens als Black-Box-Verfahren angewendet, mit einer Spezifikation als Testbasis. Im Fokus steht die funktionale Eignung des Testobjekts. Nichtfunktionale Qualitätsmerkmale wie Performanz, Zuverlässigkeit, Gebrauchstauglichkeit oder Robustheit können auch eine Rolle spielen, aber hier hilft metamorphes Testen eher indirekt als Generator von Testfällen [Zh16], [Zh21].

Die Effektivität der Fehlerfindung hängt wesentlich von den metamorphen Relationen ab. Die Analyse der metamorphen Relationen ist damit eine essenzielle Aufgabe und kreative Tätigkeit des Testers. Für das Finden geeigneter metamorpher Relationen werden in der Literatur folgende Kriterien genannt [Se16]: Kenntnis der Anwendungsdomäne, Ableitung der metamorphen Relationen aus einzelnen Teilen des Systems, formale Beschreibung der metamorphen Relationen und möglichst große Unterschiede zwischen den Codeabschnitten, die von Ausgangs- bzw. Folgetestfällen durchlaufen werden. Für die Analyse gibt es zwei übliche Ansätze. Beim eingabeorientierten Ansatz untersucht man, welche Transformationen von Eingabedaten zu einer überprüfaren Änderung der Ausgabedaten führen. Der ausgabeorientierte Ansatz geht von musterhaften Relationen zwischen den Ausgaben verschiedener Testfälle aus. Solche Muster sind Identität, Äquivalenz (z.B. Listen, die die gleichen Elemente aber in beliebiger Reihenfolge enthalten), Teilmengen, disjunkte Mengen oder feste Unterschiede [Se18].

Metamorphes Testen ist am effizientesten, wenn alle Aktivitäten nach der Testanalyse automatisiert sind, wie z. B. die Generierung von Ausgangs- und Folgetestfällen, Generierung von Testdaten, die Testausführung und die Prüfung der metamorphen Relation.

Wie jedes Testverfahren hat metamorphes Testen seine spezifischen Einschränkungen. Die prominenteste ist, dass es das Orakelproblem zwar lindern, aber nicht vollständig lösen kann. Bestandene Tests belegen nur, dass das Verhalten des Testobjekts bei den Testfällen gewisse Konsistenzbedingungen erfüllt, aber nicht, ob es korrekt ist.

6 Schlussfolgerung

Metamorphes Testen ist ein relativ neues Testverfahren, dessen Effektivität bei der Fehlerfindung empirisch belegt ist. Besonders geeignet ist dieses Verfahren bei Anwendungstypen, die ein relevantes Orakelproblem aufweisen. Heute steht es an der Schwelle zum breiteren Praxiseinsatz. Da es ein großes Potenzial zur Automatisierung aufweist, hängt seine Marktverbreitung auch von der künftigen Verfügbarkeit der Werkzeugunterstützung ab.

Literaturverzeichnis

- [Ch98] Chen, T.Y. et al.: Metamorphic testing: A new approach for generating next test cases. Dept. Comput. Sci. Hong Kong Univ. of Science and Technology, Techn. Rep. HKUSTCS98-01, 1998.
- [CT19] ISTQB®, Certified Tester Foundation Level Syllabus v3.1, 2019.
- [CT21] ISTQB®/GTB Standardglossar der Testbegriffe, <https://glossary.istqb.org/de>, Stand: 13.06.2021.
- [IS20] ISO/IEC, Technical Report TR-29119-11 Software and systems engineering - Software testing - Part 11: Guidelines on the testing of AI-based systems, 2020.
- [Li15] Lindvall, M. et al.: Metamorphic model-based testing applied on NASA-DAT – An experience report. In Proc. IEEE/ACM 37th Int. Conf. Softw. Eng., S. 129-138, 2015.
- [Li17] Lindvall, M. et al.: Metamorphic model-based testing of autonomous systems. In Proc. IEEE/ACM 39th Int. Conf. Softw. Eng., S. 35-41, 2017.
- [Se16] Segura, S. et al.: A Survey on Metamorphic Testing. IEEE Trans. Softw. Eng. 42/09, S. 805-824, 2016.
- [Se18] Segura, S. et al.: Metamorphic Testing of RESTful APIs. IEEE Trans. Softw. Eng. 44/11, S. 1083-1099, 2018.
- [Se20] Segura, S. et al.: Metamorphic Testing: Testing the Untestable. IEEE Software 20/04, S. 46-53, 2020.
- [Ti18] Tian, L. et al.: DeepTest: Automated testing of deep-neural-network-driven autonomous cars. In Proc. IEEE/ACM 40th Int. Conf. Softw. Eng., S. 303-314, 2018.
- [Zh16] Zhou, Z.Q. et al.: Metamorphic Testing for Software Quality Assessment: A Study of Search Engines. IEEE Trans. Softw. Eng. 42/03, S. 264-284, 2016.
- [Zh21] Zhou, Z.Q. et al.: Metamorphic Robustness Testing: Exposing Hidden Defects in Citation Statistics and Journal Impact Factors. IEEE Trans. Softw. Eng. 47/06, S. 1164-1183, 2021.

Autorinnen- und Autorenverzeichnis

A

Abecker, Andreas, 221, 333
Akharath, Philipp, 251
Aksoy, Alperen, 807
Akyol, Ali, 33
Albers, Christoph, 1455
Alekosai, Emal M., 1729
Algergawy, Alsayed, 565
Alhamwy, Yasin, 1047
Alnahawi, Nouri, 907
Al-Nashash, Doha, 45
Altenhoener, Albrecht, 1351
Altkrüger, Jaqueline, 251
Altschaffel, Robert, 1759
AndreaHirschel, 1669
Andrian-Werburg, Stefan von, 613
Annanias, Yves, 695
Apel, Anika, 1165
Auer, Daphne, 547
Auth, Gunnar, 867

B

Babalou, Samira, 581
Baberowski, David, 1645
Bach, Katja, 1635
Bade, Korinna, 1665, 1669
Bajramovic, Edita, 1781
Bakli, Mohamed, 565
Banse, Marvin, 57
Barakat, Raman, 1789
Barev, Torben Jan, 981
Barz, Björn, 417
Basener, Kai-Uwe, 1859
Bauer, Maximilian, 1047
Bauer, Thomas, 1503

Baumann, Michaela, 1121
Bauspieß, Pia, 907
Becker, Andreas, 1341
Becker, Ilja, 1107
Becker, Jürgen, 1247
Beckh, Katharina, 229
Beckmann, Helmut, 1477, 1487, 1531
Beer, Sonja, 1591
Behrendt, Tanja, 67
Behrens, Grit, 359, 375
Bender, Christian, 723
Bendlin, Dirk, 79
Benndorf, Jörg, 705
Berendsohn, Walter, 539
Bertram, Florian, 1325
Beskorovajnov, Wasilij, 905
Bettina, Waldvogel, 1657
Bielefeldt, Jens, 1859
Bilo, Michael, 623
Binder, Frank, 709
Biniaminov, Sergey, 1247
Birk, Sönke, 647
Bischof, Christian, 607
Bodesheim, Paul, 443, 481, 499, 547
Böhlke, Julia, 481
Böhner, Alexander, 1035
Böhnke, Stephanie, 1665, 1669
Brand, Alexander, 1603
Brockner, Annabell, 1661
Brockmann, Carsten, 1263, 1351
Brodthage, Michael, 879
Bruns, Julian, 221, 333
Brust, Clemens-Alexander, 417
Bucher, Solveig Franziska, 499

Budde, Ina, 313
Budde, Matthias, 269, 301, 333
Bunse, Christian, 1701
Bunzel, Niklas, 1059
Bürgl, Kim, 709
Bürmann, Paula, 103

C

Carl, K. Valerie, 757
Catal, Faruk, 1789
Cermann, Holger, 1325
Chen, Johannes, 1255
Chircu, Alina, 1283
Chodyko, Piotr, 1165
Cioffica, Paul, 1149
Conrad, Norbert, 605
Cornelia, Rüdüsüli, 1657
Costa, David Schellenberger, 581
Cox, Sean, 1265
Cremers, Lukas, 157
Creutzburg, Reiner, 375
Cyrus, Fabian, 147
Czarnecki, Christian, 1263, 1303

D

Dalm, Marinus, 705
Dammel, Frank, 607
Davis, Marlen, 623
Denninger, Oliver, 1825
Denzler, Joachim, 417, 443, 481, 499,
547
Dibbern, Bettina, 623
Dickhaut, Ernestine, 981
Diehl, Matthias, 359
Diel, Sarah, 789
Ding, Yongjian, 1759
Dittmann, Jana, 1759
Djeukoua, Christele Larissa Moussi,
1781

Dmitriyev, Viktor, 91
Döhmman, Indra Spiecker gen., 999
Döhmman, Indra Spiecker genannt,
963
Döring, Nico, 1177
Dreiser, Christoph, 647
Drewes, Lars, 1517
Dyck, Anne, 207

E

Ebert, André, 1165
Ehmke, Dierk, 1831
Eilers, Malte, 103
Elberzhager, Frank, 1853
Erdmann, Sven, 1573
Ertl, Christoph, 1441
Eschemann, Patrick, 147

F

Fabian, Stefan, 1047
Fähndrich, Johannes, 891
Farágó, David, 1809
Faulstich, Stefan, 229
Feike, Frederik, 607
Fellmann, Michael, 1573
Fiderer, Christian, 547
Fleiner, Christian, 1689
Floerecke, Sebastian, 1441, 1545
Förschler, Marc, 647
Franczyk, Bogdan, 743
Franziska, Mayr, 1657
Freese, Lea, 157
Freiberger, Michael, 1441
Freisleben, Bernd, 425, 461
Frerichs, Marc, 1567
Frers, Sebastian, 103
Freyberg, Julian, 1603
Frick, Raphael Antonius, 795

Friske, Mario, 1809
Frommelt, Paul, 461
Fukas, Philipp, 1369

G

Ganske, Michael, 1383
Garbers, Nicole, 879
Garmatter, Henriette, 147
Gärtig-Daug, Anja, 1611
Geihs, Kurt, 1047
Geißler, Julian, 907
Gembariski, Paul Christoph, 1361
Gerdes, Lukas, 45
Gerl, Armin, 1675
Gertz, Michael, 1603
Gibcke, Christoph, 1573
Gilch, Harald, 1713
Gleich, Theresa, 657
Gómez, Jorge Marx, 33, 45, 79, 127,
175
Gottwald, Jannis, 425
Grasmeyer, Tobias, 907
Grimm, Rüdiger, 963
Gröll, Roland, 905
Grombacher, Stefan, 965
Groß, Rainer, 1675
Großmann, Janik, 633
Grümmer, Julian, 1689
Grunz, Ariane, 517
Guatame-Garcia, Adriana, 705
Güccük, Abdullah, 115
Guinea, Alejandro Sanchez, 1047
Guiraud, Timothee, 1781
Güldali, Baris, 1809
Guldner, Achim, 287
Gumm, Dorina, 1591
Gümüs, Can, 1019
Günther, Christine, 1649

Güntsch, Anton, 539

H

Haas, Valentin, 1247
Hackel, Sascha, 1789
Häckell, Moritz Werther, 79
Hagen, Simon, 1361, 1425
Hamann, Anne, 1645
Hamburg, Matthias, 1813, 1865
Handschuh, Markus, 647
Harth, Andreas, 1689
Hasselbring, Wilhelm, 1819
Haunschmid, Ernst, 609
Heereman, Wendy, 789
Heesen, Henrik te, 251
Heimann, Thomas, 1325
Heimbach, Konrad, 1193
Heinbach, Christoph, 1397
Heinemann, Andreas, 907
Helbig, Raphaela, 667
Hempel, Ivo, 1759
Hempfen, Simon, 103
Hennig, Katharina, 45
Herbort, Volker, 251
Herget, Gebhard, 1283
Herlyn, Wilmjakob Johannes, 1313
Herrmann, Dominik, 1035
Hertweck, Philipp, 301
Herzfeldt, Alexander, 1441
Hesse, Mario, 723
Heurich, Marco, 547
Heyde, Markus von der, 1675
Hilbring, Desiree, 301
Hildebrand, Stefan, 1641
Hilger, Theresa, 45
Hintemann, Ralph, 593
Hinterholzer, Simon, 593
Hiroyasu, Kyle, 1165

Hlawitschka, Mario, 719
Höchst, Jonas, 425, 461
Hof, Hans-Joachim, 1745
Höfer, Rene, 623
Hofstedt, Petra, 1107
Hölscher, Natalia, 103
Holz, Felix, 1573
Hong, Chin-Gi, 1303
Höper, Lukas, 1623
Höper, Tobias, 127
Hornung, Gerrit, 963, 1047
Hoseini, Sayed, 1729
Höveling, Sven von, 667
Hübner, Michael, 1859

I

Ismali, Festina, 1165

J

Jahn, Hannes, 1411
Janson, Andreas, 981
Janßen, Christian, 45, 91, 103
Janßen, Julian, 321
John, André, 739
Jungermann, Imke, 1713
Jüngling, Sebastian, 313

K

Kahlert, Sebastian, 1487
Kaiser, Florian Klaus, 999
Kalloch, Benjamin, 719
Kammler, Friedemann, 1361
Kattge, Jens, 581
Kaufmann, Jens, 1729
Kazimiers, Antje, 1645
Keil, Oliver, 1759
Kellner, Domenic, 1201, 1255
Kennes, Lieven Nils, 1701
Kern, Christopher Julian, 1467

Kern, Sebastian, 1325
Kesdoğan, Doğan, 807
Khan, Siddique Reza, 1859
Kilian, Norbert, 539
Kindermann, Jörg, 229
Kirdan, Erkin, 1753
Kirsch, Nadine, 1653
Klamroth, Jonas, 1825
Klapper, Frank, 1675
Klenke, Thomas, 67, 137
Klimieczek, H.-Josef, 681
Klingauf, Uwe, 1047
Klingmann, Philipp, 1351
Knaden, Andreas, 1675
Kneuper, Ralf, 755, 1455
Knobloch, Andreas, 705
Koch, Björn, 137, 593
Koch, Christian, 1265
Koch, Werner, 633
Kohl, Tobias, 1201
Kohlbecker, Andreas, 539
Köhler, Wolfgang, 1019
König, Tim, 1603
König-Ries, Birgitta, 565, 581
Koo, Hyein, 1165
Korsch, Dimitri, 443, 481
Körschens, Matthias, 499
Kortum, Henrik, 1425
Köster, Frank, 185
Koza, Erfan, 819
Krebs, Antonia, 137
Kreiss, Jonathan, 1177
Kressner, Martin, 705
Kreten, Sandro, 287
Kroll, Franz, 1107
Kronburger, Julia, 1165
Kucewicz, Alicja Julia, 79
Kugler, Ulrike, 351

Kühne, Stefan, 1819
Kühnel, Stephan, 1729
Kuhr, Jan-Christian, 1701
Kusber, Wolf-Henning, 539
Kutzer, Karsten, 603

L

Labudde, Dirk, 847
Lahmann, Jan-Rainer, 943
Lampe, Patrick, 425, 461
Landwirth, Robert, 1059
Lange, Matthias, 1759
Lange, Olga, 1149
Lantow, Birger, 1573
Laue, Ralf, 1499, 1501
Lehner, Franz, 1545
Leible, Stephan, 1567
Leonhardt, Thiemo, 1645
Liesch, Tanja, 301
Lieser, Patrick, 461
Lindner, Karim, 657
Link, Anton, 517
Link, Norbert, 1247
Linnik, Benjamin, 1265
Löffler, Sven, 1107
Lowin, Maximilian, 1093, 1201, 1255
Ludsteck, Johannes, 1283
Ludwig, Nicole, 227
Luther, Katja, 539
Lutz, Marc-Alexander, 229

M

Mainzer, Klaus, 1153
Malzacher, Mario, 313
Marquardt, Philipp, 1247
Marschik, Grit, 1665
Marx, Erik, 1645
Marx-Gómez, Jorge, 103

Massah, Mozhdah, 1859
Mayer, Manuel, 857
Meer, Jan de, 1745
Meißner, Katherina, 1383
Meister, Vera G., 1675
Mendoza, Ricardo Carrillo, 313
Mengs, Christoph, 723
Merkel, Natalie, 705, 739
Meyer, Felix, 359
Meyer, Maurice, 867
Migliavacca, Mirco, 499
Mihale-Wilson, Cristina, 1093, 1201
Mohaupt, Marleen, 847
Möllerherm, Stefan, 705
Mrogenda, Klemens, 623
Mühlhäuser, Max, 1047
Müller, Andreas, 539
Müller, Linda, 1611
Müller, Lydia, 709
Mummert, Niclas, 1149

N

NadineBergner, 1645
Nattke, Matthias, 389
Naumann, Stefan, 221, 287
Nauss, Thomas, 425
Nenner, Christin, 1645, 1649
Neuhaus, Uwe, 1729
Neumann, Carolin, 1641
Nicole, Schumann, 1657
Niekler, Andreas, 709
Nieße, Astrid, 147
Nissen, Volker, 1517
Nitzsche, Sven, 1247
Nolte, Harm, 705
Nuhn, Helge, 1135
Nüttgens, Markus, 1567

O

Otterbach, Johannes, 1185

P

Pachideh, Brian, 1247

Pazmino, Victor, 1247

Pehlken, Alexandra, 147, 593

Pelzel, Frank, 1265

Penaherrera, Fernando, 593

Pesch, Robert, 1165

Petschke, Dmitry, 857

Pfäffel, Sebastian, 229

Pfisterer, Hans-Jürgen, 1411

Proctor, Rachel, 1325

Pufahl, Karsten, 313

R

Raab, Rodrigo, 1477

Raab-Straube, Eckhard von, 539

Rambold, Gerhard, 517

Rapp, Barbara, 157, 167

Rasche, Christoph, 1019

Raschke, Silvia, 657

Rasmussen, Karsten Boye, 197

Räuchle, Malte, 365

Rebstadt, Jonas, 1425

Reese, Kerstin, 1653

Rehm, Simone, 1675

Rehse, Jana-Rebecca, 1499

Reichelt, David Georg, 1819

Reinhardt, Lea, 709

Rennoch, Axel, 1745, 1789

Rentschler, Markus, 1797

Resch, Michael, 605

Restrepo, Diego Alejandro, 705

Reutemann, Tobias, 1149

Rieck, Julia, 1383

Riedel, Ann-Katrin, 1059

Riedel, Till, 321

Riedel1, Till, 269

Ripken, Malena, 67

Ritschel, Konstantin, 1303

Rohrmus, Dominik, 1797

Römermann, Christine, 499, 581

Roßberger, Marc, 807

Ruckdeschel, Wilhelm, 1809

Rudisch, Katharina, 313

Rudolf, Heino, 633

Ruland, Christoph, 1771

Rüter, Joachim, 359

S

Sabri, Abdelbast, 1771

Sahota, Harkiran, 251

Sander, Martin, 1325

Sauer, Fabian, 1531

Schaaf, Hylke van der, 301

Schäfer, Adeline Silva, 1853

Schäffer, Thomas, 1467

Schäfter, Sophie, 1213

Scharoba, Stefan, 1859

Schauer, Christoph, 1247

Scheerer, Max, 1825

Scheidat, Tobias, 1665

Schering, Johannes, 127

Scherr, Simon André, 1853

Scheuermann, Gerik, 695

Schindler, Josef, 1753, 1759, 1781

Schinner, Alexander, 857, 867

Schlender, Klaus, 375

Schmalriede, Florian, 57

Schmid, Ute, 1611

Schnaubelt, Marius, 1047

Schneider, Christian, 623

Schneider, Juliane, 229

Schneider, Martin A., 1789

Schomberg, Sabrina, 981
Schoormann, Thorsten, 1361, 1499
Schramm, Christina, 1661
Schreiber, Clemens, 1561
Schroeder, Ulrik, 1661
Schulte, Carsten, 1623
Schultmann, Frank, 999
Schultz, Christian, 1019
Schulz7, Michael, 1729
Schulze, Lennart, 943
Schumacher, Stephan, 1591
Schünemann, Wolf, 1603
Schuster, Thomas, 1177, 1237
Schwarzkopf, Marie-Antoinette, 79
Seifert, Stefan, 517
Selzer, Annika, 755, 773
Seßler, Matthias, 1283
Sharafeldeen, Dina, 565
Silbernagel, Selina, 1351
Simon, Anja, 1797
Simon, Frank, 1845
Sinnwell, Lukas, 333
Sokenou, Dehla, 1809
Solsbach, Andreas, 91, 167
Sommer, Heike, 623
Somplatzki, Ralf, 1839
Sonakalan, Tugay, 45
Sorge, Christoph, 963
Spetzler, Max, 45
Springob, Nils, 1635
Spychalski, Dominik, 789
Staack, Alisa, 229
Stahr, Benjamin, 103
Steden, Bettina, 175
Steinebach, Martin, 795, 1077
Steiner, Marion, 755
Steinmetz, Ralf, 425, 461
Stelzer, Dirk, 1467

Stenzel, Adina, 1303
Stephan, Peter, 607
Stepien, Leonard, 185
Stoeter, Sieke, 847
Störi, Herbert, 609
Straßburg, Sebastian, 1477, 1487
Straub, Tobias, 965
Stroscher, Jan-Philipp, 1047
Stryk, Oskar von, 1047
Sultanow, Eldar, 1263, 1265, 1283
Szemkus, Martin, 1759

T

Tanik, Milan, 197
Tellabi, Asmaa, 1771
Theel, Oliver, 57
Theesen, Cedrik, 45
Theuerkauf, René, 1729
Theurer, Lukas, 1247
Theuring, Maria, 613
Thiée, Lukas-Walter, 1221
Thimm, Heiko, 197
Thomas, Oliver, 1369, 1397, 1425
Timm, Ingo J., 773
Tomaschek, Jan, 351
Tomov, Stefan, 1193
Tontchev, Toni, 719
Tremper, Paul, 269, 321
Treu, David, 1477
Triebel, Dagmar, 517
Trollmann, Frank, 891

U

Ulrich, Josephine, 499

V

Volz, Raphael, 1177, 1237
Vosseberg, Karin, 1845

W

Wächter, Andreas, 1193
Waedt, Karl, 1745, 1753, 1759, 1771,
1781
Wahsner, Marc, 695
Waidelich, Lukas, 1237
Walgern, Julia, 229
Walther, Marina, 1603
Walz, Deborah, 1193
Wannemacher, Klaus, 1713
Warmelink, Harald, 67
Weber, Kristin, 857
Wehlitz, Robert, 743
Weigold, Matthias, 1193
Weis, Jonas, 301
Weißleder, Stephan, 1813
Welter, Sebastian, 1265
Wengler, Gregor, 657
Werner, Alexander, 1611
Wiegrefe, Daniel, 695
Wiens, Marcus, 999
Wiesbrock, Hans-Werner, 1859
Wiesmaier, Alexander, 907
Willich, Julius von, 1047

Willner, Alexander, 1797
Winau, Mona, 999
Windolph, Jonah, 743
Winter, Andreas, 57, 401
Winter, Mario, 1809, 1845
Wittig, Manuel, 1531
Wittmann, Jochen, 365
Woggon, Ulrike K., 313
Wohlgemuth, Volker, 115
Wolf, Andreas, 607
Wolf, Verena, 1653
Wollmann, Thomas, 1185
Wybrands, Marius, 207

Z

Zabloudil, Jan, 609
Zahn, Moritz von, 1255
Zander, Neele, 1059
Zeier, Alexander, 907
Ziegler, John, 1603
Zobel, Julian, 425, 461
Zöllner, Joshua, 857
Zschörnig, Theo, 743
Zvekan, Goran, 1689
Zylowski, Thorsten, 1213