

GESELLSCHAFT
FÜR INFORMATIK



Markus Böhm, Jürgen Wunderlich (Hrsg.)

AKWI Jahrestagung 2024
Wissenschaft und Praxis: Gemeinsam die digitale
Zukunft gestalten

09.-10.09.2024
Landshut, Deutschland

Gesellschaft für Informatik e.V. (GI)

Lecture Notes in Informatics (LNI) - Proceedings

Series of the Gesellschaft für Informatik (GI)

Volume P-357

ISBN 978-3-88579-801-9

ISSN 1617-5468

Volume Editors

Prof. Dr. Markus Böhm

Prof. Dr. Jürgen Wunderlich

HAW Landshut

Am Lurzenhof 1, 84036 Landshut, Germany

markus.boehm@haw-landshut.de / juergen.wunderlich@haw-landshut.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 Frank, Universität Duisburg-Essen, Germany

Barbara Hammer, Universität Bielefeld, Germany

Falk Schreiber, Universität Konstanz, Germany

Wolfgang Karl, KIT Karlsruhe, Germany

Michael Koch, Universität der Bundeswehr München, Germany

Heiko Roßnagel, Fraunhofer IAO Stuttgart, Germany

Kurt Schneider, Universität Hannover, Germany

Andreas Thor, HFT Leipzig, Germany

Ingo Timm, Universität Trier, Germany

Karin Vosseberg, Hochschule Bremerhaven, Germany

Maria Wimmer, Universität Koblenz-Landau, Germany

Dissertations

Rüdiger Reischuk, Universität Lübeck, Germany

Thematics

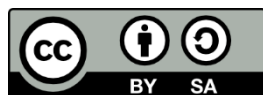
Agnes Koschmider, Universität Kiel, Germany

Seminars

Judith Michael, RWTH Aachen, Germany

© Gesellschaft für Informatik, Bonn 2024

printed by Köllen Druck+Verlag GmbH, Bonn



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

Grußwort der Sprecherin des AKWI zur 37. Jahrestagung an der HAW Landshut

Sehr geehrte Teilnehmerinnen und Teilnehmer der 37. Jahrestagung des Arbeitskreises Wirtschaftsinformatik an HAW (AKWI),

im Namen des gesamten AKWI möchte ich Sie herzlich zu unserer diesjährigen Tagung begrüßen, die vom 09. bis 10. September 2024 an der Hochschule für angewandte Wissenschaften Landshut stattfindet. Unter dem Leitthema „Wissenschaft und Praxis: Gemeinsam die digitale Zukunft gestalten“ widmen wir uns den aktuellen und zukünftigen Herausforderungen der Wirtschaftsinformatik in einer zunehmend digitalisierten Welt.

Die digitale Transformation ist nicht nur ein Schlagwort, sondern ein zentraler Treiber unserer modernen Gesellschaft. Sie stellt uns vor komplexe Fragen, die nur durch ein enges Zusammenspiel von Wissenschaft und Praxis beantwortet werden können. Unser Ziel in diesen zwei Tagen ist es daher, den Dialog zwischen Forschung und praktischer Anwendung zu fördern und gemeinsame Wege zu entwickeln, um die digitale Zukunft erfolgreich zu gestalten.

Das Programm der diesjährigen Tagung spiegelt diese Zielsetzung wider. In drei wissenschaftlichen Sessions werden hochaktuelle Forschungsergebnisse präsentiert, die neue Perspektiven auf drängende Fragen der Wirtschaftsinformatik eröffnen. Ergänzt wird der wissenschaftliche Diskurs durch praxisnahe Kurzbeiträge und Prototypen, die in einer Postersession vorgestellt werden. Diese Beiträge zeigen auf anschauliche Weise, wie theoretische Konzepte in der Praxis Anwendung finden und welche Herausforderungen dabei gemeistert werden müssen.

Besonders freue ich mich auf die beiden interaktiven Workshops, die Raum für lebendige Diskussionen bieten und den Austausch zu aktuellen Fragestellungen der Wirtschaftsinformatik intensivieren. Sie bieten uns allen die Gelegenheit, unser Wissen zu vertiefen und neue Impulse für unsere Arbeit mitzunehmen.

Abgerundet wird das Tagungsprogramm durch Keynotes und Impulsvorträge von namhaften Sprecherinnen und Sprechern aus Politik und Wirtschaft, die uns mit ihren Einsichten und Erfahrungen wertvolle Denkanstöße geben werden.

Ich wünsche Ihnen zwei inspirierende und erkenntnisreiche Tage und freue mich auf einen intensiven Austausch mit Ihnen.

Mit freundlichen Grüßen,

Vera G. Meister
Sprecherin des AKWI

Vorwort der Herausgeber

Unter dem Motto „*Wissenschaft und Praxis: Gemeinsam die digitale Zukunft gestalten*“ wollen wir den Diskurs zwischen Praktikern sowie Praktikerinnen und Wissenschaftlern sowie Wissenschaftlerinnen fördern, um gemeinsam die Herausforderungen der Digitalisierung zu meistern und ihre Potenziale zum Wohle aller zu nutzen.

Von den Hochschulen für angewandte Wissenschaften wird zunehmend mehr Forschung erwartet. Um dieser Aufgabe gerecht zu werden, werden diese auch durch entsprechende Förderprogramme (z.B. DFG Transfer HAW/TH Plus) aber auch institutionell gestärkt. Zuletzt beispielsweise durch die Vergabe des Promotionsrechts an ausgewählte bayerische Hochschulen im Rahmen des bayerischen Hochschulinnovationsgesetzes. Dies haben die Organisatoren der 37. AKWI-Jahrestagung zum Anlass genommen, auch die wissenschaftliche Qualität der Tagung zu stärken, um künftig eine hochqualitative Publikationsplattform für die angewandte Wirtschaftsinformatik zu schaffen. Aus diesem Grund erscheint dieser Tagungsband erstmalig in der Reihe der *GI-Edition: Lecture Notes in Informatics (LNI)*.

Wir durften uns über insgesamt 28 Einreichungen freuen. Dies umfasste 18 Vollbeiträge (Full Paper) abgeschlossener Forschungsarbeiten sowie 10 Kurzbeiträge laufender Forschungsarbeiten und Prototypen. Aus den 18 Vollbeiträgen wurden im Rahmen des Review-Verfahrens 9 Beiträge (50 % Annahmequote) ausgewählt welche in Teil A dieses Konferenzbandes zu finden sind. Die abgelehnten Beiträge wurden nochmals eingeladen, einen überarbeiteten Kurzbeitrag einzureichen. Von den insgesamt 19 Kurzbeiträgen wurden 10 Beiträge für den Teil B dieses Konferenzbandes angenommen (53 % Annahmequote). Das Review-Verfahren legte besonderes Augenmerk auf

- einen angemessenen neuartigen Beitrag,
- die Qualität der Argumentation,
- Forschungsmethodik und Struktur,
- Darstellung und Sprache, sowie
- das Potenzial zur Förderung einer Diskussion

Die Reviewer wurden motiviert, die zwei zentralen Funktionen des Peer-Review zu beachten: (1) Qualitätssicherung und (2) Beitragsentwicklung. Unter letzter wird vor allem verstanden, dass die Kommentare der Reviewer Anregungen zur Weiterentwicklung des Beitrags sowie des Forschungsvorhabens geben sollen, unabhängig davon, ob der Beitrag zur Annahme oder Ablehnung empfohlen wird.¹

¹ Garcia-Costa, D. et al.: Measuring the developmental function of peer review: a multi-dimensional, cross-disciplinary analysis of peer review reports from 740 academic journals. PeerJ. 10, e13539 (2022). <https://doi.org/10.7717/peerj.13539>.

Die AKWI lebt vom intensiven kollegialen Austausch. Um diesen zu stärken, wird neben dem üblichen Vortragsformat für die vollständigen Beiträge auch ein Poster-Forum angeboten. In Teil B dieses Konferenzbandes erhält der Leser einen Überblick über diese laufenden Forschungsarbeiten und Prototypen. Auch diese Kurzbeiträge wurden im Rahmen des Review Verfahrens überprüft.

Der vorliegende Tagungsband spiegelt die Vielfalt der aktuellen Forschung im Bereich digitaler Technologien und deren Anwendung in verschiedenen Sektoren wider. Die Beiträge thematisieren sowohl die Potenziale als auch die Herausforderungen, die mit der Integration und Nutzung fortschrittlicher Technologien verbunden sind. Ein zentrales Thema ist dabei die Anwendung von Künstlicher Intelligenz (KI) und maschinellem Lernen (ML) für die Analyse und Verarbeitung strukturierter und semi-strukturierter Daten.

Die in den Beiträgen adressierten Domänen reichen von der von der Alten-/Krankenpflege über den Einzelhandel, die Energieversorgung und Finanzmärkte bis hin zur fertigenden Industrie. Ebenso wird auch die Domäne der Hochschullehre adressiert. Diese Arbeiten zeigen, wie die Digitalisierung in spezifischen Kontexten genutzt werden kann, um Prozesse zu überwachen, und zu optimieren, oder neue Lösungen für bestehende Probleme zu entwickeln. Hierbei wird deutlich, dass digitale Transformation nicht nur technologische, sondern auch organisatorische und kulturelle Veränderungen erfordert. Die Beiträge untersuchen dabei immer wieder auch die Akzeptanz und Nutzung neuer Technologien, wie etwa im Bereich von digitalen Gesundheitsanwendungen, der Remote-Arbeit oder im Einzelhandel. Zudem wird auch der Datenschutz und die Regulatorik von KI-Anwendungen thematisiert.

Ein weiteres Thema der Beiträge ist die Digitalisierung von Geschäfts- und Verwaltungsprozessen. Hierbei werden Ansätze wie Process Mining, Simulationen und digitale Zwillinge vorgestellt, die das Potenzial haben, Prozesse nicht nur zu verbessern, sondern auch neu zu gestalten. Die Forschung zu IT-Sicherheit in kleinen und mittleren Unternehmen sowie die strategische Implementierung von KI in Unternehmen unterstreicht die Bedeutung eines ganzheitlichen und verantwortungsvollen Ansatzes im Umgang mit neuen Technologien. Insgesamt bietet der Tagungsband eine umfassende Perspektive auf die aktuellen Entwicklungen und zeigt auf, wie interdisziplinäre Ansätze und methodische Vielfalt zur Lösung komplexer Herausforderungen beitragen können.

Landshut, im September 2024

Markus Böhm & Jürgen Wunderlich

Sponsoren

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

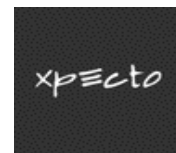
Alexander Thamm GmbH



Msg systems ag



xpecto



bayernwerk netz



DUSCHL



Zuken GmbH



audeering



SEMP CONSULTING



GradiatData



Hochschule für Angewandte Wissen-
schaften Landshut



Tagungsleitung

Gesamtleitung: Jürgen Wunderlich, HAW Landshut
Leitung des Programmkomitees: Markus Böhm, HAW Landshut

Programmkomitee / Reviewer

Wolfgang Alm	TH Aschaffenburg
Markus Böhm	HAW Landshut
Christian Czarnecki	FH Aachen
Sebastian Dünnebeil	HS München
Ingo Elsen	FH Aachen
Dorothee Feldmüller	HS Bochum
Frank Herrmann	OTH Regensburg
Norbert Ketterer	HS Fulda
Claudia Lemke	HWR Berlin
Alexander Luebbe	TH Wildau
Birte Malzahn	HTW Berlin
Christof Mosler	HFT Stuttgart
Jörg Puchan	HS München
Thomas Specht	HS Mannheim
Mike Steglich	TH Wildau
Mathias Walther	TH Wildau

Wissenschaftliches Organisationsteam

Markus Böhm	HAW Landshut
Johannes Busse	HAW Landshut
Markus Jakob	HAW Landshut
Markus Schmidtner	HAW Landshut
Dagmar Schuller	HAW Landshut
Holger Timinger	HAW Landshut
Jürgen Wunderlich	HAW Landshut

Studentisches Organisationsteam

Toni Ebner	HAW Landshut
Linda Huber	HAW Landshut
Jonathan Schuster	HAW Landshut
Andre Sinyuk	HAW Landshut

Inhaltsverzeichnis

Teil A: Vollständige Beiträge

Irene Weber

Large Language Models are Pattern Matchers: Editing Semi-Structured and Structured Documents with ChatGPT 3

Florian Schatz, Daniel Hennschen, Christian Krauss

Wolkenfrei – Ein Design Thinking Ansatz zur Entwicklung einer Sprachsteuerung in der Pflege ohne Cloud 19

Christian Reichhardt, Manuel Martin, Patrick Baier

Open Set Gestenerkennung zur Interaktion zwischen Passanten und Fahrzeugen mit Automatisierungsfunktionen 31

Lukas Koch, Can Adam Albayrak, Dirk Schreiber

Auswirkungen von Mergers und Acquisitions auf kollaborative Cloud-Services 47

Niklas Briem, Benjamin Pottkamp, Joachim Gerlach, Jessica Rövekamp, Alexander Bade

Segmentale Geschäftsmodell-Evaluation am Beispiel der digitalen Mehrwert-Anwendung Energiemanagement 61

Carolín Gotsch, Jörg Puchan

Harmonising innovation and governance: A lifecycle model for high-risk AI systems under the European AI Act 79

Sören Aguirre Reid, Frank Kammer, Dominik Kröll, Timon Pellekooorne, Harald Ritz, Markus Siepermann, Jonas Wölfer

Digital Classroom – A Communication Platform for Online Teaching 97

Can İgde, Sandy Eggert

Quantitative Analyse der Auswirkungen von Remote-Arbeit 115

Corinna Huber, Lukas Rissmann, Sebastian Meißner

Implementierung eines simulationsbasierten Digitalen Prozess-Zwillings als Entscheidungsunterstützungssystem 129

Teil B: Kurzbeiträge und Prototypen

Lena Ulrich, Birte Malzahn

Einfluss von Gamification auf die Akzeptanz von digitalen Gesundheitsanwendungen (DiGA) – Ergebnisse einer qualitativen Studie 145

Martin Przewloka, Nassim Abuhaltam, Michael Lüttebrandt, Ingrid Profitlich, Markus Maria Profitlich

Mobile Applikation zur Selbsttherapieunterstützung chronisch Erkrankter mittels Bewegungsmotivation 151

Martin Przewloka, Markus Häfele

Einsatz von Generativer KI im Self-service Micro-retailing 159

Paul Helstab, Christof Mosler, Daniel Kaiser

Researching Financial Market Dynamics through Algorithmic Trading Agents 169

Thomas Wiedemann

Simulationsgestützte Analyse der neuen dynamischen Strompreise und Alternativen zur Anreiz-Optimierung für Endkunden 173

Pia Schwägerl, Anna Schidek, Holger Timinger, Maren Martens

Modellierung von Process Mining im agilen Projektmanagement am Beispiel Scrum 181

Marianne Schmolke, Muhammad Jawad Javaid

A Use Case Analysis assessing Process Mining Model Quality for Simulation 187

Jana Feiler, Tobias Walter

Applying Predictive Process Monitoring to Predict User Behavior with Click Data 193

Martin Przewloka

A 3 phases model towards an industrialization of (generative) artificial intelligence developed for corporate environments 201

Dennis Buckenauer, Sandy Eggert

Sicherheitsmaßnahmen für den ERP-Betrieb 207

Teil A

Vollständige Beiträge

Large Language Models are Pattern Matchers: Editing Semi-Structured and Structured Documents with ChatGPT

Full Paper

Irene Weber ¹


Abstract: Large Language Models (LLMs) offer numerous applications, the full extent of which is not yet understood. This paper investigates if LLMs can be applied for editing structured and semi-structured documents with minimal effort. Using a qualitative research approach, we conduct two case studies with ChatGPT and thoroughly analyze the results. Our experiments indicate that LLMs can effectively edit structured and semi-structured documents when provided with basic, straightforward prompts. ChatGPT demonstrates a strong ability to recognize and process the structure of annotated documents. This suggests that explicitly structuring tasks and data in prompts might enhance an LLM's ability to understand and solve tasks. Furthermore, the experiments also reveal impressive pattern matching skills in ChatGPT. This observation deserves further investigation, as it may contribute to understanding the processes leading to hallucinations in LLMs.

Keywords: LLM, large language model, document processing, pattern matching, prompt engineering

1 Introduction

Large Language Models (LLMs) are extensive artificial neural networks trained on vast amounts of textual data to generate coherent continuations of given prompts. The initial training, which is time-consuming and computationally intensive, is typically followed by additional training phases. Fine-tuning with specific tasks and example responses enables LLMs to solve particular types of problems, while Reinforcement Learning with Human Feedback focuses them on delivering high-quality and socially preferred responses. Research has shown that LLMs can not only produce correct natural and formal language texts conveying plausible contents, but are also capable of reasoning, planning, and simulating other forms of intelligent behaviors. Thus, LLMs offer a wide range of potential applications, the extent of which is still not fully explored.

Frequently, LLMs are applied for creating and processing texts, for communicating, planning, and computer programming. LLMs require that all tasks and inputs are provided in a textual format. For many applications, LLMs are prompted with freely phrased, natural language text or program code. Yet, they are also capable of processing texts that are structured such that they represent data or formatted documents.

¹ Kempten University of Applied Sciences, Faculty of Mechanical Engineering, Bahnhofstraße 61, 87435 Kempten, Germany, irene.weber@hs-kempten.de,  <https://orcid.org/0000-0003-2743-1698>

The term *unstructured document* refers to textually encoded information lacking explicit organization, such as natural-language text without a defined context or fixed format. *Structured data* refers to information with an explicit and strict regular structure, like data originating from database management systems. In structured data, the meaning of a data element is defined by the structure in which it is registered, and the order of data elements, in general, is not meaningful. *Semi-structured documents* fall between unstructured documents and structured data. They have a flexible structure, often combining heterogeneous textual contents, such as short, potentially ungrammatical text fragments, longer free-text, and markup tags. [MBZ13]

There are various methods for indicating structure in texts, including markup languages like Markdown or HTML, data exchange formats like XML, JSON, and YAML, formalized languages, and tabular formats as e.g., comma-separated value (CSV) data. Specialized formalisms often build upon generic formats like XML or JSON, such as formalisms for representing process models or other types of graphs. Documents containing formatting markup, such as HTML or LaTeX, are generally considered semi-structured [MBZ13]. XML, JSON, and similar formalisms can represent semi-structured documents as well as structured data, depending on the presence and flexibility of an underlying schema. For the remainder of the paper, we will not separately name semi-structured texts where it is not relevant, but will instead understand structured texts to subsume semi-structured texts.

It is known that LLMs can handle structured inputs, having encountered the common formalisms during their basic training. Many studies explore how effectively LLMs can create structured documents from natural language text. In contrast, this paper focuses on the ability of LLMs to process already structured texts. We do not aim to convert natural language descriptions of, e.g., graphs or processes, into representations structured according to some formalism. Rather, we investigate how well LLMs can process or restructure inputs that are already structured.

Restructuring structured documents has practical applications, particularly in writing documents that include formatting and layout information, such as Markdown, HTML, or LaTeX. By inserting or adjusting such formatting, LLMs can support authoring activities beyond merely generating new content. Further applications include converting between different document formats, which is essential when data needs to be reformatted for automatic processing. In software development, the capabilities of an LLM can replace traditionally programmed conversion routines, which are often expensive to develop and test. Integrating an LLM can reduce software development costs and enable more flexible and powerful solutions than would be achievable with classical programming. However, this approach incurs ongoing operational costs if a paid LLM-as-a-Service is utilized.

Although the tasks performed by the LLM in editing structured documents may seem less demanding than other currently researched tasks, they can still bring significant labor savings and efficiency gains. The prerequisite for this to be useful is that the application

of the LLM for these tasks incurs little effort. This paper addresses the following research question:

(RQ) Can LLMs be applied for editing structured or semi-structured documents with little effort?

By 'little effort,' we mean that simple, quickly designed prompts should suffice, and the outputs of the LLM should be of high quality, requiring minimal manual post-processing. 'Editing semi-structured documents' refers to modifying their structure rather than their semantic content. To our knowledge, this question has not yet been investigated in research.

2 Related work

This study offers a qualitative exploration of an LLM's ability to transform structured inputs or convert structured inputs from one format to another. No previous work explicitly investigating this topic was identified.

The most closely related work focuses on LLM table understanding. For example, Singha et al. [Si23] and Sui et al. [Su24] conduct benchmark tests to evaluate LLM performance in interpreting structural tables. These studies present tables in various formats, including HTML, JSON, or Markdown to a range of LLMs, which then answer questions about the table data or table structure in natural language. These tests are conducted on a large scale, with performance assessed automatically. We also reviewed several applications of LLMs that operate on or produce structured outputs similar to those investigated here, as summarized in Tab. 1. However, an extensive literature review of such applications is beyond the scope of this paper.

Wu et al. [Wu20] present an application for co-reference resolution, a common task in Natural Language Processing (NLP). Their application queries an LLM twice. The first query tags a natural language input with XML tags, while the second query consumes this semi-structured result as input and yields a structured output. Two further applications use LLMs for extracting data into a queryable, highly structured tabular format. One processes various types of semi-structured documents (e.g., HTML, TXT, XML) [Ar23], while the second scans scientific articles, i.e., natural language texts, to retrieve cooling rates of metallic glasses [PM24]. An extensive overview of applications of LLMs for tasks encountered in NLP reports works where LLMs produce structured outputs from unstructured inputs [Mi23]. Several papers focus on processing graphs with LLMs. One study describes the geometric structure of graphs in natural language and then utilizes the LLM to perform graph tasks, specifically node classification [Ye24]. In [Ch24], an LLM is employed for generating structured training data to train a Graph Neural Network for node classification, thus avoiding the high costs of using the LLM for node classification directly. Jiang et al. [Ji23b] present StructGPT, a system that interfaces with various structured data pools, specifically, databases and knowledge graphs. StructGPT retrieves data from the

Ref	Purpose	Input	Output
[Wu20]	Identify co-reference	NL	SEM (XML)
[Wu20]	Resolve co-reference	SEM (XML)	STRUC
[Ar23]	Extract data	SEM (HTML, TXT, XML)	STRUC (DB)
[PM24]	Extract data	NL (scientific articles)	STRUC (DB)
[Mi23]	NLP tasks	NL	Diverse (SEM, NL, etc.)
[Ye24]	Perform graph tasks	SEM (graph)	NL (e.g., a category)
[Ch24]	Create training data	STRUC (graph)	STRUC (training data)
[Ji23b]	Answer questions	STRUC (Diverse)	NL or STRUC (queries)
[FFK23]	Draw diagrams	NL	STRUC (JSON diagrams)
[He23]	Create math exercises	NL	STRUC (LaTeX math)
[He23]	Phrase math formulae	STRUC (LaTeX math)	NL
[He23]	Create drawings	NL	STRUC (TikZ code)
[Xi24]	Create documents	NL + STRUC (an example)	STRUC (like the example)
[La23]	Edit (not create) texts	NL	NL
exp1	Edit (not create) docs	SEM (LaTeX)	SEM (LaTeX)
exp2	Edit (not create) docs	STRUC (RIS)	STRUC (OPUS XML)

Tab. 1: Applications processing structured (STRUC) or semi-structured (SEM) texts. NL indicates natural language, DB indicates database entries. ‘Create’ means ‘generate new content’. exp1 and exp2 refer to the case studies conducted in this paper.

pools and passes it to an LLM, which is tasked to answer questions based on this structured data. The LLM either provides the answer directly or generates a database query that can retrieve the answer.

The capability of ChatGPT-4 to generate entity-relationship diagrams, business process models in BPMN, and UML class diagrams from descriptions phrased in natural language is evaluated in [FFK23]. The models and diagrams are generated using representations based on JSON. In [He23], LaTeX is proposed as a means to communicate mathematical concepts and create drawings with an LLM. The LLM is tasked to generate mathematical exercises and corresponding solutions in LaTeX. It is also applied to translate LaTeX formulas into natural language, which can be read aloud to visually impaired persons. Furthermore, the LLM is tasked with creating drawings using TikZ commands, a language for producing vector graphics in LaTeX documents. Xia et al. [Xi24] contribute a benchmark dataset designed to evaluate the capabilities of LLMs in producing structured outputs across a range of application domains and document formats. Their benchmark dataset comprises prompts which instruct an LLM to create a document in a specific format with the format specified by an example. A further LLM is applied to assess whether the evaluated LLMs successfully generated documents in the required format. Laban et al. [La23] employ LLMs for editing (not generating) unstructured natural language texts. Their system aims to assist authors in writing. While these studies involve LLMs processing structured inputs or producing structured outputs, none of them investigates the capability of LLMs for reformatting or restructuring structured documents.

3 Method

To address the research question, we conduct experiments using various document formats. The research adopts a qualitative rather than a quantitative approach. The number of experiments is deliberately kept low, and the results are reviewed and evaluated "by hand". This approach allows for identifying details and making unexpected observations that automated tests with large datasets might overlook, as they typically provide only percentages of correctness as, e.g., in [Xi24]. The research aims to investigate tasks that closely resemble real-world scenarios. We conducted two series of experiments. The first series involves documents formatted with LaTeX, a widely-used typesetting language familiar to ChatGPT. The second series uses less common document formats. Here, ChatGPT is tasked with converting RIS records into an XML format used by OPUS. RIS is a standardized markup format for exchanging bibliographic information between literature management programs, while OPUS is a software used by institutions to set up publication databases [Ko26]. To ensure meaningful insights and avoid introducing unintentional biases, we use realistic sample documents. In the first series of experiments, a LaTeX-formatted table taken from a research paper [We24a] was processed. The highly specific technical terms originally presented in this table were replaced with more neutral terms using ChatGPT, without altering the structure of the table. Example documents for the second series are obtained from real university servers.

Experiments were conducted using ChatGPT (then based on GPT-3.5) through OpenAI's chat interface on April 29 and May 1, 2024. Each experiment's prompt was input into the interface, and the model's response was then analyzed externally. Chat history was cleared after each experiment to ensure independent processing. The input documents, prompts, and outputs are available online in an electronic appendix [We24b].

4 Experiment Series 1: Restructuring and reformatting LaTeX

4.1 Sample Data and Prompts

This experiment series comprises four steps in which the LaTeX table is progressively edited. The prompt for each step consists of an instruction and a table in LaTeX format, with the chat history cleared after each LLM query. Tab. 2 lists the prompts. Fig. 1 depicts the table and a piece of its LaTeX definition before the first edit.

To show the generated LaTeX tables and test the generated LaTeX commands, we manually inserted the LLM-generated tables into LaTeX documents such that a PDF could be created. The resulting tables are depicted in Fig. 2 to 5. Protocols of the experiments, along with prompts and complete versions of the input and output LaTeX tables, can be found in the electronic appendix [We24b].

Index	Topic	Course	Literature
0	Python 101	Introductory Course	[8, 7]
1	Java Fundamentals	Object-oriented Programming	[4], Code
2	Java Fundamentals	GUI Development	[4], Code
3	DataScienceBasics	Data Analysis	[2]
4	Data Science Basics	Machine Learning Models	[2]
5	Data Science Basics	Data Visualization	[2]
6	Cyber Security Fundamentals	Network Security Fundamentals	[12]
7	Cyber Security Fundamentals	Ethical Hacking	[12]
8	Web Development 101	Frontend Development	[3]
9	Mobile App Development	App Development Basics	[9]
a,b	Mobile App Development	Backend Integration	[9]
A	Machine Learning Mastery	Neural Networks	[13]
B	Machine Learning Mastery	Reinforcement Learning	[13]
C	Cloud Computing Essentials	Cloud Infrastructure	[5]
D	Database Management Systems	Database Design	[1]
E	Database Management Systems	Query Optimization	[1]
F	Game Development Basics	Graphics Programming	[6]
G	Cyber Security Professional	Incident Response	[10, 11]
H	Cyber Security Professional	Penetration Testing	[10, 11]
I	Cyber Security Professional	Threat Intelligence	[10, 11]
J	Cyber Security Professional	Cryptography	[10, 11]

```

\begin{tabular}{llll}
Index & Topic & Course & Literature \\
\hline
0 & Python 101 & Introductory Course & \cite{smithPython101Introduction2023a, smithPython101Ad} \\
1 & Java Fundamentals & Object-oriented Programming & \cite{jonesJavaFundamentalsBeginners2023}, \\
2 & Java Fundamentals & GUI Development & \cite{jonesJavaFundamentalsBeginners2023}, \href{https:

3 & DataScienceBasics & Data Analysis & \cite{brownDataScienceBasics2023} \\

```

Fig. 1: Sample LaTeX table used for experiments

No	Prompt	Result
1	I will give you a LaTeX table. Please delete the first column. '''	Fig. 2
2a	I will give you a LaTeX table. Please swap the two last columns. '''	
2b	I will give you a LaTeX table. Please swap the "Course" and "Literature" columns'''	Fig. 3
3a	I will give you a LaTeX table. I want you to reduce the number of lines as follows. Some lines only differ in the last column. Please collapse these lines in one line. Collect their last column data.	Fig. 4
3b	identical to 3a	
4a	I will give you a LaTeX table. Please format the entries in the "Course" column in Italics. Keep the formatting of separating commas as it is. '''	
4b	I will give you a LaTeX table. Please format the entries in the "Courses" column in Italics. There may be multiple entries in one cell, separated by commas. Keep the formatting of separating commas as it is. '''	
4c	I will give you a LaTeX table. Please format the entries in the "Courses" column in Italics. There may be multiple entries in one cell, separated by commas. Spare the commas out. '''	Fig. 5
4d	I will give you a LaTeX table. Please format the entries in the "Course" column in Italics excluding the commas. '''	

Tab. 2: Prompts for LaTeX restructuring experiments. Complete prompts and results are available in [We24b].

4.2 Results

In all experiments, ChatGPT generated tables in correct LaTeX syntax that the LaTeX compiler processed without issues. It was able to make all desired changes, although in some experiments, this was achieved only after modifying the prompts, as reported below. The results were not consistently reproducible, meaning that identical queries with cleared chat history sometimes, but not always, produced different outputs. This variability might stem from ChatGPT's temperature settings.

Prompt 1 produced the desired result, see Fig. 2. Prompt 2a returned the input table nearly unchanged with only a subtle modification in one cell: the content of the last column of the third row ("[8, 7]") were replaced by the content of the cell above it ("[4], Code"). Prompt 2b produced the desired result, as shown in the Fig. 3. Prompt 3a successfully restructured the table as requested. It merged rows 3 to 5, despite differences in the spelling of the "Topic" column, and adopted the spelling "Data Science Basics", as illustrated in Fig. 4. Additionally, it added a dividing line before the last table row. A second query with an unchanged prompt 3b also correctly restructured the table, but this time it adopted the spelling "DataScienceBasics" when merging rows 3 to 5 and did not generate an additional dividing line.

Topic	Course	Literature
Python 101	Introductory Course	[8, 7]
Java Fundamentals	Object-oriented Programming	[4], Code
Java Fundamentals	GUI Development	[4], Code
DataScienceBasics	Data Analysis	[2]
Data Science Basics	Machine Learning Models	[2]
Data Science Basics	Data Visualization	[2]
Cyber Security Fundamentals	Network Security Fundamentals	[12]
Cyber Security Fundamentals	Ethical Hacking	[12]
Web Development 101	Frontend Development	[3]
Mobile App Development	App Development Basics	[9]
Mobile App Development	Backend Integration	[9]
Machine Learning Mastery	Neural Networks	[13]
Machine Learning Mastery	Reinforcement Learning	[13]
Cloud Computing Essentials	Cloud Infrastructure	[5]
Database Management Systems	Database Design	[1]
Database Management Systems	Query Optimization	[1]
Game Development Basics	Graphics Programming	[6]
Cyber Security Professional	Incident Response	[10, 11]
Cyber Security Professional	Penetration Testing	[10, 11]
Cyber Security Professional	Threat Intelligence	[10, 11]
Cyber Security Professional	Cryptography	[10, 11]

Fig. 2: LaTeX table generated by Prompt 1

In step 4, ChatGPT was instructed to format specific table contents using various prompt variants, as shown Tab. 2. Specifically, certain table cells' texts were to be printed in italics, excluding commas. In all queries, the specified table contents were reformatted in Italics. However, ChatGPT only succeeded in skipping the commas as requested in some queries. Repeated queries with identical prompts sometimes succeeded and sometimes failed. The result of a successful query using Prompt 4c is depicted in Fig. 5. In some step 4 queries,

Topic	Literature	Course
Python 101	[8, 7]	Introductory Course
Java Fundamentals	[4], Code	Object-oriented Programming
Java Fundamentals		GUI Development
DataScienceBasics	[2]	Data Analysis
Data Science Basics	[2]	Machine Learning Models
Data Science Basics	[2]	Data Visualization
Cyber Security Fundamentals	[12]	Network Security Fundamentals
Cyber Security Fundamentals	[12]	Ethical Hacking
Web Development 101	[3]	Frontend Development
Mobile App Development	[9]	App Development Basics
Mobile App Development	[9]	Backend Integration
Machine Learning Mastery	[13]	Neural Networks
Machine Learning Mastery	[13]	Reinforcement Learning
Cloud Computing Essentials	[5]	Cloud Infrastructure
Database Management Systems	[1]	Database Design
Database Management Systems	[1]	Query Optimization
Game Development Basics	[6]	Graphics Programming
Cyber Security Professional	[10, 11]	Incident Response
Cyber Security Professional	[10, 11]	Penetration Testing
Cyber Security Professional	[10, 11]	Threat Intelligence
Cyber Security Professional	[10, 11]	Cryptography

Fig. 3: LaTeX table generated by Prompt 2b

Topic	Literature	Course
Python 101	[8, 7]	Introductory Course
Java Fundamentals	[4], Code	Object-oriented Programming, GUI Development
Data Science Basics		Data Analysis, Machine Learning Models, Data Visualization
Cyber Security Fundamentals	[12]	Network Security Fundamentals, Ethical Hacking
Web Development 101	[3]	Frontend Development
Mobile App Development	[9]	App Development Basics, Backend Integration
Machine Learning Mastery	[13]	Neural Networks, Reinforcement Learning
Cloud Computing Essentials	[5]	Cloud Infrastructure
Database Management Systems	[1]	Database Design, Query Optimization
Game Development Basics	[6]	Graphics Programming
Cyber Security Professional	[10, 11]	Incident Response, Penetration Testing, Threat Intelligence, Cryptography

Fig. 4: LaTeX table generated by Prompt 3a

ChatGPT added extra LaTeX commands. Specifically, it embedded the provided LaTeX text fragment in a LaTeX table environment (a structure that allows controlling the placement of the table and the inclusion of a caption and label) or even provided a complete LaTeX document (excluding the bibliography).

Topic	Literature	Courses
Python 101	[8, 7]	<i>Introductory Course</i>
Java Fundamentals	[4], Code	<i>Object-oriented Programming, GUI Development</i>
Data Science Basics	[2]	<i>Data Analysis, Machine Learning Models, Data Visualization</i>
Cyber Security Fundamentals	[12]	<i>Network Security Fundamentals, Ethical Hacking</i>
Web Development 101	[3]	<i>Frontend Development</i>
Mobile App Development	[9]	<i>App Development Basics, Backend Integration</i>
Machine Learning Mastery	[13]	<i>Neural Networks, Reinforcement Learning</i>
Cloud Computing Essentials	[5]	<i>Cloud Infrastructure</i>
Database Management Systems	[1]	<i>Database Design, Query Optimization</i>
Game Development Basics	[6]	<i>Graphics Programming</i>
Cyber Security Professional	[10, 11]	<i>Incident Response, Penetration Testing, Threat Intelligence, Cryptography</i>

Fig. 5: LaTeX table generated by Prompt 4c

5 Experiment Series 2: Converting structured documents

5.1 Sample Data and Prompts

The second series of experiments investigates the capability of the LLM in converting structured documents between different formats. We use RIS and OPUS XML data originating from the OPUS servers of Landshut University of Applied Sciences² (HAWL) and Technical University Rosenheim³ (THR) for the case study. Both servers offer the option to export stored publications in RIS and in XML format. Tab. 3 gives an overview over the data used for the experiments. The differing numbers of fields show that the RIS and XML exports are not as uniform as might be expected.

Ref.	Id	Source	Conf.	1-shot example	RIS	XML
[Se22]	SEEHUBER	HAWL	3. Symp ESI	X	17	38
[MFM22]	MUENCH	HAWL	3. Symp ESI		16	35
[Zu21]	ZUGSCHWERT	HAWL	- na -		17	23
[SH24]	SELIGER	THR	CIPS 2024		18	36

Tab. 3: Data for series 2 of experiments. The table shows the number of fields in the RIS exports, the number of fields in the OPUS export Seehuber.xml, which serves as an example, and the number of XML fields generated by the LLM (printed in *italics*).

All publications were exported in RIS format and SEEHUBER also in XML format. SEEHUBER.RIS and SEEHUBER.XML serve as one-shot prompt examples for ChatGPT. Fig. 6 shows SEEHUBER.RIS, and Fig. 7 illustrates some fields of SEEHUBER.XML. SEEHUBER.RIS and MUENCH.RIS are conference contributions to the same conference and therefore contain several identical fields. ZUGSCHWERT.RIS and SELIGER.RIS, which are also conference contributions, do not contain all fields in SEEHUBER.RIS, but they do contain additional fields not present in SEEHUBER.RIS.

² <https://opus4.kobv.de/opus4-haw-landshut>

³ <https://opus4.kobv.de/opus4-rosenheim>

```
TY - CONF
A1 - Seehuber, Stefan
A1 - Crämer, Peter
A1 - Kipfelsberger, Stefan
A1 - Versen, Martin
A2 - Artem, Ivanov
A2 - Marc, Bicker
A2 - Peter, Patzelt
T1 - EtherCAT Gateway für eine [...] Visualisierung [...]
T2 - Tagungsband 3. Symposium Elektronik und Systemintegration ESI
N2 - Die [...]
Y1 - 2022
UR - https://opus4.kobv.de/opus4-haw-landshut/frontdoor/index/index/docId/366
UR - https://nbn-resolving.org/urn:nbn:de:bvb:860-opus4-3666
SN - 978-3-9818439-6-5
SP - 98
EP - 106
ER -
```

Fig. 6: SEEHUBER.RIS as exported from the HAWL OPUS server

The prompt for ChatGPT is constructed using a one-shot pattern. It includes a brief instruction, SEEHUBER.RIS and SEEHUBER.XML at positions %%1 and %%2, and one of MUENCH.RIS, ZUGSCHWERT.RIS and SELIGER.RIS as a task at position %%3:

```
I will input a ris-document. Please convert it to Opus-XML. First, you
will be provided with an example input and output.
Here is the example input: ''' %%1 '''
Here is the example output: ''' %%2 '''
Here is the ris-document you must convert:''' %%3 '''
```

Consequently, for each experiment, the prompt contained SEEHUBER.RIS and SEEHUBER.XML as the example, and the RIS of one publication. ChatGPT was queried once with each of the resulting prompts. The XML it generated was then compared to the original corresponding RIS and to the example SEEHUBER.XML.

5.2 Additional prompts

ChatGPT was also prompted to convert a RIS into OPUS XML with a zero-shot prompt, i.e., a prompt lacking an example. The zero-shot prompting yielded a syntactically correct XML with a plausible structure and plausibly named fields, but differing from an actual OPUS XML export. This indicates that ChatGPT did not learn these formats or their interconnections during its training. ChatGPT was also asked about details of the publication by Seehuber et al. [Se22] and by Zugschwert et al. [Zu21] and stated not to know them as follows: I don't have access to specific publications or writings

```

<id>366</id>
<completedYear>2022</completedYear>
<language>deu</language>
<pageFirst>98</pageFirst>
<pageLast>106</pageLast>
<type>conferenceobject</type>
<title language="deu">EtherCAT Gateway für eine Arduino ... </title>
<abstract language="deu">Die Luftqualität i....</abstract>
<parentTitle language="deu">
  Tagungsband 3. Symposium Elektronik und Systemintegration ESI
</parentTitle>
<identifier type="isbn">978-3-9818439-6-5</identifier>
<identifier type="doi">10.57688/366</identifier>
<identifier type="urn">urn:nbn:de:bvb:860-opus4-366</identifier>
<author>Stefan Seehuber</author>
<author>Peter Crämer</author>
<author>Stefan Kippelsberger</author>
<author>Martin Versen</author>
<collection role="collections" number="">
  Tagungsband 3. Symposium Elektronik und Systemintegration ESI 2022: Fachbeiträge; ISBN 978-3-9818439-4-1
</collection>
<thesisPublisher>Hochschule für Angewandte Wissenschaften Landshut</thesisPublisher>
<file>
  https://opus4.kobv.de/opus4-haw-landshut/files/366/3ESI2022_Tagungsband_Seehuber.pdf
</file>

```

Fig. 7: Excerpt of SEEHUBER.XML as exported from the HAWL OPUS server

by SeeHuber, Crämer, and Kippelsberger in 2022 regarding Luftqualität (air quality). [...] my last update in January 2022.

5.3 Results

ChatGPT generated the XML format for all prompts without any syntactic errors. Fig. 8 shows excerpts of the output generated for SELIGER.RIS. The complete outputs of all experiments can be found in the electronic appendix [We24b]. ChatGPT correctly created all XML fields present in the example XML. For author fields occurring in varying numbers, it created the correct number of fields in the XML and filled them correctly with the authors' names as values. The names occurring in the format "Lastname, Firstname" in RIS documents were transferred to XML as "Firstname Lastname" matching the provided example. The RIS files do not contain language information. ChatGPT added this information to match the actual language of the publication, replacing "deu" with "eng", for example, `<title language="deu">` → `<title language="eng">`, according to the provided XML example. Fields that were present in the example SEEHUBER.XML but not in the example SEEHUBER.RIS were correctly filled in the generated XML documents; e.g.,
 PU VDE VERLAG GMBH → `<publisherName>VDE VERLAG GMBH</publisherName>`
 CY Düsseldorf → `<publisherPlace>Düsseldorf</publisherPlace>`

RIS fields of type KW (keywords) that were present in the new RIS documents, e.g., in SELIGER.RIS, but not in the example SEEHUBER.RIS and SEEHUBER.XML, were not added

to the generated XML, i.e., `SELIGER.XML`, meaning that no new field identifiers for keywords were hallucinated. For RIS fields of type A2 (editors of conference proceedings) that were not provided in the `ZUGSCHWERT.RIS` or the `SELIGER.RIS`, ChatGPT did not generate entries in the XML, hence editor names were neither copied nor hallucinated.

```
<id>2386</id>
<completedYear>2024</completedYear>
<language>eng</language>
<pageFirst>297</pageFirst>
<pageLast>303</pageLast>
<type>conferenceobject</type>
<publisherName>VDE VERLAG GMBH</publisherName>
<publisherPlace>Düsseldorf</publisherPlace>
<title language="eng">
  A high-frequency performance and degradation study of adhesive conductive EMI shielding tapes during High-Te
</title>
<abstract language="eng">The vari...</abstract>
<identifier type="isbn">978-3-8007-6288-0</identifier>
<identifier type="issn">0341-3934</identifier>
<identifier type="doi">10.57688/2386</identifier>
<identifier type="urn">urn:nbn:de:bvb:860-opus4-2386</identifier>
<author>Norbert Seliger</author>
<author>Cordula Helmbrecht</author>
<collection role="institutes" number="">Rosenheim University of Applied Sciences</collection>
<collection role="collections" number="">
  Proceedings CIPS 2024 - 13th International Conference on Integrated Power Electronics Systems; ISBN 978-3-8007-
</collection>
<thesisPublisher>Rosenheim University of Applied Sciences</thesisPublisher>
<file>
  https://opus4.kobv.de/opus4-rosenheim/files/2386/CIPS2024_Proceedings_Seliger.pdf
</file>
```

Fig. 8: Excerpt from `SELIGER.XML` as generated by ChatGPT

For fields not present in the example `SEEHUBER.RIS` but having more or less matching fields in the example `SEEHUBER.XML`, ChatGPT constructed appropriate entries in the XML documents. In constructing these values, ChatGPT worked in a very detailed manner. We analyze three occurrences more closely: First, ChatGPT correctly derived document IDs, presumably by extracting them from the URLs provided in the RIS tag `UR`. Notably, RIS documents do not comprise a tag for the document ID, while the XML format comprises a dedicated `<id>` field (compare Fig. 6 and Fig. 7). Second, ChatGPT constructed the string value for the `<collection role="collections">` XML field, presumably from the values provided in the `A1`, `T2`, `Y1` and `SN` RIS tags, where the last part of this string and the `SN` value do not completely match in the `SEEHUBER` example provided to ChatGPT. It should be noted that the RIS documents contain multiple `A1` fields (the authors), and ChatGPT consistently selected the value of the first `A1` in all experiments. Third, it constructed the download links for the `<file>` XML fields from multiple parts taken from several RIS fields and the document ID, which it had to extract from the `UR`. It also adapted the year 2021 in the download link in `ZUGSCHWERT.XML`. Tab. 4 juxtaposes the RIS fields containing the provided information and the constructed strings for each publication.

It should be emphasized that ChatGPT constructed the strings in the `f` and `c` fields from scratch, and that these strings were embedded within the complete XML by a single query to the LLM, rather than being constructed and positioned separately. Not all the constructed

Values from SEEHUBER.RIS and SEEHUBER.XML				
Y1	2022	SN	978-3-9818439-6-5	
A1	Seehuber, Stefan			
A1	Crämer, Peter			
A1	Kipfelsberger, Stefan			
A1	Versen, Martin			
T2	Tagungsband 3. Symposium Elektronik und Systemintegration ESI			
UR	https://opus4.kobv.de/opus4-haw-landshut/frontdoor/index/index/docId/366			
c	Tagungsband 3. Symposium Elektronik und Systemintegration ESI 2022: Fachbeiträge; ISBN 978-3-9818439-4-1			
f	https://opus4.kobv.de/opus4-haw-landshut/files/366/3ESI2022_Tagungsband_Seehuber.pdf			
Values from MUENCH.RIS and generated by ChatGPT				
Y1	2022	SN	978-3-9818439-6-5	
A1	Münch, Andreas			
A1	Frauenschläger, Tobias			
A1	Mottok, Jürgen			
T2	Tagungsband 3. Symposium Elektronik und Systemintegration ESI			
UR	https://opus4.kobv.de/opus4-haw-landshut/frontdoor/index/index/docId/365			
c	Tagungsband 3. Symposium Elektronik und Systemintegration ESI 2022: Fachbeiträge; ISBN 978-3-9818439-4-1			
f	https://opus4.kobv.de/opus4-haw-landshut/files/365/3ESI2022_Tagungsband_Muench.pdf			
Values from ZUGSCHWERT.RIS and generated by ChatGPT				
Y1	2021	SN	- na-	
A1	Zugschwert, Christina			
A1	Göschl, Sebastian			
A1	Ibanez, Federico Martin			
A1	Pettinger, Karl-Heinz			
T2	- na -			
UR	https://opus4.kobv.de/opus4-haw-landshut/frontdoor/index/index/docId/303			
c	Tagungsband 3. Symposium Elektronik und Systemintegration ESI 2021: Fachbeiträge; ISBN 978-3-9818439-4-1			
f	https://opus4.kobv.de/opus4-haw-landshut/files/303/3ESI2021_Tagungsband_Zugschwert.pdf			
Values from SELIGER.RIS and generated by ChatGPT				
Y1	2024	SN	978-3-8007-6288-0	
A1	Seliger, Norbert			
A1	Helmbrecht, Cordula			
T2	Proceedings CIPS 2024 - 13th International Conference on Integrated Power Electronics Systems			
UR	https://opus4.kobv.de/opus4-rosenheim/frontdoor/index/index/docId/2386			
c	Proceedings CIPS 2024 - 13th International Conference on Integrated Power Electronics Systems; ISBN 978-3-8007-6288-0			
f	https://opus4.kobv.de/opus4-rosenheim/files/2386/CIPS2024_Proceedings_Seliger.pdf			

Tab. 4: Constructing strings. Y1, SN, T2, UR are RIS tags and values. c and f signify the <collection> and the <field> XML fields. SEEHUBER.XML fields are part of the example XML included in the prompt. The remaining c and f values are generated by the LLM (printed in italics).

values are "correct" in the real world. While the derived document IDs are valid, the constructed links are not. The link in `MUENCH.XML` is a near miss, as the actual link only differs in the spelling "Muench" versus "Münch" from the generated one.

6 Discussion and Conclusion

This paper contributes a qualitative investigation on the original research question whether an LLM can successfully edit semi-structured documents and transform structured documents when prompted with basic and straightforward instructions. We conducted two case studies comprising multiple experiments, one restructuring a LaTeX table, and one converting RIS documents to OPUS XML format. Our results indicate that the research question has a positive answer. In all experiments, the LLM produced syntactically correct documents which could be further processed without issues. The research followed a qualitative approach, conducting a limited number of experiments. Additional and broader experiments are needed to determine if this finding generalizes to other restructuring tasks and LLMs.

While related studies on LLM capabilities [Si23, Su24, Xi24, Zh23] conduct massive tests and evaluate results automatically (e.g., through LLMs [Xi24]), our qualitative approach includes a comprehensive and in-depth manual analysis of the results. This allows us to contribute the following detailed observations. In the LaTeX experiments described in Sect. 4, the LLM was tasked with restructuring a LaTeX table. We found that the LLM understood concepts related to tables such as "row", "column" and "cell" very well. Referring to table columns by their titles worked better than referencing them by their position (i.e., "last"). While the LLM reliably recognized and handled the structure explicated by LaTeX annotations, it struggled with recognizing commas as structure indicators. These observations lead to the hypothesis that explicit structural annotations (such as LaTeX commands) may enhance an LLM's understanding of tasks and data provided in prompts, thereby yielding better outputs. Specifically, they might improve the LLM's instruction-following and format-following capabilities [Xi24, Zh23], which are crucial when developing LLM-integrated applications [We24a]. Further experiments exploring this hypothesis will be valuable.

The RIS→XML experiments in Sect. 5 reveal that the LLM has impressive pattern matching skills, which become evident in the strings it constructed, compare Tab. 4. It seems plausible that its working principle involves identifying relationships (i.e., patterns) between RIS and XML elements in the example documents and replicate these in the documents it was tasked with generating. Some data elements generated are correct with respect to the real world, while other data elements are near misses or completely deviating, as elaborated in Sect. 5.3. However, it does not seem appropriate to label the latter data elements as "hallucinated", as the process that generated them is comprehensible and reasonable, albeit overgeneralizing to some extent. This pattern matching behavior deserves further investigation, as it may constitute a novel approach to understanding the processes leading to hallucinations in LLMs [Ji23a].

References

- [Ar23] Arora, S.; Yang, B.; Eyuboglu, S.; Narayan, A.; Hojel, A.; Trummer, I.; Ré, C.: Language Models Enable Simple Systems for Generating Structured Views of Heterogeneous Data Lakes. 2023, DOI: 10.48550/arXiv.2304.09433.
- [Ch24] Chen, Z.; Mao, H.; Wen, H.; Han, H.; Jin, W.; Zhang, H.; Liu, H.; Tang, J.: Label-Free Node Classification on Graphs with Large Language Models (LLMs). 2024, DOI: 10.48550/arXiv.2310.04668.
- [FFK23] Fill, H.-G.; Fettke, P.; Köpke, J.: Conceptual Modeling and Large Language Models: Impressions From First Experiments With ChatGPT. *Enterprise Modelling and Information Systems Architectures (EMISA)* 18, pp. 1–15, 2023, DOI: 10.18417/emisa.18.3.
- [He23] Helfrich-Schkarbanenko, A.: *Mathematik und ChatGPT: Ein Rendezvous am Fuße der technologischen Singularität*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2023, ISBN: 978-3-662-68209-8.
- [Ji23a] Ji, Z.; Lee, N.; Frieske, R.; Yu, T.; Su, D.; Xu, Y.; Ishii, E.; Bang, Y. J.; Madotto, A.; Fung, P.: Survey of Hallucination in Natural Language Generation. *ACM Comput. Surv.* 55 (12), 248:1–248:38, 2023, DOI: 10.1145/3571730.
- [Ji23b] Jiang, J.; Zhou, K.; Dong, Z.; Ye, K.; Zhao, W. X.; Wen, J.-R.: StructGPT: A General Framework for Large Language Model to Reason over Structured Data. 2023, DOI: 10.48550/arXiv.2305.09645.
- [Ko26] Kooperative Bibliotheksverbund Berlin-Brandenburg (KOBV): Was Ist OPUS? - OPUS 4 Handbuch, <https://www.opus-repository.org/userdoc/introduction.html>, Freitag, 29. September 2023, 11:40:26, visited on: 04/29/2024.
- [La23] Laban, P.; Vig, J.; Hearst, M. A.; Xiong, C.; Wu, C.-S.: Beyond the Chat: Executable and Verifiable Text-Editing with LLMs. 2023, DOI: 10.48550/arXiv.2309.15337.
- [MBZ13] Madani, A.; Boussaid, O.; Zegour, D. E.: Semi-Structured Documents Mining: A Review and Comparison. *Procedia Computer Science* 22, pp. 330–339, 2013, DOI: 10.1016/j.procs.2013.09.110.
- [MFM22] Münch, A.; Frauenschläger, T.; Mottok, J.: Integration of a Security Gateway for Critical Infrastructure into Existing PKI Systems. In: *Tagungsband 3. Symposium Elektronik Und Systemintegration ESI*. Pp. 88–97, 2022, ISBN: 978-3-9818439-6-5, DOI: 10.57688/365.
- [Mi23] Min, B.; Ross, H.; Sulem, E.; Veyseh, A. P. B.; Nguyen, T. H.; Sainz, O.; Agirre, E.; Heintz, I.; Roth, D.: Recent Advances in Natural Language Processing via Large Pre-trained Language Models: A Survey. *ACM Computing Surveys* 56 (2), 30:1–30:40, 2023, ISSN: 0360-0300, DOI: 10.1145/3605943.
- [PM24] Polak, M. P.; Morgan, D.: Extracting Accurate Materials Data from Research Papers with Conversational Language Models and Prompt Engineering. *Nature Communications* 15 (1), p. 1569, 2024, ISSN: 2041-1723, DOI: 10.1038/s41467-024-45914-8.
- [Se22] Seehuber, S.; Krämer, P.; Kipfelsberger, S.; Versen, M.: EtherCAT Gateway für eine Arduino basierte Luftqualitätsmessung zur Visualisierung an eine Beckhoff SPS. In: *Tagungsband 3. Symposium Elektronik und Systemintegration ESI*. Pp. 98–106, 2022, DOI: 10.57688/366.
- [SH24] Seliger, N.; Helmbrecht, C.: A High-Frequency Performance and Degradation Study of Adhesive Conductive EMI Shielding Tapes during High-Temperature Storage. In: *Proceedings CIPS 2024 - 13th International Conference on Integrated Power Electronics Systems*. VDE VERLAG GMBH, pp. 297–303, 2024.

- [Si23] Singha, A.; Cambronero, J.; Gulwani, S.; Le, V.; Parnin, C.: Tabular Representation, Noisy Operators, and Impacts on Table Structure Understanding Tasks in LLMs. 2023, doi: 10.48550/arXiv.2310.10358.
- [Su24] Sui, Y.; Zhou, M.; Zhou, M.; Han, S.; Zhang, D.: Table Meets LLM: Can Large Language Models Understand Structured Table Data? A Benchmark and Empirical Study. In: Proceedings of the 17th ACM International Conference on Web Search and Data Mining. WSDM '24, Association for Computing Machinery, New York, NY, USA, pp. 645–654, 2024, doi: 10.1145/3616855.3635752.
- [We24a] Weber, I.: Large Language Models as Software Components: A Taxonomy for LLM-Integrated Applications, 2024, doi: 10.48550/arXiv.2406.10300.
- [We24b] Weber, I.: Weberi/2024_AKWI_structured_gpt_experiments, https://github.com/weberi/2024_AKWI_structured_gpt_experiments, 2024, visited on: 07/06/2024.
- [Wu20] Wu, W.; Wang, F.; Yuan, A.; Wu, F.; Li, J.: CorefQA: Coreference Resolution as Query-based Span Prediction. In (Jurafsky, D.; Chai, J.; Schluter, N.; Tetreault, J., eds.): Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics, Online, pp. 6953–6963, 2020, doi: 10.18653/v1/2020.acl-main.622.
- [Xi24] Xia, C.; Xing, C.; Du, J.; Yang, X.; Feng, Y.; Xu, R.; Yin, W.; Xiong, C.: FOFO: A Benchmark to Evaluate LLMs’ Format-Following Capability. 2024, doi: 10.48550/arXiv.2402.18667.
- [Ye24] Ye, R.; Zhang, C.; Wang, R.; Xu, S.; Zhang, Y.: Language Is All a Graph Needs. 2024, doi: 10.48550/arXiv.2308.07134.
- [Zh23] Zhou, J.; Lu, T.; Mishra, S.; Brahma, S.; Basu, S.; Luan, Y.; Zhou, D.; Hou, L.: Instruction-Following Evaluation for Large Language Models. 2023, doi: 10.48550/arXiv.2311.07911.
- [Zu21] Zugschwert, C.; Göschl, S.; Ibanez, F. M.; Pettinger, K.-H.: Development of a Multi-Timescale Method for Classifying Hybrid Energy Storage Systems in Grid Applications. In. Pp. 1–7, 2021, doi: 10.57688/303.

Wolkenfrei – Ein Design Thinking Ansatz zur Entwicklung einer Sprachsteuerung in der Pflege ohne Cloud

Vollständiger Beitrag

Florian Schatz¹, Daniel Hennschen², Christian Krauss³

Abstract: Die vorliegende Arbeit untersucht die Entwicklung und Implementierung eines offline-fähigen Sprachsteuerungssystems in Pflegeeinrichtungen. Angesichts des demografischen Wandels und des zunehmenden Fachkräftemangels im Gesundheitswesen bietet die Technologie die Möglichkeit, Pflegeprozesse zu optimieren und die Selbstständigkeit der Patienten zu fördern. Unter Anwendung der Design Thinking Methode wurden Anforderungen direkt von den Endnutzern erhoben und in die Gestaltung des Systems integriert. Die Ergebnisse zeigen, dass solche Systeme die Interaktion der Patienten mit ihrer Umgebung verbessern können, ohne deren Datenschutz zu gefährden.

Keywords: Sprachsteuerung, Pflegetechnologie, Datenschutz, Offline-Computing, Smart Home, User-centered Design, Design Thinking, Künstliche Intelligenz

1 Einleitung

Gesundheitliche Beeinträchtigungen können das tägliche Leben von Menschen erheblich beeinflussen und führen oft zu erheblichen Einschränkungen in ihrer Handlungsfähigkeit. Der Verlust an Selbstständigkeit und die daraus resultierende Abhängigkeit von der Hilfe anderer stellen sowohl physische als auch psychische Herausforderungen dar. In dieser kritischen Lebensphase bieten Pflegeeinrichtungen die notwendige Unterstützung, stehen jedoch selbst vor komplexen Herausforderungen. Der demografische Wandel [BA23] führt zu einer immer älter werdenden Bevölkerung, und der Fachkräftemangel im Pflegebereich [D19] verschärft diese Situation weiter. Die daraus resultierende Belastung des Pflegepersonals wird voraussichtlich in den kommenden Jahren weiter zunehmen, insbesondere durch einen höheren Anteil älterer Beschäftigter in diesem Sektor.

Angesichts dieser komplexen Herausforderungen stellt sich im Zeitalter der Digitalisierung die Frage, welchen Beitrag innovative Technologien zur Verbesserung der Situation für Pflegebedürftige und Pflegepersonal leisten können. Digitale Technologien

¹ Fachhochschule Kiel, Fachbereich Medien, Grenzstr. 3, 24149 Kiel, florian.schatz@fh-kiel.de

² Muthesius Kunsthochschule Kiel, Legienstraße 35, 24103 Kiel, daniel.hennschen@mkh-mail.de

³ Fachhochschule Kiel, Fachbereich Wirtschaft, Sokratesplatz 2, 24149 Kiel, christian.krauss@fh-kiel.de

finden bereits Anwendung im Pflege- und Gesundheitsbereich, oft mit einem Fokus auf die Optimierung administrativer Prozesse. Jedoch werden technische Assistenzsysteme, wie Smart Home Technologien, die das Potenzial bieten, Routineaufgaben zu übernehmen und die Lebensqualität zu verbessern, nur selten genutzt [BD22].

Diese Arbeit untersucht daher, wie Smart Home Systeme ohne Cloud-Anbindung im Pflegekontext eingesetzt werden können, um die Autonomie der Pflegebedürftigen zu erhöhen und gleichzeitig das Pflegepersonal zu entlasten. Dabei wird der Fokus auf die Gestaltung eines benutzerfreundlichen und datenschutzkonformen Systems gelegt, das den rechtlichen und praktischen Anforderungen in Pflegeeinrichtungen gerecht wird.

2 Grundlagen

2.1 Smart Home und Ambient Assisted Living

Der Begriff Smart Home wird in der Literatur vielfältig verwendet und kann verschiedene Konzepte umfassen. Smart Homes nutzen demnach Internet of Things (IoT), Big Data, Cloud Computing und Informationstechnologie, um durch die Vernetzung von Geräten Dienste wie die Steuerung von Beleuchtung, Heizung und Sicherheitssystemen zu ermöglichen.

Ambient Assisted Living (AAL) bezeichnet speziell für ältere Menschen oder Menschen mit gesundheitlichen Einschränkungen entwickelte Smart Home-Systeme. Diese Technologien sollen die Autonomie und Lebensqualität ihrer Nutzer erhöhen und sind ein spezifischer Anwendungsbereich von Smart Homes [QSA15]. Sie integrieren Technologien, die Sicherheit, Komfort, Gesundheit und Energiemanagement verbessern und sind besonders im Kontext des demografischen Wandels und der steigenden Anzahl pflegebedürftiger Menschen von Bedeutung.

Die Integration digitaler Technologien in der Pflege ist ein fortschreitendes Feld, das zahlreiche Potenziale birgt, jedoch auch mit bedeutenden Herausforderungen verbunden ist. In der wissenschaftlichen Diskussion wird oft kritisiert, dass neue Systeme häufig ohne ausreichende Berücksichtigung der tatsächlichen Anforderungen und Bedingungen in Pflegeeinrichtungen entwickelt werden [Sc16]. Studien, die sich mit dem Einsatz von Technologien im realen Pflegeumfeld beschäftigen, sind rar, und oft beschränken sich die Tests auf künstliche Bedingungen oder Modellwohnungen. Insbesondere im Bereich der Robotik ist zwar eine Vielzahl von Forschungsarbeiten zu verzeichnen, jedoch wird die Integration und das Feedback der tatsächlichen Endnutzer – Pflegekräfte und Pflegebedürftige – oft vernachlässigt [MJK22].

Die Studie von Eggert und Stanke [ES20] zeigt auf, dass bei der Entwicklung von digitalen Assistenzsystemen, wie Sprachassistenten, nicht nur funktionale Anforderungen wie Terminplanung oder Dokumentation der Behandlungen eine Rolle spielen. Auch

nichtfunktionale Aspekte wie Systemstabilität, -integration, Lautstärkeregelung und die Erkennung schwer verständlicher Sprachbefehle müssen berücksichtigt werden. Die Autoren heben hervor, dass der Einsatz solcher Technologien zu einer effizienteren Nutzung der Zeit führen und mehr Raum für zwischenmenschliche Interaktionen schaffen sollte. Zudem sollten Patienten die Möglichkeit haben, die Nutzung solcher Assistenzsysteme frei zu wählen.

Muckenhuber et al. [MJK22] betonen, dass die Digitalisierung zwar in einigen Bereichen, wie der Pflegedokumentation, als unterstützend wahrgenommen wird, jedoch auch Bedenken hinsichtlich des Verlustes menschlicher Interaktionen bestehen. Technologien, die potenziell den direkten Umgang mit Patienten beeinflussen könnten, werden teilweise kritisch gesehen. Pflegekräfte sind grundsätzlich offen für den Einsatz neuer Technologien, solange diese sicher im Alltag integriert werden können und nicht die Kernaspekte der Pflege – insbesondere die zwischenmenschlichen Beziehungen – beeinträchtigen. Bedenken bezüglich Datenschutz und Finanzierung spielen ebenfalls eine wesentliche Rolle.

Die genannten Studien verdeutlichen, dass eine enge Einbindung der Endnutzer in den Gestaltungs- und Entwicklungsprozess essentiell ist, um Technologien zu schaffen, die nicht nur technisch fortschrittlich, sondern auch im pflegerischen Kontext sinnvoll und akzeptiert sind. Dieser integrative Ansatz kann dazu beitragen, die digitale Transformation in der Pflege effektiver zu gestalten und gleichzeitig sicherzustellen, dass die Technologie die menschlichen Aspekte der Pflege unterstützt und bereichert, anstatt sie zu ersetzen.

2.2 Rechtliche Aspekte

Die zunehmende Integration von Smart Home Technologien in den Pflegebereich bringt erhebliche Datenschutzfragen mit sich, insbesondere aufgrund der Notwendigkeit einer ständigen Internetverbindung für die meisten dieser Systeme. Diese Abhängigkeit führt zu Bedenken hinsichtlich der Sicherheit und Privatsphäre der Nutzer, da sensible Gesundheitsdaten oft in Cloud-Systemen gespeichert werden. Dies steht im Kontrast zu den strikten Vorgaben der Datenschutzgrundverordnung (DSGVO), die spezifische Regelungen zur Verarbeitung besonders sensibler Kategorien personenbezogener Daten, wie Gesundheitsdaten, vorsieht. Laut Artikel 9 der DSGVO ist eine derartige Verarbeitung nur unter strengen Bedingungen zulässig und bedarf oft der expliziten Zustimmung der betroffenen Personen.

Die Herausforderung besteht darin, dass viele der aktuellen Smart Home Systeme nicht eindeutig darlegen, wie und von wem die Nutzerdaten verarbeitet werden [EK22]. Die Komplexität dieser Systeme erschwert es, klare Verantwortlichkeiten zu definieren, was eine Risikoquelle für Datenschutzverletzungen darstellt. Technologien, die im medizinischen Bereich eingesetzt werden, müssen zudem besondere Vorsicht bei der Datenverarbeitung walten lassen, um den hohen Anforderungen der DSGVO gerecht zu werden.

Zusätzlich ist der Einsatz derartiger Technologien generell von den Bedenken möglicher Anwender abhängig, welche sich nach Kim et al. vor allem auf Sicherheits-, Datenschutz- und Vertrauens- sowie praktische und ergonomische Aspekte der Benutzerfreundlichkeit beziehen [KCJ20].

Die Integration von Datenschutzmaßnahmen schon während der Entwicklungsphase der Technologie, ein Konzept das als "Privacy by Design" bekannt ist, wird durch Artikel 25 der DSGVO gefordert. Dies beinhaltet, dass Systeme so gestaltet werden müssen, dass sie Datenschutzgrundsätze wie Datensparsamkeit von Anfang an einbeziehen und standardmäßig datenschutzfreundliche Voreinstellungen ("Privacy by Default") nutzen [KAG19]. Doch trotz dieser Vorschriften gibt es bei vielen existierenden Systemen, die oft auf eine ständige Internetverbindung angewiesen sind, erhebliche Schwierigkeiten. Die standardmäßige Konfiguration dieser Systeme ermöglicht oft eine weite Verarbeitung und Analyse von Nutzerdaten, was die Gefahr birgt, dass sensible Informationen ohne ausreichende Sicherheitsmaßnahmen gehandhabt werden [REB21].

Zudem wirft die Verwendung von nicht europäischen Anbietern zusätzliche juristische Probleme auf, insbesondere wenn Daten auf Server außerhalb der Europäischen Union gespeichert werden, was häufig der Fall ist. Solche Praktiken stehen oft im Widerspruch zu den strengen Datenschutzanforderungen der EU und können durch internationale Gesetze wie den U.S. Patriot Act weiter kompliziert werden, der von US-basierten Unternehmen verlangen kann, Daten an Regierungen weiterzugeben [A01].

Abschließend lässt sich sagen, dass trotz der vorhandenen gesetzlichen Rahmenbedingungen und der erkennbaren Notwendigkeit für einen starken Datenschutz die Praxis oft hinter den Erwartungen zurückbleibt. Eine Umgestaltung bestehender Systeme, die eine Dezentralisierung der Datenverarbeitung und eine stärkere lokale Datenkontrolle ermöglicht, könnte hier Abhilfe schaffen und die Akzeptanz von Smart Home Technologien in sensiblen Bereichen wie der Pflege verbessern. Dies würde nicht nur die rechtlichen Bedingungen erfüllen, sondern auch das Vertrauen der Nutzer in die Technologie stärken.

2.3 Design Thinking Ansatz

Der Design Thinking Prozess, der in dieser Arbeit zur Anwendung kommt, orientiert sich am fünfphasigen Modell des Hasso Plattner Institute of Design at Stanford [PMW09]. Dieses Modell betont besonders die Phasen „Ideation“ und „Prototype“ und ist bekannt für seine Flexibilität und Anpassungsfähigkeit an realweltliche Herausforderungen. Die iterative Natur des Prozesses erlaubt es, Lösungen zu entwickeln, die praxisnah und auf die Bedürfnisse der beteiligten Stakeholder zugeschnitten sind. Der Ablauf wird in der Abbildung 1 visualisiert:

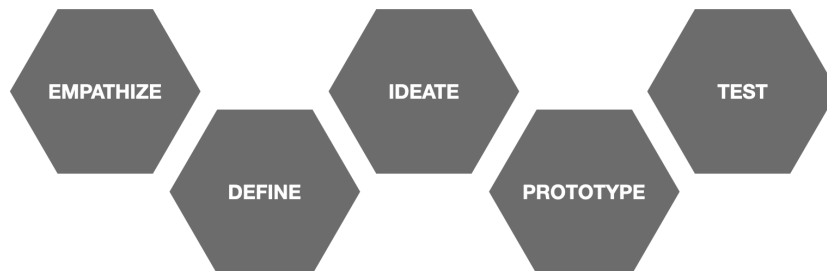


Abbildung 1: DT-Prozessmodell nach HPI of Design at Stanford

Der Design Thinking Prozess fördert eine tiefe Partizipation aller relevanten Stakeholder, was für die Entwicklung sinnvoller und akzeptierter Lösungen im Pflegebereich essentiell ist. Das Hospiz Kieler Förde dient hierbei als zentraler Partner und Teststandort, wodurch realitätsnahe Einblicke in die Bedürfnisse von Pflegepersonal und Pflegebedürftigen möglich sind.

Empathize-Phase: In der ersten Phase des Design Thinking, der Empathize-Phase, wurde durch Literaturrecherche und Interviews ein tiefes Verständnis für die täglichen Herausforderungen und Bedürfnisse aller Stakeholder entwickelt. Besonders wichtig waren hierbei unstrukturierte Interviews mit der Pflegeleitung, Pflegekräften und Pflegebedürftigen des Hospizes, um Einblicke in die individuellen Erfahrungen und Anforderungen zu gewinnen.

Define-Phase: Die gewonnenen Erkenntnisse führten in der Define-Phase zur präzisen Formulierung der Problemstellung aus Sicht der verschiedenen Stakeholder. Die Anwendung von Methoden wie der Entwicklung von Personas und Empathy Maps half dabei, die Probleme und Bedürfnisse klar zu definieren und verständlich zu machen.

Ideate-Phase: In der Ideate-Phase wurde ein kreativer Ideenfindungsprozess initiiert, um eine Vielzahl an Lösungsmöglichkeiten zu generieren. Dies geschah durch Brainstorming und die Verwendung von „How might we...?“-Fragen, die direkt auf die Herausforderungen der zuvor definierten Personas abzielten.

Prototype-Phase: Die vierte Phase, die Prototype-Phase, umfasste die Entwicklung von Prototypen, die auf den Ideen der vorherigen Phase basierten. Verschiedene Prototyping-Methoden, darunter Role-Playing und digitale sowie Papier-Prototypen, wurden genutzt, um die Interaktionen innerhalb des Pflegekontextes simulieren und visualisieren zu können.

Test-Phase: Abschließend wurde in der Test-Phase das direkte Feedback der Nutzer eingeholt, um die Prototypen iterativ zu verbessern und an die realen Bedingungen und Bedürfnisse anzupassen. Die Testphase überlappt sich oft mit der Prototype-Phase, da hierbei kontinuierlich Feedback gesammelt und umgesetzt wird.

2.4 Datenanalyse

Die gesammelten Daten aus Interviews, Beobachtungen und Workshops wurden mittels qualitativer Inhaltsanalyse nach Mayring [MF19] ausgewertet. Diese Methode erlaubt eine systematische, theoriegeleitete Analyse des gesammelten Materials, um Muster und Themen zu identifizieren, die für die Entwicklung des Sprachsteuerungssystems relevant sind. Die qualitative Inhaltsanalyse wurde in mehreren Phasen durchgeführt:

1. **Transkription der Interviews:** Die Aufnahmen aller unstrukturierten mündlichen Interviews mit der Pflegeleitung, Pflegekräften und Pflegebedürftigen des Hospizes Kieler Förde wurden vollständig transkribiert. Dies ermöglichte eine detaillierte Analyse der Gesprächsinhalte und eine systematische Kategorisierung der Aussagen.
2. **Kategorisierung:** Die transkribierten Daten wurden in einem iterativen Prozess kategorisiert. Diese Kategorien umfassten Herausforderungen im Pflegealltag, Erwartungen an digitale Lösungen, bisherige Erfahrungen mit diesen sowie spezifische Anforderungen an ein Smart Home System. Über alle Zielgruppen hinweg wurde zudem eine barrierefreie Gestaltung der Interaktion mit digitalen Systemen thematisiert.
3. **Themenidentifikation und Mustererkennung:** Durch die systematische Analyse der kategorisierten Daten wurden zentrale Themen und Muster identifiziert, die für die Entwicklung des Sprachsteuerungssystems relevant sind. Diese Themen wurden in Zusammenhang mit den Phasen des Design Thinking Prozesses analysiert, um deren Einfluss auf die Lösungsentwicklung zu verstehen.
4. **Integration von Feedback aus Prototyping und Testphasen:** Die Analyse beinhaltete die Integration des Feedbacks, das während der Prototyping- und Testphasen gesammelt wurde. Hierzu wurden sowohl die analogen als auch digitalen Prototypen dem Pflegepersonal vorgestellt und anhand von simulierten Situationen im Pflegealltag ausgewertet. Das hieraus abgeleitete Feedback wurde genutzt, um die Prototypen iterativ zu verbessern und an die realen Bedingungen und Bedürfnisse anzupassen.
5. **Berücksichtigung von Datenschutz und Barrierefreiheit:** Die Analyse umfasste auch die Betrachtung von Datenschutz- und Barrierefreiheitsaspekten im Kontext der bisherigen digitalen Lösungen im Pflegebereich. Diese Themen wurden insbesondere in den Interviews mit den Pflegekräften und Pflegebedürftigen thematisiert, um sicherzustellen, dass das entwickelte Sprachsteuerungssystem diesen Anforderungen gerecht wird.

3 Ergebnisse

3.1 Empathie und Interviewergebnisse

Der Design Thinking-Prozess in dem Projekt zielte darauf ab, die Bedürfnisse der Pflegebedürftigen und des Pflegepersonals tiefgehend zu verstehen. Die Empathiephase brachte wichtige Erkenntnisse, die durch Literaturrecherche, Beobachtungen und insbesondere Experten-Interviews gesammelt wurden:

Leitungsebene:

Die Pflegeleitung betonte die psychische Belastung der Pflegebedürftigen durch den Verlust ihrer Selbstständigkeit bei alltäglichen Aufgaben wie dem An- und Ausschalten des Lichts. Smart Home Systeme könnten diese Autonomie unterstützen und somit das Wohlbefinden und die Lebensqualität der Gäste verbessern. Gleichzeitig wurde die psychologische Hürde angesprochen, dass Gäste das Pflegepersonal nicht bei vermeintlich kleinen Aufgaben stören möchten: „Bei einem vorigen Krankenhausaufenthalt halten sich die Menschen oft mit Hilfsanfragen zurück, um den Personal nicht zur Last zu fallen.“ Die Pflegeleitung erkannte auch den Nutzen einer erweiterten Systemsteuerung, die über Sprachbefehle hinausgeht, um Barrieren für nicht deutschsprachige oder verbal eingeschränkte Nutzer zu minimieren.

Pflegeebene:

Das Pflegepersonal bestätigte die Bedeutung des Autonomieverlustes für die psychische Belastung der Pflegebedürftigen. Es wurde eine „Selbstzurückhaltung“ festgestellt, bei der Patienten, oft aus einer vorangegangenen Krankenhauserfahrung, zögern, das Personal für alltägliche Aufgaben zu rufen. Das System könnte auch hier Abhilfe schaffen, indem es den Pflegebedürftigen mehr Unabhängigkeit gibt und gleichzeitig das Pflegepersonal entlastet, wodurch mehr Zeit für zwischenmenschliche Interaktionen bleibt. Zugleich wurde eine physische Entlastung als weiterer positiver Aspekt hervorgehoben, da „weniger von A nach B gerannt werden müsste“.

3.2 Definition der Kernprobleme (POV) und Lösungsansätze

Basierend auf den Forschungsergebnissen der Empathiephase wurden zwei Hauptpersonas entwickelt:

Die Pflegebedürftigen: Sie leiden unter dem Verlust der Autonomie und fürchten den Verlust zwischenmenschlicher Kontakte durch eine zu starke Abhängigkeit von Technologie.

Das Pflegepersonal: Es sieht das Potenzial technischer Lösungen, ist jedoch durch rechtliche Bedenken und die Möglichkeit der emotionalen Distanzierung durch Technologie besorgt.

Daraus ergeben sich folgende Anforderungen an das Smart Home System:

- Die Systembedienung muss einfach und intuitiv sein, um von allen Nutzern leicht erlernt zu werden.
- Das System sollte in mehreren Sprachen verfügbar sein und die Fähigkeit haben, undeutliche Sprachbefehle zu erkennen.
- Die Pflegebedürftigen sollten durch das Personal in der Nutzung des Systems geschult werden, um sowohl die technologische als auch die emotionale Akzeptanz zu fördern.
- Die Schnittstelle muss klar kommunizieren, dass das System eine Ergänzung, aber kein Ersatz für menschliche Pflege ist.

Die Ergebnisse unterstreichen die Notwendigkeit, Technologie als Hilfsmittel zu betrachten, das die Autonomie der Pflegebedürftigen fördert und gleichzeitig das Pflegepersonal in seinen täglichen Aufgaben unterstützt. Die Integration von Smart Home Systemen sollte sorgfältig geplant werden, um sicherzustellen, dass sie die zwischenmenschlichen Beziehungen nicht ersetzen, sondern ergänzen. Durch kontinuierliche Schulungen und eine klare Kommunikation kann das Vertrauen in die Technologie gestärkt und deren Akzeptanz verbessert werden. Die Weiterentwicklung des Systems muss sich an den realen Bedürfnissen und Bedenken aller Beteiligten orientieren, um eine erfolgreiche und nachhaltige Implementierung in Pflegeeinrichtungen zu gewährleisten.

3.3 Software-Modelle und Hardware-Prototyp

In der Prototype-Phase wurden in mehreren Iterationen sowohl verschiedene Anwendungen der Künstlichen Intelligenz zur Sprachsynthese, Intent-Detection konfiguriert, trainiert.

Eine umfassende Analyse der verfügbaren KI-Modelle sowie der Hard- und Softwarekomponenten wurde durchgeführt, wobei spezifische Hardware für die Tests angeschafft wurde. Verschiedene Methoden wurden in der Software evaluiert, und umfangreiche Tests zur Offline-Sprachsynthese (Ton zu Text) wurden durchgeführt. Diese Tests umfassten auch Experimente zur Intent-Erkennung, um die Intention hinter den gesprochenen Befehlen präzise zu identifizieren. Die gewonnenen Testdaten und Trainingsmodelle zur Speech-to-Text und Intent-Verarbeitung wurden auf verschiedenen KI-Tools trainiert, das Ergebnis der funktionsfähigen Architektur ist in Abbildung 2 zu sehen.

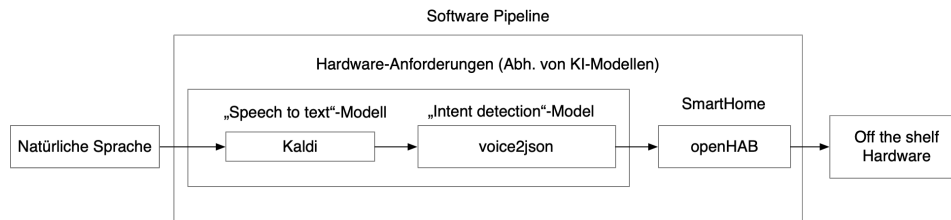


Abbildung 2: Iterationen der Software-Modelle und Hardware

Das Training und die Evaluation von Systemen war entscheidend, um eine maximale Effizienz bei minimaler Rechenleistung und geringer Falsch-Positiv-Rate der Intent-Erkennung zu erreichen. Dies war besonders wichtig, da im Pflegekontext, wie zum Beispiel in Krankenzimmern, die Toleranz und vor allem die Anlernzeit der Nutzenden sehr gering sind. Gleichzeitig wird der Ansatz verfolgt, dass die Nachrüstung der Räume ausschließlich durch ergänzende Hardware erfolgen soll, um die Kosten und die Komplexität einer nachträglichen Installation zu reduzieren, da bauliche Maßnahmen einer breiten Umsetzung im Wege stehen.

Schließlich wurde eine Hausautomatisierungs-Software implementiert, und die Steuerung verschiedener Sensoren und Aktoren wurde getestet. Verschiedene Anbieter und Sensoren/Aktoren wurden evaluiert, um diejenigen zu identifizieren, die eine zuverlässige Offline-Funktionalität ohne Cloud-Anbindung bieten und sich problemlos integrieren lassen.

Das Projekt wurde in drei Hauptbereiche unterteilt, in denen entscheidende Erkenntnisse für den Einsatz der Steuerung in der Praxis gewonnen wurden. Diese Bereiche sind in Abbildung 3 detailliert dargestellt und bieten einen umfassenden Überblick über die verschiedenen Aspekte der Systementwicklung und -integration.

Im Bereich User-Interaktion wurde sich damit befasst, die Anwendungsfelder an Hand der Bedürfnisse zu erheben. Weiterhin fällt in den Bereich die Evaluation der Quality of Voice also sowohl der Sensitivität als auch Spezifität des Systems, so dass das Benutzererlebnis gut ist. Im Gebiet Interaktionsdesign wurde darauf geachtet, wie und in welchem Kontext die Menschen mit dem System interagieren. Hierzu gehören auch die Mitarbeiterinnen und Mitarbeiter, die durch das System eine Veränderung ihrer Arbeitsweise erfahren.

Im Bereich KI-Steuerung sind die verschiedenen Abstraktionslevel der Komponenten des Systems zu sehen. Die Pipeline von Sprache zu einer Aktion, z.B. einer angeschalteten Lampe erfolgt in den vier Schritten.

Um aus der Nutzung zu lernen, werden im dritten Bereich lokal Daten über die Interaktionen gesammelt. Hierbei wurde darauf geachtet, dass diese nicht personenbezogen sind und keine Sprache gespeichert wird. So werden lediglich die erkannten Wortkombinationen des Befehlswortschatzes erfasst.

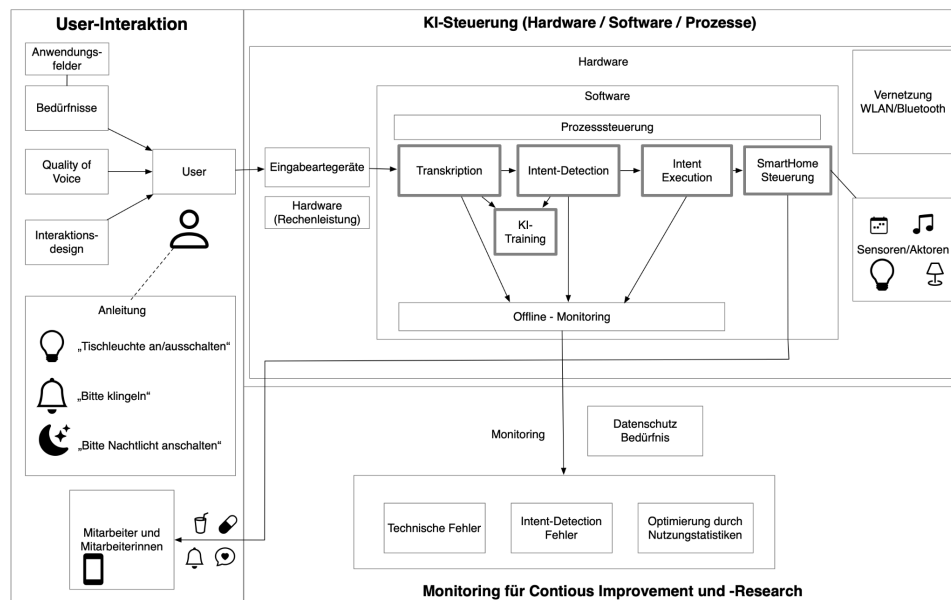


Abbildung 3: Forschungsfelder

4 Diskussion und Ausblick

Die in dieser Arbeit verwendeten partizipativen Methoden ermöglichten es, detaillierte funktionale und nicht-funktionale Anforderungen für ein Smart Home System im Pflegebereich zu identifizieren, das bewusst auf eine Cloud-Anbindung verzichtet. Die tiefgehenden Einblicke, die durch die Kombination aus Literaturrecherche und qualitativen Interviews gewonnen wurden, zeigten klare Wertvorstellungen und Bedürfnisse der Stakeholder, insbesondere der Pflegebedürftigen und des Pflegepersonals.

Die Ergebnisse legen nahe, dass das Smart Home System verschiedene Eingabemöglichkeiten unterstützen sollte, um allen Nutzern gerecht zu werden, was die Bedürfnisse nach einfacher Bedienbarkeit und Barrierefreiheit hervorhebt. Besonders wichtig ist dabei das Design, das Stigmatisierung vermeiden und die individuelle Anpassbarkeit fördern sollte, um nicht nur technische, sondern auch soziale und psychologische Barrieren zu minimieren.

Eine zentrale Erkenntnis war die Bedeutung der freiwilligen Nutzung. Es muss sichergestellt werden, dass Pflegebedürftige die Technologie als unterstützend und nicht als aufdringlich oder ersetzend empfinden. Das System sollte so gestaltet sein, dass die Nutzer selbst entscheiden können, wann und wie sie es verwenden möchten, was die Akzeptanz und den Komfort erheblich verbessert.

Basierend auf den gesammelten Erkenntnissen wurden mehrere Empfehlungen für die Integration eines Smart Home Systems entwickelt. Diese Empfehlungen adressieren verschiedene Aspekte der Systemnutzung:

- Verschiedene Interaktionsmöglichkeiten: Das System sollte über Touch, Sprache und möglicherweise auch über Gesten bedienbar sein, um Nutzern mit unterschiedlichen Fähigkeiten gerecht zu werden.
- Simple verbale Befehle und grafische Interaktion: Die Benutzeroberfläche sollte einfach und intuitiv sein, um Nutzern mit unterschiedlichem technischem Verständnis den Zugang zu erleichtern.
- Akustisches und visuelles Feedback: Klare Rückmeldungen nach jeder Aktion helfen den Nutzern zu verstehen, ob ihre Eingaben erfolgreich waren.
- Individuelle Anpassbarkeit/Flexibilität: Einstellungen sollten personalisierbar sein, um individuellen Bedürfnissen und Fähigkeiten gerecht zu werden.
- Keine Stigmatisierung: Das Design sollte so neutral wie möglich sein, um keine negativen Konnotationen oder Gefühle zu erwecken.
- Unterschiedliche Lernmethoden: Schulungen und intuitive Lernhilfen sollten verfügbar sein, um allen Nutzern den Umgang mit dem System zu erleichtern.

Der Ansatz zur nachhaltigen Entwicklung des Systems umfasst ökonomische, ökologische und soziale Überlegungen. Ökonomisch kann das System durch Effizienzsteigerungen und die Entlastung des Personals zur Kosteneinsparung beitragen. Ökologisch sollte das System auf minimale Umweltbelastung ausgelegt sein, indem es energieeffizient arbeitet und langlebige Komponenten verwendet. Sozial muss das System faire und inklusive Zugangsmöglichkeiten bieten und darf die zwischenmenschliche Kommunikation nicht ersetzen, sondern sollte diese fördern.

Obwohl die Studie umfassende Einblicke bietet, ist sie durch die begrenzte Anzahl von Interviews und die spezifische Einrichtung, in der sie durchgeführt wurde, limitiert. Zukünftige Forschungen sollten eine breitere Datenbasis nutzen und das System in verschiedenen Pflegesettings evaluieren. Langzeitstudien könnten besonders aufschlussreich sein, um die langfristigen Auswirkungen und die praktische Nutzbarkeit des Systems zu untersuchen.

Insgesamt bieten die Ergebnisse dieser Arbeit eine Basis für die Weiterentwicklung von Smart Home Systemen im Pflegebereich, die sowohl den technischen als auch den menschlichen Aspekten gerecht werden. Die empfohlenen Gestaltungsprinzipien sollten weiterhin verfeinert und an die sich ändernden Bedürfnisse der Nutzer angepasst werden, um eine optimale Integration und Nutzung zu gewährleisten.

Literaturverzeichnis

- [A01] Act, Patriot: Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism (USA PATRIOT ACT) Act of 2001. Public Law 107 (2001): 56, 2001.
- [BA23] Statistik der Bundesagentur für Arbeit, Berichte: Blickpunkt Arbeitsmarkt – Arbeitsmarktsituation im Pflegebereich, Nürnberg, Mai 2023.
- [BD22] Brönneke, J.B.; Debatin, J.F.: Digitalisierung im Gesundheitswesen und ihre Effekte auf die Qualität der Gesundheitsversorgung. Bundesgesundheitsbl 65, 342–347, 2022.
- [D19] Destatis: Bevölkerung im Erwerbsalter sinkt bis 2035 voraussichtlich um 4 bis 6 Millionen. Pressemitteilung 242/19, Juni 2019.
- [EK22] Vgl. Eggert, M.; Kreuzer, V.: Sprachsteuerung im Gesundheitswesen – Anforderungen und Auswahl geeigneter Anbieter, HMD Praxis der Wirtschaftsinformatik 59, Nr. 6, S. 1641–42, 2022.
- [ES20] Eggert, M.; Stanke, M.: Adoption of Integrated Voice Assistants in Health Care – Requirements and Design Guidelines, in WI2020 Zentrale Tracks, GITO Verlag, S. 1156–71, 2020.
- [KAG19] Krämer, N.; Artelt, A.; Geminn, C. et al.: KI-basierte Sprachassistenten im Alltag: Forschungsbedarf aus informatischer, psychologischer, ethischer und rechtlicher Sicht. Universität Duisburg-Essen, Universitätsbibliothek, 2019.
- [KCJ20] Kim, M.; Cho, M.; Jun, H.: Developing Design Solutions for Smart Homes Through User-Centered Scenarios, Frontiers in Psychology, Artikel 335, S. 4, 2020.
- [MF19] Mayring, P.; Fenzl, T.: Qualitative Inhaltsanalyse. In: Baur, N., Blasius, J. (eds) Handbuch Methoden der empirischen Sozialforschung. Springer VS, Wiesbaden, 2019.
- [MJK22] Muckenhuber, J.; Janschitz, G.; Klebel, T.: Pflege 2.0? Ausgestaltung und Auswirkungen der Digitalisierung auf die Arbeitsbedingungen und die Art der Tätigkeiten im Bereich der Pflege. In: Rußmann, U.; Aubke, F.; Ortiz, D.; Pezenka, I.; Schulz, A.; Schweiger, C. (Hrsg.): Zukunft verantwortungsvoll gestalten. Forschung und Praxis an der FH Wien der WKW. Wiesbaden, Springer Gabler, S. 185–197, 2022.
- [PMW09] Plattner, H.; Meinel, C.; Weinberg, U.: Design thinking. Landsberg am Lech: Mi-Fachverlag, 2009.
- [QSA15] Queirós, A.; Silva, A.; Alvarelhão, J. et al.: Usability, accessibility and ambient-assisted living: a systematic literature review. Univ Access Inf Soc 14, 57–66, 2015.
- [REB21] Reichmann, H.; Elson, M.; Borgert, N. et al.: Erfahrbarer Datenschutz und IT-Sicherheit in Smart Home-Anwendungen. Datenschutz Datensich 45, 259–264, 2021.
- [Sc16] Schelisch, L.: Technisch unterstütztes Wohnen im Stadtquartier. Potentiale, Akzeptanz und Nutzung eines Assistenzsystems für ältere Menschen. Springer VS (Wiesbaden) 2016.

Open Set Gestenerkennung zur Interaktion zwischen Passanten und Fahrzeugen mit Automatisierungsfunktionen

Vollständiger Beitrag

Christian Reichhardt¹, Manuel Martin ² und Patrick Baier¹


Abstract: Die Zunahme an intelligenten Steuerungssystemen in Fahrzeugen wirft zunehmend die Frage auf, inwiefern ein Fahrzeug mit seiner Umgebung automatisch interagieren kann. Eine wichtige Aufgabe besteht dabei in der Kommunikation und Interaktion mit Passanten am Fahrbahnrand, welche eine anwendungsspezifische Gestenerkennung von Personen voraussetzt. Für die Erkennung solcher Gesten existieren mittlerweile Ansätze auf Basis maschineller Bilderkennungsverfahren, welche einfache Gesten erkennen können. Zwei wichtige Fragestellungen sind dabei, welche Aktionsklassen durch ein Fahrzeug zuverlässig erkennbar sind und wie mit nicht explizit trainierten Aktionsgesten umgegangen werden kann. Unter diesen Aspekten untersucht die vorliegende Arbeit die Fähigkeiten und Grenzen der Gestenerkennung im Kontext relevanter Aktionen von Passanten im Straßenverkehr. Hierzu wurde als Teil einer Verarbeitungskette das häufig verwendete tiefe neuronale Netz *ST-GCN++* betrachtet und mit Gelenkpunkten von Personen in unterschiedlichen Detailgraden trainiert. Um mit nicht explizit trainierten Gesten umgehen zu können, wurde eine Restklasse eingeführt, welche auf im Training ungesehenen Gesten hohe Genauigkeiten erzielte.

Keywords: Gestenerkennung, Passanten, Fahrzeug, Open Set Recognition, ST-GCN++

1 Einleitung

Nonverbale Kommunikation gehört zum Alltag im Straßenverkehr und ist maßgebend für den reibungslosen Verkehrsablauf auf deutschen und internationalen Straßen. Dies wird insbesondere in Situationen klar, in denen Fahrzeuge und Passanten aufeinandertreffen. So zeigte Šucha [Šu14] in einer Studie, dass Passanten mit Autofahrern sowohl über Augenkontakt, Handgesten als auch über Kopfbewegungen interagieren. Dass bestimmte Gesten einen Einfluss auf das Fahrverhalten haben können, wurde bereits mehrfach durch Studien [CVL11; ZW14] belegt, die den Einfluss von Handgesten auf die Rate anhaltender Fahrzeuge untersuchten. Obwohl in der Industrie bereits mehrere Ansätze [MB24; NM15; SA24] existieren, Intentionen automatisierter Fahrzeuge Passanten bekannt zu geben, ist es bislang nicht üblich, dass Gesten von Passanten erkannt, interpretiert und darauf automatisiert reagiert werden kann. Dies könnte potentiell zu Missverständnissen mit schwerwiegenden Folgen führen.

¹ Hochschule Karlsruhe - Technik und Wirtschaft, Fakultät für Informatik und Wirtschaftsinformatik, Moltkestr. 30, 76133 Karlsruhe, Deutschland, rech1028@h-ka.de; patrick.baier@h-ka.de

² Fraunhofer-Institut für Optronik, Systemtechnik und Bildauswertung IOSB, Human-AI Interaction, Fraunhoferstraße 1, 76131 Karlsruhe, Deutschland, manuel.martin@iosb.fraunhofer.de,  <https://orcid.org/0000-0001-9473-8804>

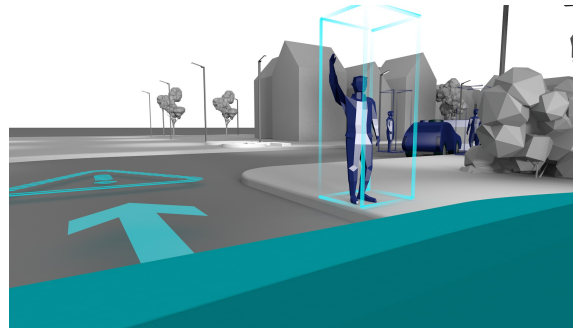


Abb. 1: Szenario des Forschungsprojekts *INITIATIVE* [PI21]. Interaktion zwischen Fahrzeug und Passant, wobei der Passant das Fahrzeug durch eine Geste vor einer Gefahrenstelle warnen möchte. [PIS21]

Aus diesem Grund liegt das Ziel aktueller Forschungen [PIG21] darin, einen durch maschinelles Lernen gestützten Interaktionsprozess zwischen Fahrzeugen mit Automatisierungsfunktionen und anderen Verkehrsteilnehmern zu etablieren. Stanciu et al. [St18] führen in ihrer Arbeit aus, dass solche Fahrzeuge zukünftig in der Lage sein müssen, sowohl Strategien aller Verkehrsteilnehmer zu erkennen, als auch zu interpretieren. Um eine Interpretation der Intentionen von Passanten zu ermöglichen, ist jedoch eine Gestenerkennung Voraussetzung. Abbildung 1 zeigt ein solches Szenario, in dem ein Passant das entgegenkommende Fahrzeug durch Winken auf eine Gefahrenstelle aufmerksam machen möchte. Eine besondere Herausforderung stellt dabei die Diversität genutzter Gesten dar. Wie von Najamuddin [Na19] gezeigt, bestehen außerdem große Unterschiede was die Nutzung von Gesten in verschiedenen kulturellen Umfeldern angeht. Zur Lösung von Problemen dieser Art, schlagen Stanciu et al. [St18] in ihrer Arbeit eine Standardisierung von Kommunikationsstrategien im Straßenverkehr vor.

Um einen Überblick zu bieten, welche Aktionsgesten hierfür verwendbar sind, wurden in der vorliegenden Arbeit für diesen Anwendungsfall spezifische Aktionsklassen gewählt und auf deren Erkennbarkeit geprüft. Hierzu wurde die von Duan et al. [Du22] beschriebene Architektur *Spatial-Temporal Graph Convolutional Network ST-GCN++* zur Gestenerkennung in unterschiedlichen Konfigurationen trainiert und ausgewertet. Diese erhält zweidimensionale Gelenkpunkte von Personen über eine vordefinierte Zeit als Eingabe, um Aktionsgesten zu klassifizieren. Intern werden diese Gelenkpunkte als mehrdimensionale Graphen in Form kinematischer Skelette über einen vordefinierten Zeitraum dargestellt. Um für den Anwendungsfall einen sinnvollen Graphen zu finden, wurden zwei Graphtopologien untersucht und in jeweils zwei Konfigurationen trainiert. Diese Topologien unterscheiden sich jeweils in der Anzahl betrachteter Knotenpunkte. Zu diesem Zweck wurden aus einem öffentlichen Datensatz Gelenkpunkte von Personen in unterschiedlichen Detailgraden für das Training generiert. Üblicherweise wurde die Gestenerkennung im Straßenverkehr bisher unter einer geschlossenen Anzahl von Aktionsklassen betrachtet, wie beispielsweise bei der

Erkennung von Verkehrszeichen autorisierter Verkehrslotsen [Mi21]. Diese Annahme ist für die Interaktion mit Passanten aufgrund der Vielzahl möglicher Gesten nicht praktikabel. Aus diesem Grund wurden Gesten im Kontext unbekannter Bewegungen erfasst, was auch als *Open Set Recognition Problem* [Sc12] bezeichnet wird. Da sich dieser Ansatz grundlegend von üblichen Trainings- und Evaluationsmethoden von Gestenerkennungs-Algorithmen unterscheidet, wurde ein öffentlicher Datensatz umfassend an diese Anforderungen angepasst. Zur Auswertung wurden zwei unterschiedliche Ansätze zur Manipulation der Trainingsdaten verfolgt und die Ergebnisse eingeordnet. Dabei konnten hohe Genauigkeiten bezüglich im Training ungesehener Aktionsgesten der hierfür eingeführten Restklasse beobachtet werden. Explizit trainierte Aktionsklassen mit einem hohen Bewegungsanteil lieferten ebenfalls sehr hohe Genauigkeiten, während Handgesten vergleichsweise geringere Genauigkeiten erzielten.

2 Stand der Forschung

Im Bereich der Gestenerkennung gehören tiefe neuronale Netze aktuell zum Standard und erzielen in unterschiedlichen Ausprägungen auch bei einer höheren Anzahl an Aktionsgesten gute Resultate. Dies zeigen unter anderem Ergebnisse von Duan et al. [Du22]. Aufgrund der dynamischen Natur vieler Gesten ist es bei der Erkennung menschlicher Aktivitäten oft unerlässlich, neben den räumlichen Informationen zudem zeitliche Informationen zu berücksichtigen. Gegenüber videobasierten Methoden [St20; Su17; Wa16; Zh21] existieren hierfür skelettbasierte Ansätze, wobei Gelenkpunkte als Knoten von Graphen genutzt werden. Für deren Klassifikation eignen sich neben Faltungsnetzen wie *PoseConv3D* [Du22] auch Graphfaltungsnetze [KW16] besonders gut. Diese ermöglichen durch die Repräsentation eines Graphen als Adjazenzmatrix Verbreitung und Faltungen von Graphen im Netz. Architekturen wie das von Duan et al. [Du22] vorgeschlagene *ST-GCN++* versprechen dabei echtzeitfähige Laufzeit bei einer hohen Genauigkeit. Da als Eingang solcher Netze meist eine Posenschätzung benötigt wird, wird in der Praxis häufig eine *Top-Down* Verarbeitungskette aus Objekterkennung, Posenschätzung und anschließender Gestenerkennung genutzt. Dabei bieten sich neuronale Netze, beispielsweise basierend auf den Architekturen *R-CNN* [Gi14] und *HRNet* [Su19] an.

Bei bisherigen Untersuchungen zur Gestenerkennung im Straßenverkehr [He20; MAA16; Mi21; Po19] wird a priori angenommen, dass alle zur Laufzeit auftretenden Gesten bekannt sind. Die Vielzahl an möglichen Aktionen im Straßenverkehr stellt jedoch die Anforderung an eine *Open Set Recognition (OSR)* [Sc12], welche darin besteht, zur Inferenzzeit im Training ungesehene Aktionsklassen von gesehenen unterscheiden zu können. Um zu erkennen, ob ein Datenelement einer trainierten Aktion entspricht, existieren in der Literatur verschiedene Ansätze [BIG18; BYK21; Ro20; Sh18]. Einen einfachen und effizienten Ansatz bietet die von Zhang und LeCun [ZL17] beschriebene *Dustbin*-Methode. Dabei wird eine zusätzliche Restklasse mit Samples einer Menge nicht beschrifteter, d.h. nicht explizit zu erkennender, Aktionen trainiert. Sollte eine Klasse zur Inferenzzeit mit höchster Wahrscheinlichkeit der

Restklasse zugeordnet werden, so wird diese als unbekannte Geste eingestuft. Nach Zhang; LeCun [ZL17] kann sich dieses Vorgehen durch einen Regularisierungseffekt außerdem positiv auf das Training auswirken. Im Gegensatz zu bisherigen Untersuchungen wird in dieser Arbeit daher die Gestenerkennung im Straßenverkehr unter dem Gesichtspunkt der *OSR* mithilfe der *Dustbin*-Methode betrachtet und ausgewertet.

3 Modell

In dieser Arbeit werden Modelle des tiefen neuronalen Netzes *ST-GCN++* auf unterschiedlichen Daten trainiert, wobei die Architektur, die Vorverarbeitung sowie die Netzkonfiguration der Implementierung von Duan et al. [Du22] folgt. Die Modelle unterscheiden sich dabei lediglich in den Graphtopologien, beschrieben in Abschnitt 4.1, was zu einer Änderung der Adjazenzmatrix führt. Die Gelenkpunkte des Datensatzes, welche durch zweidimensionale Koordinaten gegeben sind, wurden zunächst relativ zur Bildmitte auf den Bereich $[-1, 1]$ normiert und anschließend mit deren Erkennungswahrscheinlichkeiten konkateniert. Fünfdimensionale Tensoren der Form $N \times M \times T \times V \times C$ wurden als Eingabe verwendet. Diese bestehen aus N Samples pro Batch mit T Frames pro Sample. Dabei existieren pro Sample M detektierte Personen aus jeweils V Knotenpunkten mit C Dimensionen. In diesem Fall ist $C = 3$, wobei die dritte Dimension der Knotenpunkte die Erkennungswahrscheinlichkeit der Posenschätzung ist. Zu beachten ist, dass jedes geladene Video mit $T = 100$ gesamplet wird. Zu diesem Zweck wurden Videos in $n = 100$ Segmente der selben Länge geteilt und pro Segment zufällig ein Frame gewählt. Für den Fall, dass der Eingang kleiner als T ist, startet die Indizierung an einem zufälligen Frame des Videos und dieses wird solange wiederholt, bis die festgelegte Anzahl der Frames erreicht wurde. Die Zahl erkannter Personen wird insgesamt auf $M = 2$ beschränkt. Die zeitliche Faltung des trainierten Modells *ST-GCN++* wird mit einer Kernelgröße $\Gamma \times 1$ und $\Gamma = 3$ für jeden der vier *Convolutional Streams* durchgeführt. Dies gilt ebenfalls für die Kernelgröße des *MaxPooling Streams*. Als Partitionierungsstrategie für *ST-GCN++* wurde die von Duan et al. [Du22] beschriebene *Spatial Configuration* gewählt, mit Knotenpunkt null als Körperschwerpunkt für beide untersuchten Graphtopologien.

4 Experiment

Die beschriebene Architektur wurde in mehreren Konfigurationen trainiert und evaluiert, wobei zwei Graphtopologien für die Abbildung von Gesten miteinander verglichen wurden. Hierfür wurde ein öffentlich verfügbarer Datensatz umfassend entsprechend den Anforderungen einer *OSR* modifiziert. Um Modelle ebenfalls auf im Training ungesesehenen Gesten evaluieren zu können, wurden zwei verschiedene Versionen eines Trainingsdatensatzes erstellt und jeweils unterschiedliche Modelle trainiert. Als Basis für die Implementierung diente der Programmcode der Veröffentlichung von Duan et al. [Du22].

Klassendefinitionen					
ID	Klasse	NTU RGB+D	ID	Klasse	NTU RGB+D
0	Restklasse	-	13	Fallen	A43
1	Brille abnehmen	A19	14	Auf Person zeigen	A54
2	Winken mit einer Hand	A23	15	Ball prellen	A64
3	Hüpfen auf einem Bein	A26	16	„Leise“-Zeichen	A67
4	Aufspringen	A27	17	Daumen hoch	A69
5	Telefonieren	A28	18	Daumen runter	A70
6	Handy/ Tablet bedienen	A29	19	„OK“-Zeichen	A71
7	Mit Finger zeigen	A31	20	„Victory“-Zeichen	A72
8	Kopf nicken	A35	21	Faust schütteln	A93
9	Kopf schütteln	A36	22	Beide Hände oben	A95
10	Salut	A38	23	Armkreisen	A97
11	Arme kreuzen („Stopp“)	A40	24	Armschwingen	A98
12	Taumeln	A42			

Tab. 1: Trainierte Klassen mit deren neuen Labels (ID) sowie deren Bezeichnung im *NTU RGB+D 120* Datensatz.

4.1 Datensatz

Aufgrund des Mangels an frei verfügbaren Videodatensätzen im Straßenverkehr wurde als Basis für den finalen Datensatz der Trainings- und Testdatensatz des *Cross-Subject* Evaluationsprotokolls von *NTU RGB+D 120* [Li19] verwendet. Dieser wurde an die Anforderung einer *OSR* sowie für die Erkennung spezifischer Gesten angepasst.

Der finale Datensatz muss zunächst für den Anwendungsfall relevante Gesten annotieren, die das Modell explizit erkennen soll. Aus diesem Grund dienten als Grundlage für die Auswahl betrachteter Aktionsgesten Untersuchungen zur Nutzung von Gesten im Straßenverkehr von Zhuang; Wu [ZW14] und Brand; Schmitz [BM23]. Die dort untersuchten Gesten wurden mit den im Ausgangsdatsatz verfügbaren Klassen abgeglichen. Vergleichbare Gesten, sowie einige weitere Klassen des Datensatzes, wurden nach Ermessen der Autoren aus dem *NTU RGB+D 120* Datensatz ausgewählt und mit neuen IDs versehen. Tab. 1 liefert eine Übersicht über alle Klassen des finalen Datensatzes sowie deren ursprüngliche Bezeichnung im Ausgangsdatsatz. Während sich davon Manche primär über die Position der Gelenkpunkte der Arme definieren (Klassen ID 1, 2, 5, 6, 7, 10, 11, 14, 15, 16, 21, 22, 23 und 24), beziehen sich Andere auf Bewegungen des gesamten Körpers (3, 4, 12, 13). Weitere Gesten zeichnen sich hauptsächlich durch Kopfbewegungen (8, 9) oder Fingerstellungen (17, 18, 19, 20) aus. Auch der Grad der Bewegung in den Datenproben variiert teilweise stark zwischen einzelnen Klassen. Besonders hohe Bewegungsgrade sind bei den Gesten *Aufspringen*, *Fallen*, *Armkreisen* sowie *Armschwingen* zu beobachten.

Zur Umsetzung der *OSR*, in Anlehnung an die Methodik der *Dustbin*-Methode von Zhang; LeCun [ZL17], wurde eine zusätzliche Restklasse mit ID null erstellt. Dieser

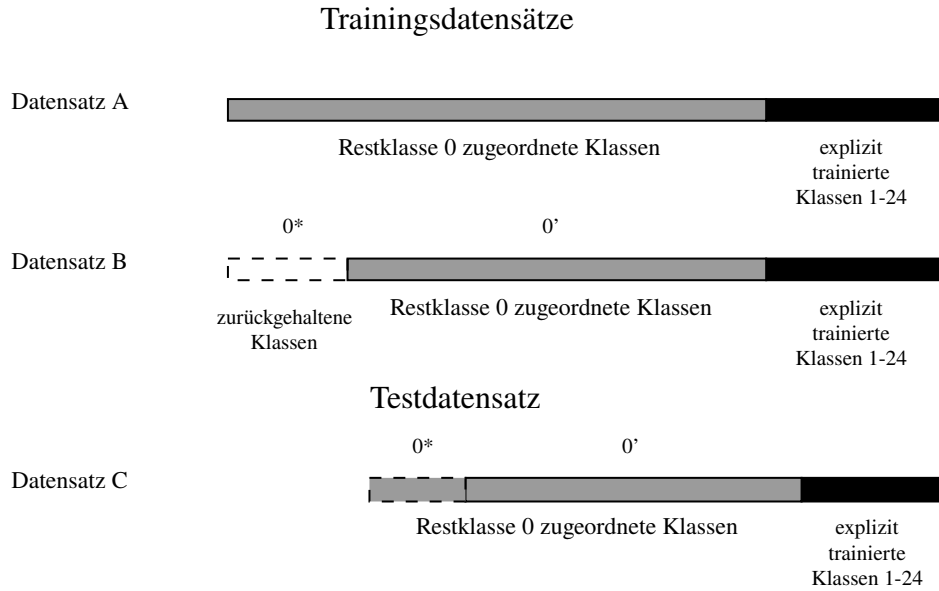


Abb. 2: Modifikationen des *NTU RGB+D 120* Datensatzes. Als Ausgangspunkt dienten die Trainings- und Testdatensätze des *Cross-Subject* Evaluationsprotokolls.

Klasse werden alle Datenproben des *NTU RGB+D 120* Datensatzes zugeordnet, deren Erkennung nicht den anderen Klassen in Tab. 1 entsprechen. Insgesamt wurden durch Modifikationen des *Cross-Subject* Evaluationsprotokolls zwei Trainingsdatensätze und ein Testdatensatz generiert, deren Aufbau schematisch in Abb. 2 gezeigt wird. Datensatz A weist allen Samples im Trainingsdatensatz, welche nicht den Klassen 1–24 zugeordnet wurden, der Restklasse zu. Das Training mit Datensatz A wird im Folgenden als Training *ohne Rückhaltung* bezeichnet. Diese Methodik ermöglicht jedoch keine Evaluation bezüglich im Training ungesehener Gesten auf dem Testdatensatz. Aus diesem Grund wurde eine weitere Modifikation der Trainingsdaten vorgenommen. In Datensatz B wurde zusätzlich eine Untermenge 0^* der Aktionsgesten der Restklasse für das Training entfernt. 0^* beinhaltet die Videosamples der Klassen *A1–A18* sowie *A20* im *NTU RGB+D 120* Datensatz. Das Training mit Datensatz B wird weiterhin als Training *mit Rückhaltung* bezeichnet. Die Modifikationen des Testdatensatzes C entsprechen den Modifikationen von Datensatz A und somit beinhaltet dieser Datenproben einer Menge von Klassen 0^* , die in Datensatz B nicht enthalten sind.

Für die Datensätze A bis C wurden jeweils zweimal Merkmale generiert, einmal 17 und einmal 65 zweidimensionale Koordinaten von Gelenkpunkten inklusive deren Erkennungswahrscheinlichkeit pro Person und Frame. Die Merkmale beschreiben dabei die Knotenpunkte der Graphen, die das Modell zur räumlich-zeitlichen Faltung verwendet. Insgesamt wurden zwei Graphtopologien in separaten Modellen erstellt, die sich in der

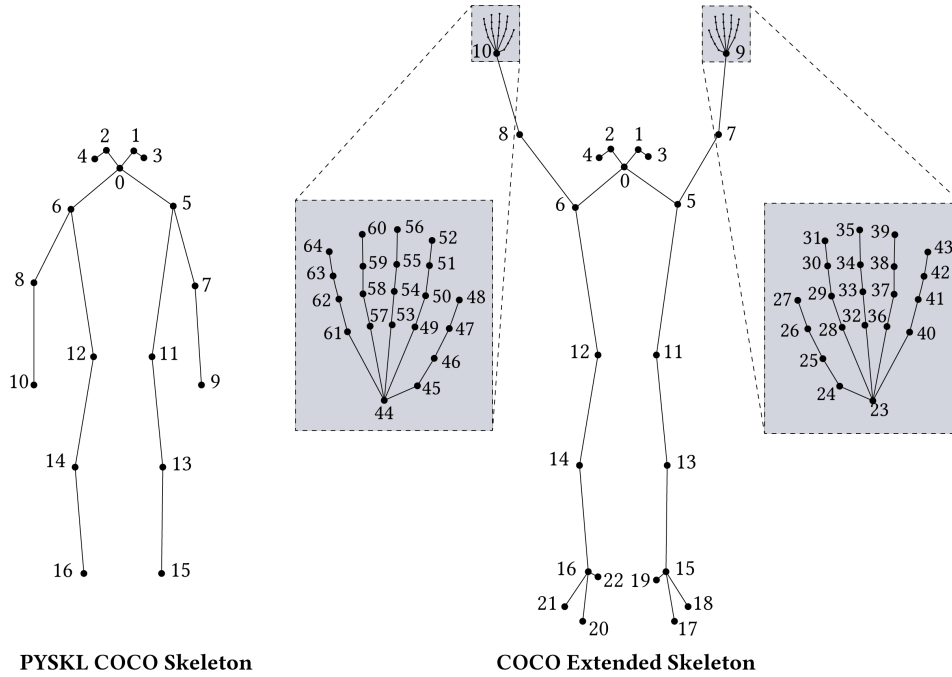


Abb. 3: Räumliche Komponenten der in dieser Arbeit untersuchten Graphtopologien sowie Indizierung der Knotenpunkte.

räumlichen Ebene durch die Anzahl an Knotenpunkten unterscheiden und in Abb. 3 schematisch dargestellt sind. Die erste Topologie entspricht derjenigen von Duan et al. [Du22], welche auf Gelenkpunkten des *COCO*-Datensatzes [Li14] basiert. Hierfür wurden Merkmale genutzt, die in der Implementierung von Duan et al. [Du22] zur Verfügung gestellt sind. Die zweite Topologie verwendet eine Teilmenge derjenigen Gelenkpunkte, die in der *COCO-Wholebody* Annotation [Ji20] definiert sind. Um diese Gelenkpunkte zu erhalten, wurden für den gesamten Datensatz zweidimensionale *WholeBody* Gelenkpunkte mithilfe der vortrainierten *MMPose* [MM20] Modelle *Faster R-CNN R50* [Re15] zur Objekterkennung sowie *HRNet w48* [Su19] zur Posenschätzung aus den RGB-Videodaten extrahiert. Anschließend wurden Samples mit mangelhaften Posenschätzungen automatisiert aus dem Trainingssatz entfernt. Da die Entfernung der Probanden zu den Kameras im Datensatz zwischen 2 und 4.5 Metern beträgt [Li19], gewährleistet dies eine ausreichende Auflösung, um Gelenkpunkte der Extremitäten zu generieren. Die resultierenden Gelenkpunkte wurden auf räumlicher Ebene ähnlich des *PYSKL COCO Skeleton* miteinander verbunden, jedoch inklusive Daten für Hände und Füße. Der finale Graph wird im Folgenden als *PYSKL Extended Skeleton* bezeichnet und ist auf räumlicher Ebene in Abbildung Abb. 3 rechts dargestellt.

4.2 Training

Um die Funktionalität der *Dustbin*-Methode gegenüber im Training ungesehenen Klassen quantitativ evaluieren zu können, wurde zunächst ein *ST-GCN++* Modell in *PYSKL COCO Skeleton* sowie in *COCO Extended Skeleton* Topologie auf Datensatz B 100 Epochen trainiert³. Zusätzlich dazu wurden in beiden Topologien *ST-GCN++* Modelle auf Datensatz A trainiert⁴, mit dem Ziel einer höheren Generalisierung. Diese Modelle sollen später für Feldversuche genutzt werden. Deren Training bestand insgesamt aus 500 Epochen. Alle Modelle wurden “from scratch“ trainiert. In Anlehnung an die Implementierung von Duan et al. [Du22] diente als Grundlage der Validierung aller Trainingsdurchläufe der verwendete Testdatensatz C.

Im Training wurde *Stochastic Gradient Descent* mit *Nesterov Momentum* zur Optimierung genutzt sowie die Lernrate während des Trainings durch *Cosine Annealing* über alle Epochen angepasst, mit einer minimalen Lernrate von null. Als Verlustfunktion diente *Cross-Entropy Loss*. Da beide Varianten der Trainingsdatensätze ohne weitere Maßnahmen im Training zu einer Überrepräsentation der Samples der Restklasse gegenüber Samples der anderen trainierten Klassen führen würden, wurde im Dateneinzug eine automatisierte Balancierung implementiert. Diese berechnet mithilfe der normierten Häufigkeit der Datenproben im Datensatz eine Auswahlwahrscheinlichkeit pro Klasse. Anhand dieser Wahrscheinlichkeit werden pro Epoche Samples aus den jeweiligen Klassen gezogen. Da die Gesten der Restklasse als eine gemeinsame Klasse betrachtet werden, wird in diesem Fall zufällig aus Klasse null anhand der Auswahlwahrscheinlichkeit der Restklasse gezogen. Dabei wird nicht zwischen Gesten innerhalb der Restklasse unterschieden.

5 Numerische Resultate

Weiterhin werden die Ergebnisse des Experiments besprochen. Für die Evaluation aller Modelle wurden jeweils die Gewichte der Epoche mit der höchsten Top-1 Accuracy auf den Validierungsdaten genutzt. Für die Testdaten wird kein Sampling im Dateneinzug anhand der Klassenhäufigkeit verwendet.

5.1 Evaluation der mit Rückhaltung trainierten Modelle

Zunächst werden zur Bewertung der *Dustbin*-Methode die mit Rückhaltung trainierten Modelle untersucht. Die Ergebnisse sind in Tab. 2 gezeigt. Eine Top-1 Accuracy von 1.0 auf im Training ungesehenen Klassen, hier gekennzeichnet durch 0*, würde bedeuten, dass alle Samples dieser Menge korrekt der Restklasse 0 zugeordnet wurden.

³ Siehe Abschnitt 5.1 für die Evaluation.

⁴ Siehe Abschnitt 5.2 für die Evaluation.

Klassen IDs	PYSKL COCO Skeleton			COCO Extended Skeleton		
	0'	1-24	0*	0'	1-24	0*
Top-1 Accuracy	0.9252	0.8203	0.8263	0.9580	0.8635	0.8745

Tab. 2: *Cross-Subject* Evaluation der mit Rückhaltung trainierten Modelle auf dem Testdatensatz C. Gezeigt wird für beide Graphtopologien jeweils die gemittelte Top-1 Accuracy der explizit trainierten Klassen eins bis vierundzwanzig (1–24), im Training der Restklasse zugeordneten Klassen (0') sowie der im Training ungesehenen Klassen (gekennzeichnet durch 0*).

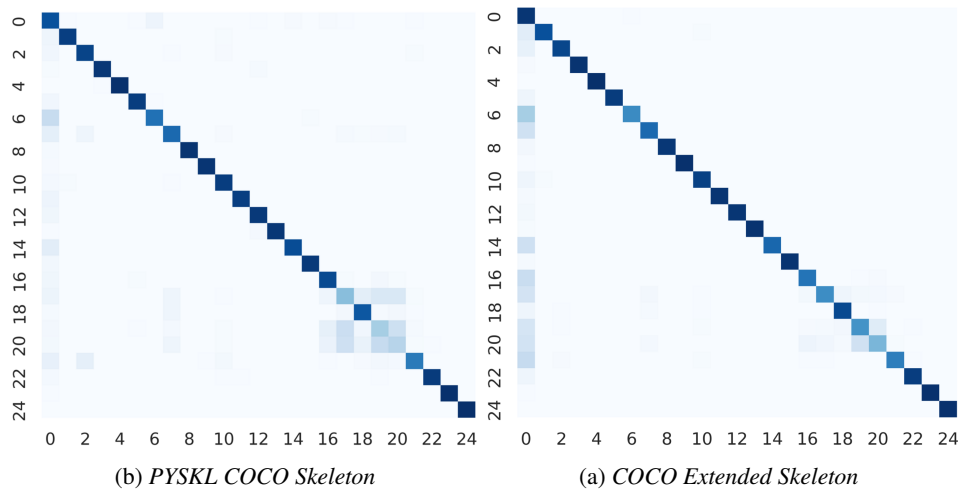


Abb. 4: Zeilenweise normalisierte Konfusionsmatrizen der mit Datensatz A trainierten *ST-GCN++* Modelle in beiden Topologien. Dunkelblau entspricht einem Wert von 1.0. Wahre Label IDs sind auf der y-Achse abgebildet.

Insgesamt ergaben die gemittelten Top-1 Accuracies über allen explizit trainierten Klassen; Samples der Menge 0* sowie den im Training zurückgehaltenen Klassen in der *COCO Extended Skeleton* Topologie höhere Werte als in der *PYSKL COCO Skeleton* Topologie. Ohne einen spezifischen Schwellwert der Erkennungswahrscheinlichkeit für die Restklasse gegenüber explizit trainierter Klassen erreichte die Restklasse in beiden Topologien eine höhere Top-1 Accuracy. In über 82 % der Fällen für die *PYSKL COCO Skeleton* Topologie und in über 87 % für die *COCO Extended Topologie* wurden im Training ungesehene Klassen korrekt der Restklasse zugeordnet.

5.2 Evaluation der ohne Rückhaltung trainierten Modelle

Die Ergebnisse der *Cross-Subject* Evaluierung der mit Datensatz B trainierten Modelle werden in Tab. 3 gezeigt. Die zugehörigen Konfusionsmatrizen sind in Abb. 4 dargestellt.

Klasse	Evaluationsergebnisse	
	ST-GCN++	ST-GCN++
	PYSKL COCO Skeleton	COCO Extended Skeleton
0 Restklasse	0.8698	0.9740
1 Brille abnehmen	0.9416	0.8768
2 Winken mit einer Hand	0.9234	0.9167
3 Hüpfen auf einem Bein	0.9673	0.9819
4 Aufspringen	0.9928	1.0000
5 Telefonieren	0.9455	0.9529
6 Handy/ Tablet bedienen	0.7455	0.6486
7 Mit Finger zeigen	0.7790	0.7790
8 Kopf nicken	0.9819	0.9709
9 Kopf schütteln	0.9782	0.9928
10 Salut	0.9457	0.9348
11 Arme kreuzen („Stopp“)	0.9420	0.9783
12 Taumeln	0.9601	0.9746
13 Fallen	0.9709	0.9855
14 Auf Person zeigen	0.8913	0.7862
15 Ball prellen	0.9632	0.9826
16 „Leise“-Zeichen	0.8970	0.7361
17 Daumen hoch	0.4278	0.6441
18 Daumen runter	0.8487	0.9045
19 „OK“-Zeichen	0.3496	0.6181
20 „Victory“-Zeichen	0.2974	0.4618
21 Faust schütteln	0.7153	0.6997
22 Beide Hände oben	0.9547	0.9461
23 Armkreisen	0.9913	0.9948
24 Armschwingen	0.9983	0.9948
Top-1 Accuracy	0.8598	0.9496
Top-5 Accuracy	0.9931	0.9946
Balanced Accuracy	0.8511	0.8694

Tab. 3: Ergebnisse der untersuchten Konfigurationen von *ST-GCN++* in *Cross-Subject* Evaluation auf Datensatz C. Die Modelle wurden auf Datensatz A trainiert. Die Accuracies einzelner Klassen werden für die jeweilige Konfiguration sowie die Top-1 Accuracy, Top-5 Accuracy und Balanced Accuracy über allen Klassen gezeigt. Der größte Wert pro Zeile ist jeweils hervorgehoben.

Insgesamt erreichten beide Modelle auf den Testdaten hohe Accuracies, die *COCO Extended Skeleton* Topologie sogar über allen Klassen eine durchschnittliche Top-1 Accuracy von über 94 %. Die Accuracies der Restklassen liegen in beiden Modellen über der gemittelten Top-1 Accuracy. Größere Konfusionen mit der Restklasse wurden beispielsweise in Klasse sechs beobachtet. Schwierigkeiten bei der Erkennung bereiteten insbesondere die Klassen 16 bis 20, deren Aktionen sich weitestgehend durch Hand- bzw. Fingerstellungen definieren. Wie Abb. 4 zeigt, entstanden in diesem Bereich besonders in der *PYSKL COCO Skeleton*

Topologie, welche keine Gelenkinformationen der Hände beinhaltet, einige Verwechslungen. Die Accuracies dieser Klassen fallen in *COCO Extended Skeleton* Topologie jedoch ebenfalls geringer aus als die gemittelte Top-1 Accuracy. Die besten Ergebnisse erzielten in beiden Modellen die Klassen vier, 22 sowie 23.

6 Diskussion

Insgesamt lieferte *ST-GCN++* als Modell zur Gestenerkennung vielversprechende Accuracies auf den Testdaten in allen untersuchten Konfigurationen. Wie aufgrund zusätzlicher Gelenkinformationen erwartet, konnte im *COCO Extended Skeleton* Layout gegenüber *PYSKL COCO Skeleton* die Erkennung von Handgesten verbessert werden. Hierzu gehörten die Aktionen *Daumen hoch*, *Daumen runter*, „OK“-Zeichen und „Victory“-Zeichen. Auch bezüglich der Genauigkeit der Restklasse übertraf *COCO Extended Skeleton* die Topologie *PYSKL COCO Skeleton* mit weniger Keypoints und lieferte damit insgesamt eine höhere Top-1 Accuracy. Mit einer Genauigkeit von über 99% bei beiden Topologien sind besonders die Gesten *Aufspringen*, *Armkreisen* und *Armschwingen* positiv hervorzuheben. Diese Aktionen zeichnen sich verglichen mit anderen Gesten durch einen hohen Bewegungsgrad aus, wobei sich die Bewegung zudem auf beide Arme erstreckt. Betrachtet man andere Klassen mit einem hohen Bewegungsgrad wie *Hüpfen auf einem Bein* oder *Ball prellen*, so erreichten diese ebenfalls vergleichsweise hohe Genauigkeiten. Demgegenüber erzielten Gesten geringerer Bewegungen wie *Handy/Tablet bedienen* oder *Mit Finger zeigen* geringere Werte. Die Gesten *Kopf nicken* und *Kopf schütteln* konnten in beiden Topologien mit mehr als 97% Accuracy größtenteils korrekt klassifiziert werden.

Bezüglich der *Dustbin*-Methode wurden über allen Konfigurationen im Test hohe Accuracies für die Restklasse in *Cross-Subject* Evaluation erreicht. Trotz einer vergleichsweise geringen Anzahl Trainingsepochen erzielten die Modelle mit Rückhaltung auf den im Training vorenthaltenen Aktionen über 82% Accuracy, in *COCO Extended Skeleton* Topologie sogar über 87%. Die Ergebnisse der Evaluation sprechen demnach für die hier untersuchte *Dustbin*-Methode zum Umgang mit ungesesehenen Aktionen. Ein weiteres Argument für diese Methodik ist deren einfache Implementierung, verglichen mit auf Wahlen basierenden Frameworks [Ro20]. Es besteht zudem die Möglichkeit, dass sich die Erkennung durch eine größere Auswahl an Klassen der Menge 0^* noch verbessern ließe. Hierzu könnten ergänzende Datensätze genutzt werden, um eine höhere Variation im Training der Restklasse zu erreichen. Wie Abb. 4 zeigt, bestehen bei einigen Klassen geringe Konfusionen gegenüber der Restklasse, wobei die Anpassung eines Schwellenwertes für die Erkennungswahrscheinlichkeit der Restklasse diese Unterscheidung möglicherweise noch verbessern könnte.

Sowohl der Einfluss des Abstandes von Personen zur Kamera als auch die Bewegung der Kamera wurden in dieser Arbeit nicht quantitativ evaluiert. Zudem ist zu beachten, dass sich die verwendeten Videosamples von den Lichtverhältnissen und dem Versuchsaufbau stark von Videodaten im Straßenverkehr unterscheiden, wo zudem kontinuierliche Videodaten



Abb. 5: Typische Probleme der Posenschätzungen bei qualitativen Versuchen im Straßenverkehr bei 3 Meter (a) sowie 5 Meter (b) Entfernung zur Kamera. Im Fall (a) konnte die Geste „Victory“-Zeichen im *COCO Extended Skeleton* nicht mit ausreichender Wahrscheinlichkeit zugeordnet werden. Trotz Fehler in der Posenschätzung wurde im Fall (b) die Geste *Arme kreuzen* („Stopp“) korrekt erkannt.

vorliegen. Der Mangel an öffentlich verfügbaren Datensätzen erschwert jedoch eine quantitative Evaluation. Aus diesem Grund wurden mit den mit Datensatz A trainierten Modellen qualitative Versuche auf kontinuierlichen Videodaten stationärer Kameras im Straßenverkehr durchgeführt. Probanden wurden dabei in unterschiedlichen Entfernungen gefilmt. Zu jedem Frame t wurden die Ergebnisse der Posenschätzung in einer Warteschlange der Größe⁵ $T = 59$ gespeichert und für jedes t die letzten T Frames für die Gestenerkennung von *ST-GCN++* verwendet. Die *Dustbin*-Methode konnte auch hier nicht explizit trainierte Gesten zuverlässig zuordnen. In den Ergebnissen wurde außerdem eine Abhängigkeit zwischen der Erkennbarkeit von Gesten auf größere Entfernungen und dem Grad der Bewegung des gesamten Körpers beobachtet. Besonders gut erkennbare Klassen auf eine Distanz bis zu 20 Metern waren unter anderem die Aktionen *Aufspringen* und *Hüpfen auf einem Bein*. Handgesten konnten in beiden Topologien selbst auf kurze Distanzen unter herrschenden Lichtverhältnissen nicht korrekt vorhergesagt werden, wobei bereits in der Gestenerkennung größere Fehler auftraten. Beispiele hierfür werden in Abb. 5 gezeigt. Dennoch schien der erweiterte Graph in der qualitativen Analyse Vorteile für die Gestenerkennung von Passanten zu bringen, insbesondere bei fehlenden Gelenkpunkten im nahen Sichtbereich. Eine Kombination beider Graphtopologien in verschiedenen Entfernungen wäre demnach zukünftig möglich.

7 Zusammenfassung und Ausblick

Unter dem Aspekt der Interaktion zwischen Passanten und automatisierten Fahrzeugen wurden in dieser Veröffentlichung gut erkennbare Aktionsklassen herausgestellt sowie Aussagen über die Vor- und Nachteile unterschiedlich detaillierter Graphtopologien aufgezeigt. Dabei wurden sowohl im Training als auch in der Evaluation Gegebenheiten einer *OSR*

⁵ Dies entspricht dem Median der explizit trainierten Klassen der Trainingsdatensätze.

berücksichtigt und ein verfügbarer Datensatz entsprechend dieser Anforderungen angepasst. Hierfür wurden zwei Methodiken vorgestellt und verfolgt.

Das Modell *ST-GCN++* lieferte sowohl auf den Testdaten als auch auf kontinuierlichen Videodaten bei Ganzkörperposen vielversprechende Ergebnisse. Die Erkennung von Handgesten konnte auf den Testdaten durch die Nutzung einer erweiterten Graphtopologie mit 65 Gelenkpunkten zwar verbessert werden, Ergebnisse auf Testdaten sowie qualitative Tests der untersuchten Modelle zeigen jedoch Probleme bei der Nutzung von Handgesten für den Anwendungsfall auf. Besonders gute Ergebnisse wurden bei Gesten erreicht, welche einen hohen Bewegungsgrad sowie eine Bewegung mehrerer Körperpartien beinhalteten. Hierzu gehörten die Gesten *Aufspringen*, *Armkreisen* und *Armschwingen*. Zur Lösung des *OSR* Problems wurde die *Dustbin*-Methode untersucht, welche auf im Training ungesehenen Aktionsklassen eine Accuracy von über 82% und im erweiterten Layout über 87% erreichte. In einem Training ohne Rückhaltung ungesehener Aktionsklassen der Restklasse erreichte die erweiterte Topologie auf trainierten Aktionsgesten eine Top-1 Accuracy von fast 95% sowie über 97% bezüglich der Restklasse.

Die Ergebnisse dieser Arbeit lassen auch auf zukünftige Herausforderungen schließen. Wie Untersuchungen von Brand; Schmitz [BM23] zeigen, variieren im Straßenverkehr genutzte Gesten bislang stark. Die vorliegende Arbeit belegt, dass auch die Erkennbarkeit einzelner Aktionen stark divergieren kann. Daher wird die Vermutung gestützt, dass Passanten zukünftig ein Aktionsalphabet zur Interaktion mit Verkehrsteilnehmern mit Automatisierungsfunktionen erlernen müssen. Um eine bestmögliche Selektion von Aktionsklassen zu gewährleisten, können zukünftige Forschungen an die hier präsentierten Ergebnisse anknüpfen.

Literatur

- [BIG18] Busto, P. P.; Iqbal, A.; Gall, J.: Open set domain adaptation for image and action recognition. *IEEE transactions on pattern analysis and machine intelligence* 42/2, S. 413–429, 2018.
- [BM23] Brand, T.; M., S.: Analysis of Pedestrian Gestures in a Virtual Traffic Environment, Paper presented at the 1st Würtual Reality XR Meeting, Würzburg, 11.-13.04.2023. 2023.
- [BYK21] Bao, W.; Yu, Q.; Kong, Y.: Evidential deep learning for open set action recognition. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. S. 13349–13358, 2021.
- [CVL11] Crowley-Koch, B. J.; Van Houten, R.; Lim, E.: Effects of pedestrian prompts on motorist yielding at crosswalks. *Journal of applied behavior analysis* 44/1, 2011.

- [Du22] Duan, H.; Wang, J.; Chen, K.; Lin, D.: PYSKL: Towards Good Practices for Skeleton Action Recognition. In: Proceedings of the 30th ACM International Conference on Multimedia. 2022.
- [Gi14] Girshick, R.; Donahue, J.; Darrell, T.; Malik, J.: Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2014.
- [He20] He, J.; Zhang, C.; He, X.; Dong, R.: Visual Recognition of traffic police gestures with convolutional pose machine and handcrafted features. *Neurocomputing* 390/, 2020.
- [Ji20] Jin, S.; Xu, L.; Xu, J.; Wang, C.; Liu, W.; Qian, C.; Ouyang, W.; Luo, P.: Whole-Body Human Pose Estimation in the Wild. In: Proceedings of the European Conference on Computer Vision (ECCV). 2020.
- [KW16] Kipf, T. N.; Welling, M.: Semi-Supervised Classification with Graph Convolutional Networks. *arXiv preprint arXiv:1609.02907/*, 2016.
- [Li14] Lin, T.-Y.; Maire, M.; Belongie, S.; Hays, J.; Perona, P.; Ramanan, D.; Dollár, P.; Zitnick, C. L.: Microsoft COCO: Common Objects in Context. In: *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6–12, 2014, Proceedings, Part V* 13. Springer, 2014.
- [Li19] Liu, J.; Shahroudy, A.; Perez, M.; Wang, G.; Duan, L.-Y.; Kot, A. C.: NTU RGB+D 120: A Large-Scale Benchmark for 3D Human Activity Understanding. *IEEE transactions on pattern analysis and machine intelligence* 42/10, 2019.
- [MAA16] Mueid, R. M.; Ahmed, C.; Ahad, M. A. R.: Pedestrian activity classification using patterns of motion and histogram of oriented gradient. *Journal on Multimodal User Interfaces* 10/, 2016.
- [MB24] Mercedes-Benz Australia/Pacific Pty Ltd: The Mercedes-Benz F 015 Luxury in Motion, URL: <https://www.mercedes-benz.com.au/passengercars/campaigns/mercedes-benz-f-015.html>, Stand: 12.03.2024.
- [Mi21] Mishra, A.; Kim, J.; Cha, J.; Kim, D.; Kim, S.: Authorized Traffic Controller Hand Gesture Recognition for Situation-Aware Autonomous Driving. *Sensors* 21/23, 2021.
- [MM20] MMPose Contributors, 2020, URL: <https://github.com/open-mmlab/mmpose>, Stand: 25.04.2023.
- [Na19] Najamuddin, A.: The Meaning of Gesture in Social Cultural Context. *El-Tsaqafah: Jurnal Jurusan PBA* 18/1, 2019.
- [NM15] Nissan Motor Co., Ltd.: Nissan IDS Concept: Nissan’s vision for the future of EVs and autonomous driving, 2015, URL: <https://usa.nissannews.com/en-US/releases/nissan-ids-concept-nissan-s-vision-for-the-future-of-evs-and-autonomous-driving?selectedTabId=releases>, Stand: 12.03.2024.

- [PI21] Projekt INITIATIVE: Initiative Projekt – Intelligente Mensch-Technik-Kommunikation im gemischten Verkehr, 2021, URL: <https://www.initiative-projekt.de>, Stand: 21.03.2024.
- [PIG21] Projekt INITIATIVE: Gesamtziel des Vorhabens, 2021, URL: <https://www.initiative-projekt.de/gesamtziel-des-vorhabens/>, Stand: 21.03.2024.
- [PIS21] Projekt INITIATIVE: Szenarien Katalog, 2021, URL: <https://www.initiative-projekt.de/szenarien-katalog/>, Stand: 21.03.2024.
- [Po19] Pop, D. O.; Rogozan, A.; Chatelain, C.; Nashashibi, F.; Bensrhair, A.: Multi-Task Deep Learning for Pedestrian Detection, Action Recognition and Time to Cross Prediction. *IEEE Access* 7/, 2019.
- [Re15] Ren, S.; He, K.; Girshick, R.; Sun, J.: Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *Advances in neural information processing systems* 28/, 2015.
- [Ro20] Roitberg, A.; Ma, C.; Haurilet, M.; Stiefelhagen, R.: Open Set Driver Activity Recognition. In: 2020 IEEE Intelligent Vehicles Symposium (IV). IEEE, 2020.
- [SA24] Semcon AB: The Smiling Car, URL: <https://semcon.com/smilingcar/>, Stand: 12.03.2024.
- [Sc12] Scheirer, W. J.; de Rezende Rocha, A.; Sapkota, A.; Boult, T. E.: Toward Open Set Recognition. In: Bd. 35. 7, IEEE, 2012.
- [Sh18] Shu, Y.; Shi, Y.; Wang, Y.; Zou, Y.; Yuan, Q.; Tian, Y.: Odn: Opening the deep network for open-set action recognition. In: 2018 IEEE international conference on multimedia and expo (ICME). IEEE, S. 1–6, 2018.
- [St18] Stanciu, S. C.; Eby, D. W.; Molnar, L. J.; St. Louis, R. M.; Zanier, N.; Kostyniuk, L. P.: Pedestrians/Bicyclists and Autonomous Vehicles: How Will They Communicate? *Transportation research record* 2672/22, 2018.
- [St20] Stroud, J.; Ross, D.; Sun, C.; Deng, J.; Sukthankar, R.: D3D: Distilled 3D Networks for Video Action Recognition. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 2020.
- [Šu14] Šucha, M.: Road users' strategies and communication: driver-pedestrian interaction. *Transport Research Arena (TRA)* 1/, 2014.
- [Su17] Sun, L.; Jia, K.; Chen, K.; Yeung, D.-Y.; Shi, B. E.; Savarese, S.: Lattice Long Short-Term Memory for Human Action Recognition. In: *Proceedings of the IEEE international conference on computer vision*. 2017.
- [Su19] Sun, K.; Xiao, B.; Liu, D.; Wang, J.: Deep High-Resolution Representation Learning for Human Pose Estimation. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2019.
- [Wa16] Wang, L.; Xiong, Y.; Wang, Z.; Qiao, Y.; Lin, D.; Tang, X.; Van Gool, L.: Temporal Segment Networks: Towards Good Practices for Deep Action Recognition. In: *European conference on computer vision*. Springer, 2016.

- [Zh21] Zha, X.; Zhu, W.; Xun, L.; Yang, S.; Liu, J.: Shifted Chunk Transformer for Spatio-Temporal Representational Learning. *Advances in Neural Information Processing Systems* 34/, 2021.
- [ZL17] Zhang, X.; LeCun, Y.: Universum Prescription: Regularization using Unlabeled Data. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Bd. 31. 1, 2017.
- [ZW14] Zhuang, X.; Wu, C.: Pedestrian gestures increase driver yielding at uncontrolled mid-block road crossings. *Accident Analysis & Prevention* 70/, 2014.

Auswirkungen von Mergers und Acquisitions auf kollaborative Cloud-Services

Vollständiger Beitrag

Lukas Koch¹, Can Adam Albayrak²  und Dirk Schreiber³

Abstract: Merger-und-Acquisitions-Prozesse gelten als ein wesentliches Thema der strategischen Unternehmensausrichtung und haben erhebliche Auswirkungen auf die betroffenen Unternehmen. Aufgrund der gestiegenen Relevanz von kollaborativen Cloud-Services in den letzten Jahren untersucht dieser Beitrag anhand von Experteninterviews die Auswirkungen von Merger-und-Acquisition-Prozesse auf die kollaborativen Cloud-Services. Im Rahmen dieses Beitrags wird eine Übersicht über die Hintergründe zu Mergers und Acquisitions, kollaborativen Cloud-Services sowie der Herausforderungen bei der Integration gegeben. Die Ergebnisse zeigen praxisnah, welche Auswirkungen ein Merger-und-Acquisition-Prozess auf kollaborative Cloud-Services hat. Dabei werden insbesondere das Vorgehen, die Migration sowie Benutzer und Stakeholder betrachtet.

Keywords: Mergers und Acquisitions, kollaborative Cloud-Services, Cloud-Services, Kollaboration

1 Einleitung und Motivation

Im Zuge der strategischen Unternehmensausrichtung spielt das Wachstum des Unternehmens eine wesentliche Rolle. Einer Hays-Studie zufolge gaben 29 % von 987 befragten betrieblichen Entscheidern als strategisches Ziel ‚Marktanteile gewinnen‘ an, weitere 44 % der Befragten nannten Umsatzsteigerung. [Ha22] Beide Ziele lassen sich u. a. durch Unternehmenstransaktionen wie Zukäufe oder Fusionen ermöglichen, die auch als Mergers & Acquisitions bezeichnet werden. In den letzten Jahren gab es immer wieder neue Rekorde hinsichtlich der Anzahl und des Volumens von Mergers & Acquisitions. Im Jahr 2021 wurde nach dem Global M&A Report von Bain & Company, einer weltweit tätigen Unternehmensberatung, erneut ein Rekordhoch für das Volumen der Mergers & Acquisitions weltweit vermeldet. [Ba22] Unabhängig davon steigt auch der Einsatz von Kollaborationssoftware in Unternehmen. So hat sich in den letzten sechs Jahren der

¹ Lukas Koch, M.Sc., TH Köln, Institut für Distance Learning & Further Education, Steinmüllerallee 1, 51643 Gummersbach, koch92@outlook.com

² Prof. Dr. Can Adam Albayrak, Institut für Distance Learning & Further Education, Steinmüllerallee 1, 51643 Gummersbach, Can.Adam.Albayrak@th-koeln.de sowie Hochschule Harz, Friedrichstr. 57-59, 38855 Wernigerode, <https://orcid.org/0000-0002-7698-9797>

³ Prof. Dr. Dirk Schreiber, Institut für Distance Learning & Further Education, Steinmüllerallee 1, 51643 Gummersbach sowie Hochschule Bonn-Rhein-Sieg, Grantham-Allee 20, 53757 Sankt Augustin, dirk.schreiber@h-brs.de

globale Umsatz von Kollaborationssoftware mehr als verdoppelt, [PMT22] zeitgleich stieg die weltweite Nutzung von Kollaborationstools durch Arbeitnehmer zwischen 2019 und 2021 von 55 % auf 79 %. [Ga21] Diese zwei Statistiken zeigen das enorme Wachstum des Marktes in den letzten Jahren. Im gleichen Zeitraum gab es ebenso eine Zunahme der Nutzung von Kollaborationssoftware innerhalb von Unternehmen. Bei allen Unternehmenstransaktionen handelt es sich um komplexe Prozesse, die zwangsläufig Auswirkungen auf die IT der Unternehmen haben. [BES09] Die Integration eines Unternehmens bedeutet auch immer eine Integration der IT. Durch die gestiegene Verbreitung kollaborativer Tools kann dies jetzt und zukünftig ein wesentlicher Bestandteil innerhalb des Mergers-&-Acquisitions-Prozesses werden, der Stand 2024 ggf. nicht ausreichend berücksichtigt wird.

2 Hintergrund

2.1 Mergers & Acquisitions

Für den Begriff *Mergers & Acquisitions*, der aus dem angelsächsischen Sprachraum stammt, hat sich in der Literatur noch keine allgemeingültige Definition durchgesetzt. Auch in den USA, wo der Begriff seit etwa dem Ende des 19. Jahrhunderts Anwendung findet, gibt es keine konsistente Verwendung des Begriffes. Im weiteren Sinne beschreibt das Begriffspaar alle mit dem Erwerb oder der Veräußerung von Unternehmen bzw. Unternehmensanteilen zusammenhängende Vorgänge. Die Änderung von Eigentumsverhältnissen am Eigenkapital vom Unternehmen bzw. von Unternehmensanteilen ist dabei ein maßgebliches Merkmal von Mergers & Acquisitions. [Wi17] Die zwei Begriffe *Mergers* und *Acquisitions* werden teilweise als Synonyme betrachtet, oder ihnen werden unterschiedliche Bedeutungen zugeschrieben. Wenn die Begriffe nicht als Synonyme betrachtet werden, findet die Unterscheidung bei Mergers und Acquisitions anhand der rechtlichen Selbstständigkeit der Beteiligten statt. [Wi17]

Der Ablauf eines Mergers-&-Acquisitions-Prozesses wird üblicherweise in Phasen eingeteilt. In der Literatur finden sich verschiedenste Ausprägungen und Darstellungen, wie ein Merger-&-Acquisition-Prozess in Phasen eingeteilt werden kann. Eine Einteilung in eine übergeordnete Gliederung mit den drei Phasen Pre-Merger, Transaktion/Durchführung und Post-Merger-Phase ist allen Darstellungen und Ausprägungen gemein. [Wi17]

Der in Abb. 1 dargestellte Prozess orientiert sich an Jansen und gliedert den Prozess in die drei Phasen Pre-Merger, Transaktionsphase und Post-Merger-Phase. Zu beachten ist, dass diese Phasen inklusive der zugehörigen Prozesse integrativ und iterativ zu sehen sind und nicht zwangsläufig einer chronologischen Reihenfolge entsprechen.[Ja16]

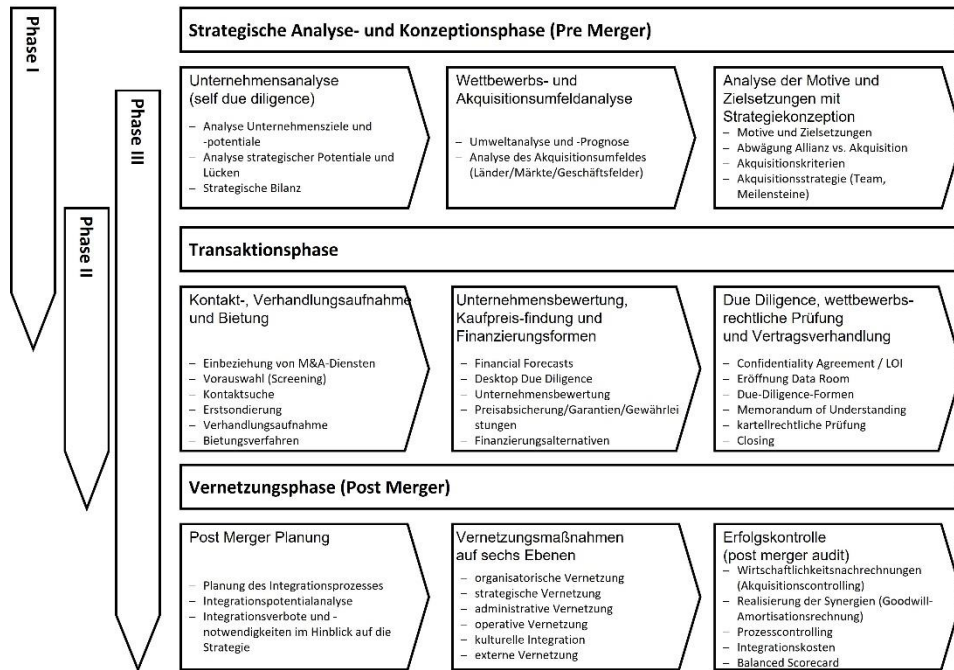


Abb. 1 : Idealtypische Phasen eines Mergers-&-Acquisitions-Prozesses [Ja16]

2.2 Kollaborative Cloud-Services

Der Begriff *Kollaboration*, oder das meistens aus dem Englischen übernommene *collaboration*, ist häufig nicht exakt definiert und wird in der Praxis für alle möglichen Kontexte verwendet. Der lateinische Begriff *laborare*, der in *Kollaboration* enthalten ist, steht für ‚arbeiten‘ oder ‚sich abmühen‘, womit *kollaborieren* bedeutet, dass zwei Objekte miteinander arbeiten, um etwas herzustellen oder etwas zu erreichen. In verschiedenen Definitionen wird auch von einer synchronen Eigenschaft, also dem zeitgleichen Arbeiten, gesprochen. Dillenbourg beschreibt in *Collaborative learning. Cognitive and computational approaches*, dass auch von Kollaboration gesprochen werden kann, wenn die Aufgaben nur mittelbar synchron stattfinden. [Di99] Die elektronische Zusammenarbeit (E-Kollaboration) ist eine Form der Kollaboration. Das gemeinsame Arbeiten an einer Aufgabe unter dem Einsatz elektronischer Technologien wird im weitesten Sinne als E-Kollaboration bezeichnet. Dies umfasst beispielsweise Chat-Tools, E-Konferenztools, E-Mail und kollaborative Schreibwerkzeuge. [Ko01] E-Kollaboration wird durch die Nutzung von Kollaborationssoftware durchgeführt.

Dabei kann Kollaborationssoftware sowohl als sogenannte *On-Premise-Software* als auch als Cloud-Services bereitgestellt werden. Abb. 2 zeigt die Marktanteile der Hersteller von Kollaborationssoftware aus dem Jahr 2021. Die Lösungen der drei größten Hersteller für Kollaborationssoftware, Microsoft, Zoom und Google, sind entweder reine Cloud-Lösungen oder werden als On-Premise-Produkte vermarktet, die jedoch Cloud-Technologien benötigen, zum Beispiel Zoom.

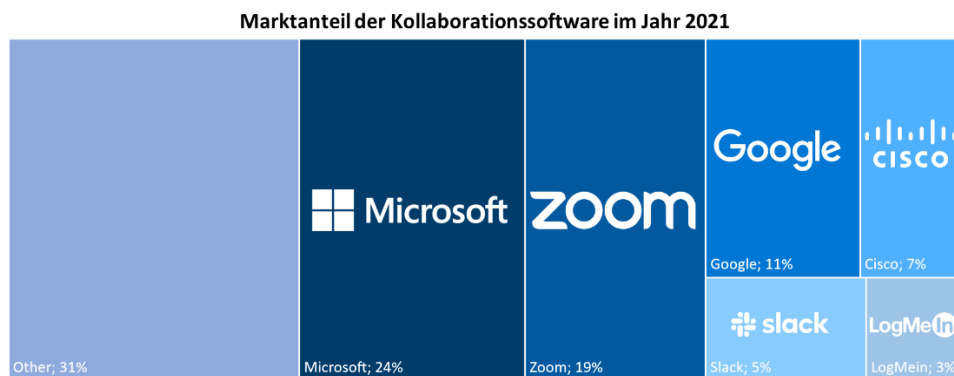


Abb. 2 : Marktanteile der Kollaborationssoftware in 2021 [St22]

Beispielsweise stellt Microsoft mit *Microsoft 365* E-Mail-, Kollaborations- und Office-Anwendungen sowie weitere Serverdienste als Cloud-Service im Rahmen eines Software-as-a-Service bereit. [MM21] Google als Marktführer für Internetsuchmaschinen bietet innerhalb der Google Cloud das Software-as-a-Service-Produkt *Google Workspace* an, das vormals als *G Suite* bekannt war.

2.3 Herausforderungen bei der Integration

Bei einer Mergers-&Acquisitions-Transaktion kann es von Bedeutung sein, Daten aus zwei getrennten Kollaborationsumgebungen in eine zusammenzuführen. Die Migration der Daten in eine Zielumgebung steht dabei immer in Abhängigkeit zu den eingesetzten Kollaborationsservices. So besteht die Möglichkeit, die Daten innerhalb eines kollaborativen Cloud-Services (beispielsweise von *Google Workspace* zu *Google Workspace*) oder zwischen zwei verschiedenen kollaborativen Cloud-Services (von *Google Workspace* zu *Microsoft 365*) zu migrieren. Die Möglichkeiten einer Migration stehen auch in Abhängigkeit zu den vorhandenen kollaborativen Cloud-Services.

Die Migration von beispielsweise Microsoft-365-Services ist aktuell noch begrenzt möglich, und es gibt verschiedene Bedingungen und Einschränkungen. Ebenso kann der Einsatz von Zusatztools für die Migration erforderlich sein. Da sich Microsoft 365 auch dauerhaft weiterentwickelt und neue Services und Funktionen bereitgestellt werden, muss ein entsprechender Migrationspfad zur Durchführung der Migration geprüft werden. Das

Verschieben von E-Mail-Postfächern und OneDrive-Daten, zwischen zwei Microsoft-365-Umgebungen, wurde im November 2022 veröffentlicht und muss über eine Add-On-Lizenz freigeschaltet werden.⁴ Gerade für *Microsoft Teams*, das eines der wesentlichen Tools für die kollaborative Zusammenarbeit ist, gibt es aktuell keine Bordmittel, die bei einer Migration zwischen zwei Umgebungen unterstützen. Da innerhalb der Microsoft-365-Umgebungen nicht nur Dateien liegen, sondern auch viele Berechtigungen, Freigaben und weitere Informationen gespeichert werden, müssen bei der Migration weitere Informationen entsprechend übertragen werden. Sollten gewisse Daten nicht übertragbar sein, so kann es zu Einschränkungen in der Benutzbarkeit und der Akzeptanz durch den Benutzer nach der Migration kommen. Zwischen den Microsoft-365-Services gibt es verschiedene Abhängigkeiten. Daher ist es von großer Bedeutung, erst die Services und Daten zu migrieren, die als Voraussetzung für weitere Services benötigt werden.

Die am Beispiel von Microsoft 365 dargestellten Abhängigkeiten und Auswirkungen, zeigen, dass die Konsolidierung von Kollaborationsplattformen ein komplexer Prozess ist. Daher muss die Integration einer Konsolidierung der Kollaborationsplattformen innerhalb eines Merger-&-Acquisitions-Prozess genau beleuchtet werden. Dieser Beitrag beschäftigt sich daher mit der Forschungsfrage:

Welche Auswirkungen haben Merger-&-Acquisition-Prozesse auf kollaborative Cloud-Services?

In einer Meta-Analyse von 81 empirischen Studien, die zwischen 1991 und 2016 veröffentlicht wurden, betrachten die Autoren Henningsson, Yetton und Wanne (2018) die Literatur zur Integration von Informationssystemen (IS) bei Mergers & Acquisitions. Die Autoren argumentieren, dass die Integration von Informationssystemen ein entscheidender Erfolgsfaktor für M&A-Transaktionen ist, da sie nicht nur betriebliche Effizienzen und Synergien ermöglicht, sondern auch strategische Vorteile realisiert. Die Autoren identifizieren mehrere Herausforderungen, darunter kulturelle Unterschiede, technologische Heterogenität und die Notwendigkeit der Anpassung an dynamische Geschäftsanforderungen.

Der Artikel stellt verschiedene Integrationsstrategien vor, die von vollständiger Konsolidierung bis hin zu koexistierenden Systemen reichen. Diese Strategien hängen stark vom spezifischen Kontext und den Zielen der M&A-Transaktion ab. Die Autoren betonen die Wichtigkeit eines flexiblen und anpassungsfähigen Ansatzes, um den sich ständig ändernden Anforderungen gerecht zu werden. Sie heben hervor, dass Governance-Strukturen und die Unterstützung durch das Top-Management wesentliche Faktoren für eine erfolgreiche Integration der Informationssysteme sind.

Zusätzlich diskutieren die Autoren methodische Ansätze und Modelle, die zur Planung und Durchführung der IS-Integration verwendet werden können. Sie fordern eine intensivere Forschung in diesem Bereich, um theoretische Modelle weiter zu verfeinern

⁴ Vgl. [Mi23].

und praktische Leitlinien zu entwickeln, die Unternehmen bei zukünftigen M&A-Transaktionen unterstützen können.

Abschließend betonen Henningsson, Yetton und Wanne, dass trotz erheblicher Fortschritte in der Forschung noch viele offene Fragen bestehen. Sie rufen dazu auf, zukünftige Studien durchzuführen, um ein tieferes Verständnis für die komplexen Dynamiken der Integration von Informationssystemen bei M&A zu gewinnen und somit die Erfolgsrate solcher Transaktionen zu erhöhen. Dieser Beitrag greift das Thema für den Bereich der kollaborativen Cloud-Services auf. [HYW18]

3 Forschungsdesign

Im Rahmen dieses Beitrags wurden vier Experteninterviews durchgeführt, um die Forschungsfrage zu beantworten. Dazu wurden ein Interviewleitfaden mit zwölf Fragen erstellt der in die folgenden Abschnitte unterteilt wurde.

- Allgemeine Fragen zu Mergers-&-Acquisitions-Prozessen
- Auswirkungen von Mergers & Acquisitions auf kollaborative Cloud-Services
- Durchführung von Mergers-&-Acquisitions-Prozessen im Kontext kollaborativer Cloud-Services
- Trends und Entwicklungen von Mergers & Acquisitions

Die vier Experten weisen Expertise im Bereich Mergers & Acquisitions auf und haben in verschiedensten Rollen innerhalb von Mergers-&-Acquisitions-Prozessen mitgewirkt. Zwei der Experten waren innerhalb verschiedenster Mergers-&-Acquisitions-Projekte für die Migration kollaborativer Cloud-Services im Microsoft- oder Google-Umfeld verantwortlich. Ein weiterer Experte ist als Geschäftsführer an mehreren Mergers-&-Acquisitions-Prozessen beteiligt gewesen. Der vierte Experte hat verschiedenste Unternehmen, hauptsächlich Start-ups, dabei unterstützt, von anderen Firmen aufgekauft zu werden.

Die Methode der qualitativen Inhaltsanalyse wurde verwendet, um die Daten der Experteninterviews auszuwerten. Dabei wurde sich bei der Durchführung der qualitativen Inhaltsanalyse an Mayring orientiert [Ma16], jedoch wurden auch weitere Autoren wie Kuckartz berücksichtigt. [KR22] Die Darstellung der Ergebnisse orientiert sich dabei an den in der qualitativen Inhaltsanalyse erstellten Hauptkategorien.

4 Ergebnisse

4.1 Vorgehen und Auswirkungen von Mergers-&-Acquisitions-Prozesses

Mergers & Acquisitions hat verschiedenste Auswirkungen auf die Kollaboration im Unternehmen. Zu Beginn eines Mergers-&-Acquisitions-Prozesses, vor der Homogenisierung der Kollaborationsplattform, muss die Kollaboration zwischen den bestehenden Plattformen sichergestellt werden. Durch den Einsatz kollaborativer Cloud-Services ist dies bei gleicher Kollaborationsplattform durch technische Mittel relativ einfach möglich. Die Experten beschrieben, dass die Kollaboration zwischen zwei Unternehmen über technische Lösungen einfach sichergestellt werden kann und dies die Kollaboration entsprechend ermöglicht, ohne eine vollständige technische Integration eines Unternehmens durchgeführt zu haben. Allerdings gibt es entsprechende Voraussetzungen und Limitierungen, beispielsweise kann bei Microsoft 365 der Zugriff auf das Adressbuch des anderen Unternehmens nicht ermöglicht werden.

Die Etablierung der Kollaboration von unterschiedlichen Kollaborationslösungen ist hingegen deutlich komplexer und bedarf beispielsweise der Bereitstellung von Gast-Konten für einen Teil der Anwender, um die Kollaboration mit dem anderen Unternehmen zu ermöglichen.

Die Antworten der befragten Experten zeigen, dass im Rahmen des Mergers-&-Acquisitions-Prozesses zur Integration der IT hauptsächlich die Absorptionsmethode in der Praxis eingesetzt wird. Bei der Absorptionsmethode (Konsolidierung) werden die Systeme des übernehmenden Unternehmens verwendet, um eine schnelle und effiziente Umwandlung zu ermöglichen. Wenn beide Unternehmen, die gleiche Kollaborationsplattform verwenden, wird die Konsolidierung der Kollaborationsplattform als unkritisch von den befragten Experten gesehen. Bei unterschiedlichen Kollaborationsplattformen wird dies allerdings als herausfordernd beschrieben. Die Konsolidierung von zwei unterschiedlichen Kollaborationsplattformen birgt technische Herausforderungen, die Auswirkungen auf im Unternehmen eingesetzte Prozesse und Tools haben. Im Unternehmen eingesetzte Anwendungen und Schnittstellen müssen ggf. an die neue Kollaborationsplattform angepasst werden, was den Abschluss einer Konsolidierung der Kollaborationsplattform verzögern kann. Während der Anpassungsphase kommt es in der Praxis häufig zu einem parallel Betrieb der Altumgebung bis die Anpassung der entsprechenden Anwendungen und Schnittstellen abgeschlossen ist. Der Weiterbetrieb der Altumgebung sorgt für entsprechende Betriebsaufwände und kann somit Aufwände und Kosten verursachen, die nicht innerhalb des Mergers-&-Acquisitions-Prozesses eingeplant sind.

Ein weiterer Aspekt, der von den befragten Experten genannt wurde und der im Rahmen der Konsolidierung von unterschiedlichen Kollaborationsplattformen eine Herausforderung ist, ist die emotionale Verbundenheit der Mitarbeiter zur jeweiligen Kollaborationsplattform. Als geeignete Maßnahme, um die Akzeptanz der neuen

Kollaborationsplattform von Beginn an sicherzustellen und Resistenz gegen den übergeordneten Mergers-&-Acquisitions-Prozess zu vermeiden, ist die Ausarbeitung einer langfristigen Strategie inklusive einer geschlossenen Kommunikation an die Mitarbeiter sowie eine entsprechende Unterstützung durch das Management.

Bei der Integration eines Unternehmens, das bis dahin keine kollaborativen Cloud-Services verwendet hat, ist eine Anpassung der technischen Komponenten an die Nutzung der kollaborativen Cloud-Services zu berücksichtigen. Dies umfasst beispielsweise die Anpassung der Internetanbindung und der Hauswareausstattung.

Ein weiterer zentraler Punkt innerhalb des Mergers-&-Acquisitions-Prozesses, zusätzlich zur technischen Integration des Unternehmens, umfasst die kulturelle und organisatorische Integration des Unternehmens. Dies ist einem Experten nach Voraussetzung, damit eine nachgelagerte technische Integration gelingen kann.

4.2 Migration im Rahmen von Mergers & Acquisitions

Innerhalb des Migrationsvorgehen, ist der Umfang der zu migrierenden Daten von Bedeutung und kann in Abhängigkeit zum eingesetzten Kollaborationstool stehen. Die befragten Experten führten an, dass Daten wie E-Mails, persönliche Speicherorte und Gruppenspeicherorte bei der Migration im Fokus stehen. Bei Kollaborationstools, die Chat-Nachrichten im Fokus haben, wie beispielsweise *Slack*, steht die Migration von Chats deutlich mehr im Fokus der Migration. Ebenso muss im Rahmen der Migration geprüft werden, ob die Daten grundsätzlich migriert werden dürfen. Beispielsweise kann es Einschränkungen aufgrund von datenschutzrechtlichen Themen oder Verträgen geben, sodass Kundendaten nicht migriert werden dürfen.

Sollte es im Rahmen der Migration zu einer Phasenmigration kommen, d.h. die Umstellung findet nicht an einem Tag oder Wochenende statt, sondern in mehreren Phasen, so ist eine entsprechende Aufteilung der zu migrierenden Daten in die einzelnen Phasen vorzunehmen. Dies ist gerade bei der Migration von großen Unternehmen notwendig, da hier entsprechend große Datenmengen migriert werden müssen. Bei der Aufteilung in sogenannte Migrationsbatch oder Cluster ist es dem Experten nach von Bedeutung, welche Benutzer zusammen migriert werden, um die Zusammenarbeit während der Migrationsphasen entsprechend zu ermöglichen. Dabei sollen auch weitere Daten wie Berechtigungen auf dem persönlichen Postfach oder gemeinsam verwendete Gruppenpostfächer entsprechend betrachtet werden, um eine passende Migrations-einteilung vorzunehmen und die Kollaboration entsprechend aufrechtzuerhalten. Die Experten sehen gerade in den gemeinsam genutzten Speicherorten eine gewisse Komplexität, wenn eine Person bei der täglichen Arbeit das alte und das neue System parallel verwenden muss.

Drei der befragten Experten gaben an, dass der Einsatz von Cloud-Services im Rahmen eines Mergers-&-Acquisitions-Prozesses von Vorteil sein kann. Ein Experte führte an, dass gerade im Umfeld von Google Workspace, welches nur cloudbasiert bereitgestellt

wird, die Konsolidierung einfach durchgeführt werden kann. Diese besteht aus dem Aufnehmen der Identität in das Google Directory und anschließende Bereitstellen der Zugangsdaten für die Benutzer. Allerdings wurde auch darauf verwiesen, dass der Einsatz von Cloud-Services nur von Vorteil ist, wenn diese aktuell schon im Unternehmen eingesetzt werden. Einen Mergers-&-Acquisitions-Prozess zu nutzen, um ein Unternehmen in die Cloud zu migrieren, wird nicht als sinnvoll betrachtet.

Im Rahmen der Interviews verwies ein Experte noch auf weitere Vorteile bei der Durchführung eines Carve-Outs eines Unternehmens durch die Nutzung von Cloud-Services. Die Zeit für die Suche, die Beschaffung und den Aufbau eines eigenen Rechenzentrums entfällt. Stattdessen können entsprechende vorgefertigte Designs von Hyperscalern wie beispielsweise Microsoft verwendet werden und das neue Unternehmen aufgebaut werden.

4.3 Benutzer und Stakeholder

Bei der Betrachtung von Mergers & Acquisitions geht es auch um Personen innerhalb der betroffenen Unternehmen. Dabei werden zum einen die Benutzer der Kollaborationstools betrachtet, zum anderen aber auch entsprechende Personen und Personengruppen, die im Rahmen des Mergers-&-Acquisitions-Prozesses von besonderer Bedeutung sind und Stakeholder innerhalb des Prozesses sein können.

Dabei stehen die Auswirkungen auf die Benutzer in Abhängigkeit zu den eingesetzten Kollaborationsplattformen. Wenn alle beteiligten Unternehmen die gleiche Kollaborationsplattform verwenden, sind die Auswirkungen auf die Benutzer gering, was zwei der befragten Experten unabhängig voneinander erwähnten. Wie in Kapitel 2.1 bereits angeführt, ist die Änderung der Kollaborationsplattform für einige Mitarbeiter ein emotionales Thema und kann aufgrund der Verbundenheit mit der Kollaborationsplattform zu Widerständen führen. Zur Vermeidung von Widerständen im Unternehmen empfahlen zwei Experten die Nutzung von Change-Management-Maßnahmen. Die Relevanz von Change-Maßnahmen wird ebenfalls deutlich, da ein Experte die Einbindung von im Unternehmen vorhandenen *Change Agent Networks* im Rahmen des Mergers-&-Acquisitions-Prozesses empfiehlt. Dabei wird der Zusammenschluss von Personen, die innerhalb der Organisation den technischen Wandel vorantreiben und begleiten als *Change Agent Network* bezeichnet. Als Maßnahmen zur Kommunikation und Unterstützung der Anwender beim Wechsel der Kollaborationsplattform werden zum einen sog. *T-Minus-Kommunikationen* empfohlen, bei denen die Mitarbeiter zu gewissen Zeitpunkten über Migration und zugehörige Aktivitäten informiert werden, zum anderen aber auch sog. *Awareness Sessions*, wo die Mitarbeiter einige Tage vor der Umstellung in einem Termin eingeladen werden und entsprechende Informationen zur Umstellung und den notwendigen Schritten erhalten. Hintergrund ist, dass die E-Mail-Kommunikation laut dem Experten häufig von den Mitarbeitern ignoriert wird und daher relevante Informationen die Mitarbeiter nicht erreichen und zu Tickets nach der Umstellung führen.

Als eines der wesentlichen Probleme innerhalb des Change-Managements im Rahmen von Mergers & Acquisitions nannte ein Experte, dass die Benutzer nicht über die organisatorischen und technischen Hintergründe der Maßnahmen informiert wurden und seitens der Unternehmensführung keine Informationen zum Mergers-&-Acquisitions-Prozess kommuniziert wurden, die ggf. für Akzeptanz bei den Mitarbeitern gesorgt hätten.

Die von den Experten konkreter genannten Stakeholder umfassen zum einen Ansprechpartner aus dem Bereich der IT-Systeme innerhalb des Unternehmens. Hier wurden die folgenden Stakeholder genannt:

- Infrastrukturverantwortliche
- Workplace-Verantwortliche
- Device-Management (Fat Clients/Mobile)
- Identity Services
- Softwaredistribution
- Kollaborationsverantwortliche
- Anwendungsverantwortliche
- IT-Security

Auffällig ist dabei, dass nur IT-Security von mehr als einem Experten genannt wurde. Alle anderen Stakeholder-Gruppen wurden jeweils nur von einem Experten genannt. Hierbei ist allerdings zu beachten, dass die Stakeholder-Gruppen auch in Abhängigkeit vom jeweiligen Unternehmen ggf. überschneidend sind. So können Device-Management, Softwaredistribution und Kollaborationsverantwortliche auch zu den Workplace-Verantwortlichen gezählt werden.

Als Stakeholder außerhalb der IT wurden von den Experten genannt:

- Top-Management
- Lokationsverantwortliche
- Datenschützer
- Change-Management
- Unternehmenskommunikation

Die Stakeholder-Gruppe Change-Management wurde von zwei der vier Experten genannt, während die anderen Stakeholder nur von einem der jeweiligen Experten genannt wurden. Die Stakeholder-Gruppen Change-Management und Unternehmenskommunikation werden in den Interviews in der Kommunikationsschnittstelle zu den Benutzern betrachtet.

5 Zusammenfassung und Ausblick

Im Rahmen dieses Beitrages wurden die Auswirkungen von Mergers & Acquisitions auf kollaborative Cloud-Services analysiert, was ein Thema von zunehmender Relevanz darstellt. Anhand der vier durchgeführten Experteninterviews und der anschließenden qualitativen Inhaltsanalyse wurde ein umfassendes Verständnis für die Herausforderungen und Chancen eines Mergers-&-Acquisitions-Prozesses sowie die Bedeutung der kollaborativen Cloud-Services in diesem Umfeld beschrieben. Die verschiedensten Auswirkungen auf die kollaborativen Cloud-Services, die auf den Aussagen der Experten beruhen, wurden dargestellt. Dies zeigt, dass kollaborative Cloud-Services unbedingt im Rahmen des Mergers-&-Acquisitions-Prozesses berücksichtigt werden müssen.

Der Erfolg und die Ausprägung einer Konsolidierung kollaborativer Cloud-Services sind abhängig von vielen verschiedenen Faktoren und müssen in der Planung und Durchführung entsprechend berücksichtigt werden. Wesentliche Auswirkungen auf den Umfang und die Intensität der Kommunikations- und Change-Management-Maßnahmen, die im Rahmen des Mergers-&-Acquisitions-Prozesses angewendet werden, haben die Ausgangsplattformen der Unternehmen.

Die dargestellten Ergebnisse basieren auf den Antworten der vier Experteninterviews, die im Rahmen dieses Beitrages dargestellt wurden. Im Zuge der Arbeit kamen qualitative Forschungsmethoden zum Einsatz, weshalb die Ergebnisse in Abhängigkeit zu den Erfahrungen eines kleinen Personenkreises stehen. In einer zukünftigen Forschung könnte der Kreis der Experten vergrößert werden, wodurch die getroffenen Aussagen verifiziert werden könnten. Hierbei sollten auch weitere Einflussparameter berücksichtigt werden und gezielt Experten zu anderen IT-Integrationsmodellen befragt werden. Zusätzlich könnte auf Basis der gewonnenen Erkenntnis in einem weiteren Schritt eine quantitative Befragung durchgeführt werden, um weitere Schlussfolgerungen im Rahmen dieser Forschung zu treffen und weitere praktische Impulse für Unternehmen zu liefern, die vor der Herausforderung eines Mergers-&-Acquisitions-Prozesses stehen.

Literaturverzeichnis

- [Ba22] Bain & Company Germany, Inc.: M&A-Geschäft bricht alle Rekorde. Globaler Report von Bain zu Fusionen und Übernahmen. <https://www.bain.com/de/ueber-uns/presse/pressemitteilungen/germany/2022/ma-geschäft-bricht-alle-rekorde/>, Stand: 27.06.2023.
- [BES09] Buchta, D.; Eul, M.; Schulte-Croonenberg, H.: Strategisches IT-Management. Gabler, Wiesbaden, 2009.
- [Di99] Dillenbourg, P. Hrsg.: Collaborative learning. Cognitive and computational approaches. Pergamon, Amsterdam, 1999.

- [Ga21] Gartner: Usage of digital workplace technologies among workers worldwide in 2019 and 2021.
<https://www.gartner.com/en/newsroom/press-releases/2021-08-23-gartner-survey-reveals-44-percent-rise-in-workers-use-of-collaboration-tools-since-2019>, Stand: 16.11.2023.
- [Ha22] Hays: Welches sind die wichtigsten strategischen Ziele Ihres Unternehmens oder Ihrer Organisation? (Unternehmen/Organisationen in der DACH-Region).
<https://de.statista.com/statistik/daten/studie/1293994/umfrage/wichtigste-strategische-ziele-von-unternehmen/>, Stand: 20.11.2023.
- [HYW18] Henningsson, S.; Yetton, P. W.; Wynne, P. J.: A Review of Information System Integration in Mergers and Acquisitions. *Journal of Information Technology* 4/33, S. 255–303, 2018.
- [Ja16] Jansen, S. A.: Mergers & Acquisitions. Unternehmensakquisitionen und -kooperationen. Eine strategische, organisatorische und kapitalmarkttheoretische Einführung. Springer Fachmedien Wiesbaden, Wiesbaden, 2016.
- [Ko01] Kock, N. et al.: E-collaboration: A look at past research and future challenges. In (Chiong, D. R. Hrsg.): *Journal of Systems and Information Technology*. o. V., S. 1–8, 2001.
- [KR22] Kuckartz, U.; Rädiker, S.: *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung. Grundlagentexte Methoden*. Beltz Juventa, Weinheim, Basel, 2022.
- [Ma16] Mayring, P.: *Einführung in die qualitative Sozialforschung*. Beltz, Weinheim, 2016.
- [Mi23] Microsoft Corporation: Mandantenübergreifende Postfachmigration.
<https://learn.microsoft.com/de-de/microsoft-365/enterprise/cross-tenant-mailbox-migration?view=o365-worldwide>, Stand: 21.11.2023.
- [MM21] Mercurio, R.; Merrill, B.: *Beginning Microsoft 365 Collaboration Apps. Working in the Microsoft Cloud*. Apress, New York, 2021.
- [PMT22] Pang, A.; Markovski, M.; Trifunovski, M.: Top 10 Collaboration Software Vendors, Market Size and Market Forecast 2021-2026.
<https://www.appsruntheworld.com/top-10-collaboration-software-vendors-and-market-forecast/>, Stand: 30.11.2023.
- [St22] Statista: *Technology Market Outlook. Software - Market data analysis & forecasts*, 2022.

- [Wi17] Wirtz, B. W.: Mergers & Acquisitions Management. Strategie und Organisation von Unternehmenszusammenschlüssen. Springer Gabler, Wiesbaden, 2017.

Segmentale Geschäftsmodell-Evaluation am Beispiel der digitalen Mehrwert-Anwendung Energiemanagement

Vollständiger Beitrag

Niklas Briem¹, Benjamin Pottkamp², Joachim Gerlach³, Jessica Rövekamp⁴ und
Alexander Bade⁵

Abstract: Mehrwert-Anwendungen auf Basis intelligenter Messsysteme sollen die Digitalisierung der Energiewende akzeptanzbasiert vorantreiben. Als Mehrwert-Anwendungen können im privaten Elektrizitätssektor nicht-energiewirtschaftlich relevante Zusatz-Leistungen verstanden werden, die eine weitere Systemkomponente in der Infrastruktur des intelligenten Messsystems erfordern. Ein Beispiel hierfür liefert das optionale Energiemanagement. In dieser Arbeit wird zunächst die Methodik der „Segmentalen Geschäftsmodell-Evaluation“ (SGME) vorgestellt. Anschließend wird diese auf die Evaluation von Mehrwert-Anwendungen angewendet. Als Ergebnis zeigt sich für die Mehrwert-Anwendung „Energiemanagement“, dass von allen betrachteten Kundensegmenten vor allem Prosumer mit Wärmepumpe und ohne Batteriespeicher mögliche Zielkunden sind, aber auch Haushalte mit Wärmepumpe ohne Erzeugungs- und Speichertechnologien relevant sind. Das über alle betrachteten Kundensegmente dieser Mehrwert-Anwendung ermittelte Marktpotenzial von 329 Mio. €/a dürfte durch die Verbreitung von Wärmepumpen zunehmen. Die Nachrüstung bestehender PV-Anlagen mit Batteriespeichern würde sich hemmend auf das Marktpotenzial auswirken, da sich hierdurch der durch die Mehrwert-Anwendung „Energiemanagement“ erzielbare monetäre Kundennutzen reduziert.

Keywords: Digitalisierung der Energiewende, Smart Energy, Energiewirtschaft, digitale Geschäftsmodelle, Energiemanagement, Home Energy Management System, HEMS, intelligente Messsysteme, iMSys, Smart-Meter-Gateway, SMGW, Mehrwert-Anwendungen, MWA

1 Einführung

Mit dem „Gesetz zur Digitalisierung der Energiewende“ wurden 2016 die notwendigen regulatorischen Grundlagen für die Transformation des Energienetzes von zentralisierten, nachfrageorientierten Großkraftwerken hin zu zunehmend dezentralen, volatilen Klein-Anlagen und angebotsgesteuertem Verbrauch gelegt. [GN23] Eine Schlüsseltechnologie zur Realisierung dieser Transformation liefern dabei intelligente Messsysteme (iMSys),

¹ Hochschule Albstadt-Sigmaringen, eMpowerSYS, Anton-Günther-Str. 51, 72488 Sigmaringen, briem@hs-albsig.de

² Hochschule Albstadt-Sigmaringen, eMpowerSYS, Poststraße 6, 72458 Albstadt, pottkamp@hs-albsig.de

³ Hochschule Albstadt-Sigmaringen, INF, Poststraße 6, 72458 Albstadt, gerlach@hs-albsig.de

⁴ Hochschule Albstadt-Sigmaringen, BSM, Anton-Günther-Str. 51, 72488 Sigmaringen, rovekamp@hs-albsig.de

⁵ Hochschule Albstadt-Sigmaringen, BSM, Anton-Günther-Str. 51, 72488 Sigmaringen, bade@hs-albsig.de

die die analogen Ferraris-Stromzähler durch digitale Messeinrichtungen ersetzen und über Smart-Meter-Gateways (SMGW) eine informationstechnisch sichere Anbindung der Netzanschlusspunkte an energiewirtschaftliche Akteure über das Weitverkehrsnetz (WAN) gewährleisten. [Bu21b] Über die CLS⁶-Proxy-Funktionalität des SMGW können weitere Systemkomponenten (CLS-Komponenten) im Heimnetz (HAN) der Liegenschaft sicher angebunden werden. [Bu23b] Hierdurch können zusätzlich zu den Grundfunktionalitäten des iMSys weitere Datenerfassungs- und Datenverarbeitungskapazitäten als Edge-Technologien beim Endkunden integriert und dort z. B. lokal angebundene Verbrauchs- und Erzeugungseinheiten geregelt und optimiert werden.

Um das Sicherheitsniveau der iMSys-Infrastruktur konsequent fortzuführen, wurde vom Bundesamt für Sicherheit in der Informationstechnik für die Ausgestaltung dieser CLS-Komponenten mit der technischen Richtlinie TR 03109-5 ein umfangreicher Anforderungskatalog bereitgestellt. Dessen Umsetzung wird im Rahmen eines Zertifizierungsverfahrens für neue CLS-Komponenten geprüft. [Bu23b] Mit dieser sowohl technisch als auch regulatorisch abgesicherten Erweiterung des informationstechnischen Energienetzes ergeben sich neue digitale Geschäftsmodelle, die deutlich über die Möglichkeiten des klassischen, rein analogen Netzes hinausgehen.

Die Notwendigkeit, mit den CLS-Komponenten separate Hardware beim Endkunden bereitstellen zu müssen, sowie der erhöhte Entwicklungsaufwand durch die Zertifizierung stellen jedoch Hemmnisse sowohl anbieterseitig beim Markteintritt als auch verbraucherseitig bei der Nutzerakzeptanz für die Markteinführung dieser Geschäftsmodelle dar. Das Ziel des Verbundvorhabens „eMpowerSYS“ ist es, diese Hemmnisse zu minimieren, indem mit dem Mehrwertmodul (MWM) eine CLS-Komponente als Plattformlösung entwickelt und bis zur Feldtestreihe gebracht wird. Das MWM bietet eine Hardwareplattform in der Endkundenliegenschaft, die die notwendigen Schnittstellen für die protokollbasierte Datenerfassung und die Regelung angebundener Verbrauchs- und Erzeugungseinheiten sowie die sichere Kommunikation über das SMGW zur Verfügung stellt. Nachdem diese Plattform einmalig beim Endkunden in Betrieb genommen ist, können lokale Softwarelösungen, die für die verschiedenen digitalen Geschäftsmodelle notwendig sind, als separate Anwendungen über das WAN auf das MWM aufgespielt werden. Dies vereinfacht den technischen Anbindungsprozess des Kunden. Dadurch kann mittelfristig eine zunehmende Akzeptanz der iMSys-Technologie erreicht werden. Dies geschieht durch die Erschließung kundenseitiger Nutzenaspekte über die Umsetzung von Mehrwerten, die über die systemrelevanten Aufgaben des iMSys hinausgehen. [Fo24]

⁶ CLS bezeichnet das „Controllable Local System“, also steuerbare Verbrauchs- oder Erzeugungseinrichtungen des Endkunden, die über einen abgesicherten Kanal mit den relevanten Akteuren im Weitverkehrsnetz angebunden sind. In diesem Fall agiert das SMGW als TLS-Proxy und terminiert sowohl die WAN-Verbindung als auch die HAN-Verbindung. [Bu21b]

Als Partner des Verbundvorhabens „eMpowerSYS“ liegt der Fokus der Hochschule Albstadt-Sigmaringen u. a. in der Eruierung, Evaluation und potenzialorientierten Exploration dieser Mehrwert-Anwendungen (MWA). [Ho24] Entsprechend der Ziele des Verbundvorhabens wird dabei auf digitale Geschäftsmodelle eingegangen, die sich in der Umsetzung als lokale, softwarebasierte MWA auf dem MWM abbilden lassen. Im Rahmen der Projektarbeiten wurde hierfür eine als „Segmentale Geschäftsmodell-Evaluation“ (SGME) bezeichnete Methodik zur Evaluation digitaler, privatkundenorientierter Geschäftsmodelle im Energiebereich entwickelt. In der vorliegenden Publikation wird gezeigt, wie MWA identifiziert werden können und die SGME auf diese adaptiert werden kann. Anschließend erfolgt eine beispielhafte Anwendung der Methodik zur Kundensegmentierung von MWA.

2 Gegenstand der Evaluation: Mehrwert-Anwendungen

Im Rahmen der vorliegenden Analyse werden potenzielle Geschäftsmodelle für MWA daraufhin untersucht, ob sie aus Anbietersicht lohnend sind. Hierzu ist es zunächst erforderlich, dass Kandidaten isoliert werden, die sich als digitale Geschäftsmodelle im privaten Energiebereich anbieten. Diese müssen anschließend auf diejenigen Kandidaten eingegrenzt werden, die sich gemäß dem Projektziel als MWA abbilden lassen. Bevor diese Eingrenzung vorgenommen werden kann, ist es daher zunächst erforderlich zu definieren, was konkret unter einem „digitalen Geschäftsmodell“ bzw. einer MWA verstanden wird.

2.1 Digitale Geschäftsfelder im privaten Elektrizitätssektor auf Basis von iMSys

Im Rahmenwerk „Geschäftsmodell-Framework Energiewirtschaft“ des „Kopernikus-Projekts Systemintegration“ werden 69 Geschäftsmodell-Prototypen (GM-Prototypen) multidimensional typisiert, welche die Energiewirtschaft umfassend abbilden. [Gi19] Hieraus sind in der vorliegenden Arbeit insbesondere privatkundenfokussierte GM im Elektrizitätssektor auf Basis von iMSys, der digitalen Mess- und Kommunikationsinfrastruktur, relevant. Daher werden nur GM-Prototypen mit den Ausprägungen „Haushalte“, „Prosumer“ (Erzeuger und Verbraucher in einer Person) oder „[Elektro-]Autonutzer“ im Kundensegment sowie den Technologieausprägungen „Smart Meter“ und „sonstige Mess- und Steuerungstechnik“ berücksichtigt. [Gi19]

In Abb. 1 ist der Selektionsprozess von links nach rechts dargestellt, mit welchem sich aus den 69 GM-Prototypen anhand der obigen Merkmalsausprägungen 26 für potenzielle MWA relevante GM bzw. 12 übergeordnete Geschäftsfelder (GF) identifizieren lassen. Letztere gewährleisten weiter gefasst eine für diese Betrachtung optimierte Abgrenzung der teils noch nicht im Markt etablierten GM untereinander. Die Zusammenfassung der GM zu GF führt zu einer effektiveren Bewertung, da sich bei artverwandten GM vor dem

Hintergrund der teils mangelnden Datenlage kaum Bewertungsunterschiede ergäben. Unter den GF lassen sich die 26 GM, wie über die Pfeile visualisiert, subsumieren. Gründe hierfür sind etwa technische Voraussetzungen, direkte Substitutions- oder unterstützende Ergänzungsleistungen der GM zu den GF. Die GF sind anhand eines aussagekräftigen GM-Prototypen benannt, d. h. unter den GF sind auch die identisch bezeichneten GM-Prototypen enthalten. [Hu14], [Gi19] Die Differenzierung in Basis-Anwendungen, transzendente und native MWA wird nachfolgend erläutert.

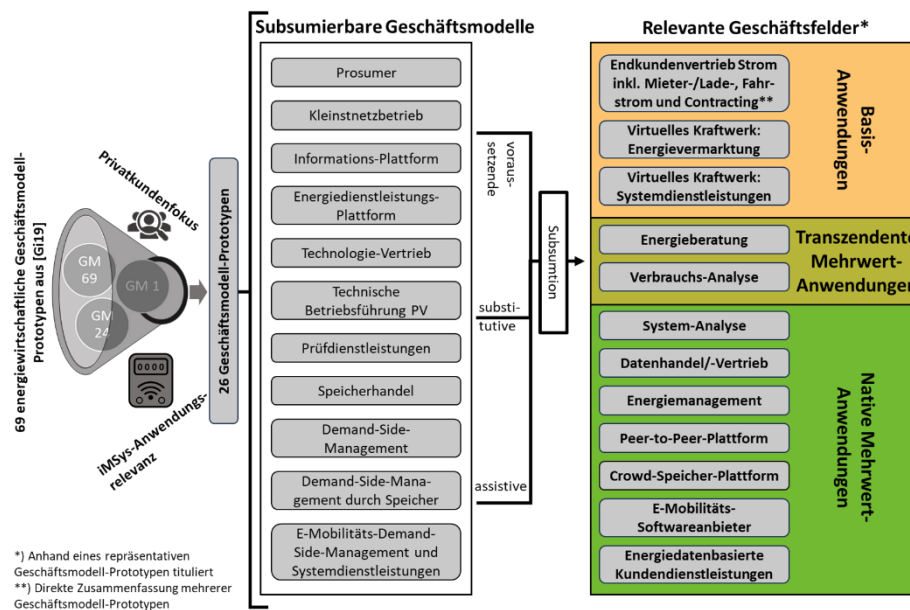


Abb. 1: Selektionsprozess der für MWA relevanten GF, mit Daten aus: [Gi19], [Ms23], [Bu21a]

2.2 Definition des Begriffs der Mehrwert-Anwendungen

„Wert“ bzw. „Wertschöpfung“ ist im Kundennutzen integraler Bestandteil eines GM. [Sc14] Eine „Mehrwert-Anwendung“ muss entsprechend einen größeren Kundennutzen als die übliche Wertstufe aufweisen. Das MsbG definiert in § 34 „Standard-, auf Verlangen verpflichtende und optionale Zusatz-Leistungen“ des Messstellenbetreibers auf Basis von iMSys. Aufgrund ihrer Quasi-Angebotsverpflichtung stellen diese „Standard- und verpflichtenden Zusatz-Leistungen“ also zunächst Basis-Anwendungen ohne Mehrwert dar. Konkret muss etwa die Datenkommunikation an den Energielieferanten zur stromtariflichen Abrechnung („Zählerstände“) vom Messstellenbetreiber offeriert werden. Ein direkter Mehrwert durch dieses „Angebot“ besteht also nicht. Bei MWA handelt es sich somit nicht um Anwendungen, mit denen bilanzierungs-, netz-, oder abrechnungsrelevante Mess- oder Steuerungsleistungen erbracht werden, die auf

„energiewirtschaftlich relevanten Daten“ (ERD)⁷ gemäß § 19, Abs. 2 MsbG basieren und energiewirtschaftlich zwingend erforderlich sind. [Ms23], [Bu23a] MWA auf Basis von iMSys sind daher „betrieblich“⁸ und dergestalt als „energieversorgungsfremd“ anzusehen. [Ms23], [Bu23a], [We23a] Jede der energiedatenbasierten Leistungen aus § 34 MsbG lässt sich anhand der „energiewirtschaftlichen Anwendungsfälle“ (EAF) des Bundesamts für Sicherheit in der Informationstechnik, welche energiewirtschaftliche sowie systemrelevante Leistungen der iMSys formulieren, [Bu21a] einordnen. Energiewirtschaftliche und systemrelevante Leistungen sind etwa die Erhebung (nicht) abrechnungsrelevanter Elektrizitätsdaten (EAF 0.2 und EAF 0.1) oder die netzdienliche Steuerung von Verbrauchseinrichtungen in der Niederspannung nach § 14a EnWG (EAF 1), also die Reduktion der Strombezugslast bei Netzengpässen. Daher beruhen mit Ausnahme der „optionalen Zusatz-Leistungen“ alle EAF der „Standard- und Zusatz-Leistungen“ auf ERD. [Bu21a], [Ms23]

MWA, die Privatkunden fokussieren, lassen sich also aus den betrieblichen EAF im Elektrizitätssektor ableiten, welche nicht die „Standard- und verpflichtenden Zusatz-Leistungen“ umfassen. Der Mehrwert von als MWA bezeichneten GM besteht also inhärent dadurch, dass es sich um ein optionales Angebot handelt. Unter MWA werden also GM verstanden, welche vom Messstellenbetreiber angeboten werden können, aber nicht müssen, und vom Kunden genutzt werden können, aber nicht müssen. Hierfür ist eine zusätzliche Systemkomponente in der Infrastruktur des iMSys erforderlich, etwa das MWM, da die MWA nach aktuellem Stand nicht Teil des Funktionsumfangs des SMGW sind. Auf dieser eingangs erläuterten Systemarchitekturlösung lassen sich digitale GM als lokale, softwarebasierte MWA implementieren. Sie können in native und transzendente MWA unterschieden werden. Native MWA, wie das über die netzorientierte Dimmung nach § 14a EnWG [En24] hinausgehende „Energiemanagement“ (EAF 2), weisen einen inhärenten Mehrwert zu den verpflichtenden Basis-Anwendungen auf. Transzendente MWA, wie z. B. die „Verbrauchs-Analyse“, können entweder den Basis-Anwendungen (z. B. der „TRuDi-Transparenz- und Displaysoftware“) zugeordnet werden oder zu MWA transzendieren. Beispielsweise überschreitet die Ausgestaltung der „Verbrauchs-Analyse“ durch Verwendung von hochfrequenten Messwerten (EAF 14) zur Geräte-Identifikation den Basis-Bereich und stellt damit eine MWA dar. [Gä21], [Ms23], [Bu21a] Die übrigen MWA wurden ebenfalls anhand dieses definitorischen Rahmenwerks klassifiziert und grenzen so den Analysefokus im Verbundprojekt „eMpowerSYS“ ein.

Zusammengefasst lassen sich die GF im privaten Elektrizitätssektor unter den Basis-Anwendungen, transzenten MWA oder nativen MWA einordnen. *Basis-Anwendungen* beinhalten verpflichtende bzw. „energiewirtschaftlich relevante“ Leistungen. Demgegenüber sind *native MWA* optional und lassen sich aus den „betrieblichen“ EAF ableiten. Konkrete GM der als *transzendente MWA* kategorisierten GF sind entweder eine

⁷ ERD weisen aufgrund ihrer Relevanz für das Energiesystem eine besondere Schutzbedürftigkeit auf. Die Datenkommunikation der ERD darf ausschließlich über die per „Smart-Metering-Public-Key-Infrastruktur“ (SM-PKI) gesicherte WAN-Verbindung erfolgen. [Bu23a]

⁸ Die Bundesnetzagentur definiert „betriebliche Daten“ als Gegenstück zu den ERD. [Bu23a]

Basis-Anwendung oder eine MWA. Daneben existieren nicht-energiekostenbasierte MWA, welche das Smart-Meter-Gateway (SMGW), z. B. zur Fernüberwachung von Rauchmeldern, als Kommunikationsplattform nutzen und hier nicht betrachtet werden. [Bo18], [Hä22]

3 Segmentale Geschäftsmodell-Evaluation (SGME)

Um die zuvor identifizierten MWA zu evaluieren, wird im Folgenden die „Segmentale Geschäftsmodell-Evaluation“ (SGME) als auf Kundensegmentierung basierte Methodik vorgestellt und angewendet. Ziel der SGME ist es, ein potenzielles Geschäftsfeld daraufhin zu untersuchen, für welche Kundensegmente es welchen Nutzen aufweist und hieraus für den Anbieter ein mögliches Marktpotenzial abzuschätzen. Hierzu ist es zunächst nötig, dass der Kreis potenzieller Kunden in einzelne Segmente aufgeteilt wird. Die Kundensegmente sollten dabei untereinander disjunkt und in sich weitgehend homogen sein. Als Segmentierungskriterien eignen sich daher Eigenschaften, die im Hinblick auf die gewählten Kennzahlen, die zum möglichen Marktpotenzial führen, unterschiedliche segmentale Bewertungen erwarten lassen. Um die Aussagekraft der Kennzahlen sowohl für die Segmente als auch für die Gesamtbetrachtung zu erhöhen, ist deren Anordnung im Kennzahlenbaum sinnvoll. Im Ergebnis zeigt sich in Abhängigkeit der gewählten Kennzahlen, welche Kundensegmente attraktiv sind, Randkunden darstellen oder vernachlässigt werden können. [Hu14]

Im Folgenden wird zunächst aufgezeigt, wie die hierfür notwendigen Kundensegmente im privaten Elektrizitätssektor gebildet werden. Anschließend werden die relevanten Kennzahlen für die Evaluierung von MWA identifiziert und vorgestellt. Hierdurch ist es grundsätzlich möglich, alle MWA zu bewerten und somit deren Marktpotenzial für jedes einzelne Kundensegment abzuschätzen. Im 4. Kapitel erfolgt dann ein konkretes Rechenbeispiel für eine exemplarische MWA.

3.1 Kundensegmente

Da der Nutzen, der den Kunden durch die verschiedenen GM entsteht, von individuellen Faktoren abhängt und zwischen Kunden stark variieren kann, ist es notwendig Kundensegmente zu bilden, für die ein repräsentativer Musterkunde existiert.

Für die Privatkundenfokussierten GM im Elektrizitätssektor ist der maßgebende Faktor, der die Nutzenkalkulation beeinflusst, die Kombination an „Energiewendetechnologien“⁹ in Form von Photovoltaik-Anlage (PV), Batteriespeicher (BS) oder Wärmepumpe (WP), die eingesetzt werden. Der Grund hierfür ist, dass diese Technologien große Auswirkungen auf den für mehrere MWA nutzenrelevanten Lastgang und Strombezug haben. Beispielsweise unterscheidet sich der Nutzen der MWA

⁹ Aufgrund unzureichender Daten konnten E-Autos nicht berücksichtigt werden.

„Energiemanagement“ oder „Peer-to-Peer-Anwendung“ stark in Abhängigkeit davon, ob ein Kunde über einen Batteriespeicher verfügt oder nicht. Auch der Nutzen anderer, nicht auf physischen Stromflüssen basierender MWA, ist technologieabhängig. Beispielsweise ist die MWA der „System-Analyse“ (Predictive Maintenance) nur dann von Nutzen, wenn ein Kunde überhaupt über technische Anlagen wie PV, BS oder WP verfügt. [Gi19] Daher wurde bei der Kundensegmentierung zur Evaluierung von MWA ein technologiebasierter Ansatz auf Basis des „KfW-Energiewendebarmeters“ gewählt. [Rö23]

Es gibt in Deutschland rund 39,3 Mio. Haushalte, davon 16,5 Mio. Eigentums-Haushalte (EH) und 22,8 Mio. Miet-Haushalte. [St23a] Die SGME lässt sich grundsätzlich auf alle denkbaren Kundengruppen anwenden. Aufgrund der Datenlage zum Zeitpunkt der Betrachtung werden im Folgenden allerdings lediglich EH betrachtet.

Aus der Kombination der verschiedenen Technologien unter der Randbedingung, dass BS zur Direktverbrauchsoptimierung nur von Prosumern eingesetzt werden, ergeben sich sechs Kundensegmente. Diese sind in Tab. 1 dargestellt. Basierend auf Berechnungen mit Daten aus [St23a], [Rö23] wurde das jeweilige segmentale Absatzpotenzial bestimmt. Mit 14,9 Mio. der insgesamt rund 16,5 Mio. EH in Deutschland deckt die Methodik somit etwa 90% der deutschen EH ab.

Kunden-Segment	EH	Technologiemix			Bezeichnung	Absatzpotenzial
		PV	BS	WP		
0	Ja	-			Unberücksichtigte EH	1,6 Mio.
1	Ja	Nein	Nein	Nein	EH	11,3 Mio.
2	Ja	Ja	Nein	Nein	EH mit PV	1,3 Mio.
3	Ja	Ja	Ja	Nein	EH mit PV und BS	0,6 Mio.
4	Ja	Ja	Ja	Ja	EH mit PV, BS und WP	0,2 Mio.
5	Ja	Ja	Nein	Ja	EH mit PV und WP	0,4 Mio.
6	Ja	Nein	Nein	Ja	EH mit WP	1,1 Mio.
EH Total						16,5 Mio.
Berücksichtigte EH						14,9 Mio.

Tab. 1: Absatzpotenziale der Kundensegmente nach Technologiemix [Rö23], [St23a]

Die Modellparameter der segmentspezifischen, repräsentativen Musterkunden werden im Rahmen der Anwendung in Kap. 4 erläutert.

3.2 Kennzahlen

Damit für potenzielle Anbieter von MWA Handlungsimplicationen abgeleitet werden können, müssen generische, primär quantitative Analysen mit betriebswirtschaftlichem Fokus durchgeführt werden. Solche Analysen liegen insbesondere für energiewirtschaftliche Anwendungen kaum vor. [Ge20] Da auch für noch nicht im Markt etablierte GM

kaum aussagekräftige Daten vorliegen, ist darüber hinaus ein Evaluationssystem auf Basis weniger, wesentlicher Kriterien sinnvoll.

Im Kennzahlenbaum in Abb. 2 sind die sachlogischen und kalkulatorischen Zusammenhänge der nachfolgend ausgeführten Kriterien dargestellt. Das Oberkriterium des Marktpotenzials ist entsprechend der Pfeilrichtung das Produkt aus den Unterkriterien des maximalen monetären Kundennutzens und des Absatzpotenzials. Wie oben erläutert, kann jedes Kriterium für jedes Kundensegment getrennt betrachtet und anschließend kumuliert werden. Dies wird über die farblich visualisierten Kundensegmente, die entsprechend der „Plus-Notation“ beliebig kombiniert und für das kumulierte Marktpotenzial addiert werden können, in den Kriterien verdeutlicht.

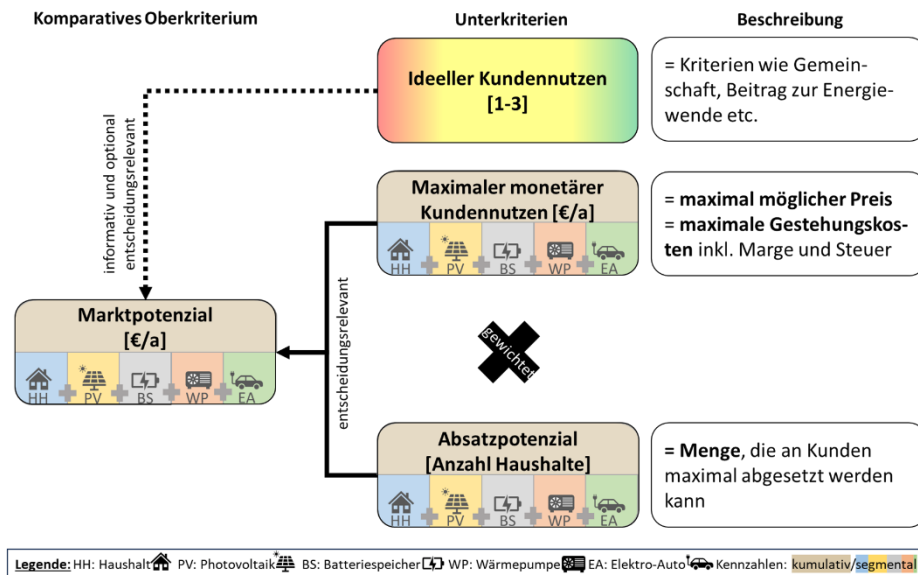


Abb. 2: Kennzahlenbaum der SGME, mit Daten aus: [Rö23], [Ge20], [Sc14], [Kn14]

Der Kundennutzen ist integraler Bestandteil eines GM. [Sc14] Auch wenn hierunter häufig die Kostenminimierung oder die Erlösmaximierung verstanden wird, tragen auch ideelle Faktoren zur Nutzenmaximierung bei. [Sh21] Daher wird der Kundennutzen für die nachfolgende Bewertung sowohl quantitativ als auch qualitativ berücksichtigt. Die quantitative Betrachtung umfasst den maximalen jährlichen monetären Kundennutzen der einzelnen Kundensegmente, während der qualitative Nutzen auf ideelle Faktoren abzielt. Unter der Annahme, dass die MWA i. d. R. nur abgesetzt wird, wenn der monetäre Kundennutzen größer als die Aufwendungen inkl. Marge und Steuern ist, kann dieser auch als maximal erzielbare Preisobergrenze angesehen werden. Die Selbst- bzw. Gestehungskosten inkl. Marge und Steuern dürfen also nicht den monetären Kundennutzen übersteigen. Somit wird gewährleistet, dass die mit der Evaluation verbundenen strategischen Entscheidungen über die Implementierung der MWA vor dem Hintergrund

der unternehmensindividuellen Kostenkalkulation getroffen werden können. Neben dem monetären Kundennutzen fließt, wie oben erwähnt, auch der ideelle Nutzen in die Nutzenbetrachtung der Kunden ein. Allerdings sind solche ideellen Faktoren nur schwer quantifizierbar. Dieses Problem wird durch die aktuell noch kaum vorhandene Datenbasis zusätzlich erschwert. Darüber hinaus handelt es sich bei den meisten MWA um Anwendungen, die lediglich im Hintergrund agieren und keine direkte „Nutzer-Experience“ versprechen. Somit dürfte es kaum möglich sein, mittels ideeller Produktdifferenzierung die Kaufentscheidung zu beeinflussen. Diese wird folglich durch monetäre Aspekte dominiert.

Aus diesen Gründen wird im Folgenden davon ausgegangen, dass der monetäre Kundennutzen allein entscheidungsrelevant ist, wohingegen die ideelle Betrachtung, insbesondere beim Vergleich mehrerer MWA, lediglich informativ einfließt.

Als Produkt aus maximalem Preis und Absatzpotenzial nach Tab. 1 ergibt sich das Oberkriterium des Marktpotenzials in Euro pro Jahr. Dies stellt eine Obergrenze dar und ist gleichbedeutend zu dem in einem reifen Markt maximal realisierbaren Marktvolumen bei einer Marktsättigung von 100 %. Für eine Abschätzung des tatsächlichen (und nicht maximalen) Marktvolumens bedarf es adäquater Abschläge, die sich etwa aus dem Produktlebenszyklus-Modell ergeben können. [Ba13]

Durch Addition des Marktpotenzials der jeweiligen Kundensegmente ergibt sich das kumulierte Ergebnis. Evaluiert werden ausschließlich die durch die MWA erzielbaren monetären Vorteile, z. B. von einem EH mit PV und MWA zum identischen EH mit PV ohne MWA. Durch die segmentale Betrachtung wird gewährleistet, dass der monetäre Kundennutzen explizit auf die MWA, losgelöst der zu Grunde liegenden Technologien, zurückgeführt werden kann. Aus den segmentalen und kumulativen Kriterien lassen sich differenzierte Handlungsimplicationen ableiten, wie in Kap. 4.2 anhand der MWA „Energiemanagement“ beispielhaft gezeigt wird.

4 Anwendung der Evaluationsmethodik am Beispiel des Energiemanagements

Die oben beschriebene und generisch auf MWA adaptierte Bewertungsmethodik der SGME wurde im Rahmen des Verbundvorhabens „eMpowerSYS“ auf mehrere MWA angewendet. Im Folgenden werden beispielhaft die Ergebnisse der MWA „Energiemanagement“ dargestellt, da dieser Anwendungsfall besonders praxisrelevant ist. Anschließend werden diese Ergebnisse evaluiert und Handlungsimplicationen abgeleitet. Grundlage für alle Berechnungen sind Daten aus dem Jahr 2023. Die Bewertung wird als Ex-post-Analyse durchgeführt.

4.1 Modellierung der Musterkunden und des Energiemanagements

Der Jahresstromverbrauch eines durchschnittlichen EH ohne weitere Technologie wird in der folgenden Modellierung mit 3.170 kWh/a angenommen. [St24], [St23b], [St23a] Rund die Hälfte hiervon entfällt auf die bedarfs- und programmorientierten Prozesswärme-Anwendungen wie Ofen, Herd oder Trockner und die zustandsorientierten Prozesskälte-Verbraucher, etwa Kühl- oder Gefrierschränke.¹⁰ [Ka14], [Ze12], [BD24a] Bei den EH mit WP, welche mit 5.000 kWh/a dimensioniert sind, liegt der Anteil zustandsorientierter Verbraucher am Gesamtverbrauch (8.170 kWh/a) wesentlich höher. [Bu23c], [Gü24] Für die zum Energiemanagement prädestinierten programm- und insbesondere zustandsorientierten Verbraucher (42 % Anteil am Gesamtverbrauch ohne WP) können Verschiebungsanteile zwischen 25 % und 35 % des Gerätegesamtverbrauchs angenommen werden. [KI07], [Ka14] Hieraus ergeben sich maximale Lastverschiebungsanteile zur Gesamtlast von etwa 15 % bei EH ohne WP respektive 27 % bei den EH mit WP.¹¹ In den Analysen zeigte sich, dass sich der monetäre Kundennutzen approximativ proportional zum Lastverschiebungsanteil verhält.

Für Prosumer ist die direktverbrauchsorientierte Lastverschiebung und für reine Konsumenten die preisorientierte Ausgestaltung, hier in Form eines stündlich-dynamischen Tarifs, relevant. [Ke20], [Hi14] In der Modellierung wird die Annahme getroffen, dass Prosumer die klassische Stromtarifizierung mit einem statischen, zeitunabhängigen Strompreis präferieren. Dies liegt daran, dass bei diesen der Strombezug aus dem Netz überwiegend in den eigenerzeugungsarmen Monaten im Winter und Herbst stattfindet. Somit können sie kaum von den vergleichsweise niedrigen Strompreisen und vorteilhafteren Preisverläufen im Frühjahr und Sommer profitieren. [Li23], [Sm24a], [Sm24b]

Der Jahresgesamtverbrauch bleibt sowohl bei reinen Konsumenten als auch bei Prosumern unverändert. Der monetäre Kundennutzen ergibt sich allein aus der Lastverschiebung. [Hi14] Beim direktverbrauchsorientierten Energiemanagement ergibt er sich aus dem Produkt des tatsächlich zum Direktverbrauch verschobenen Verbrauchs und dem Bezugsstrompreis abzüglich der Einspeisevergütung als Opportunitätskosten. Bei Kunden ohne PV-Eigenerzeugung ergibt sich der monetäre Nutzen durch die Differenz der Strombezugspreise aufgrund der Verschiebung des Verbrauchs in Stunden mit niedrigerem Strompreis. Hierzu werden verschiedene Preisbereiche innerhalb eines Tages definiert. Im teuersten Drittel des Tages findet die Lastreduktion statt, welche im günstigeren Drittel „nachgeholt“ wird. Im mittleren Preisbereich wird infolge marginaler Verschiebungsanreize vereinfachend von einem unveränderten Verbrauch ausgegangen.

¹⁰ Bedarfsorientierte Verbraucher, z. B. die Beleuchtung, werden nur in Bedarfssituationen und i. d. R. manuell gesteuert, wohingegen programmorientierte Verbraucher, darunter Waschmaschinen, teil-automatisiert, etwa über Zeitprogrammierungen, geregelt werden können. Zustandsorientierte Verbraucher sind überwiegend thermische Anwendungen, z. B. Wärmepumpen oder Kühl-/Gefrierschränke, deren Steuerung sich voll-automatisiert an definierten Temperatur-Zuständen ausrichtet. [Ka14], [Ze12]

¹¹ Bei EH ohne WP ergibt sich der Lastverschiebungsanteil wie folgt: $15 \% = 42 \% \cdot 35 \%$.

Mit WP muss dieser mit den Verbrauchsanteilen mengengewichtet um den der WP (35 %) ergänzt werden ($27 \% = 15 \% \cdot (3.170 \text{ kWh/a} / 8.170 \text{ kWh/a}) + 35 \% \cdot (5.000 \text{ kWh/a} / 8.170 \text{ kWh/a})$).

Die stündliche Lastverschiebung wird durch die näherungsweise angenommene Grund- bzw. Minimallast der Standardlastprofile in Höhe von 100 W, [Li23], [Fr23] die maximalen Lastverschiebungsanteile und bei Prosumern zusätzlich durch die Erzeugungsmenge begrenzt.

Für die stundenbasierte Ex-post-Analyse fließen die Standardlastprofile H0 ein, die je nach Kundensegment um den WP-Verbrauch komplettiert werden. [Li23], [Fr23] Das PV-Erzeugungsprofil basiert auf modifizierten deutschlandweiten Daten. [Sm24b] Der mit einer angenommenen installierten Leistung der PV-Anlage von rund 7,7 kWp [Ma24b] erzielbare Ertrag wird mit 7.700 kWh/a beziffert. Bei einer maximalen Lade-/Entladeleistung von 5,3 kW wird der BS mit rund 8,4 kWh Speicherkapazität, [Ma24a], [We23b] einer maximalen Entladetiefe von 95 % sowie einem Wirkungsgrad von ebenfalls rund 95 % modelliert. [We23b] Zu den stündlichen Börsenpreisen werden typische Steuern, Abgaben und Umlagen addiert, um so stündliche Brutto-Strompreise zu erhalten. Daraus ergibt sich ein durchschnittlicher Strompreis von rund 29 ct/kWh. [Sm24a], [BD24b] Analog wird dies als Verbrauchspreis der klassischen Tarifierung angesetzt und kann als repräsentativ für Neukundenpreise in der zweiten Hälfte des Jahres 2023 angesehen werden. [NV24] Die PV-Einspeisevergütung beträgt 8,2 ct/kWh. [Bu24]

4.2 Evaluationsergebnisse und anbieterseitige Handlungsimplicationen

Die Evaluationsergebnisse der vorgestellten Methodik sind in Abb. 3 zusammengefasst.

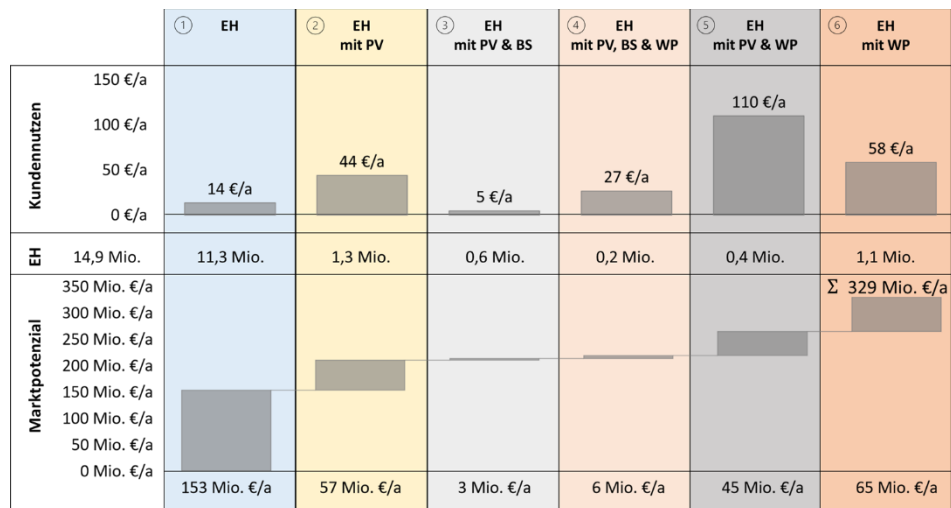


Abb. 3: Evaluations-Matrix der MWA „Energiemanagement“, mit Daten aus [Rö23], [St23a], [Fr23], [Li23], [Sm24b], [Sm24a], [BD24b]

Generell weist die MWA „Energiemanagement“ einen höheren monetären Kundennutzen in den verbrauchsstarken Segmenten mit WP auf, sofern diese über keinen BS verfügen (Kundensegmente 5 und 6), da hier auch absolut betrachtet größere Lasten verschoben werden können. Jene mit BS (Kundensegment 4) können allerdings kaum profitieren, da sowohl BS als auch das betrachtete Energiemanagement den Direktverbrauch maximieren. Dieser kann durch die MWA „Energiemanagement“ lediglich marginal erhöht werden, da er durch den BS bereits auf einem hohen Niveau liegt. Kundensegment 6 zeichnet sich durch einen geringeren Kundennutzen aus als Kundensegment 5, da die durch den betrachteten stündlich-dynamischen Tarif abschöpfbare Preisdifferenz finanziell weniger attraktiv als die Verschiebung zu Gunsten eines höheren Direktverbrauchs ist.

Bei den verbrauchsschwachen EH (Kundensegmente 1 bis 3) liegen zwar zum Teil hohe segmentale Marktpotenziale vor, es ist jedoch fraglich, ob die Kaufschwelle erreicht werden kann, da z. B. beim Kundensegment 1 (EH ohne PV, BS und WP) der jährliche maximale Kundennutzen bei lediglich 14 €/a liegt. Dies dürfte kaum genügen, um das Produkt unternehmensseitig zu finanzieren. Insgesamt ist erkennbar, dass sich der monetäre Kundennutzen in erster Linie durch eine Lastverschiebung der Wärmepumpe ergibt, sofern diese nicht bereits über einen BS stattfindet. Die Anzahl der EH, bei denen dies möglich ist (Kundensegmente 5 und 6), ist mit insgesamt rund 1,5 Mio. zwar vergleichsweise gering, allerdings führt der hohe Kundennutzen zu einem hohen Marktpotenzial. Dieses ist mit rund 110 Mio. €/a etwa drei Viertel so hoch wie bei Kundensegment 1 (153 Mio. €/a), obwohl es über 11 Mio. EH ohne Energiewendetechnologien gibt. Durch den ideellen Kundennutzen können zwar ggf. höhere Produktpreise erzielt werden, allerdings sind diese, wie oben beschrieben, schwierig zu quantifizieren.

Resultierend lassen sich folgende anbieterseitigen Handlungsimplicationen auf Basis der quantitativen Evaluation formulieren:

Zielmarkt: Das Marktpotenzial der MWA „Energiemanagement“ ist allein in den betrachteten Kundensegmenten mit rund 329 Mio. €/a attraktiv. Synergieeffekte, etwa in kombinierter Ausgestaltung mit der oben angeführten „Verbrauchs-Analyse“ als transzendente MWA, könnten dieses zusätzlich potenzieren. Wie oben erläutert, dürfte derzeit das tatsächlich realisierbare Marktvolumen in Abhängigkeit der Marktsättigung allerdings deutlich unter dem Marktpotenzial liegen.

Zielkunden: Als Zielkunden lassen sich anhand des monetären Kundennutzens eindeutig Prosumer mit WP identifizieren, sofern diese über keinen BS verfügen (Kundensegment 5). Werbemittel können diesbezüglich effizient allokiert werden. Auch wenn die Kunden des Segments 6 unattraktiver sind, bestehen auch hier Einsparungen, die den Kauf hinreichend incentivieren dürften. Die Kundenakquise der anderen betrachteten Segmente erscheint nur bedingt zielführend. Gegenwärtig dürfte der Großteil der EH daher nicht akquiriert werden können, da der geringe Kundennutzen i. d. R. unterhalb der Kosten liegen dürfte.

Randkunden: Sofern die unternehmensindividuelle Kostensituation dies zulässt, könnte durch nutzenabhängige Preisdifferenzierung auch, zumindest partikular, das Kundensegment der verbrauchsarmen Prosumer ohne BS (Kundensegment 2) erschlossen werden.

Kostenkalkulation: Für die Gestehungskosten muss ein Zielwert von deutlich unter 110 bzw. 58 €/a je Kunde angestrebt werden, sodass auch nach Margenaufschlag und Mehrwertsteuer ein für die Kaufschwelle relevanter Netto-Kundennutzen verbleibt.

Marktwachstum: Insgesamt dürfte infolge der zunehmenden Verbreitung von WP auch das Marktpotenzial der MWA „Energiemanagement“ stark wachsen. Ein signifikant wachstumstreibender Effekt aufgrund der PV-Verbreitung ist hingegen ohne Großverbraucher nicht anzunehmen, da der Netto-Kundennutzen kaum über der Kaufschwelle liegen dürfte. Wachstumshemmende Implikationen könnten sich dergestalt ergeben, dass viele Prosumer BS nachrüsten, wodurch die MWA „Energiemanagement“ bei diesen praktisch obsolet wird.

5 Zusammenfassung und Ausblick

Bei Mehrwert-Anwendungen handelt es sich um Zusatz-Leistungen, die sich aus den „betrieblichen“ energiewirtschaftlichen Anwendungsfällen ableiten lassen. Sie gehen jedoch über den gesetzlich verpflichtenden Leistungsumfang hinaus und erfordern eine zusätzliche CLS-Komponente. Bei nativen MWA, wie dem „Energiemanagement“, ist dies inhärent und bei transzendenten MWA, etwa der „Verbrauchs-Analyse“, ausgestaltungsabhängig. [Ms23], [Bu21a], [Gä21] In der vorliegenden Arbeit wurden sieben native und zwei transzendente MWA im privaten Elektrizitätssektor aus dem Rahmenwerk „Geschäftsmodell-Framework Energiewirtschaft“ [Gi19] abgeleitet.

Um ein Geschäftsmodell, das auf einer solchen MWA basiert, zu bewerten, wurde die Methodik der „Segmentalen Geschäftsmodell-Evaluation“ (SGME) entwickelt. Diese basiert auf der technologiebasierten Kundensegmentierung, [Rö23] wobei die entscheidungsrelevanten quantitativen Kriterien segmental zu evaluieren sind. Anhand des maximalen monetären Kundennutzens, der Absatzpotenziale sowie dem sich multiplikativ daraus ergebenden Marktpotenzial können differenzierte Handlungsimplicationen abgeleitet werden.

Angewendet auf die MWA „Energiemanagement“ zeigten sich erhebliche Nutzendivergenzen der betrachteten Kundensegmente. Der höchste monetäre Kundennutzen ergibt sich mit 110 €/a für Prosumer mit Wärmepumpe, sofern diese über keinen Batteriespeicher verfügen. Der Kundennutzen entspricht dem maximal erzielbaren Preis. Die maximalen Gestehungskosten für Anbieter entsprechen dem zum Kauf notwendigen Netto-Kundennutzen abzüglich Marge und Steuern. Da der monetäre Kundennutzen bei den anderen Kundensegmenten größtenteils deutlich unter 50 €/a liegt, dürfte es schwierig sein, diese zu akquirieren und gleichzeitig die unternehmensseitigen Kosten zu decken. Das Marktpotenzial für diese MWA konnte in den betrachteten Kundensegmenten auf insgesamt

rund 329 Mio. €/a quantifiziert werden. Allein bei den relevanten Kundensegmenten mit WP liegt es bei insgesamt etwa 110 Mio. €/a. Das größte Potenzial für das Marktwachstum wird durch die Zunahme verbrauchsstarker Technologien wie Wärmepumpen getrieben. Eine Nachrüstung bestehender PV-Anlagen mit Batteriespeichern dürfte sich hingegen hemmend auf das Marktwachstum auswirken, da sich für diese infolge marginaler monetärer Einsparungen kaum Anreize durch die MWA „Energiemanagement“ ergeben.

In der vorliegenden Arbeit wurde gezeigt, wie die vorgestellte „Segmentale Geschäftsmodell-Evaluation“ zur Entscheidungsunterstützung bei der Vermarktung von MWA genutzt werden kann. Allerdings beschränkt sich die beispielhafte Analyse auf die beschriebenen Kundensegmente. Insbesondere Kundensegmente abseits von EH bedürfen zukünftig einer vertiefenden Analyse. Darüber hinaus kann die SGME auf weitere MWA angewendet und in für Anbieter relevante Prognoseszenarien bis zum Jahr 2030 überführt werden. Letzteres wird von den Autoren gegenwärtig bearbeitet.

Danksagung

Diese Arbeit wurde gefördert durch das Bundesministerium für Wirtschaft und Klimaschutz (BMWK) im Rahmen des Projekts „eMpowerSYS – Aufbau einer Mehrwertplattform für die digitale Energiewende“ (Förderkennzeichen 03EI6074C).

Literaturverzeichnis

- [Ba13] Baum, H.-G.; Coenenberg, A. G.; Günther, T.: Strategisches Controlling. Schäffer-Poeschel Verlag für Wirtschaft Steuern Recht GmbH, Stuttgart, 2013.
- [BD24a] BDEW, <https://www.bdew.de/service/daten-und-grafiken/stromverbrauch-der-haushalte/>, Stand: 05.01.2024.
- [BD24b] BDEW, <https://www.bdew.de/service/daten-und-grafiken/bdew-strompreisanalyse/>, Stand: 24.03.2024.
- [Bo18] Bogensperger, A. et al.: Smart Meter - Umfeld - Technik - Mehrwert, München, 2018.
- [Bu21a] Bundesamt für Sicherheit in der Informationstechnik; Bundesministerium für Wirtschaft und Energie: Stufenmodell zur Weiterentwicklung der Standards für die Digitalisierung der Energiewende - Energiewirtschaftliche Anwendungsfälle, 2021.
- [Bu21b] Bundesamt für Sicherheit in der Informationstechnik: Technische Richtlinie BSI TR-03109-1 - Anforderungen an die Interoperabilität der Kommunikationseinheit eines intelligenten Messsystems, Bonn, 2021.
- [Bu23a] Bundesnetzagentur: BK6-22-253. Positionspapier zu energiewirtschaftlich relevanten Mess- und Steuerungsvorgängen nach § 19 Absatz 2 MsbG, 2023.

- [Bu23b] Bundesamt für Sicherheit in der Informationstechnik: Technische Richtlinie BSI TR-03109-5. Kommunikationsadapter, Bonn, 2023.
- [Bu23c] Bundesverband Wärmepumpe (BWP) e. V.: Branchenstudie 2023: Marktentwicklung - Prognose - Handlungsempfehlungen, Berlin, 2023.
- [Bu24] Bundesnetzagentur: Bestimmung der anzulegenden Werte für Solaranlagen für die Kalendermonate November 2022 bis Januar 2024. (Excel), 2024.
- [En24] Energiewirtschaftsgesetz. EnWG 2024.
- [Fo24] Forschungszentrum Jülich GmbH: Verbundvorhaben: eMpowerSYS – Aufbau einer Mehrwertplattform für die digitale Energiewende: Teilvorhaben Theben: Implementierung von Applikationen auf einem Mehrwertmodul für Smart Meter Gateways, <https://www.enargus.de/pub/bscw.cgi/?op=enargus.eps2&q=empowersys&v=10&id=10556926>, Stand: 30.03.2024.
- [Fr23] Frolova, K.; te Heesen, H.: Synthese von Wärmelastprofilen nach BDEW. (Excel), Hochschule Trier, 2023.
- [Gä21] Gährs, S. et al.: Potenziale der Digitalisierung für die Minderung von Treibhausgasemissionen im Energiebereich. Abschlussbericht, Dessau-Roßlau, 2021.
- [Ge20] Germanus, N. et al.: Entwicklung eines generischen Bewertungsmodells für Geschäftsmodelle der Energiewirtschaft. Zeitschrift für Energiewirtschaft 4/44, S. 285–299, 2020.
- [Gi19] Giehl, J. F. et al.: Data Documentation - Vollaufnahme und Klassifikation von Geschäftsmodellen der Energiewende. Zenodo, 2019.
- [GN23] Gesetz zum Neustart der Digitalisierung der Energiewende. GNDEW, 2023.
- [Gü24] Günther, D. et al.: Abschlussbericht – Wärmepumpen in Bestandsgebäuden – Ergebnisse aus dem Forschungsprojekt „WPsmart im Bestand“. Felduntersuchung optimal abgestimmter Wärmepumpenheizungssysteme in Bestandsgebäuden beim Betrieb im konventionellen sowie im intelligenten Stromnetz (Smart Grid), Freiburg, 2024.
- [Hä22] Häbig, P.; Hufendiek, K.: Business Model Language „Value Network“ für smarte Geschäftsmodelle: Fallstudie für eine Plattform zum Handel von Flexibilität. Zeitschrift für Energiewirtschaft 2/46, S. 87–106, 2022.
- [Hi14] Hillemacher, L.: Lastmanagement mittels dynamischer Strompreissignale bei Haushaltskunden. Diss., 2014.
- [Ho24] Hochschule Albstadt-Sigmaringen, <https://www.hs-albsig.de/forschungsdetailseite/empowersys/>, Stand: 24.03.2024.
- [Hu14] Hungenberg: Strategisches Management in Unternehmen. Springer Fachmedien Wiesbaden, Wiesbaden, 2014.
- [Ka14] Kamper, A.: Dezentrales Lastmanagement zum Ausgleich kurzfristiger Abweichungen im Stromnetz. Diss., Karlsruher Institut für Technologie (KIT), 2014.

- | | |
|---------|---|
| [Ke20] | Keppinger, P. et al.: Autonomes Demand Side Management verteilter Energiespeicher. Ein Überblick über die Forschungsarbeiten der Forschungsgruppe "Energiesysteme und -komponenten" an der Fachhochschule Vorarlberg. e & i Elektrotechnik und Informationstechnik 1/137, S. 52–58, 2020. |
| [KI07] | Klobasa, M.: Dynamische Simulation eines Lastmanagements und Integration von Windenergie in ein Elektrizitätsnetz auf Landesebene unter regelungstechnischen und Kostengesichtspunkten. Diss., Zürich, 2007. |
| [Kn14] | Knab, S.; Rohrbeck, R.; Konnertz, L.: Kooperative Geschäftsmodellentwicklung für systemische Nachhaltigkeitsinnovationen: Eine Fallstudie im deutschen Smart Energy-Markt. In (Schallmo, D. R. A. Hrsg.): Kompendium Geschäftsmodell-Innovation. Grundlagen, aktuelle Ansätze und Fallbeispiele zur erfolgreichen Geschäftsmodell-Innovation. Springer Gabler, Wiesbaden, Germany, S. 283–317, 2014. |
| [Li23] | Liesenklas, L.; te Heesen, H.: Synthese von Standardstromlastprofilen nach BDEW. (Excel), Hochschule Trier, 2023. |
| [Ma24a] | Marktstammdatenregister,
<a 24.03.2024.<="" a="" einheit="" einheiten="" erweiterteoeffentlicheeinheitenuebersicht?filter="Bruttoleistung%20der%20Einheit~lt~%2720%27~and~Energietr%C3%A4ger~eq~%272495%27~and~Name%20des%20Anlagenbetreibers%20(nur%20Org.)~ct~%27nat%3BCrliche%20Person%27~and~Betriebs-Status~eq~%2735%27~and~Inbetriebnahmedatum%20der%20Einheit~lt~%2713.04.2023%27," href="http://www.marktstammdatenregister.de/MaStR/Einheit/Einheiten/ErweiterteOeffentlicheEinheitenuebersicht?filter=Inbetriebnahmedatum%20der%20Einheit~lt~%2713.04.2023%27~and~Betriebs-Status~eq~%2735%27~and~Name%20des%20Anlagenbetreibers%20(nur%20Org.)~ct~%27nat%3BCrliche%20Person%27~and~Energietr%C3%A4ger~eq~%272496%27~and~Bruttoleistung%20der%20Einheit~lt~%2720%27~and~Nutzbare%20Speicherkapazit%C3%A4t%20in%20kWh~lt~%2720%27~and~Bruttoleistung%20der%20Einheit~nnn~%27%27, Stand: 24.03.2024.</td></tr> <tr> <td>[Ma24b]</td><td>Marktstammdatenregister,
 |
| [Ms23] | Messstellenbetriebsgesetz. MsbG, 2023. |
| [NV24] | Norddeutscher Rundfunk; Verifox, https://www.ndr.de/nachrichten/info/Strompreis-aktuell-So-viel-kosten-die-Kilowattstunden,strompreis182.html , Stand: 24.03.2024. |
| [Rö23] | Römer, D.; Salzgeber, J.: KfW-Energiewendebarmeter 2023. Energiewende im Spannungsfeld zwischen Handlungsbedarfen und finanziellen Möglichkeiten, Frankfurt am Main, 2023. |
| [Sc14] | Schallmo, D.R.A. Hrsg.: Kompendium Geschäftsmodell-Innovation. Grundlagen, aktuelle Ansätze und Fallbeispiele zur erfolgreichen Geschäftsmodell-Innovation. Springer Gabler, Wiesbaden, Germany, 2014. |
| [Sh21] | Shamon, H. et al.: Smart Energy in Haushalten. Technologien, Geschäftsmodelle, Akzeptanz und Wirtschaftlichkeit. Forschungszentrum Jülich GmbH, Jülich, 21. |
| [Sm24a] | Smard: Großhandelspreise DE/LU. (Excel), https://www.smard.de/home/downloadcenter/download- |

marktdaten/?downloadAttributes=%7B%22selectedCategory%22:3,%22selectedSubCategory%22:8,%22selectedRegion%22:%22DE-LU%22,%22selectedFileType%22:false,%22from%22:1672527600000,%22to%22:1704063599999%7D, Stand: 24.03.2024.

- [Sm24b] Smard: Realisierte Stromerzeugung DE/LU. (Excel), <https://www.smard.de/home/downloadcenter/download-marktdaten/?downloadAttributes=%7B%22selectedCategory%22:1,%22selectedSubCategory%22:1,%22selectedRegion%22:%22DE-LU%22,%22selectedFileType%22:%22XLSX%22,%22from%22:1672527600000,%22to%22:1704063599999%7D>, Stand: 24.03.2024.
- [St23a] Statistisches Bundesamt: Wohnen in Deutschland. Zusatzprogramm des Mikrozensus 2022. (Excel), Wiesbaden, 2023.
- [St23b] Statistisches Bundesamt: Statistischer Bericht. Umweltökonomische Gesamtrechnung (UGR) - Private Haushalte und Umwelt - Berichtszeitraum 2000-2021. (Excel), Wiesbaden, 2023.
- [St24] Statistisches Bundesamt, <https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Umwelt/UGR/private-haushalte/Tabellen/stromverbrauch-haushalte.html>, Stand: 24.03.2024.
- [We23a] Weiss, M.-L.: Beschlussempfehlung und Bericht des Ausschusses für Klimaschutz und Energie (25. Ausschuss), 2023.
- [We23b] Weniger, J. et al.: Stromspeicher-Inspektion 2023, Berlin, 2023.
- [Ze12] Zeilinger, F.; Einfalt, A.: Modell für hochauflösende synthetische Haushaltslastprofile. Conference Paper, Graz, 2012.

(Stand: März 2024)

Harmonising innovation and governance: A lifecycle model for high-risk AI systems under the European AI Act

Full Paper

Carolin Gotsch¹, Jörg Puchan²

Abstract: The rapid advancement of Artificial Intelligence (AI) technologies has sparked a discussion within organisations, questioning whether regulatory frameworks like the European Artificial Intelligence Act (EU AI Act) pose obstacles or opportunities for innovation and growth. In response to this ongoing discourse, this paper introduces a comprehensive lifecycle model tailored for high-risk AI systems. On the basis of a literature review, state-of-the-art Machine Learning (ML)/AI and software development lifecycles were identified to establish a foundational framework. By examining the requirements of the AI Act, specific to high-risk AI systems, actionable steps were extracted and integrated into the lifecycle. The resulting framework was developed iteratively, incorporating adaptations from identified lifecycles and mapping essential compliance steps. Expert interviews provided valuable insights for refinement, leading to a universal and future-proof lifecycle. The proposed framework not only ensures compliance with regulatory standards but also fosters innovation and development in the AI landscape.

Keywords: Artificial Intelligence; AI Lifecycle; Reference Model; AI Act; High-Risk AI Systems

1 Introduction

While Artificial Intelligence (AI) systems provide a lot of opportunities regarding efficiency and productivity to businesses, experts from the field also warn about the potential risks of the technology and call for a regulation. [KS23; An23] In order to address these threats, the European (EU) Parliament and the EU member states reached an agreement on the European Artificial Intelligence Act (EU AI Act) on the 9th December 2023, which is the first ever comprehensive regulation on AI. [Vo23; Eu23a] The aim of the act is to ensure AI in Europe being safe and respecting fundamental rights, while also thriving innovation and the expansion of businesses. [Eu24] Therefore, certain applications are banned by the law, whereas others are subject to specific obligations based on their potential risk. [Eu24] Regulatory sandboxes and real-world testing should foster innovation and are especially promoted to small and medium enterprises enabling development and training of AI before placing it on the market.

¹ Hochschule München, Fakultät für Wirtschaftsingenieurwesen, Lothstraße 64, 80335 München, caro.gotsch@gmail.com

² Hochschule München, Fakultät für Wirtschaftsingenieurwesen, Lothstraße 64, 80335 München, puchan@hm.edu

[Eu23a] A new institution, the so-called AI Office, has already been established within the EU on 21st February 2024 and will play a key role in implementing the regulation by developing tools, methodologies and benchmarks. [Eu24] After the entry into force, which is likely to be by end of May 2024, the act becomes fully applicable after a transition period of two years, with certain provisions coming into effect sooner (AI Act, Art. 85). As fines for non-compliance will be up to €35 million or 7% of the companies' turnover, whichever is higher [Eu23a], it can be seen that the impact of non-compliance with the AI Act may have a severe impact on providers or deployers of AI systems. In addition, there are already several opinions on the question of whether or not the act will lead to overregulation and thus to an innovation brake. [KNS23] This discussion also motivates the underlying work of this paper, the aim of which was to examine the impact of the EU AI Act on the development of AI systems. The following sections therefore provide an understanding of the theoretical background (chapter 2) and the methodology (chapter 3). Afterwards, an AI systems lifecycle, which is compliant to the AI Act, is presented and serves as the research questions solution in form of a reference model (chapter 4). Finally, the work is concluded and limitations are revealed (chapter 5).

2 Background

2.1 AI Act

For the purpose of this paper, the most recent version of the AI Act is used, which is the text of the provisional agreement, dated on 2nd February 2024. Within this paper it is often referred to the Annexes II, III and VII of the regulation, consequently this term always indicates a reference to the AI Act. When citing specific parts of the law, the related article of the AI Act is specified through the abbreviation “Art.” in parentheses.

The act applies to AI systems placed on the European market, put into service or used in the EU. Therefore, obligations of providers, deployers and other parties, such as distributors and importers are in the scope of the act (Art. 2). In the abovementioned version of the act an AI system is defined as “[...] a machine-based system designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments” (Art. 3 (1)). However, not the technology itself is regulated by the act but the application that the AI system serves. Therefore, a risk-based approach categorises them into four risk levels: AI systems, that pose an “unacceptable risk”, a “high risk”, a “limited risk” and a “minimal or no risk”.

Practices that are in scope with AI systems with an unacceptable risk are prohibited, for instance: real-time remote biometric identification, social scoring and emotion recognition in the workplace and educational institutions (Art. 5).

The focus of the law is on the regulation of high-risk systems as most of the text addresses the requirements for this risk category. There are two ways in which AI systems can fall into this risk category. Firstly, the system is used as a product or as a safety component across a product spectrum including civil aviation, vehicle security, marine equipment, toys, lifts, pressure equipment and personal protective equipment (Annex III). Furthermore, it must be subject to third-party conformity assessment in accordance with that Annex II legislation (Art. 6(1)). Secondly, the system is automatically classified as high-risk if it falls within one of the following uses cases: non-banned biometrics, critical infrastructure, education and vocational training, employment and workers management and access to self-employment, access to and enjoyment of essential public and private services, law enforcement, migration and asylum and border control management, administration of justice and democratic processes (Annex III; Art. 6(2)). However, exceptions apply to systems within these use-cases, which are not likely to lead to a significant risk to the health, safety or fundamental rights and may circumvent the requirements of the high-risk category. Additionally, AI systems profiling individuals are always considered high-risk (Art. 6(2a)). The majority of obligations fall on providers of high-risk AI systems for example the establishment of a risk and quality management system, conduction of data governance, design of a record keeping functionality and implementation of human oversight (Title III, Chapter 2 of the AI Act). All requirements are described in detail in chapter 4 of this paper.

Lastly, AI systems, which interact with humans, like chatbots, emotion recognition and biometric categorisation systems and AI systems that generate or manipulate image, audio or video content (deep fakes) are in the scope of limited risk systems. They underly obligations for the purpose of informing the user about him interacting with an AI system and in case of AI generated content it must be disclosed the automated creation of the content. Further, a voluntary code of conduct for none high-risk systems can be created (Art. 52). The risk of AI systems, which do not fall under one of the aforementioned applications is classified as minimal or none-existent and the systems are not subject to intervention by the AI Act. So-called general-purpose AI (GPAI) models show a significant level of generality and are designed to perform a wide range of tasks across different domains rather than being specialized in one specific area. [Art. 2(44b); Tr24] Providers of such GPAI models need to cooperate with the high-risk AI systems providers to enable compliance. Furthermore, they are obliged to create technical documentation on the model's training, testing and evaluation. Downstream providers are supplied with information regarding model integration and copyright compliance (Art. 52c). Systemic GPAI models, identified by the cumulative amount of compute used for its training being greater than 10^{25} floating point operations, (Art. 52a) are required to be notified to the Commission by their providers. They need to undergo additional measures, such as adversarial testing, risk assessment, incident tracking, and ensuring cybersecurity protection (Art. 52d).

2.2 State of the art ML/AI lifecycles and software development lifecycles (SDLCs)

In order to derive requirements for the design of the problem's solution, the state of the art regarding established SDLCs, AI/ML lifecycles and process models in data mining projects was reviewed. Through key term search as well as reference mining of relevant sources, the following models were identified as a suitable base for this work.

First of all, the CRISP-DM (Cross Industry Standard for Data Mining), may be seen as the most common industry-independent methodology for data mining, analytics, and data science projects even more than twenty years after its release. [SKM21] It is divided into six sequential phases, starting with business understanding, followed by data understanding, data preparation, modelling, evaluation and finishing of with the deployment phase. [Ch00; WH00] During the literature review it has been identified, that the development of ML applications is mainly based on these phases. However, as the CRISP-DM is limited in managing requirements of current technologies [ADP18], further models were taken into account. More flexible methodologies have been introduced by big IT companies [Am19], for instance the TDSP (Team Data Science Process) by Microsoft, which has a more agile and iterative approach in order to deliver predictive analytics solutions. [Ma21] Continuous integration (CI) and continuous delivery (CD) paradigm is the current industry standard for automating and streamlining the process of building, testing and deploying code changes. [La22; KAA20] Thus, the integration of DevOps as one of the mostly used agile processes is integrated into the solution delivered by this paper. The DevOps process focuses on collaboration between development and operation teams and includes the steps: plan, code, build, test, release, deploy, operate, monitor and feedback. [KAA20; LCB20]

3 Methodology

The research objective has been elaborated with the Design Science Research (DSR) approach, which is a research paradigm that focuses on the development and validation of prescriptive knowledge in an iterative way. [Pe06] Due to its applicability to application-oriented problem solving in the intersection of IT and organization [Pe06], it is suitable for the research question of this work. The DSR process includes six steps in order to develop and evaluate a so-called artefact. [Pe07]

Problem identification and motivation

As already pointed out, the requirements of the AI Act are considered as extensive, especially for high-risk AI systems. Also, the formulation of these requirements in the act is rather theoretical. Thus, there is a need to investigate the explicit impact of the upcoming legislation across the lifecycle of a high-risk AI system.

Objectives of the solution

The solution process requires the generation of an artefact, which in this case represents the lifecycle of these systems. It must include the functional requirement of assuring compliance to the AI Act whilst developing and operating the system efficiently. As the features of the state of the art SDLCs and ML lifecycles are well established in practice, they were considered as structural requirements for the developed artefact. This includes the AI Act compliant lifecycle to be divided into specific phases with belonging activities, to ensure an agile approach through feedback loops between phases and to be application-neutral.

Design and development of the artefact

Based on the objectives, the artefact has been developed combining the deductive and inductive approach for creating a reference model. [Fe14] Therefore, constructs from generally accepted frameworks, already presented in 2.2, have been adopted to map the necessary process steps across the lifecycle of an AI system. Afterwards, the requirements of the AI Act for high-risk AI systems were extracted from the latest draft of the act and implemented into this lifecycle in order to ensure the AI systems compliance to the regulation.

The complete lifecycle is presented within chapter 4. However, an example of how the AI Acts requirements have been mapped onto the lifecycle is given at this point for reference. Therefore, the requirement of Data and Data Governance derived from Art. 10 of the AI Act has been chosen: Mapped mainly in the Data Management and Machine Learning Workflow sections, the requirements for robust data governance practices are placed where data-related processes are dominant. This includes data acquisition, preprocessing, and model training stages. Ensuring data relevance, representativeness, freedom from errors, completeness, and addressing biases are crucial during these stages to build reliable and fair AI systems. Processing personal data, the compliance with GDPR must be ensured. It was assumed, that data preprocessing plays a crucial role in achieving a model with a performance as high as possible. Recently the realisation arose, that in many AI projects, the leverage for improving model performance lies in the curation of the training data used, which is also known as data-centric AI. [Ja24] Hence, the data and data governance requirements were not mapped as an additional activity, as the text in the law is rather vague and refers to the provider acting appropriately regarding the intended purpose. Furthermore, bias mitigation measures must take place during the operation of the system, when the model is retrained with new data. The mapping here emphasizes the need for thorough handling and scrutiny of data, which is fundamental to the training and validation of AI models. The underlying master thesis of this paper points out how the mapping of the remaining requirements has been conducted.

Through the back iteration by the activity “demonstration and evaluation” the deductively developed generic lifecycle was inductively revised processing the empiric findings from this phase.

Demonstration and evaluation

In this work the steps demonstration and evaluation have been combined as the underlying regulatory framework was not in force at the time of this work and the model may not have been directly demonstrated in specific use cases or experiments. However, with the growing popularity of AI, the need of AI governance has found its way into science as well as into the industry, even before the upcoming legislation. [La22] Therefore, interviews have been conducted with experts from the field of AI development evaluating the developed artefact on the one hand and on the other hand gaining valuable insights into AI governance practices the companies already have in place. Six interviewees representing both companies with more than 200.000 employees and start-ups counting less than 50 employees have been selected. This ensures, that potential biases regarding the subjective perception of the AI Acts impact on the organisation were identified and balanced. They were selected from the interviewer's environment and through referrals from previous interviewees. As the topic requires both expertise in AI development and familiarity with the regulation, the profiles of the informants differ in order to cover both areas.

Qualitative data was collected through semi-structured interviews in order to adapt the questions, which arose during the interview. [Mi19] The developed framework of the AI Act compliant lifecycle is structured into six major phases, that serve as thematic complexes [Mi19], each of which was presented individually and feedback was then obtained from the interviewees. As semi structured interviews are characterized by a certain openness, the analysis requires a method that makes it possible to address both predetermined and spontaneous themes and statements. [BLM14] The qualitative content analysis allows an open exploration of the data without having fixed hypotheses or assumptions in advance and is therefore a suitable method in achieving the goal of processing the valuable insights of the interviewees. [BLM14] The transcripts were categorised by marking meaningful statements, segmenting the data into significant units and classifying them into categories. [GL10; Mi19; Ma15] While the interviews were conducted according to the structure of the six major phases of the developed lifecycle, the initial categorisation of the data followed this structure as well. During the encoding of the data, subcategories under each of the lifecycle phases arose iteratively. [GL10]

The findings regarding the second phase of the lifecycle are allocated for reference in the following and are illustrated by Figure 1 Analog to the mapping of the AI Acts' requirements, the complete data is analysed in detail in the underlying thesis, which also provides the transcripts of the interviews. The informants brought up that the availability of suitable data resources is a key prerequisite for ML and AI projects. However, as this phase requires a relatively high effort one of the interviewees mentioned, that they conduct a proof-of-concept prior running through all the data management process steps. The informant described their practice as follows: "Simply have the data made available to us, as large data sets as possible, so that we can throw them into [...] different models to see if we get meaningful results." (Interviewee 1) All of the experts addressed the topic of bias in the data. Bias mitigation measures should be explicitly included in the

lifecycle. One interview partner provided insights on how their practices look like: “we do this on a correlation basis. What we analyse is how the performance correlates with different parameters of the model. And then we try to correct those biases in the model performance.” (Interviewee 6) At the same time, it was emphasised that the database must be as transparent as possible, even when it comes to mitigating bias. The negative example of Google’s AI model ‘Gemini’ was mentioned, which failed in generating historically accurate pictures of people in an attempt at racial and gender diversity. In terms of data security, the importance of data ownership and security of data must be considered and a reference to GDPR could be included into the lifecycle, as this regulation applies at this point and must be complied to especially when processing personal data.

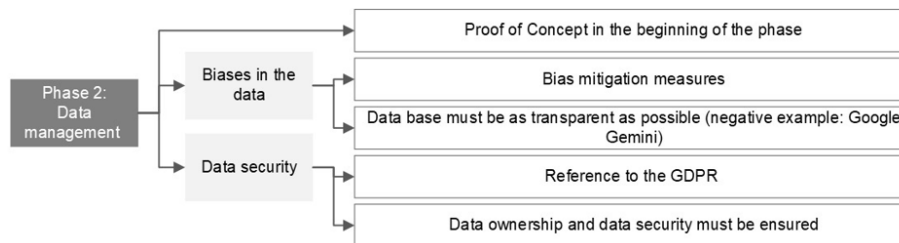


Figure 1: Results of the data analysis categorising the interview data (white boxes) according to the lifecycle phase (grey box)

With the findings from the interviews the artefacts quality was improved by iterating back to the design and development phase. [Pe07]

Communication

The problems’ relevance and its solution in the form of an AI Act compliant lifecycle for AI Systems is published through this paper and its underlying master thesis.

4 Development of an AI Act compliant lifecycle for AI systems

In the following sections, the reference model of an AI Act compliant lifecycle for AI systems is presented. It consists of five phases: business understanding (4.1), data management (4.2), ML/AI modelling (4.3), application development (4.4), deployment and operation (4.5). After each phase is explained individually, their interactions and phase agnostic parts of the lifecycle are discussed (4.6). The model language Business Process Modelling Notation (BPMN) was chosen, as it features comprehensibility for all relevant stakeholders with different backgrounds. [Ke13]

4.1 Phase 1: business understanding

The business understanding phase, illustrated in Figure 2, is initiated by the definition of the use case, followed by identifying and analysing stakeholder needs, defining the targets and assessing whether or not and how ML or AI technologies is suitable for the use case. [Am19] These activities may be run through iteratively. However, at some point it must be checked, if the system will be in scope of the AI Act and subject to its obligations. Therefore, the regional scope of the EU market must be given. Furthermore, the acts' definition of an AI system must include the use case (Art. 3(1)). If the system is not in scope of the act, it does not need to be compliant and is not in scope of this work. As the act follows a risk-based approach, a risk classification needs to be conducted to find out, which obligations of the act apply. Firstly, it needs to be clarified, whether the system is in scope of the GPAI models definition (Art. 2(44b)) and it must be differentiated between GPAI models posing a systemic risk (Art. 52a) and those that do not. However, neither the development of GPAI systems nor their use is in scope of this work.

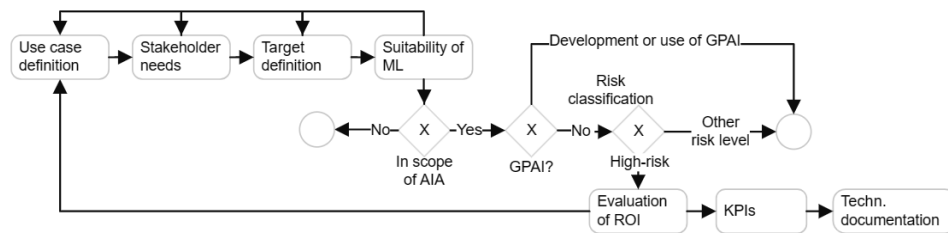


Figure 2: Phase 1 of the lifecycle - Business understanding

Subsequently, the risk classification is performed, which initiates by confirming, that the system will not pose an unacceptable risk and will therefore be prohibited (Art. 5). It is verified if the system is falling under one of the two possible paths of being classified as a high-risk system (Art. 6). Firstly, it is investigated whether the system falls under the Union Harmonisation Legislation listed in Annex II. Fulfilling this first condition, it is checked if a third-party conformity assessment is necessary under the specific legislation (Art. 6(1)). With these two checks being positive, the system is classified as high-risk. In cases, where the first condition is not fulfilled, it needs to be checked, if the system belongs to the use cases listed in Annex III (Art. 6(2)). Additionally, there must be a significant risk (Art. 6(2a)) in order that the system will become a high-risk system. Next, the costs of the system including costs of compliance and operating costs to estimate the return on invest (ROI) need to be evaluated. At this point, back-iteration is also possible, as the cost estimations could indicate non-profitability of the use case. After defining key performance indicators (KPIs) of the system, the first part of the technical documentation required by the AI Act (Art. 9) is created. This includes information regarding: functions, purposes, target groups, application contexts and performance features of the system.

4.2 Phase 2: data management

As data preprocessing is crucial for the efficiency and accuracy of training AI or ML models, and due its complexity and time consuming nature [Ta24; Kr16; BB01], a proof of concept may demonstrate the feasibility of the use case before starting with the activities in this phase, visualized in Figure 3. During the data acquisition, internal or external data sources are identified and the data is collected. [Ch00] Initial explorative analyses help in understanding the data and discover possible problems. [Ch00; SM21] Depending on the use case the AI Act provides exceptions, which permit the usage of particular categories of personal data, for instance to eliminate bias (Art. 10(5)). Therefore, it needs to be checked if such data is acquired to ensure compliance with the General Data Protection Regulation (GDPR). Before the data preparation, quality requirements must be defined, which the data sets need to fulfil. The data preparation step includes multiple activities, which are highly dependent on the use case and on the model that they should be trained on. Therefore, the entire second phase is also strongly connected to the third phase (see 4.3), where the modelling takes place. However, data cleansing, identification and elimination of bias, data labelling, feature selection and data splitting into training, test and validation data sets are necessary steps during data preparation with this list not being complete. [Am19; SM21] Providers are obliged to establish “data governance and management practices appropriate for the intended purpose of the AI system” (Art. 10). Data quality is controlled in accordance with the predefined quality requirements and in case of failure, back iteration to data preparation and data acquisition is possible. If the data quality requirements have been met, data versioning enables to track changes made to the data over time. Finally, the technical documentation regarding data sources and processing, description of the data and data governance practices is created (Art. 11).

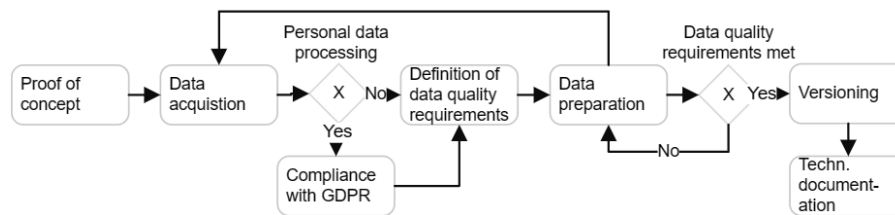


Figure 3: Phase 2 of the lifecycle - Data management

4.3 Phase 3: ML/AI modelling

This phase, featured in Figure 4, starts by the selection of ML/AI model techniques, which need to fit to the problem statement of the use case. [LCB20] Additionally, this activity also has an impact on making the underlying model of the system explainable. [Pa23] It can be summarized, that the act focuses in a holistic manner on ensuring an appropriate use of the system through a transparency requirement being achieved by the

provision of relevant documentation (Art. 13). However, techniques for explainable AI (XAI) are not explicitly addressed by the act but may facilitate human oversight (Art. 14), and the transparency requirement (Art. 13). [Pa23; Ar20] Therefore, at this point of the lifecycle the model could be developed interpretable, so that inner workings are available and understandable to humans. This could be achieved by a Transparency by Design approach [Fe20], meaning, amongst others, the avoidance of black box models by rather choosing model types, of which the decision-making process is replicable and easy to step through, e.g. decision trees and linear and case-based reasoning models. [Pa23; Fe20] Depending on the use case and the required model type, this is not always possible though.

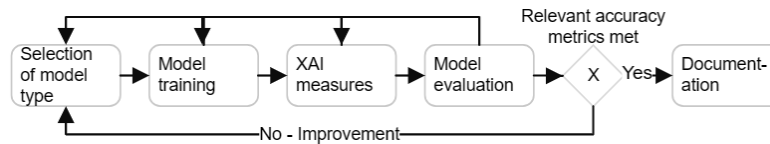


Figure 4: Phase 3 of the lifecycle – ML/AI modelling

The next activity is the training of the model with the training data set including the tuning of hyperparameters to achieve the highest possible output on the validation data set. After the training, the XAI approach may provide techniques in order to make the decision-making process explainable to humans overseeing the system. [Pa23] In contrast to the interpretable AI approach, this approach includes techniques, that may be model agnostic and allow explanation for black-box models. [Pa23] However, this so-called post-hoc explainability comes with its challenges and limitations and is subject to complex research [Pa23; Ar20]. During the model evaluation, the performance using the test data set is measured by testing against metrics such as accuracy, precision, recall and F-score. [LCB20; SFR23] As an “appropriate level of accuracy, robustness, and cybersecurity” of the system is required (Art. 15), the testing against the accuracy metrics needs to be included in the model evaluation. Further information on benchmarks and measurement methodologies addressing these metrics, are announced to be developed in cooperation with relevant stakeholders and organisations but are not yet available (Art. 15). Based on identified deficiencies and patterns, the model is improved for example by adapting feature engineering, parameters or the model selection. [SFR23] This highly iterative process may be stepped through multiple times until the model evaluation is satisfying and is then concluded by the documentation of this phase. [Am19; Ch00; SM21]

4.4 Phase 4: application development

The ML/AI model developed and trained in the previous phase is deployed and somehow serves an application or system, the development of which takes place in this phase, see Figure 5. [SM21]

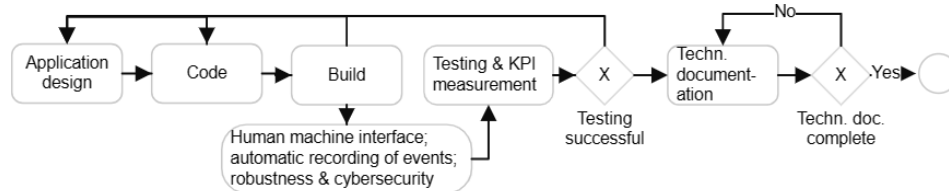


Figure 5: Phase 4 of the lifecycle - Application development

Even though, the use case has already been defined previously and designs may exist through conducting the proof of concept, the application design activity initiates this phase in order to ensure an agile way of working. During the coding and build activities, the functionality of the software is created and functionalities of all components should be orchestrated in such a way that the trained model operates correctly with the rest of the system. [LCB20; SFR23] Within these activities, various requirements of the AI Act must be implemented. Ensuring the system can be overseen at any time during its use by a natural person, who can also intervene to prevent and minimize risks, human oversight measures need to be implemented (Art. 14) Either these measures are built in before placement on the market or it is guaranteed, that the implementation is appropriate to be done by the user. Furthermore, the automatic recording of events (Art. 12) is crucial in order to facilitate post-market monitoring and to identify potential risks. Therefore, especially for high-risk systems, which are under the use cases of Annex III paragraph 1, point (a), the logs should include time of the systems use, input data as well as the database against which these data has been checked and, in some cases, even further information. As already mentioned, accuracy, cybersecurity, resilience and robustness are also requirements to be considered and appropriate measures need to be implemented (Art. 15). Robustness in AI systems means, that they are resilient against adversarial attacks and that they are able to withstand perturbations in input data and model parameter, so that the likelihood of systems failure is reduced [OE19, Fe24]. Suitable measures for achieving robustness depend on the type of the model and system and need to be comprehensive. Literature describes solutions such as adversarial training, which involves training neural networks against adversarial examples [Fe24; Ma17]. However, navigating the techniques and strategies of robustness may be seen as a research topic in itself, but it should be mentioned, that at this point in the lifecycle, it may be necessary to iterate back to the modelling phase (4.3) in order to be able to implement appropriate measures accordingly. Prior to the testing of the system, planning of the tests and definition of test cases must take place in order to determine the systems functionality effectively [FK23], for instance, by conducting integration tests and acceptance testing [FK23]. With regard to software testing, it is distinguished between manual and automated tests. Automated tests are an important part of CI/CD and may be integrated into the agile DevOps workflow [FK23], though, the testing process and tools need to be appropriate for the specific use case. Analogous to the accuracy (chapter 4.3), benchmarks and measurement methodologies for robustness levels are being developed by the Commission and will be decisive for the testing activity soon (Art. 15). With

finalizing the development of the application, several parts of the technical documentation need to be completed (Art. 11), including information regarding components required by the AI Act, that were implemented in this phase. A detailed checklist of the information required in the technical documentation is provided in the work on which this paper is based, hence why it is not discussed in detail here. Furthermore, instructions for use (Art. 13) need to be created with the purpose of not only serving the user of the system but also ensure compliance across the value chain.

4.5 Phase 5: deployment and operation of the system

Before placement on the market the high-risk AI system must undergo a conformity assessment (Art. 43). Depending on the use case of the system, it is distinguished between different types of conformity assessments, which are summarized in Table 1. As Figure 6 illustrates, the CE marking (Art. 49) is affixed and an EU declaration of conformity (Art. 48) must declare the conformity with Chapter 2 of the act, which comprises all the AI Act requirements mapped across the lifecycle. Systems that are within the use cases of Annex III (see 2.1) must be registered in an EU database (Art. 51). High-risk systems with the exception of use cases in scope of critical infrastructure (Annex III point 2) need to be registered at national level. Finally, the system is placed on the market. At this point responsibilities are transferred from the provider to the roles of importers, distributors, and deployers. The presented lifecycle pays primary attention on the obligations of providers. In the underlying work of this paper requirements for the other roles are outlined to make the lifecycle applicable for further participants across the value chain.

Type of system (use case)	Type of conformity assessment
<u>1.</u> : Annex III points 2-8; point 1 (only if harmonised standards or common specifications are applied (Art. 40&41))	Internal conformity assessment (Annex VI)
<u>2.</u> : Annex III point 1 use cases, where no/only parts of harmonised standards or specifications are available/were applied	Conformity assessment with the involvement of a notified body (Annex VII)
<u>3.</u> : System under Union harmonisation legislation (Annex II) part A	Conformity assessment under specific act with inclusion of Chapter 2 (AI Act) requirements and Annex VII, point 4.3, 4.4, 4.5, and 4.6§5

Table 1: Types of conformity assessments under the AI Act

While operating the system, recording of the logs should take place, which are essential for model monitoring and risk management. Also, generated data and collection of feedback is processed by iterating back into previous phases through feedback loops, which is further discussed in 4.6.

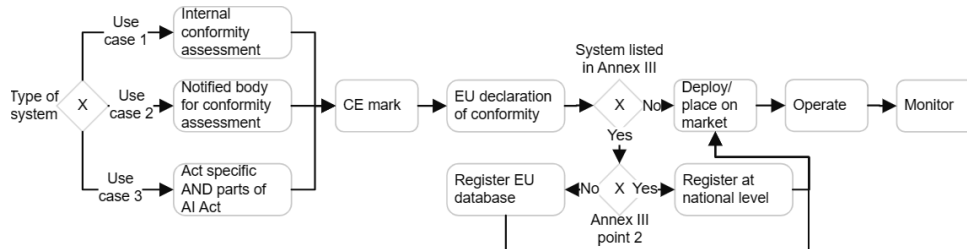


Figure 6: Phase 5 of the lifecycle - Deployment and operation

4.6 General parts of the lifecycle

Interactions in every direction between the single phases of the lifecycle ensure proper transfer of data, other components of the system and feedback. [LCB20; SFR23] This is visualised in Figure 7. It fosters an agile working and development style, which is state of the art, as pointed out in chapter 2.2. If, for instance, data is generated during the operation of the system (phase 5), which is used for model training (phase 3) in order to adapt the model to changing conditions, the data needs to go through various steps of the data management (phase 3). In this case measures against possible bias must be taken (Art. 15). Depending on the impact of the feedback and potential change while iterating backwards, obviously some process steps may be redundant and can be skipped. However, it should always be considered, how a change may influence other components of the system and which further adaptations might be required, especially with regard to compliance with the AI Act. Across the entire lifecycle of the high-risk AI system a risk management system is being operated (Art. 9). It forms the backbone of requirements mentioned above. [PZ24] Especially in terms of the obligations regarding transparency, data governance, accuracy, robustness and cybersecurity, there may be competing priorities, which force the provider to make compromises in reaching the highest possible levels in each of them [Pa23]. Therefore, providers must analyse potential risks, that are reasonable and foreseeable as well as ones, that may arise based on the post market monitoring data (Art. 9). Subsequently, risk management measures are being developed in accordance with a three step-approach of “eliminate or reduce”, “mitigate and control”, and “inform” (Art. 9(4)). The European Committee for Standardization (CEN) and the European Committee for Electrotechnical Standardization (CENELEC) do currently work on developing standards [Eu23b] in order to provide practical guidance for manufacturers by translating the regulation into actionable steps. [PZ24] However, as an already existing standard the ISO/IEC 23894 serves as a guidance on managing risks of AI systems [PZ24] and the National Institute of Standards and Technology (U.S) (NIST) released an Artificial Intelligence Risk Management Framework (AI RMF 1.0). Following this framework, the risk management is split into four phases “govern”, “map”, “measure” and “manage”, each of which entailing specific actions and outcomes. [Ta23]

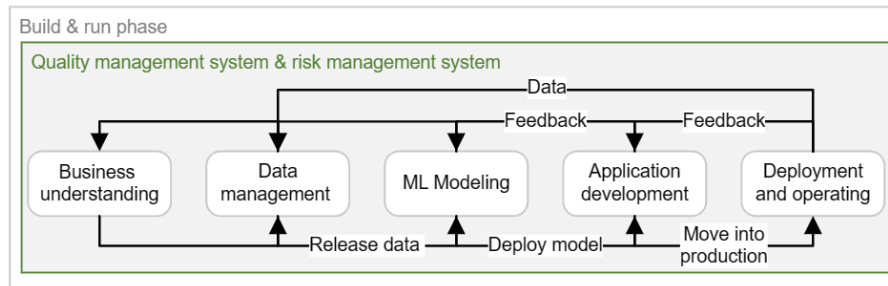


Figure 7: High-level view of the lifecycle

5 Limitations and conclusion

In this paper the impact of the AI Act on the lifecycle of high-risk AI system was analysed by developing a reference model, which is application and sector agnostic. It was focused in particular on the requirements set out in Title III, Chapter 2 of the AI Act affecting mainly the provider of the system. Starting the development of the model from existing state of the art models for ML/AI lifecycle and software development, it could be found, that the existing lifecycles must be revised in order to ensure compliance with the regulation. Although, providers and further participants of the value chain do have up to 2 years of time until the regulation becomes fully applicable, some actions within in the presented lifecycle may already be taken now in order to be prepared, for instance to manage risks of the AI system. Conducting the interviews with experts from the field, it was noticeable, that there are mixed opinions on the impact of the AI Act especially regarding competitiveness with international competitors not being affected by the EU regulation and workload, which may impact time-to-market negatively. Further research could therefore analyse the gap between start-ups and larger companies. However, in order to do this, the development of standards and more concrete interpretations of the AI Act may be necessary. This also had a limiting effect on this work and may be balanced by iterating the developed reference model again as soon as more tangible information concerning the AI Act do exist. Furthermore, deep dives in each of the lifecycle's steps are possible and necessary, as the existing model focuses on the whole lifecycle and is scraping the surface of further profound research topics. The decommissioning phase of the high-risk AI system is also not considered within this work and could be part of further research.

In conclusion, the insights presented in this paper may foster future discussion on the implementation of the AI Acts requirements into the lifecycle of AI systems, leading to the creation of best practices for ensuring compliance with the regulation.

6 Bibliography








- [ADP18] Ahmed, B.; Dannhauser, T.; Philip, N.: A Lean Design Thinking Methodology (LDTM) for Machine Learning and Modern Data Projects: 2018 10th Computer Science and Electronic Engineering (CEECE), pp. 11–14, 2018.
- [Am19] Amershi, S. et al.: Software Engineering for Machine Learning: A Case Study: 2019 IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP). IEEE, pp. 291–300, 2019.
- [An23] Anger, H. et al.: Warum KI-Unternehmen strengere Gesetze fordern und die Politik nicht liefert. Handelsblatt, 2023.
- [Ar20] Arrieta, A. B. et al.: Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion* 58, pp. 82–115, 2020.
- [BB01] Banko, M.; Brill, E.: Scaling to very very large corpora for natural language disambiguation: Proceedings of the 39th Annual Meeting on Association for Computational Linguistics. Association for Computational Linguistics, USA, pp. 26–33, 2001.
- [BLM14] Bogner, A.; Littig, B.; Menz, W.: Interviews mit Experten. Eine praxisorientierte Einführung. Springer Fachmedien Wiesbaden, pp.71-75, 2014.
- [Ch00] Chapman, P. et al.: CRISP-DM 1.0 Step-by-step data mining guide, 2000.
- [Eu23a] Artificial Intelligence Act: deal on comprehensive rules for trustworthy AI, 2023.
- [Eu23b] European Commission: Implementing Decision on a standardisation request to the European Committee for Standardisation and the European Committee for Electrotechnical Standardisation in support of Union policy on artificial intelligence. C(2023)3215, 2023.
- [Eu24] European Commission: European AI Office, 2024.
- [Fe14] Fettke, P.: Eine Methode zur induktiven Entwicklung von Referenzmodellen. In (Kundisch, D. Ed.): Tagungsband Multikonferenz Wirtschaftsinformatik 2014 (MKWI 2014). 26.-28. Februar 2014 in Paderborn. Univ, Paderborn, 2014.
- [Fe20] Felzmann, H. et al.: Towards Transparency by Design for Artificial Intelligence. *Science and Engineering Ethics* 26, 2020.
- [Fe24] Ferrara, E.: The Butterfly Effect in artificial intelligence systems: Implications for AI bias and fairness. *Machine Learning with Applications* 15, p. 100525, 2024.
- [FK23] Forgács, I.; Kovács, A.: Modern software testing techniques. A practical guide for developers and testers. Apress, New York, NY, 2023.
- [GL10] Gläser, J.; Laudel, G.: Experteninterviews und qualitative Inhaltsanalyse. Lehrbuch. VS Verlag, Wiesbaden, pp. 198-215, 2010.

- [Ja24] Jakubik, J. et al.: Data-Centric Artificial Intelligence. Business & Information Systems Engineering, 2024.
- [KAA20] Karamitsos, I.; Albarhami, S.; Apostolopoulos, C.: Applying DevOps Practices of Continuous Automation for Machine Learning. Information 7/11, p. 363, 2020.
- [Ke13] Kelemen, Z. D. et al.: Selecting a Process Modeling Language for Process Based Unification of Multiple Standards and Models, 2013.
- [KNS23] Klöckner, J.; Neuerer, D.; Scheppe, M.: Unternehmen wollen mit KI radikal umsteuern. Handelsblatt, 2023.
- [Kr16] Krishnan, S. et al.: ActiveClean: An Interactive Data Cleaning Framework For Modern Machine Learning: Proceedings of the 2016 International Conference on Management of Data. Association for Computing Machinery, New York, NY, USA, pp. 2117–2120, 2016.
- [KS23] Koch, M.; Scheuer, S.: Der Frankenstein-Moment: Wenn wir Künstliche Intelligenz nicht kontrollieren, kontrolliert sie uns. Handelsblatt, 2023.
- [La22] Laato, S. et al.: AI governance in the system development life cycle: Crnkovic (Hg.) 2022 – Proceedings of the 1st International Conference on AI Engineering: Software Engineering for AI, pp. 113–123.
- [LCB20] Lwakatare, L. E.; Crnkovic, I.; Bosch, J.: DevOps for AI – Challenges in Development of AI-enabled Applications: 2020 International Conference on Software, Telecommunications and Computer Networks (SoftCOM). IEEE, pp. 1–6, 2020.
- [Ma15] Mayring, P.: Qualitative Inhaltsanalyse. Grundlagen und Techniken. Beltz, Weinheim, Basel, 2015.
- [Ma17] Madry, A. et al.: Towards Deep Learning Models Resistant to Adversarial Attacks, 2017.
- [Ma21] Martínez-Plumed, F. et al.: CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories. IEEE Transactions on Knowledge and Data Engineering 8/33, pp. 3048–3061, 2021.
- [Mi19] Misoch, S.: Qualitative Interviews. De Gruyter Oldenbourg, Berlin, Boston, 2019.
- [OE19] OECD: Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/00449, 2019.
- [Pa23] Panigutti, C. et al.: The role of explainable AI in the context of the AI Act: 2023 ACM Conference on Fairness, Accountability, and Transparency. ACM, New York, NY, USA, pp. 1139–1150, 2023.
- [Pe06] Peffers, K. et al.: The design science research process: A model for producing and presenting information systems research. Proceedings of First International Conference on Design Science Research in Information Systems and Technology DESRIST, 2006.
- [Pe07] Peffers, K. et al.: A Design Science Research Methodology for Information Systems Research. Journal of Management Information Systems 3/24, pp. 45–77, 2007.
- [PZ24] Pouget, H.; Zuhdi, R.: AI and Product Safety Standards Under the EU AI Act, CARNEGIE Endowment for International Peace, 2024.

- [SFR23] Steidl, M.; Felderer, M.; Ramler, R.: The pipeline for the continuous development of artificial intelligence models—Current state of research and practice. *Journal of Systems and Software* 199, p. 111615, 2023.
- [SKM21] Schröer, C.; Kruse, F.; Marx Gómez, J.: A Systematic Literature Review on Applying CRISP-DM Process Model. *Procedia Computer Science* 181, pp. 526–534, 2021.
- [SM21] Schreckenberger, F.; Moroff, N. U.: Developing a maturity-based workflow for the implementation of ML-applications using the example of a demand forecast. *Procedia Manufacturing* 54, pp. 31–38, 2021.
- [Ta23] Tabassi, E.: Artificial Intelligence Risk Management Framework (AI RMF 1.0), NIST Trustworthy and Responsible AI, National Institute of Standards and Technology (U.S.), Gaithersburg, MD, 2023.
- [Ta24] Tawakuli, A. et al.: Time-series data preprocessing: A survey and an empirical analysis. *Journal of Engineering Research*, 2024.
- [Tr24] Triguero, I. et al.: General Purpose Artificial Intelligence Systems (GPAIS): Properties, definition, taxonomy, societal implications and responsible governance. *Information Fusion* 103, p. 102135, 2024.
- [Vo23] Volkery, C. et al.: EU beschließt umfangreichstes KI-Gesetz der Welt – das sind die wichtigsten Punkte. *Handelsblatt*, 2023.
- [WH00] Wirth, R.; Hipp, J. Eds.: CRISP-DM: Towards a Standard Process Model for Data Mining, 2000.
- European Parliament: PROVISIONAL AGREEMENT RESULTING FROM INTERINSTITUTIONAL NEGOTIATIONS Subject: Proposal for a regulation laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts 2021/0106(COD) (COM(2021)0206 – C9-0146(2021) – 2021/0106(COD)); 02.02.2024

Digital Classroom – A Communication Platform for Online Teaching

Full Paper








Sören Aguirre Reid ¹, Frank Kammer ¹, Dominik Kröll ¹, Timon Pellekooorne ¹,
Harald Ritz ¹, Markus Siepermann ¹, and Jonas Wölfer ¹

Abstract: In an increasingly digitally influenced teaching and with the desire to work together despite spatial separation, it is becoming more and more important that communication can take place quickly and flexibly between all those involved in teaching and learning. Unfortunately, the opportunities of on-site tutorials, in particular quick changes between group communication and global communication as well as organization issues between lecturer and tutors are poorly represented by existing video conferencing solutions. In this paper we use the design science approach to develop the concept of a digital classroom (DC), a design artifact that aims to map the requirements of academic group work and group discussion into the online context while taking advantage of the online setting.

Keywords: Online Teaching, Communication Requirements, Digital Classroom, Ticket System, Design Science

1 Introduction

One of the main problems in distance learning and online teaching is that students miss the interaction and connection with other students and lecturers and therefore feel disconnected [Be15]. Online group discussions and group work are two learning approaches that seem suitable to cope with this problem and can help to reconnect students. They even foster students' communication and collaboration skills, which are highly valued by firms [A111; Ro13]. The first step in group discussion involves students discussing a problem and possible solutions together. The second step, group work, entails solving the problem based on the approaches discussed. Moreover, these two teaching approaches have a very positive effect on understanding and knowledge acquisition [De19; El18; HA13; ML07; RAC15] due to higher learning motivation and engagement [Bö21; GOA22; Wu13] and self-efficacy [Mu21a]. However, group discussions and group work can only be successful when several

¹ Technische Hochschule Mittelhessen, Fachbereich Mathematik, Naturwissenschaften und Informatik, Wiesenstraße 14, 35390 Gießen, soeren.aguirre.reid@mni.thm.de,  <https://orcid.org/0000-0002-3229-423X>; frank.kammer@mni.thm.de,  <https://orcid.org/0000-0002-2662-3471>; dominik.kroell@mni.thm.de,  <https://orcid.org/0009-0009-0593-042X>; timon.pellekooorne@mni.thm.de,  <https://orcid.org/0009-0008-7069-5932>; harald.ritz@mni.thm.de,  <https://orcid.org/0009-0002-1916-7364>; markus.siepermann@mni.thm.de,  <https://orcid.org/0000-0002-5626-9725>; jonas.woelfer@mni.thm.de,  <https://orcid.org/0000-0002-6502-6960>

requirements are met. First of all, timely feedback is crucial for the acceptance and success of online teaching and learning. Otherwise, students easily become frustrated because of response delays when posing questions or exchanging ideas [Ha00; Ke10; LC19; Ro13].

The support provided by lecturers strongly influences the students' learning experience and their perception of the learning effectiveness in a positive way [Bu22]. In addition, the feedback improves students' understanding and deepens their knowledge [Ki14; LHC17].

Although many students think that group discussions should take place in offline settings [Ch21], students favor online group work over the offline setting, as in the online setting it is easier to exchange ideas with other students than the ones of their direct peer group [GH22]. Despite students' readiness for online learning post-pandemic, existing technical tools such as Zoom and Adobe Connect fall short for group work and discussions due to their lack of specific functionalities, such as flexible tutor coordination, quick group switching, and effective document sharing. Therefore, there is a critical need for a concept that adequately supports these teaching methods online, which this paper aims to address.

RQ: How can academic peer group discussions and group work be transferred into the online context?

2 Method

To answer our research question, we used the design science research approach (DSR) introduced by Hevner et al. [He04]. For this, we orientated to the DSR processes of Conboy et al. [CGC15] and Peffers et al. [Pe12] as well as to the recommendations given by Gregor; Hevner [GH13]. Hence, the remainder of this paper is organized as follows: After the problem identification in this section, we review the related literature in the field of communication systems for group work and group discussion in Section 3. In Section 4, we identify the characteristics of the teaching and learning methods group work and group discussion. Based on these characteristics and the literature review, we then derive requirements for digital tools that aim at supporting group work and group discussion in online settings. Subsequently, the concept of an online platform for group work and group discussion is developed with the help of these requirements. To evaluate the concept, a prototype was developed and used in a specific setting for assessment purposes. The evaluation process included two main approaches: qualitative feedback from users, such as lecturers and students, which was integrated into the prototype in real-time and is not extensively reported in this document, and a quantitative survey conducted to gauge the perceptions of students, lecturers, and tutors regarding the prototype. The paper presents the initial findings from the quantitative survey. The document concludes with a discussion on these results, the implications they suggest for the future, and the necessary steps to further refine the prototype and concept.

3 Related Work

For the development of our artifact, we conducted a literature review to identify its key aspects. Our literature review is based on the PRISMA approach of Page et al. [Pa21]. Before we started the literature search, we analysed existing literature reviews in the research field ([AG15; Mu21b; NC19]). With the help of the two literature analyses, we developed our keywords and were able to identify suitable databases (IEEExplore, Sciencedirect, ACM) for the search. We used the following search strings between 1990 and 2023: “virtual teams” AND (“digital communication” OR “distance education” OR “learning communities” OR “group work” OR “group collaboration” OR “digital classroom”) to identify relevant publications. Using the keywords, we were able to identify 608 papers. For this reason, we have added a criterion for the literature screening process: We excluded all studies that did not examine computer-supported media communication in group work within a conference system. By adding this criterion, we reduced the literature to 12 studies. In addition, we applied a backward approach and were thus able to identify three further relevant studies.

Step	Search and filtering criteria	IEEExplore	Sciencedirect	ACM	Total
1	Search string	19	173	132	324
2	Filtering by computer-supported media communication in group work within a conference systems	0	9	3	12
3	Apply a backward approach				15

Tab. 1: Search results

The first five publications investigated group work communication with asynchronous text based communication. The first study by Hollingshead et al. [HMO93] explored the asynchron computer-mediated communication via forum and group messenger and task type on group task performance with two theoretical models. The findings suggest face-to-face groups performed better on intellectual tasks. The study of Potter; Balthazard [PB02] compared virtual teams and conventional teams with regard to their interaction style (e.g., constructive, passive), decision performance (e.g., solution acceptance, satisfaction) and process outcomes. The students were able to communicate with each other via a forum. The results show that the interaction styles of virtual teams are consistent with those of conventional face-to-face groups. The study by Geister et al. [GKH06] investigated the effects of process feedback on motivation, satisfaction and performance in virtual teams. To do this, students were divided into several teams to develop two solutions to the problem. These teams were then randomly assigned to receive online feedback or non-online feedback. The students of this study certified that they had communicated exclusively by electronic communication tools (e.g., e-mail) and had no face-to-face contact. Virtual teams that used the online feedback system showed an increase in performance compared to the control group. Oliveira et al. [OTP11] focused on online group work patterns. The lecturer divided the participants into eight groups (with 4–5 participants each), and the participants had to develop a project that would be presented to the whole class. The results show that the successful groups had a higher level of communication (e.g., participants actively

contributed with feedback and elaborations). The study of Smith et al. [Sm11] analyzed overcoming student resistance to group work in the context of online versus face-to-face group work. The communication in the online group work context was based on a text-based forum. The results revealed that students in the online group were less satisfied than on-site group because of fewer communication channels and the lack of immediacy of face-to-face meetings. The publications show that communication between students also takes place in a similar way in digital media and forms the basis for successful learning. However, electronic means of communication cannot prevail over face-to-face events.

Five further studies therefore addressed communication between group members (e.g. students) using a synchronous approach. The first study by Walther [Wa92] examined the effects of time and communication channels among computer-mediated (computer conferencing system with a microphone to communicate) versus face-to-face groups. The results showed that computer-mediated groups approach the level of face-to-face groups regarding several dimensions (e.g., formality). Jang et al. [JSP02] investigated a web-based collaborative system called TeamSCOPE to support the awareness needs of virtual teams. This was evaluated with a survey. All teams appreciated the shared file space, but no information was reported how students perceived the integrated live chat function. The study by Robinson [Ro13] analyzed the interrelationship of emotion and cognition when students undertake collaborative group work online. The study design requires students to collaborate online, in small groups of 4–8 students by using text-based forums and live chats. The study revealed that students who engage in online group work should adopt verbal immediacy behaviors so their peers can get to know and trust them [Ro13]. The second study utilized the Adobe Connect Meeting Platform system to investigate group work activities. In the digital group room, students could display documents, share their screens, broadcast webcam and voice, and exchange text chat and files. The results indicate that students were more engaged in group work than in pure online classes [Bo07]. The study of Jia et al. [Ji23] applied design-based research to develop and evaluate design principles for a fully online flipped classroom to support students' learning outcomes. In this study, students communicate asynchronously via Moodle and synchronously via Zoom. Moreover, the lecturer checks each group during online discussions to provide guidance or feedback as necessary. Results show that the fully online flipped model was as effective as the conventional flipped model in promoting student learning outcomes [Ji23]. The studies show that collaboration between students can be successful not only offline, but also through synchronous digital communication.

Four of the remaining five studies investigated communication among all participants (lectures and students) within a digital classroom system. Sabin; Higgs [SH07] investigated teaching and learning in a digital classroom with the so-called iLinc communications system. Students can participate in an online lecture without commuting to the university. But the communication among all participants in the investigated system does not enable synchronous discussions. Students and lecturers can only communicate via a text-based forum. Gonçalves et al. [Go14] investigated the impact of direct communication architecture and virtual communication architecture on the user perspective. In the case of the virtual communication architecture, the end-user can see and hear an interlocutor on the other side

via computer-mediated face-to-face communication. Whereas the virtual communication architecture uses an avatar or software agent representation of the interlocutor's image, and the end-user cannot see or hear the real image of the human interlocutor. 60 different students tested both architectures in an experimental setting. The result indicated that the virtual communication architecture does not affect end-user communication. Chang et al. [Ch21] investigated the difference of learning effectiveness between a physical and a digital classroom in an oral pathology and diagnosis course. From the students' perspective, the study revealed a better learning effectiveness, especially in terms of effective usage of time or convenient learning methods in the digital classroom. However, the investigated digital classroom only provides a chat communication function for students [Ch21]. Petchamé et al. [Pe23] designed a smart classroom system and investigated the students' and lecturers' perceptions of the artifact. The system supported hybrid learning, which allowed for flexibility in the location of both lecturers and students. However, it lacked features for fully integrating remote students, such as chat or a virtual hand for questions. Zolfaghari Mashhadi; Reza Kargozari [ZR11] discussed the advantages and disadvantages of digital classrooms but did not provide any information on developing requirements, an artifact, or empirical studies. Nevertheless, they highlighted the potential of digital classrooms to enhance traditional education methods.

The following key aspects for the development of the artifact can be derived from the existing literature.

1. Most studies focus on text-based communication, either via chats or via forums. Only Bower [Bo07], Jia et al. [Ji23], and Petchamé et al. [Pe23] used face-to-face communication. We also consider communication to be a central aspect, as it promotes the exchange of information and understanding [Ba77], which supports the transfer of knowledge and work in groups (e.g. [Ha00; Ki14; OTP11; Ro13; SH07]). Therefore, we consider communication one of the most important pillars for the artifact (see subchapter 4.2).
2. Former studies indicate that, besides communication, feedback is a crucial aspect for the learning success of students (e.g., [Bo07; GO08]). For instance, feedback provides students with specific information on what they are doing well and where they need improvement. Therefore, we consider various features to enable student feedback (see subchapter 4.2).
3. Except for Petchamé et al. [Pe23], most studies relied on existing videoconferencing systems such as Adobe Connect or Zoom, which have known weaknesses for online teaching, such as limited organizational features (slow change between rooms, keeping track of requests to speak, assigning tutors to work groups, etc.). The smart classroom artifact of Petchamé et al. [Pe23] is limited to live video with communication between participants and playback of recorded video. Regarding the artifact, we consider the organizational requirements a crucial part of the artifact's development, providing lecturers with the necessary support to manage their courses effectively (see subchapter 4.2). For instance, features allowing lecturers to track group-building processes or monitor students' learning

processes can significantly streamline course management. The administrative workload can be substantial in large courses, but targeted functionalities can simplify these tasks.

4 Artifact: The Concept of a Digital Classroom

4.1 Group Work and Group Discussion in the Scope of Application

When using group discussion and group work for teaching and learning, two different roles can be distinguished: The students who are the active learners in this scenario and the lecturers who set up this learning scenario for the students to teach them. The lecturers can be divided into two subgroups: The main lecturer(s) who is(are) responsible for the course and tutors who support the lecturers. Consistent with existing research we establish three categories of functionality alongside the views of the two roles: communication (C), feedback (F), and organization (O).

Students' view. Students come together in groups either self-organized (O) or predetermined by the lecturer, to work out a solution to a problem. For this purpose, the students communicate (C) with each other in order to coordinate, exchange information and thoughts, and jointly create documents, presentations, drawings, texts, etc., in this way successively creating the solution [OTP11]. If the students need support from the lecturers, e.g., because questions arise or they are unsure whether they are on the right track, the students call the attention of the lecturers through hand signals or voice. One of the lecturers will then usually come to the group in a timely manner to provide assistance. If several groups ask for support at the same time, the lecturer must decide on the sequence and remember it. Since group work often takes place in the same room and only at different tables, students can also listen to questions from other groups that they find interesting, as well as the lecturer's response. In this way, students receive feedback (F) from the lecturer on their questions as quickly and briefly as possible [GKH06].

Lecturers' view. The role of the lecturer includes organizing group work, supervising students, and managing their requests for assistance. This involves assigning students to groups, if not self-organized, and communicating with groups individually or collectively to address common questions. When working with tutors, lecturers coordinate their use for answering student questions and providing feedback, especially when multiple groups have simultaneous inquiries. Lecturers also oversee the work within each group, ensuring effective supervision and guidance.

4.2 Requirements

With regard to the categories established in subsection 4.1, we derive the following requirements for the artifact with a closer look at the categories:

Communication. The better students can communicate and the more communication

opportunities they can use, the better are their learning outcomes [OTP11]. Hence, a digital tool for group work and group discussion has to support both user groups, students as well as lecturers in their communication tasks. In particular, the following communication requirements (C1-C4) should be met by a digital tool:

C1. All kinds of communication forms. Learning needs bidirectional visual/aural communication [Da69; Ki14; LHC17]. Tutorials also need to consider the different forms of communication (gestures and facial expressions, verbally, (hand) sketches, etc.). There is substantive evidence showing that dialogue-rich teaching promotes student achievement and cognitive development [HA13; ML07; RAC15].

C2. Easy communication between students. First of all, students need to get to know and interact with their peers [SH07]. During collaborative work, conflicts are an important element that must be managed through communication [OTP11; Ro13].

C3. Easy communication between lecturer/tutor and students. Students need direct interaction with the lecturer [SH07] and the lecturer needs to choose exactly which students (plenary, groups, individuals) the interaction is directed at [Bo07].

C4. Quick follow-up with the lecturer/tutor. Students become frustrated if communication is not fast and easy [Ha00; Ke10; SKK91]. Furthermore, in tutorials, a quick question often helps to prevent students from working in the wrong direction. If communication is not fast and easy, students avoid asking their questions when they are almost sure.

Feedback. Also the feedback part concerns both user groups [HT07]. As group work and group discussion is a guided and supervised teaching method, asking for and receiving feedback is an inherent part [Bo07; GO08]. In detail, the following feedback requirements (F1 – F4) shall be met:

F1. Draw attention to a question. Often, students are reluctant to pose questions. Hence, the obstacle to asking a question should be as low as possible [Ch21].

F2. Asking questions without interrupting the lecturer. Students prefer to comment or ask questions to a lecturer without interrupting him/her [Bo07; Ch21; GH22].

F3. Fast Feedback from a Lecturer. The students prefer to raise questions in a physical class. The reason for this is that in the offline context they get immediate feedback [Ch21]. Hence, a digital tool should support feedback to be provided as fast as possible.

F4. See Questions of Other Students and Hear Answers. It happens that students have similar questions about the content of a lecture or exercise. In addition, students also find the questions raised by others are also very interesting as well as a summary by the professor and his answers. Students often listen to such a discussion very carefully or even participate as they like this kind of interactivity [GH22].

Organization. The organizational requirements (O1-O5) concern only the lecturer who is responsible for the organization and setup of the course [Co20].

O1. Student/Lecturer controlled group building/changing. Often the group building process is a crucial part of the teaching method. In general, students appreciate the opportunity to choose their groups [Bo07] or work alone [Sm11]. But working together with other students than already known peers, is also a valuable skill. Besides, lecturer controlled group building allows spontaneous and quick groupwork [Bo07].

O2. Fair coordination of the questions. In offline as well as online settings, keeping track of students questions is often difficult. Usually, questions should be handled in chronological order, which is mostly regarded as the fairest order possible. In particular, each student should receive equal support and no question should be overlooked [Co20].

O3. Coordination among several lecturers/tutors. To avoid duplicates and unnecessary allocations, lectures and tutors should know which questions are being answered by other lecturers/tutors [Al20; Ji23].

O4. Knowing who raised a question. To understand the learning progress of a student, a lecturer must know which student raised a question. Also, questions may be used in the assessment and grading process [Al20].

O5. Monitor students learning progress. Similarly, to O4, lecturers must monitor group performance during activities and in retrospect [Bo07] for assessment and grading.

4.3 Concept for Mapping the Requirements of an On-Site Tutorial Online

Based on the evolving roles of teaching and the related functional requirements for a digital tool for conducting group work and group discussions online, we develop a concept for a digital classroom with four components that cover the specified requirements.

Video conferencing system. With the integration of a video conferencing system, the different types of communication (C1) can be mapped in a digital classroom. In this way, a digital classroom supports collaborative work on documents and sharing notes. Furthermore, a suitable video conferencing system offers the possibility to ask questions in a chat without having to interrupt a lecturer during his lecture (F2).

Group Work. A digital classroom should support work in groups. To ensure easy communication between several students as well as between students and lecturers/tutors in the groups (C2, C3), a video conference is created for each group. Such a group conference can be created in a classroom by a lecturer but also by a student (O1). In order to be able to work undisturbed as a group, such a conference should be locked or switched to invisible. If a conference is visible, every user should see which users are in this conference and can join this conference (O1). A lecturer should see all conferences and join them to provide quick feedback. To make this process easier, a lecturer should see for each student which conferences they have in common, so that they can join and give feedback more quickly (F3).

Plenum Conference. In addition to group work, a plenary conference allows the lecturer to address all students at once, providing explanations for tasks, clarifying errors, or answering common questions. While students work in groups, it is crucial to maintain clear communication during the plenary. Thus, a digital classroom should include a clearly marked plenum conference as a separate video conference to ensure that students stay connected to both their group and plenary discussions, receiving all pertinent information seamlessly.

Ticket System. With the help of a ticket system, several important requirements from all areas can be managed at once. To support all kinds of communication forms (C1) a student should be able to raise a hand to call attention to a question. This is also possible in

several web conferencing systems. A problem arises when groups are in breakout rooms. They can still raise their hands in the main conference, but a lecturer does not know in which breakout room s/he has to go. Some conferencing systems offer a solution for this by allowing you to call for help from a breakout room, but this also leads to problems, because after clicking away such a call for help the lecturer has no possibility to find the group again. You also quickly lose track when there are many calls for help (O4). To solve this kind of problem, a digital classroom should offer a ticket system to better coordinate support for students (O2). This ticket system allows students to create a ticket and link it directly to a conference (C4). Once a ticket is created, lecturers and tutors receive an auditory signal and a push notification to alert them that the ticket has been created (F1). This ensures that s/he notices the creation of a ticket even while s/he worked on another ticket (O4). Notifying all teaching persons, lecturer and tutors, is a big advantage over classic web conferencing systems, where only the host receives notifications of help calls. Additionally, by linking group conferences directly to tickets, an lecturer can more easily track student learning (O5). Another advantage of the ticket system is that the questions become visible to others, which is not the case when raising the hand or calling for help (F4). This allows students to join an existing conference of an open ticket to get the answer from the lecturer or tutor. If, nevertheless, several tickets have been created for one particular question, a student, lecturer or tutor can easily see this from the ticket list. In this case, a lecturer or tutor could decide to answer this question in the plenum conference for everyone. To give students feedback that their ticket is being processed, a lecturer or tutor should assign a ticket to themselves or others (O3). This also allows a tutor to ask the lecturer for help if s/he cannot help the student.

A ticket system offers several advantages over an on-site tutorial. For example, the ticket system results in a fairer learning environment. It guarantees equal attention for all students, as tickets can be handled in the order in which they are created (O2). Also, lecturers or tutors cannot be intercepted by other students to answer their question first, rather than the question of the student who asked first. Another advantage is that the ticket system solves coordination problems. As tickets are assigned to a specific person, lecturer or tutor, (O3) situations where two tutors want to help in the same group because they did not know of each other, are avoided.

5 Design Evaluation of the Digital Classroom

5.1 Implementation

To evaluate the digital classroom concept, we created a browser-based prototype that includes all the intended features and provides easy access for students. We selected Big Blue Button due to its open-source nature and ability to effectively meet our web conferencing requirements. The interface is designed to provide a simple user experience for lecturers, tutors, and students, consolidating the digital classroom into a single page with

individual video conferences in separate tabs. It displays the participants (Figure 1, area 1) with the possibility to create or join a conference with a specific participant (Figure 1, area 4 and 5), the tickets for organizing the questions (Figure 1, area 3), and the existing conferences (Figure 1, area 2) for communicating with the participants and in groups.

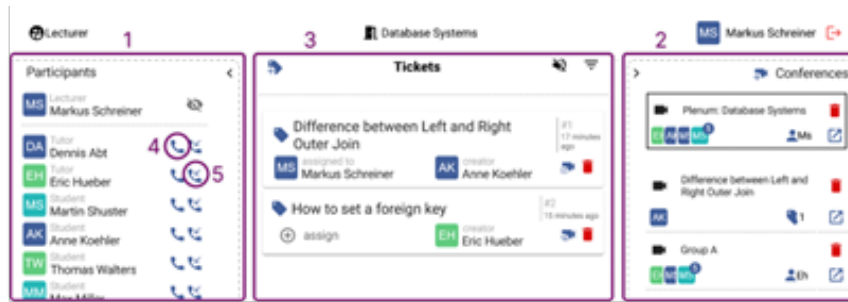


Fig. 1: Overview of the Digital Classroom interface
(Names changed by authors)

5.2 Scenario

Following the design science research approach [GH13], we evaluate the validity and utility of our proposed artifact in two ways. We evaluated the artifact using a real scenario and user feedback, complemented by a questionnaire derived from user comments and aligned with a theoretical construct. The scenario involved five database tutorials, each with a lecturer, four tutors, and about 30 students tackling SQL tasks. If needed, students could seek further clarification in online tutorials through the Digital Classroom accessed via Moodle. At the tutorial's start, a plenum conference is created for initial instructions and to address common questions, enabling simultaneous group and plenum participation. This setup facilitated efficient query resolution and allowed for direct support from lecturers and tutors during group activities, demonstrating the artifact's practical application and educational benefits. To find out whether users would use the artifact, we used the variable intention to use as a theoretical construct in the survey. The intention to use can be defined as follow: Behavioural intention to us is a measure of the strength of one's intention to perform a specified behaviour (e.g., [FA75, p. 288]). In studies with different contexts of technology acceptance (e.g. mobile banking, self-service, (online) learning systems), behavioural intention consistently precedes and predicts actual system use (e.g., [A121; Zo23]). Furthermore, Turner et al. [Tu10] found strong evidence in a meta-analysis that behavioural intention reliably predicts system usage both subjectively and objectively.

5.3 User Evaluation

During the winter term of 2021/2022, 50 students from different semesters participated in an evaluation of a database tutorial. Of these, 78.9% attended online and 21.1% attended

in person. The results showed that 91.3% of online participants were satisfied with the Digital Classroom. The platform's ease of use and effective support were highlighted. The ticket system of the artifact was praised for its simplicity in raising questions and the prompt response when a tutor joined the group room after a ticket was issued, enhancing learning efficiency and productivity. It is important to note that this feedback is subjective and may not reflect the experience of all users. However, some students found the ticket system less practical and more time-consuming than traditional methods, and a few noted technical issues like video transmission and functionality. However, students expressed their appreciation for the Digital Classroom's organizational capabilities. They indicated a preference to continue using it due to its convenience, separate group rooms, and ticket system. Additionally, they enjoyed the comfort of participating from home. The Digital Classroom (create their group room). In general, students plan to use the Digital Classroom in the future because of the separate group rooms or the ticket system. Moreover, they considered joining the class comfortable even when staying home.

Tab. 2: Overview of the evaluation of the Digital Classroom functionalities

Functionalities	Students/Lecturers
Collaborate work on shared PDF documents.	3.4 / 4.4
Share notes with fellow students and lecturers/tutors.	4.4 / 4.5
Possibility to chat with fellow students and lecturers/tutors.	4.4 / 4.4
Create a group room.	4.2 / 4.9
Lock the group room.	3.0 / 3.9
Switch group room invisible.	2.6 / 3.8
Create a ticket for my questions.	4.4 / 4.9
I can place my ticket so that I don't disturb the lecturer/tutor.	4.8 / 4.5
Coordination of student questions through the ticket system.	4.4 / 5.0
Ticket system invites lecturers/tutors to the group room.	4.0 / 4.9
I can see the processing status of my question.	4.0 / 4.8
Obtain questions from other groups.	3.8 / 4.1
Notice the lecturer's/tutor's answers to the questions of other groups.	3.8 / 4.3
Your camera image is only seen by the lecturer/tutor.	2.7 / 3.8
Set myself invisible so that the lecturer/tutor only sees me.	2.4 / 3.3
Use plenary for general information and group work Conferences.	4.0 / 4.6
Assignment of lecturers/tutors to tickets.	n.a. / 4.9
Ticket assignment overview.	n.a. / 4.9

5.4 Evaluation of the Functionalities

We applied a build-and-evaluate loop iteration with the user evaluation and the evaluation of the functionalities to develop the artifact further [MMG02]. By doing so, we invited experienced users to evaluate the current functionalities of the digital classroom in a survey. The first part of the survey collected data on the functionality and participants' usage intention. The second part focused on demographic variables (gender, age) and additional

information (e.g., traveling time to the university or students' learning preferences). The demographics of the 16 respondents show that 13 are male, 2 are female and 1 are divers. All questions were measured on a five-point Likert scale. For the measurement of the intention to use, we followed the design of Venkatesh et al. [Ve03]. The results of the descriptive analysis can be found in Table 2.

For lecturers and tutors, the key is to efficiently coordinate student queries to streamline question organization during lectures. They find features such as group creation, tracking question progress, tutor assignment, and an overview of ticket assignments vital for managing classes. Conversely, students prioritize the ability to submit questions (tickets) without disrupting tutors. They also value document sharing, chatting with peers, and the ticket system of the artifact for posing questions. The ticket coordination function is appreciated by students. Upon evaluating the intention to use the artifact, we noted a generally positive inclination among students (average: 3.6) and a more pronounced positive response from lecturers and tutors (average: 4.0).

There are discrepancies between the perspectives of students and lecturers with regard to the assessment of the functional capabilities of the digital classroom. The mean rating for the "Collaborate work on shared PDF documents" function was 4.4 for lecturers and 3.4 for students. This discrepancy could be attributed to the fact that lecturers perceive the overview and collaboration aspects to be more straightforward, whereas students may experience challenges due to the limitation of only one individual sharing and editing the screen with the document, which constrains collaboration. The other features for which the ratings differ also exhibited the lowest ratings. For instance, the room locking function was rated 3.0 by students and 3.9 by lecturers, while the adjustment of room visibility scored 2.6 from students and 3.8 from lecturers. These low ratings might stem from a preference for an open learning environment that promotes continuous interaction. Similarly, features such as the exclusive use of cameras by lecturers (students: 2.8, lecturers: 3.8) and the ability to prevent one's visibility (students: 2.4, lecturers: 3.8) were also rated poorly. These scores further indicate the importance of maintaining an open and interactive learning atmosphere. Additionally, the low evaluations may be due to a lack of perceived necessity for these features. Students and lecturers might find these functions unnecessary. Consequently, these evaluations suggest that these features do not align with the primary objective. Therefore, it may be beneficial to consider removing these features in future versions of the digital classroom.

6 Conclusion

6.1 Discussion of the Research Question

To answer our research question, we developed the concept of a digital classroom that aims at mapping on-site academic peer discussion and group work into online teaching.

Following the design science approach, several iterations of design, development, and evaluation have been run through leading to gradually improved requirements according to the needs of lecturers and students. Concerning our research question: *“How can academic peer group discussions and group work be transferred into the online context?”*, we focused on communication between all participants, feedback of lecturers, and the ticket system. An important feature in online group discussion is to have rich communication types (C1) and several communication channels in parallel as well as to be able to switch between these channels very easily (C2-C4, O1). Consequently, a lecturer can easily change to the plenum conference and interrupt all group discussions—similar to an on-site message to all students. Afterwards, students can easily continue their discussion. This is as fast as in the offline context; it is not necessary to wait until students changed any video conference channels. Another important feature is the ticket system, which fulfills the organization requirements: students can ask questions (F1, F2), see and know the questions from other students (O4, O5), can ask to join a breakout sessions where a question is answered (F5). Moreover, the tickets allow a fair order of the students’ questions (O2). This is highly valued by lecturers/tutors as well as by students. In the same way, tutors can give feedback. A separate breakout session allows quick and easy coordination between tutors and lecturer (F3) so that students get fast feedback (O3).

6.2 Implications

Evaluations show that moving on-site academic discussions and group work to an online format is not only feasible, but also well received by students. The COVID-19 pandemic has equipped both students and faculty with the necessary skills to use these digital platforms effectively, recognizing and taking advantage of their features. This adaptability paves the way for the use of digital tools to create new learning opportunities. It’s therefore recommended that proven face-to-face teaching methods be adapted to online environments in order to exploit their pedagogical potential. Such generic concepts as peer discussion and group work can be applied across disciplines, provided that lecturers are willing and able to integrate digital technologies into their teaching.

To facilitate initial adoption, collaborations across disciplines and universities are envisioned. Technically skilled academics could create sophisticated yet user-friendly systems and help faculty from different disciplines to use these platforms. In turn, these faculty could identify new applications and requirements, enabling the development of versatile systems. With the increasing specialization of professions requiring specialized courses, the importance of appropriately sized learning groups, including those in different locations, becomes paramount. Thus, the requirements outlined and the concept of the digital classroom provide valuable guidance and a viable solution framework.

The digital classroom concept developed in this work serves as the foundation for future advancements. By utilizing a free, open-source tool, we are facilitating its accessibility to a vast number of academic institutions. The incorporation of particular organizational functions

enables the application to be tailored to diverse pedagogical settings. The application is published on GitHub [II24], thereby facilitating the extension of its functionality by other individuals or project teams.

The digital classroom provides the fundamental functionality necessary for digital teaching, and it can be adapted to suit a variety of individual requirements through different implementations. The development process must be centered on the users' needs and requirements. This is the only viable approach for optimally addressing the needs of both students and lecturers. In order to enhance the appeal of the system to other educational institutions, it is essential to facilitate greater flexibility in the visual elements. This includes the option of uploading one's own logos and selecting individual colors. By integrating additional functions such as the distribution of lecture notes or exercise sheets, the digital classroom can be further developed into a comprehensive overall solution for digital teaching. It is important that all developments are made in the existing GitHub repository [II24] so that the results can also be made available to other institutions.

6.3 Limitation and Future Research

The main issue is the relatively small sample size of the evaluation. Due to this, we could only analyze the data descriptively and not with a statistical approach to identify significant relationships. However and although the evaluation can only be a starting point, the results already point in a very positive direction. Nevertheless, for future research, more participants' responses will be collected to get more reliable insights also into further usage behavior data (e.g., how participants behave during the usage of the Digital Classroom or the impact on students' final grades). Concerning future research, it will be interesting to investigate the combination of a digital classroom and its tickets with a FAQ wiki. Additionally, we think about a combination of our digital classroom with an intelligent tutoring system (ITS) for two reasons: (1) If a student does not understand the feedback given so some exercise from an ITS, then all information with the exercise and the answer of the student can be linked directly to a ticket. (2) An ITS can help to answer tickets if there was a similar ticket and a lecturer recorded an answer for that ticket.

References

- [AG15] Alzoubi, Y.; Gill, A.: Agile Global Software Development Communication Challenges: A Systematic Review. In: Proceedings - Pacific Asia Conference on Information Systems, PACIS 2014. 2015.
- [AI11] Ala-Mutka, K.: Mapping digital competence: Towards a conceptual understanding. Sevilla: Institute for Prospective Technological Studies, pp. 7–60, 2011.
- [AI20] Al-Fraihat, D.; Joy, M.; Masa'deh, R.; Sinclair, J.: Evaluating E-learning systems success: An empirical study. *Computers in Human Behavior* 102, pp. 67–86, 2020.

- [Al21] Almulla, M. A.: Technology Acceptance Model (TAM) and E-Learning System Use for Education Sustainability. *Academy of Strategic Management Journal* 20, 2021, URL: <https://api.semanticscholar.org/CorpusID:237012669>.
- [Ba77] Bandura, A.: Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review* 84 (2), pp. 191–215, 1977.
- [Be15] Beth, A. D.; Jordan, M. E.; Schallert, D. L.; Reed, J. H.; Kim, M.: Responsibility and generativity in online learning communities. *Interactive Learning Environments* 23 (4), pp. 471–484, 2015.
- [Bo07] Bower, M.: Groupwork Activities in Synchronous Online Classroom Spaces. *SIGCSE Bull.* 39 (1), pp. 91–95, 2007.
- [Bö21] Böheim, R.; Schnitzler, K.; Gröschner, A.; Weil, M.; Knogler, M.; Schindler, A.-K.; Alles, M.; Seidel, T.: How changes in teachers' dialogic discourse practice relate to changes in students' activation, motivation and cognitive engagement. *Learning, Culture and Social Interaction* 28, p. 100450, 2021.
- [Bu22] Bui, D. T.; Nhan, T. T.; Dang, H. T. T.; Phung, T. T. T.: Online learning experiences of secondary school students during COVID-19 – Dataset from Vietnam. *Data in Brief* 45, p. 108662, 2022.
- [CGC15] Conboy, K.; Gleasure, R.; Cullina, E.: Agile Design Science Research. In (Donnellan, B.; Helfert, M.; Kenneally, J.; VanderMeer, D.; Rothenberger, M.; Winter, R., eds.): *New Horizons in Design Science: Broadening the Research Agenda*. Springer International Publishing, Cham, pp. 168–180, 2015.
- [Ch21] Chang, J. Y.-F.; Wang, L.-H.; Lin, T.-C.; Cheng, F.-C.; Chiang, C.-P.: Comparison of learning effectiveness between physical classroom and online learning for dental education during the COVID-19 pandemic. *Journal of Dental Sciences* 16 (4), pp. 1281–1289, 2021.
- [Co20] Cong, L. M.: Successful Factors for Adoption of Synchronous Tools in Online Teaching at Scale. In: *Tertiary Education in a Time of Change: Disruptions, Challenges and Strategies*. Springer Singapore, Singapore, pp. 39–60, 2020.
- [Da69] Dale, E.: *Audio-Visual Methods in Teaching*. Rinehart and Winston, 1969.
- [De19] Deslauriers, L.; McCarty, L. S.; Miller, K.; Callaghan, K.; Kestin, G.: Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. *Proceedings of the National Academy of Sciences* 116 (39), pp. 19251–19257, 2019.
- [El18] Elmaadaway, M. A. N.: The effects of a flipped classroom approach on class engagement and skill performance in a Blackboard course. *British Journal of Educational Technology* 49 (3), pp. 479–491, 2018.
- [FA75] Fishbein, M.; Ajzen, I.: *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research*. Addison-Wesley, 1975.
- [GH13] Gregor, S.; Hevner, A.: Positioning and Presenting Design Science Research for Maximum Impact. *MIS Quarterly* 37, pp. 337–356, 2013.
- [GH22] Gu, M. M.; Huang, C. F.: Transforming habitus and recalibrating capital: University students' experiences in online learning and communication during the COVID-19 pandemic. *Linguistics and Education* 69, p. 101057, 2022.
- [GKH06] Geister, S.; Konradt, U.; Hertel, G.: Effects of Process Feedback on Motivation, Satisfaction, and Performance in Virtual Teams. *Small Group Research* 37 (5), pp. 459–489, 2006.

-
- [GO08] Gallien, T.; Oomen-Early, J.: Personalized Versus Collective Instructor Feedback in the Online Courseroom: Does Type of Feedback Affect Student Satisfaction, Academic Performance and Perceived Connectedness With the Instructor? *International Journal on E-Learning* 7 (3), pp. 463–476, 2008.
 - [Go14] Gonçalves, P.; Ferreira, L.; Gonçalves, J.; Putnik, G. D.; Cruz-Cunha, M. M.: Direct Communication versus Virtual Communication in Virtual Teams. *Procedia Technology* 16, CENTERIS 2014 - Conference on ENTERprise Information Systems / ProjMAN 2014 - International Conference on Project MANagement / HCIST 2014 - International Conference on Health and Social Care Information Systems and Technologies, pp. 3–10, 2014, ISSN: 2212-0173.
 - [GOA22] Gutentag, T.; Orner, A.; Asterhan, C. S.: Classroom discussion practices in online remote secondary school settings during COVID-19. *Computers in Human Behavior* 132, p. 107250, 2022.
 - [Ha00] Haythornthwaite, C.; Kazmer, M. M.; Robins, J.; Shoemaker, S.: Community Development among Distance Learners: Temporal and Technological Dimensions. *Journal of Computer-Mediated Communication* 6 (1), JCMC615, 2000.
 - [HA13] Howe, C.; Abedin, M.: Classroom dialogue: a systematic review across four decades of research. *Cambridge Journal of Education* 43 (3), pp. 325–356, 2013.
 - [He04] Hevner, A. R.; March, S. T.; Park, J.; Ram, S.: Design Science in Information Systems Research. *MIS Quarterly* 28 (1), pp. 75–105, 2004, visited on: 11/17/2022.
 - [HMO93] Hollingshead, A.; Mcgrath, J.; O'Connor, K.: Group Task Performance and Communication Technology: A Longitudinal Study of Computer-Mediated Versus Face-to-Face Work Groups. *Small Group Research - SMALL GROUP RES* 24, pp. 307–333, 1993.
 - [HT07] Hattie, J.; Timperley, H.: The Power of Feedback. *Review of Educational Research* 77 (1), pp. 81–112, 2007.
 - [II24] II, T. M.: Digital Classroom, <https://github.com/thm-mni-ii/digital-classroom>, 2024.
 - [Ji23] Jia, C.; Hew, K. F.; Jiahui, D.; Liuyufeng, L.: Towards a fully online flipped classroom model to support student learning outcomes and engagement: A 2-year design-based study. *The Internet and Higher Education* 56, p. 100878, 2023.
 - [JSP02] Jang, C.-Y.; Steinfield, C.; Pfaff, B.: Virtual team awareness and groupware support: an evaluation of the TeamSCOPE system. *International Journal of Human-Computer Studies* 56 (1), pp. 109–126, 2002, ISSN: 1071-5819.
 - [Ke10] Kear, K.: Social presence in online learning communities. In: *Proceedings of the 7th International Conference on Networked Learning* 2010. 2010.
 - [Ki14] Kim, M. K.; Kim, S. M.; Khera, O.; Getman, J.: The experience of three flipped classrooms in an urban university: an exploration of design principles. *The Internet and Higher Education* 22, pp. 37–50, 2014.
 - [LC19] Lee, B. K.; Cai, H.: Evaluation of an Online “Internationalization at Home” Course on the Social Contexts of Addiction. *Journal of Studies in International Education* 23 (3), pp. 365–388, 2019.
 - [LHC17] Lo, C. K.; Hew, K. F.; Chen, G.: Toward a set of design principles for mathematics flipped classrooms: A synthesis of research in mathematics education. *Educational Research Review* 22, pp. 50–73, 2017.
 - [ML07] Mercer, N.; Littleton, K.: Dialogue and the development of children’s thinking: A sociocultural approach. Routledge, 2007.

- [MMG02] Markus, M. L.; Majchrzak, A.; Gasser, L.: A Design Theory for Systems That Support Emergent Knowledge Processes. *MIS Quarterly* 26 (3), pp. 179–212, 2002, issn: 02767783, visited on: 03/10/2023.
- [Mu21a] Muhonen, H.; Pakarinen, E.; Rasku-Puttonen, H.; Lerkkanen, M.-K.: Educational dialogue among teachers experiencing different levels of self-efficacy. *Learning, Culture and Social Interaction* 29, p. 100493, 2021.
- [Mu21b] Muszyńska, K.: A bibliometric review of research on communication in virtual project teams. *Procedia Computer Science* 192, Knowledge-Based and Intelligent Information & Engineering Systems: Proceedings of the 25th International Conference KES2021, pp. 4770–4779, 2021, issn: 1877-0509.
- [NC19] Nadae, J. D.; Carvalho, M. M.: Communication Management and Knowledge Management in Complex Projects: A Literature Review. *Revista de Gestão e Projetos* 10 (1), pp. 19–36, 2019, issn: 2236-0972, doi: 10.5585/gep.v10i1.10910, visited on: 07/25/2024.
- [OTP11] Oliveira, I.; Tinoca, L.; Pereira, A.: Online group work patterns: How to promote a successful collaboration. *Computers & Education* 57 (1), pp. 1348–1357, 2011.
- [Pa21] Page, M. J.; McKenzie, J. E.; Bossuyt, P. M.; Boutron, I.; Hoffmann, T. C.; Mulrow, C. D.; Shamseer, L.; Tetzlaff, J. M.; Akl, E. A.; Brennan, S. E.; Chou, R.; Glanville, J.; Grimshaw, J. M.; Hróbjartsson, A.; Lalu, M. M.; Li, T.; Loder, E. W.; Mayo-Wilson, E.; McDonald, S.; McGuinness, L. A.; Stewart, L. A.; Thomas, J.; Tricco, A. C.; Welch, V. A.; Whiting, P.; Moher, D.: The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews. *Systematic Reviews* 10 (1), p. 89, 2021, issn: 2046-4053, doi: 10.1186/s13643-021-01626-4, visited on: 07/24/2024.
- [PB02] Potter, R.; Balthazard, P.: Virtual team interaction styles: Assessment and effects. *International Journal of Human-Computer Studies* 56, pp. 423–443, 2002.
- [Pe12] Peffers, K.; Rothenberger, M.; Tuunanen, T.; Vaezi, R.: Design Science Research Evaluation. In: Vol. 7286, pp. 398–410, 2012, isbn: 978-3-642-29862-2.
- [Pe23] Petchamé, J.; Iriondo, I.; Korres, O.; Paños-Castro, J.: Digital transformation in higher education: A qualitative evaluative study of a hybrid virtual format using a smart classroom system. *Heliyon* 9 (6), e16675, 2023, issn: 2405-8440.
- [RAC15] Resnick, L.; Asterhan, C.; Clarke, S.: Socializing intelligence through academic talk and dialogue. American Educational Research Association, 2015.
- [Ro13] Robinson, K.: The interrelationship of emotion and cognition when students undertake collaborative group work online: An interdisciplinary approach. *Computers & Education* 62, pp. 298–307, 2013.
- [SH07] Sabin, M.; Higgs, B.: Teaching and Learning in Live Online Classrooms. In: Proceedings of the 8th ACM SIGITE Conference on Information Technology Education. SIGITE '07, Association for Computing Machinery, New York, NY, USA, pp. 41–48, 2007, isbn: 9781595939203.
- [SKK91] Sproull, L.; Kiesler, S.; Kiesler, S. B.: Connections: New ways of working in the networked organization. MIT press, 1991.
- [Sm11] Smith, G. G.; Sorensen, C.; Gump, A.; Heindel, A. J.; Caris, M.; Martinez, C. D.: Overcoming student resistance to group work: Online versus face-to-face. *The Internet and Higher Education* 14 (2), Web mining and higher education: Introduction to the special issue, pp. 121–128, 2011.

- [Tu10] Turner, M.; Kitchenham, B.; Brereton, P.; Charters, S.; Budgen, D.: Does the Technology Acceptance Model Predict Actual Use? A Systematic Literature Review. *Information and Software Technology* 52 (5), pp. 463–479, 2010, issn: 09505849, doi: 10.1016/j.infsof.2009.11.005, visited on: 07/29/2024.
- [Ve03] Venkatesh, V.; Morris, M. G.; Davis, G. B.; Davis, F. D.: User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly* 27 (3), pp. 425–478, 2003.
- [Wa92] Walther Joseph B.;Burgoon, J. K.: Relational communication in computer-mediated interaction. *Human Communication Research* 19, pp. 50–88, 1992.
- [Wu13] Wu, X.; Anderson, R. C.; Nguyen-Jahiel, K.; Miller, B.: Enhancing motivation and engagement through collaborative discussion. *Journal of Educational Psychology* 105 (3), p. 622, 2013.
- [Zo23] Zobeidi, T.; Homayoon, S. B.; Yazdanpanah, M.; Komendantova, N.; Warner, L. A.: Employing the TAM in Predicting the Use of Online Learning during and beyond the COVID-19 Pandemic. *Frontiers in Psychology* 14, p. 1104653, 2023, issn: 1664-1078, doi: 10.3389/fpsyg.2023.1104653, visited on: 07/29/2024.
- [ZR11] Zolfaghari Mashhadi, V.; Reza Kargozari, M.: Influences of digital classrooms on education. *Procedia Computer Science* 3, World Conference on Information Technology, pp. 1178–1183, 2011.

Quantitative Analyse der Auswirkungen von Remote-Arbeit

Vollständiger Beitrag

Can Iğde¹ und Sandy Eggert²

Abstract: Digitale Technologien und neue Anforderungen von Mitarbeitenden führen zu Veränderungen in der Arbeitswelt. Dies zeigt sich u.a. in einem Wandel von traditioneller Arbeit hin zu einer zunehmenden Verbreitung von Remote-Arbeit. Um die Auswirkungen von Remote-Arbeit auf die Zusammenarbeit und Kommunikation in Teams in Unternehmen zu untersuchen, wurde eine empirische Analyse durchgeführt. Dieser Beitrag zeigt die Ergebnisse dieser Untersuchung unter Berücksichtigung zuvor aufgestellter Hypothesen, welche Aspekte wie Kommunikationseffizienz, Unternehmenskultur, Teamzusammenarbeit, Teamgeist und einen Vergleich von Unternehmensgrößen fokussieren.

Keywords: Remote-Arbeit, Teams, Zusammenarbeit, Kommunikation, Unternehmenskultur.

1 Einleitung

Die Arbeitswelt hat sich in den letzten Jahrzehnten durch technologische und soziale Innovationen verändert [Bri20]. Um wettbewerbsfähig zu bleiben, müssen Unternehmen vermehrt flexible Arbeitszeiten und -orte anbieten [Bri20]. In den traditionellen Unternehmen war diese Remote-Arbeitsweise nur für ausgewählte Gruppen möglich. Dazu gehörten Führungskräfte oder Vertriebsmitarbeitende, die diese Art der Arbeit auch nur tageweise ausübten. Zudem war es nicht gern gesehen, wenn Mitarbeitende die Möglichkeit der Remote-Arbeit in Anspruch nahmen [Bru20]. Aus der Notwendigkeit in der COVID-19-Pandemie remote arbeiten zu müssen, entstand auch nach der Pandemie der Wunsch nach mehr Flexibilität [Bru20].

Angesichts des zunehmenden Wandels von traditioneller Arbeit hin zu Remote-Arbeit ist es das Ziel in der Untersuchung, Erkenntnisse darüber zu gewinnen, ob und wie Remote-Arbeit die Art und Weise verändert, wie Teams in Unternehmen zusammenarbeiten und kommunizieren. Darüber hinaus soll ein Beitrag zum Verständnis der Auswirkungen von Remote-Arbeit auf die Arbeitsdynamik geleistet werden. Insbesondere sollen dabei Aspekte der Kommunikation und Teamarbeit in mittleren und großen Unternehmen beleuchtet werden.

¹ HWR Berlin, Badensche Str. 52, 10825 Berlin, igde00@gmail.com,

² HWR Berlin, Badensche Str. 52, 10825 Berlin, sandy.eggert@hwr-berlin.de

2 Grundlagen und Hypothesen

Der theoretische Rahmen bildet die Grundlage für die empirische Untersuchung. Neben Definitionen werden technologische, soziale und organisatorische Aspekte berücksichtigt.

2.1 Remote-Arbeit

Remote-Arbeit wird definiert als eine flexible Arbeitsvereinbarung, bei der es dem Arbeitnehmer ermöglicht wird, seine Arbeit an einem anderen Ort als dem Betriebssitz zu praktizieren. In diesem Fall kann der Arbeitsplatz jeder andere Ort sein und ist somit nicht auf den Unternehmensstandort beschränkt [OPM24]. Die zunehmende Verbreitung von Remote-Work wird maßgeblich durch die Entwicklung der globalen Wirtschaft sowie durch technologische Fortschritte beeinflusst. Erfolge, die von Unternehmen durch die Implementierung solcher Organisationsstrukturen und Arbeitsmodelle erzielt worden sind, haben wesentlich zur Förderung der Remote-Arbeit beigetragen. Ein weiterer Faktor ist der demografische Wandel in der globalen Arbeitnehmerschaft. Die Integration von Remote-Arbeitsmöglichkeiten und den damit einhergehenden Technologien am Arbeitsplatz werden von jüngeren Generationen oft als Voraussetzung gesehen. Dadurch stoßen Unternehmen an ihre Grenzen, da sie sicherstellen müssen, dass die Technologien, die neuesten Tools, die Sicherheitsvorkehrungen und der Support gewährleistet sind. Zugleich steigt die Erwartung, ein Gleichgewicht zwischen dem Berufs- und Privatleben zu schaffen, was durch die Remote-Arbeit besser gewährleistet werden kann [Bel21]. Zurückzuführen ist dies auch auf die COVID-19-Pandemie als zentraler Einflussfaktor, die eine weltweite, zwangsläufige Machbarkeitsstudie im Jahr 2020 initiierte und sich damit für viele Unternehmen als Gelegenheit erwies, eine neue Arbeitskultur, die Umstellung auf Remote-Arbeit, zu erproben.

2.2 Unternehmenskulturen

Der Begriff der Unternehmenskultur kann aufgrund seiner Vielfalt pluralistisch definiert werden. So kann sie als Zusammenspiel symbolischer Referenzpunkte und unhinterfragter Annahmen verstanden werden. Diese können als Leitfaden für die verbale und nonverbale Kommunikation innerhalb der Organisation fungieren [Ric15]. Im Laufe der Zeit können Mitarbeitende ein intuitives Verständnis dafür entwickeln, welches Verhalten zum Erfolg führen kann [Hom14]. Dabei handelt es sich um ein intuitives Verständnis für die fundamentalen Überzeugungen und Annahmen in einer Organisation. Diese prägen die Wahrnehmung, das Denken, das Fühlen und Handeln [Ric15]. Dazu gehören Verhaltensweisen und Muster wie das Erkennen angemessener Umgangsformen innerhalb der Organisation. Dabei kann zwischen formeller und informeller Kommunikation mit Mitarbeitenden und Abteilungen differenziert werden. Ebenso wird ein Bewusstsein dafür entwickelt, wie verschiedene Aspekte der Aufgaben bewertet werden, ob es sich um starke oder eher schwache Kontrollen handelt [Hom14]. Diese Verhaltensweisen und Muster sind nicht isoliert zu betrachten. Die Gesamtheit aller Verhaltensweisen der Unternehmenskultur

sind als eine umfassende Struktur zu bezeichnen. Im Zuge dessen definiert der Experte Edgar H. Schein die Unternehmenskultur als ein Muster gemeinsamer Grundprämissen. Diese entwickeln sich aus der erfolgreichen Bewältigung von Problemen der externen Anpassung und internen Integration. Einmal etabliert, werden sie für die Organisation bindend und an neue Mitglieder als der richtige Ansatz zur Lösung der Probleme weitergegeben [Hom14].

2.3 Bedeutung von Kommunikation

Der Begriff Kommunikation wird definiert als ein Prozess der Übertragung von Nachrichten, Botschaften oder Informationen. Der Austausch findet zwischen einem Sender und einem oder mehreren Empfängern statt. Im wissenschaftlichen Kontext können diese auch als Kommunikator und Rezipient bezeichnet werden [Sie24]. Kommunikation ermöglicht es, Bestrebungen, Hoffnungen, Ziele und weitere emotionale Zustände mit anderen Mitmenschen zu teilen. Dies kann sowohl bewusst als auch unbewusst geschehen [Sch16]. Durch die Betonung von Sätzen sind Menschen in der Lage, über den eigentlichen Wortsinn hinaus Informationen zu transportieren. Prosodische Merkmale wie die Geschwindigkeit beim Sprechen und/oder die verwendete Lautstärke spielen eine signifikante Rolle [Sch16]. Nonverbale Kommunikation kann unabhängig von verbaler Kommunikation stattfinden. Alle Aspekte, die einen signifikanten Einfluss auf die zwischenmenschliche Kommunikation haben, werden als Wirkfaktoren bezeichnet. Das Zeitalter der digitalen Kommunikation führt zu einer signifikanten Reduktion der verbalen und nonverbalen Kommunikation, da die digitalen Formate eine nur bedingte Übertragung derer ermöglichen. [Her21] Dennoch werden durch die Digitalisierung die Reichweite der Kommunikationsmöglichkeiten für den Einzelnen signifikant gesteigert. Die digitalen Medien ermöglichen ohne einen nennenswerten Aufwand eine ortsunabhängige Kommunikation. Daraus folgt die Entstehung neuer Kommunikationswege, -mittel und -arten [vAm21]. Durch elaborierte Einsatzkonzepte wird den Mitarbeitenden in den Unternehmen mit einer möglichst hohen Sicherheit und Zuverlässigkeit der Zugang zu essenziellen Informationen und Ressourcen gewährleistet [Her21]. Das Internet und Technologien haben die Kommunikation ohne Zeit- und Ortsbeschränkungen ermöglicht, gemacht, was zu einer permanenten Vernetzung und Erreichbarkeit führt [vAm21]. Daraus resultiert, dass physische Entfernungen für die Kommunikation und Zusammenarbeit kaum noch relevant sind [Her21].

2.4 Zusammenarbeit und Teams

Sowohl im privaten als auch im beruflichen Umfeld treffen Menschen immer wieder auf Teams bzw. Zusammenarbeit [Sch16]. Im beruflichen Umfeld setzt sich das Team aus den Mitarbeitenden zusammen [Sch16]. Diese kooperieren und arbeiten für einen bestimmten Zeitraum gemeinsam, wobei ihre individuellen Fähigkeiten und Kompetenzen synergetisch zusammenwirken. Das Team strebt die gemeinsamen Ziele an, wobei die individuellen Ergebnisse von den erbrachten Resultaten der Teammitglieder abhängen [Böh21].

Die Zusammenarbeit im Team wird als die am meisten bevorzugte Arbeitsform angesehen, da der Arbeitsaufwand reduziert und der Nutzen verdoppelt wird [Mei12]. Dies geschieht durch die Nutzung von Synergien sowie die Zusammenarbeit der Teammitglieder, die sich gegenseitig unterstützen und ihr Wissen kombinieren. Dennoch ist anzumerken, dass Teamarbeit bzw. Zusammenarbeit nicht universell zu positiven Resultaten führt. Das Ergebnis ist von den verschiedenen Teammitgliedern, deren Erwartungen und Vorstellungen abhängig. Nur durch die konsequente Verfolgung der gemeinsamen Ziele und die effektive Zusammenarbeit im Team kann ein positives Ergebnis erzielt werden [Böh21]. Das gemeinsame Arbeiten im Team bietet sowohl für die Produktivität als auch für das Arbeitsklima viele Vorteile. Eine umfassendere Sichtweise bei der Problemlösung kann durch die Nutzung der kollektiven Erfahrung und des kollektiven Wissens erreicht werden. Dies fördert die Kommunikationsdynamik in der Gruppe, wodurch diese produktiver und effizienter arbeiten kann. Durch die Zusammenarbeit können Fehlentscheidungen aufgrund einer gemeinsamen Entscheidungsfindung minimiert bzw. vermieden werden, da viele verschiedene Meinungen und Expertisen in den Entscheidungsprozess miteinfließen. Durch die Kooperation kann ein angenehmes Arbeitsklima geschaffen werden, das durch eine gegenseitige Anerkennung und Wertschätzung die Vorurteile, Spannungen und Konflikte in der Zusammenarbeit minimiert. Dies wirkt sich wiederum positiv auf die Zusammenarbeit, die Motivation und die Arbeitszufriedenheit aus, sodass effizienter gearbeitet werden kann [Mei12].

Analog zu einer Sportmannschaft, kann von einem guten Team gesprochen werden, wenn das Team durch die synergetische Integration heterogener Fähigkeiten und Rollen auf einen gemeinsamen Erfolg ausgerichtet ist. Dabei liegt die Effektivität nicht nur in der individuellen Kompetenz. Die Rollenklarheit muss klar definiert sowie auf alle Teammitglieder und deren Kompetenzen abgestimmt sein. Diese Rollenklarheit ist entscheidend für die effektive Zusammenarbeit des Teams [Sch16]. Daher ist es bei der Bildung eines guten Teams sinnvoll, Menschen mit unterschiedlichen persönlichen Eigenschaften zu einem Team zu formen. Die Zielpromotoren sind hierbei diejenigen, die den Fortschritt mit den Zielen in Einklang bringen, während die Fachpromotoren das Fachliche in den Vordergrund stellen. Die sozialen Promotoren sind hingegen die Teammitglieder, die sich um das Team kümmern und es zusammenhalten. Weitere wichtige Faktoren sind Teammitglieder mit einer ausgeprägten Kreativität. Sie tragen dazu bei, dass durch die kreative Herangehensweise an Probleme Lösungsideen effizienter gefunden werden können. Um die Effizienz zu gewährleisten, sind auch Teammitglieder mit einem ausgeprägten Finanzwissen notwendig, sodass bswp. die Kosten für ein Projekt nicht überschritten werden [Mei12]. Zudem ist eine ausgewogene Mischung verschiedener persönlicher Eigenschaften notwendig [Mei12]. Weiterhin ist die Rolle des Teamleiters von großer Relevanz, da dieser die Aufgabe hat, als koordinierende Instanz des Teams zu fungieren, um die Methode und Vorgehensweise zu definieren und zu optimieren. Eine weitere entscheidende Aufgabe des Teamleiters besteht darin, dafür zu sorgen, dass die Leistungsfähigkeit des Teams kontinuierlich aufrechterhalten wird [Sch16].

3 Methodische Vorgehensweise

Aus den erarbeiteten Grundlagen lassen sich Hypothesen ableiten, die anschließend überprüft werden sollen. Dafür erweist sich eine quantitative Forschungsmethode als zweckgemäß, um systematisch relevante Daten in Bezug auf die thematischen Schwerpunkte zu erheben und analysieren. Um die Gütekriterien: Objektivität, Reliabilität und Validität zu erfüllen [Sch14], eignet sich eine standardisierte Befragung [Möh13].

3.1 Hypothesen

Aus den zuvor erarbeiteten Erkenntnissen zur Remote-Arbeit, Kommunikation, Teamzusammenarbeit und -aufbau lassen sich sechs Hypothesen ableiten, die mithilfe eines Fragebogens näher untersucht werden sollen.

H-1: Kommunikationseffizienz: Beschäftigte, die die im Unternehmen eingesetzten Kommunikationsmittel positiv bewerten, beurteilen auch die Kommunikation während der Remote-Arbeit positiver.

H-2: Unternehmenskultur: Die Einführung von Remote-Arbeit in Unternehmen wirkt sich positiv auf die Einhaltung und Umsetzung der Werte der Unternehmenskultur aus.

H-3: Teamzusammenarbeit: Teams, die regelmäßig Remote-Arbeit praktizieren, bewerten die Qualität ihrer Zusammenarbeit höher als Teams, deren primäre Arbeitsumgebung das Büro ist.

H4: Teamgeist: Regelmäßige Teambesprechungen, während der Remote-Arbeit tragen wesentlich zur Stärkung des Teamgeistes bei.

H-5: Vergleich von Unternehmensgrößen: Im Vergleich zu mittleren Unternehmen wird die Kommunikation bei Remote-Arbeit in Großunternehmen positiver bewertet.

H-6: Durchschnittsalter: Die Kommunikationseffizienz bei Remote-Arbeit ist in Unternehmen mit einem niedrigen Durchschnittsalter höher als in Unternehmen mit einem hohen Durchschnittsalter.

3.2 Datenerhebung

Zur Überprüfung der Hypothesen erfolgte zunächst die Entwicklung eines Fragebogens entlang der Hypothesen. Der aus 26 Fragen bestehende Fragebogen, wurde mit einem digitalen Umfragetool erstellt und anschließend zur Datenerhebung auf unterschiedlichen Kommunikationskanälen, sozialen Plattformen und in unternehmensinternen Netzwerken verbreitet.

4 Ergebnisse

Nachfolgend werden die Umfragedaten detailliert analysiert. Ziel ist es, fundierte Einsichten zu gewinnen und die Hypothesen H-1 bis H-6 zu überprüfen.

4.1 Deskriptive Datenanalyse

Im Zeitraum vom 12.02.2024 bis zum 29.02.2024 haben 215 Personen an der Umfrage teilgenommen. Die am stärksten vertretene Altersgruppe in dieser Stichprobe sind Personen im Alter von 19 bis 29 Jahren (60,5 %). Die Altersgruppen der 30- bis 44-Jährigen sowie die Altersgruppe der über 45-Jährigen waren beide mit einem Anteil von jeweils 18,1 % vertreten. Die bis 18-Jährigen stellten mit 3,3 % die kleinste Gruppe der Umfrage dar. In der Geschlechterverteilung lag der Anteil weiblicher Personen bei 46 % und männlicher bei 53,5 %. Ein Teilnehmender gab an, sich keinem Geschlecht zuzuordnen.

Beschäftigungsstatus: Die Analyse des Beschäftigungsstatus zeigt, dass 45,1 % der Befragten vollzeitbeschäftigt sind. 33 % sind Studierende oder Auszubildende. Teilzeitbeschäftigte stellen 12,1 % dar, Selbstständige oder Freiberufler 7,4 %. Die kleinsten 2,3 % sind aktuell nicht berufstätig.

Berufserfahrung: Im Rahmen der Berufsdauer hatten 27,9 % der Probanden weniger als 2 Jahre Erfahrung, 25,6 % zwei bis fünf Jahre. Personen mit sechs bis zehn Jahren machten 19,1 % aus, solche mit mehr als 20 Jahren 16,3 % und mit 11 bis 20 Jahren 11,2 %.

Remote-Berufserfahrung: Bei der Frage zur Remote-Berufserfahrung gaben 29,3 % der Probanden an, keine Erfahrungen damit zu haben. Jeweils 22,8 % berichteten, dass weniger als 25 % bzw. 25–50 % ihrer Berufserfahrung Remote-Arbeit umfasst. 13 % gaben an, dass 51–75 % ihrer Erfahrung remote waren, während 12,1 % mehr als 75 % angaben.

In Frage sechs erfolgte eine Anpassung der Stichprobengröße. Aufgrund der Relevanz der Remote-Berufserfahrung beziehen sich die folgenden Prozentangaben nicht mehr auf die 215 Teilnehmenden, sondern ausschließlich auf die 152 Teilnehmenden, die Remote-Erfahrungen in ihrem beruflichen Werdegang vorweisen können.

Aktuelle Arbeitsform – Remote, Hybrid, Büro: Bei der Frage zur Erfassung der aktuellen Arbeitsform lag der größte Anteil der Probanden mit Remote-Erfahrung mit 52,6 % bei den derzeit hybrid arbeitenden Teilnehmenden. 29 % arbeiteten vollständig remote und 18,4 % gaben an, vollständig im Büro zu arbeiten. Insgesamt gaben 19,74 % an, nicht remote zu arbeiten. Der Anteil der Probanden, die einen Tag pro Woche remote arbeiten, lag bei 9,87 %, für zwei Tage pro Woche bei 11 % und für drei Tage bei 18,42 %. 9,21 % der Teilnehmenden üben ihre berufliche Tätigkeit an vier Tagen pro Woche aus. Die größte Gruppe bilden diejenigen, die vollständig remote arbeiten (fünf Tage in der Woche), mit einem Anteil von 25,65 %.

Remote-Arbeit aus einem anderen Land: Auf die Frage, ob die Teilnehmenden an ihrem derzeitigen Arbeitsplatz die Möglichkeit haben, ihre Arbeit vom Ausland aus zu verrichten, gaben 53,9 % an, dass es ihnen nicht erlaubt ist. Im Vergleich dazu meinten 46,1 % der Teilnehmenden, dass es ihnen gestattet sei. Von den 70 Befragten, die im Ausland arbeiten dürfen, gaben 14,3 % an, dies für bis zu zwei Wochen zu tun. 41,4 % dürfen zwei bis sechs Wochen, 20 % sechs bis zwölf Wochen und 24,3 % mehr als drei Monate im Ausland arbeiten.

Länderübergreifende Unternehmenskommunikation: Bei der Frage, ob die Kommunikation im Unternehmen hauptsächlich auf nationaler oder internationaler Ebene stattfindet, lag der prozentuale Anteil der internationalen Kommunikation bei 32,9 % und der nationalen bei 67,1 %. Hinsichtlich der Angabe der genutzten Kommunikationstools konnten Teilnehmende mehrere Optionen wählen. Am häufigsten wurde E-Mail von 133 Personen genutzt, gefolgt von internen Chat- und Messaging-Plattformen mit 127 Nutzern. Telefon- und Videokonferenzsysteme zählten 98 Nutzer. 39 Personen verwendeten das Intranet oder firmeninterne Netzwerke, 33 setzten Projektmanagement- und Kollaborationstools ein, und ein Teilnehmer nutzte WhatsApp.

Unternehmensgröße und Durchschnittsalter: Bei der Erfassung der Unternehmensgröße gaben 20,4 % der Teilnehmenden an, in einem kleinen Unternehmen mit einem bis 50 Mitarbeitenden angestellt zu sein. 15,1 % erwiderten, in einem mittleren Unternehmen mit 51 bis 249 Mitarbeitenden tätig zu sein. Der prozentuale Anteil der Teilnehmenden, die in einem großen Unternehmen mit 250 bis 500 Mitarbeitenden arbeiten, lag bei 25 %. Der größte Teil, genauer 27 %, ist in einem großen Unternehmen mit 501 bis 10.000 Angestellten tätig. 6,6 % der Teilnehmenden gaben an, in einem Großunternehmen mit 10.001 bis 100.000 Mitarbeitenden zu arbeiten. Nur 5,9 % sind in einem Unternehmen mit mehr als 100.000 Mitarbeitenden angestellt.

Bei der Erfassung des Durchschnittsalters gaben 3,9 % an, dass dieses unter 25 Jahren lag. 20,4 % fielen in die Altersgruppe 25 bis 34 Jahre, während 59,9 % zwischen 35 und 44 Jahre alt waren. 14,5 % berichteten ein Durchschnittsalter von 45 bis 54 Jahren und nur 1,3 % zwischen 55 und 64 Jahren. Es gab keine Probanden, die ein Durchschnittsalter von über 65 Jahren angegeben haben.

Kommunikationsmittel und die Bewertung der Kommunikation: Bei der Auswahl der vorrangig genutzten Kommunikationsmittel wurde E-Mail mit 133 Nennungen am häufigsten verwendet. Interne Chat- und Messaging-Plattformen wie Slack und Microsoft Teams folgten mit 127 Nennungen. Telefon- und Videokonferenzsysteme erhielten 98 Nennungen, während Intranet oder firmeninterne soziale Netzwerke 39 und Projektmanagement- und Kollaborationstools 33 Bewertungen erhielten. WhatsApp wurde mit nur einer Bewertung am seltensten genutzt. Bei Frage 14 zur Bewertung der Kommunikationsmittel im Unternehmen fanden 53,9 % diese effizient. Jeweils 19,7 % bewerteten sie als sehr effizient bzw. neutral. Lediglich 4,6 % sahen sie als äußerst ineffizient an, während 2 % sie als ineffizient einstufte. In der Bewertung der Kommunikation während der Remote-Arbeit bewerteten mit 20,4 % der Probanden diese als sehr effizient und 50 % als effizient.

18,4 % sahen diese als neutral, während 8,6 % dies ineffizient und 2 % als sehr ineffizient bewerteten.

Zusammenarbeit: Von den 152 relevanten Befragten fanden 17,8 % die Teamarbeit sehr effizient und 65,1 % effizient. 13,9 % bewerteten sie als neutral, während jeweils 2 % sie als ineffizient bzw. sehr ineffizient ansahen. In Frage 18 lag der Wert für die Zustimmung, dass die Einführung von Remote-Arbeit die Zusammenarbeit verbessert, bei 8,6 %, die voll und ganz zustimmten. 28,9 % stimmten nur zu. Fast die Hälfte der Teilnehmenden (46,1 %) gab an, dass sich die Zusammenarbeit durch die Einführung verbessert oder verschlechtert hat. 12,5 % stimmten nicht zu, dass sich die Zusammenarbeit nicht verbessert und 3,9 % stimmten überhaupt nicht zu. Bei Maßnahmen zur Stärkung des Teamgeists während der Remote-Arbeit waren regelmäßige virtuelle Teamtreffen mit 107 Nennungen am beliebtesten. Informelle Treffen folgten mit 33 Nennungen, und Workshops sowie das Fehlen spezieller Maßnahmen erhielten jeweils 31 und 30 Nennungen. Die Anzahl der Online-Meetings liegt für teambildende Maßnahmen mit 5,9 % bei einmal pro Tag. 33,6 % meinten, ein- bis zweimal pro Woche Online-Meetings abzuhalten. Mit 2,6 % war die Anzahl der Meetings drei- bis viermal pro Woche am niedrigsten. 15,8 % gaben an, alle zwei Wochen ein Treffen vorzunehmen, und 3,9 % haben alle drei Wochen ein Meeting. Mit 38,2 % ist der Wert bei den Teilnehmenden, die nur einmal im Monat ein Meeting zu Teambildungsmaßnahmen abhalten, am höchsten.

Unternehmenskultur: Auf die Frage, ob Remote-Arbeit die Unternehmenskultur stärkte, stimmten 4,6 % voll und 27 % teilweise zu. Die Mehrheit, 47,4 %, sah keinen Unterschied, während 14,5 % nicht zustimmten und 6,6 % vollständig ablehnten. Von den 152 Befragten waren 26,3 % überzeugt, dass Remote-Arbeit die Unternehmenswerte verbessert hat. Für 54,6 % blieb die Einhaltung der Unternehmenskultur unverändert. 11,8 % sahen eine Erschwernis, während 7,2 % Veränderungen durch Remote-Arbeit bemerkten.

Abschließend gaben 7,6 % der Teilnehmenden in der letzten Frage an, dass sie sich keine Remote-Arbeit im beruflichen Umfeld wünschen. Der Anteil der Probanden, die diese gelegentlich in Anspruch nehmen möchten, lag bei 12,1 %. 22 % der Teilnehmenden führten an, dass sie 50 % ihrer Arbeit remote erledigen wollen. Der höchste Anteil lag mit 31,1 % bei den Probanden, die sich überwiegend Remote-Arbeit wünschen, die 50,75 % der Arbeit ausmachen würde. Schließlich gaben 27,3 % an, zu 100 % remote arbeiten zu wollen.

4.2 Hypothesenprüfung

Dieser Abschnitt zeigt das statistische Vorgehen zur Untersuchung der Hypothesen sowie die anschließende Hypothesenprüfung mit den daraus resultierenden Ergebnissen auf.

Statistisches Vorgehen

Zur Überprüfung der Hypothesen wurden der Chi-Quadrat-Test, der t-Test für unabhängige Stichproben, die Pearson-Korrelationsanalyse und die einfaktorielle Varianzanalyse

verwendet, mit einer Signifikanzschwelle von 5% ($p < 0,05$). Die Voraussetzungen für diese Verfahren sind in Tabelle 1 aufgeführt.

Testverfahren	Voraussetzungen
Chi-Quadrat-Test	-Nominales Skalenniveau -Die erwartete Häufigkeit pro Zelle soll bei 20 % aller Zellen größer als 5 sein.
t-Test für unabhängige Stichproben	-Unabhängigkeit der Messungen -Abhängige Variable ist intervallskaliert -Unabhängige Variable ist dichotom -Ausreißerfrei -Normalverteilung der Variablen
Einfaktorielle Varianzanalyse (ANOVA)	-Unabhängigkeit der Messungen -Die abhängige Variable ist mindestens intervallskaliert. -Die unabhängige Variable ist unabhängig und nominalskaliert. -Die abhängige Variable ist für jede Gruppe (etwa) normalverteilt. -Ausreißerfrei
Korrelationsanalyse nach Pearson	-Beide Variablen sind mindestens intervallskaliert.

Tab. 1: Statistische Tests

Alle Hypothesen wurden mittels Chi-Quadrat-Test überprüft. Für Hypothesen 1, 2, und 7 erfolgte zusätzlich eine Pearson-Korrelationsanalyse bei einem Alpha-Niveau von $\alpha = 0,05$. Hypothese 3 wurde mit einem t-Test für unabhängige Stichproben getestet, und H-5 sowie H-6 mit einer einfaktoriellen Varianzanalyse (ANOVA). Vorab wurden Normalverteilung und Varianzhomogenität (mittels Kolmogorov-Smirnov-Anpassungstest und Levene-Test) geprüft, die nicht signifikant sein sollten ($p > 0,05$). Signifikante Mittelwertunterschiede wurden mittels Bonferroni Post-hoc-Test analysiert. Die statistische Auswertung erfolgte durch Microsoft Excel und IBM SPSS Statistics.

Ergebnisse

Im Rahmen der statistischen Auswertung werden in diesem Abschnitt die Fragen aus der Umfrage dargestellt, die miteinander verglichen werden. Für die erste Hypothese H-1 werden die Fragen 14 (Wie bewerten Sie die Kommunikationstools in Ihrem Unternehmen?) und 16 (Wie würden Sie die Kommunikation während der Remote-Arbeit in Ihrem Unternehmen bewerten?) gegenübergestellt. Bei der zweiten Hypothese H-2 werden die Aussage 22 (Die Unternehmenskultur ist durch Remote-Arbeit stärker geworden.) und die Frage 23 (Wie beeinflusst die Umstellung auf Remote-Arbeit Ihrer Meinung nach die Einhaltung und Umsetzung der Werte der Unternehmenskultur?) betrachtet. Darüber hinaus werden für Hypothese H-3 die Fragen 7 (Arbeiten Sie aktuell im Büro, hybrid oder remote (Home-Office etc.)?) und 14 (Wie bewerten Sie die Qualität der Zusammenarbeit im Team oder in ihrem Unternehmen?) statistisch ausgewertet. Folglich werden für Hypothese H-4 die Fragen 20 (Wie oft finden in Ihrem Team Online-Meetings zur Teambildung statt?) und 21 (Wie empfinden Sie diese Online-Meetings?) gegenübergestellt. Anschließend werden für Hypothese H-5 die Fragen 11 (Größe des Unternehmens: Wie viele Mitarbeiter hat ihr Unternehmen?) und 14 (Wie würden Sie die Kommunikation während der Remote-Arbeit in Ihrem Unternehmen bewerten?) betrachtet. Für die letzte Hypothese H-6 werden die Fragen 12 (Was ist das geschätzte Durchschnittsalter der

Mitarbeitenden in Ihrem Unternehmen?) und 16 (Wie würden Sie die Kommunikation während der Remote-Arbeit in Ihrem Unternehmen bewerten?) im Vergleich untersucht.

Hypothese 1: Der durchgeführte Chi-Quadrat-Test ergab keinen statistischen signifikanten Zusammenhang ($\chi^2 = 24,247$; $p = 0,084$) zwischen den Fragen 14 und 16. Der p-Wert ist nur auf 10 % signifikant. Zudem ist dieses Ergebnis kritisch zu hinterfragen, da 76 % der Zellen eine erwartete Häufigkeit < 5 aufweisen, was zu ungenauen Testergebnissen führen kann. Auch die Korrelationsanalyse nach Pearson zeigt einen Koeffizienten von 0,069, der mit $p = 0,463$ nicht signifikant ist. Daraus kann geschlossen werden, dass kein Zusammenhang zwischen den untersuchten Variablen besteht.

Hypothese 2: Der Chi-Quadrat-Test zeigt einen statistisch signifikanten Zusammenhang ($\chi^2 = 108,711$; $p = < ,001$) zwischen der Aussage 22 und der Frage 23. Es ist zu beachten, dass 55 % der Zellen eine erwartete Häufigkeit von weniger als fünf aufweisen. Dies könnte die Genauigkeit der Testergebnisse und damit die Zuverlässigkeit der Untersuchung einschränken. Ergänzend dazu zeigt die Korrelationsanalyse nach Pearson eine signifikant schwache negative Korrelation zwischen den beiden Variablen. Dies deutet darauf hin, dass der Anstieg des Wertes einer Variable mit einem Rückgang des Wertes einer anderen Variable einhergeht.

Hypothese 3: Der Chi-Quadrat-Test lässt keinen statistisch signifikanten Zusammenhang ($\chi^2 = 4,870$; $p = ,771$) zwischen der Aussage 7 und der Frage 14 erkennen. Die Voraussetzungen für den t-Test für unabhängige Stichproben sind mit Ausnahme der Normalverteilung erfüllt. Dies ergibt sich aus dem Kolmogorov-Smirnov-Anpassungstest $= < 0.001$, der eine signifikante Abweichung von der Normalverteilung anzeigt. Dies beeinträchtigt die Genauigkeit nur geringfügig, sodass die Analyse auch ohne die Normalverteilung durchgeführt werden kann. Die Analyse zeigt, dass Personen, die im Büro arbeiten, die Teamarbeit mit einem Wert von 4,04 bewerten und hybride oder remote Arbeitende durchschnittlich mit einem Wert von 3,93. Dieser Unterschied ist statistisch nicht signifikant ($T = 0,686$; $df = 150$; $p = 0,247$). Auch der Vergleich zwischen im Büro Arbeitenden und nur remote Tätigen offenbart keine Unterschiede in der Bewertung der Teamarbeit. Der Mittelwert für im Büro Arbeitende liegt bei $M = 4,04$ und für remote Arbeitende bei $M = 4,00$ ($T = ,230$; $df = 70$; $p = ,410$). Folgend resultiert aus dem Test, dass es keinen signifikanten Unterschied in der Bewertung der Teamarbeit zwischen den im Büro Arbeitenden und den remote Beschäftigten gibt.

Hypothese 4: Das Resultat des Chi-Quadrat-Tests verdeutlicht, dass kein statistisch signifikanter Zusammenhang ($\chi^2 = 28,001$; $p = ,109$) zwischen den Fragen 20 und 21 besteht.

Hypothese 5: Die Resultate des Chi-Quadrat-Tests zeigen keinen statistisch signifikanten Zusammenhang zwischen den Fragen 11 und 14. Es wurde eine einfaktorielle ANOVA durchgeführt, da die Voraussetzungen für eine weitere Überprüfung gegeben waren. Die Ergebnisse der einfaktoriellen ANOVA weisen auf keine signifikanten Mittelwertunterschiede der Bewertung von Remote-Arbeit zwischen den verschiedenen Unternehmensgrößen hin ($F = ,766$; $p = ,576$). Daraus folgt, dass die Einschätzung der Kommunikation der Remote-Arbeit unabhängig von der Unternehmensgröße ist.

Hypothese 6: Die Resultate des Chi-Quadrat-Tests deuten auf einen statistisch signifikanten Zusammenhang ($\chi^2 = 20,567$; $p = < ,001$) zwischen den Fragen 12 und 16 hin. Die ANOVA zeigt signifikante Mittelwertunterschiede in der Bewertung der Telearbeit zwischen den mittleren Altersgruppen der Unternehmen ($F = 31,07$; $p = ,030$). Die Post-hoc-Analyse mittels Bonferroni lässt erkennen, dass der Unterschied in der Bewertung besonders zwischen den Altersgruppen 3 und 4 ausgeprägt ist. Die mittlere Differenz zwischen diesen Altersgruppen beträgt $M_{diff} = -,688$ und ist mit einem $p = ,038$ signifikant.

4.3 Interpretation der Ergebnisse

H-1: Die erste Hypothese betrachtet die Effizienz der Kommunikation und postuliert, dass Mitarbeitende, die Unternehmenskommunikationswerkzeuge positiv bewerten, auch die Remote-Kommunikation besser bewerten. Die statistische Analyse widerlegt jedoch diese Annahme und zeigt keinen signifikanten Zusammenhang. Die theoretisch abgeleiteten Einflussfaktoren scheinen somit nur eine geringe Rolle zu spielen. Die Wahl des Kommunikationsmediums beeinflusst die Kommunikationseffizienz bei Remote-Arbeit nicht signifikant, vielmehr hängt dies von Faktoren wie Teamstruktur und Unternehmenskultur ab.

H-2: Die Einführung von Remote-Arbeit beeinflusst signifikant die Unternehmenskultur, jedoch nicht immer positiv. Eine Pearson-Analyse offenbart einen negativen Zusammenhang, der auf Schwierigkeiten bei der Aufrechterhaltung von Unternehmenskultur und -werten hinweist. Der Mangel an direkter Face-to-Face-Kommunikation kann den Zusammenhalt und die Verbundenheit zu den Unternehmenswerten schwächen. Dies kann eine Diskrepanz zwischen alten Werten und neuen Arbeitspraktiken verursachen und durch digitale Kommunikation können Missverständnisse auftreten, die eine effektive Vermittlung der Unternehmenswerte erschweren.

H-3: Die Untersuchung der Hypothese H-3 zeigt keinen signifikanten Unterschied in der wahrgenommenen Qualität der Teamarbeit zwischen Mitarbeitenden, die hybrid oder remote und solchen, die im Büro arbeiten. Statistisch liegen die Mittelwerte der Bewertungen nahe beieinander: Hybrid- oder Remote-Mitarbeiter mit 3,93, ausschließlich Remote-Arbeitende mit 4,00, und Büroarbeitende mit 4,04. Dies verdeutlicht, dass die Arbeitsortwahl weniger Einfluss auf die Teamarbeit hat, als andere Faktoren. Unterstützt wird diese Erkenntnis durch theoretische Überlegungen über den Aufbau eines Teams. Die Struktur und die Dynamik des Teams, sowie eine ausgewogene Rollen- und Kompetenzverteilung, sind entscheidend. Insbesondere eine balancierte Mischung aus Fach-, Sozial- und Zielpromotoren fördert die Qualität der Zusammenarbeit. Es wird auch hervorgehoben, dass der Arbeitsort nicht vollständig vernachlässigt werden sollte. Durch Remote-Arbeit geschaffene Flexibilität kann ein entspannteres Arbeitsumfeld und somit bessere Ergebnisse ermöglichen. Zukünftige Untersuchungen sollten daher eine breitere Palette von Faktoren berücksichtigen, die effektive Teamarbeit beeinflussen, anstatt sich auf einzelne Aspekte zu konzentrieren.

H4: Hypothese H-4, die einen positiven Einfluss regelmäßiger Teambesprechungen auf den Teamgeist in der Remote-Arbeit annimmt, wurde statistisch nicht bestätigt. Die Untersuchung ergab keinen signifikanten Zusammenhang zwischen der Häufigkeit der Meetings und einer Stärkung des Teamgeistes. Vielmehr scheinen Qualität der Interaktion, inhaltliche Ausrichtung und Teamdynamik entscheidendere Faktoren zu sein.

H-5: Bei der Untersuchung der Hypothese H-5, die den Vergleich von Unternehmensgrößen betrifft und besagt, dass die Kommunikation bei Remote-Arbeit in Großunternehmen positiver bewertet wird, zeigte sich, dass es keine statistisch signifikanten Mittelwertunterschiede gab. Es ist zu erkennen, dass Faktoren, wie bereits in Kapitel 2 thematisiert wurden, beispielsweise die Unternehmenskultur, die Kommunikationsstrategie, die eingesetzten Kommunikationswerkzeuge, die Teamstruktur sowie regelmäßige produktive Teambesprechungen zur Stärkung des Teamgeistes und vieles mehr, signifikanter für die Analyse der Kommunikationseffizienz unterschiedlicher Unternehmen sind. Daraus resultiert, dass trotz eigener Strategien und Ansätze, für die Unternehmen ähnliche Herausforderungen und Chancen in Bezug auf Remote-Arbeit bestehen.

H-6: Die Ergebnisse des Chi-Quadrat-Tests der ANOVA und der Post-hoc-Analyse zeigen einen signifikanten statistischen Zusammenhang zwischen dem Durchschnittsalter und der Bewertung der Kommunikation bei Remote-Arbeit im Unternehmen. Es wird deutlich, dass jüngere Mitarbeitende tendenziell positiver zur Remote-Arbeit eingestellt sind und dadurch auch die Kommunikationseffizienz steigt. Die Ergebnisse stehen im Einklang der theoretischen Betrachtung. Die Verschmelzung von Arbeits- und Privatleben verbessert das Gleichgewicht beider Bereiche, wobei das Zuhause zunehmend als effektiver Arbeitsplatz gesehen wird. Dies bietet Vorteile wie Flexibilität und Zeitersparnis durch den Wegfall des Pendelns. Für die jüngere Generation ist besonders die Möglichkeit, international zu arbeiten, attraktiv. Jedoch zeigt die Umfrage, dass nur 46,1 % der Befragten im Ausland arbeiten dürfen, häufig eingeschränkt durch rechtliche Hürden. Unternehmen müssen daher angepasste Remote-Arbeitsstrategien entwickeln, die auf die Bedürfnisse aller Altersgruppen eingehen. Für ältere Mitarbeitende könnten zusätzliche Schulungen erforderlich sein. Laut Hypothese H-6 beeinflusst das Alter signifikant die Kommunikationseffizienz, was eine gezielte Berücksichtigung der unterschiedlichen Bedürfnisse der Altersklassen zur Schaffung einer effektiven Arbeitsumgebung erfordert.

5 Kritische Würdigung

Obwohl die Umfrage wichtige Erkenntnisse liefert, zeigt die Datenerhebung Grenzen auf. Von den ursprünglich 215 ausgefüllten Fragebögen, wurden nur 152 berücksichtigt, da nur diese über relevante Erfahrungen mit Remote-Arbeit verfügten. Um repräsentativere Ergebnisse zu erzielen, sollte eine genauere Zielgruppe definiert und geeignete Kanäle zur Erreichung relevanter Probanden genutzt werden. Zudem ist die Diversifizierung der Teilnehmergruppen ein wichtiger Aspekt. Weiterhin waren in der Befragung viele Branchen vertreten. Zukünftige Untersuchungen sollten spezifischere Branchen fokussieren, um Kommunikationseffizienz und Kooperation genauer zu analysieren.

Durch präzisere Fragestellungen könnten Unterschiede zwischen den Branchen identifiziert und effiziente Lösungen für branchenspezifische Herausforderungen entwickelt werden. Auch die Antwortmöglichkeiten waren zu stark kategorisiert, was die Datenvielfalt für die statistische Auswertung limitierte. Ein konkretes Beispiel ist die Erfassung des Alters. Es wäre sinnvoller gewesen, das Alter in einem Eingabefeld zu erfassen, um deskriptive statistische Analysen durchzuführen und diese anschließend statistisch auszuwerten sowie in Bezug zu anderen Fragestellungen zu setzen. So hätte das Durchschnittsalter zu noch genaueren Ergebnissen bezüglich der Hypothese H-6 geführt. Vor diesem Hintergrund sollten sich Unternehmen stärker auf die Analyse der Bedürfnisse verschiedener Altersgruppen konzentrieren. Durch gezielte interne Befragungen können Anforderungen, Wünsche und Präferenzen erfasst werden. Besonders relevant im Kontext von Hypothese H-6 ist Frage 16, die die Einschätzung der Kommunikation im Unternehmen während der Remote-Arbeit betrifft, da sie nicht nur in H-6, sondern auch in H-1 und H-5 zur Überprüfung herangezogen wird. Es wird empfohlen, die Fragestellung für ein tieferes Verständnis der Remote-Kommunikation in spezifischere Fragen aufzugliedern. Da die interne Kommunikation von vielen Faktoren wie den emotionalen Zuständen der Beschäftigten abhängt, könnte eine detaillierte Analyse von Zufriedenheit, Einsamkeit und Stress die Effizienz der Kommunikation verbessern.

Literaturverzeichnis

- [Bel21] Belling, S.: Remotely possible: Strategic Lessons and Tactical Best Practices for Remote Work. Apress Media LLC, Fitchburg, WI, 2021.
- [Böh21] Böhlich, S.: Projektmanagement im Gesundheitswesen: Ein praxisorientierter Leitfaden. Springer Fachmedien, Wiesbaden, 2021.
- [Bri20] Britz-Averkamp, I.; Eich-Fangmeier, C.: ÜBERLEBEN NEUEN ARBEITSWELT: Desk Sharing, Open Space, Mobiles Arbeiten & Co. Survival Guide für Manager und Mitarbeiter. Walhalla u. Praetoria Verlag, Regensburg, 2020.
- [Bru20] Bruhn, P.: Homeoffice und mobiles Arbeiten im Team effektiv umsetzen: Praxisratgeber: Remote Work und Heimarbeitsplatz technisch schnell einrichten. Springer Fachmedien, Wiesbaden, 2020.
- [Her21] Herget, J.: Digitale Unternehmenskultur: Strategien für die moderne Arbeitswelt. Springer Fachmedien, Wiesbaden, 2021.
- [Hom14] Homma, N. et al.: Unternehmenskultur – eine Einleitung, in: Einführung Unternehmenskultur Grundlagen, Perspektiven, Konsequenzen. Springer Fachmedien, Wiesbaden, S. 1-14, 2014.
- [Mei12] Meier, R.: 30 Minuten Teamarbeit, 5. Aufl. GABAL Verlag, Offenbach, 2012.
- [Möh13] Möhring, W.; Schlütz, D.: Standardisierte Befragung – Messmethodik und Designs in der Medienwirkungsforschung. In: Schweiger, W. (Hrsg.): Handbuch – Medienwirkungsforschung. Springer VS, Wiesbaden, S. 565-579, 2013.
- [OPM24] OPM.gov, <https://www.opm.gov/frequently-asked-questions/telework-faq/remote-work/what-is-the-definition-of-remote-work/>, Stand: 19.03.2024, 20:55 Uhr.
- [Ric15] Richter, F. et al.: Die Unternehmenskultur als kritischer Erfolgsfaktor – Erfindung oder Wirklichkeit?, in: Frank, Richter. (Hrsg.): Wettbewerbsfaktor Unternehmenskultur: Wie Unternehmenskulturen den Erfolg beflügeln oder lähmen. Schäffer-Poeschel Verlag für Wirtschaft · Steuern · Recht GmbH, Stuttgart, S. 1-3, 2015.
- [Sch14] Scholl, A.: Die Befragung [electronic resource], 3. überarb. Aufl. UTB GmbH, 2014.
- [Sch16] Schwinning, G.: Kommunikation, Führung und Zusammenarbeit in Unternehmen: Wahre Situationen und handfeste Lösungen, 1. Aufl. Haufe-Lexware, Freiburg · München · Stuttgart, 2016.
- [Sie24] Siepermann, M.; Lackes, R.; Maier, G. W.: Kommunikation. In Gabler Wirtschaftslexikon, <https://wirtschaftslexikon.gabler.de/definition/kommunikation-37167>, Stand: 19.03.2024, 21:26 Uhr.
- [vAm21] Von Ameln, F.; Buckel, C.: Digitalisierung. Zeitschrift für Psychodrama und Soziometrie, 20(2), S. 187-194, <https://doi.org/10.1007/s11620-021-00616-x>, 2021.

Implementierung eines simulationsbasierten Digitalen Prozess-Zwillings als Entscheidungsunterstützungssystem

Vollständiger Beitrag

Corinna Huber¹, Lukas Rissmann² und Sebastian Meißner³

Abstract: This paper deals with the implementation of a simulation-based Digital Process Twin as a decision support system. The approach is based on a framework developed by Zuhr et al. The framework provides a two-part architecture consisting of a data supply unit and an information control unit, enabling the integration of operational data and supporting services such as state monitoring, process analysis, and process improvement. The digital twin is visualized through a dashboard, presenting recommendations and their implications. An implementation example demonstrates the application of this framework in a production environment. Discussions on the effectiveness and challenges of implementation, as well as prospects for further research approaches, conclude the paper.

Keywords: Digital twin, simulation, logistics process, dashboard, Industry 4.0

1 Einleitung

In der wissenschaftlichen Literatur lassen sich aufbauend auf den unterschiedlichen theoretischen Überlegungen seit der Einführung des Konzepts der Digitalen Zwillinge 2002 durch *Grieves* [Grie16] verschiedene Definitionen des Digitalen Zwillings finden. Diese Definitionen variieren in ihren Schwerpunkten abhängig von den Anwendungsbereichen und den Zielgruppen. In diesem Artikel werden Digitale Prozess-Zwillinge definiert als eine funktionale, virtuelle und dynamische Repräsentation eines realen Prozesses. Sie bestehen aus Diensten sowie relevanten Modellen und wird mit aktuellen Betriebsdaten gespeist, um Informationen für Entscheidungsunterstützung oder Prozessoptimierung bereitzustellen [Zuhr21].

Entscheidungen basieren meist auf einer Vielzahl an gleichzeitig anfallenden Informationen [FRBM17]. Die Menge an Daten und die Komplexität ihrer Zusammenhänge können Menschen beim Treffen schneller Entscheidungen teilweise überfordern. Entscheidungsunterstützungssysteme können die Anwender bei der

¹Technologiezentrum für Produktions- und Logistiksysteme, Hochschule Landshut, Bräuhausgasse 33, 84130 Dingolfing, <https://orcid.org/0009-0007-6555-5321>

²Technologiezentrum für Produktions- und Logistiksysteme, Hochschule Landshut, Bräuhausgasse 33, 84130 Dingolfing, <https://orcid.org/0000-0002-9747-7707>

³Technologiezentrum für Produktions- und Logistiksysteme, Hochschule Landshut, Bräuhausgasse 33, 84130 Dingolfing, sebastian.meissner@haw-landshut.de, <https://orcid.org/0000-0002-5808-9648>

Entscheidungsfindung unterstützen oder teilweise Entscheidungsprozesse vollständig übernehmen [BCKM09]. Entscheidungsunterstützungssysteme sind wesentliche Bestandteile der Mensch-Maschine-Systemtechnik, die sich durch die informationelle Verbindung von Maschinen und Anwendern auszeichnet [TiBa02]. Digitale Zwillinge können als Entscheidungsunterstützungssysteme angewendet werden. Sie verwenden vorhandene Daten, um zukünftige Prozesse und deren Auswirkungen im Kontext neuer Herausforderungen zu analysieren und zu prognostizieren [KoSc23].

Dieser Artikel beschreibt die beispielhafte Umsetzung des von Zuhr et al. entwickelten Frameworks zur Entwicklung eines simulationsbasierten Digitalen Prozess-Zwilling. Dieses Framework unterstützt die Integration von Betriebsdaten und nutzt simulationsbasierte Methoden zur Bewertung von Zukunftsszenarien für die Unterstützung von Entscheidungsprozessen.

Ein realisiertes Anwendungsbeispiel in der Lern- und Musterfabrik des Technologiezentrum Produktions- und Logistiksysteme (TZ PULS) belegt die Umsetzbarkeit des Frameworks und den Nutzen von simulationsbasierten Digitalen Zwillingen als Entscheidungsunterstützungssystem. Ein Dashboard dient zur Visualisierung der Simulationsergebnisse, der daraus resultierenden Handlungsempfehlungen und deren Auswirkungen. Die Veranschaulichung erleichtert die Interaktion mit dem Digitalen Zwilling, was wiederum schnelle Reaktionen und Entscheidungsfindungen auf Basis realer Daten ermöglicht.

2 Digitale Zwillinge

Die in der Literatur beschriebenen Konzepte zu Anwendung Digitaler Zwillinge unterscheiden sich stark hinsichtlich ihrer Anwendungsbereiche und der technischen Umsetzungen [JSNY20, SFHQ21]. Dabei werden unterschiedliche Konzepte und Definition von Digitalen Zwillingen beschrieben, welche nachfolgend diskutiert werden.

Korth et al. entwickelten ein Konzept simulationsbasierter Digitaler Zwillinge zur Optimierung von Logistiksystemen, mit einem Fokus auf Datenerfassung durch Sensoren und spezifischen Datenschnittstellen. Der Ansatz umfasst eine Architektur mit Ereignissteuerung, Modellierung, Simulation und Berichterstattung, die sowohl auf Personal- als auch auf Produktionsplanung angewendet wird. Ein automatisches Palettenregallager und Kommissionierbereiche in einem Verteilzentrum dienen als praktisches Beispiel, um die Vorteile einer gleichmäßigen Mitarbeiterauslastung aufzuzeigen. Der Digitale Zwilling dient hier als Entscheidungsunterstützung, indem er Anpassungen von Arbeitsplänen basierend auf vorherigen Erfahrungen durch die Anwender ermöglicht. Trotz der Präsentation von Ergebnisdaten wird die Übertragbarkeit des Konzepts auf komplexere Systeme in den Arbeiten nicht näher beschrieben [KoSZ18].

Kritzinger et al. ordnen die verschiedenen Ansätze in Kategorien wie Digitale Modelle, Digitale Schatten und Digitale Zwillinge ein, wie Abbildung 1 zeigt [KTH18].

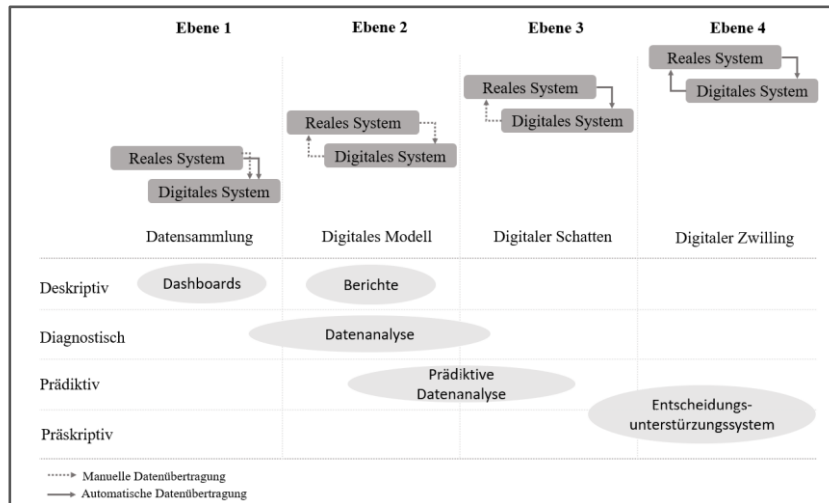


Abb. 1: Ebenen digitaler Abbildungen (eigene Darstellung nach [KKTH18])

Abbildung 1 zeigt die verschiedenen Ebenen der digitalen Abbildungen eines Realsystems. Auf der ersten Ebene, der „Datensammlung“-Ebene, werden die Daten des Realsystems auf einem Dashboard angezeigt. Unterstützen digitale Abbildungen die Generierung von Berichten sowie die retrospektive und beschreibende Untersuchung von Systemen so werden diese der zweiten Ebene zugeordnet. Solche Analysen werden zuerst nur periodisch und für bestimmte Fragestellungen durchgeführt, was vor allem in der strategischen Planung Anwendung findet. Digitale Schatten, die der dritten Ebene zuzuordnen sind, ermöglichen einen automatisierten Datentransfer von der physischen in die virtuelle Welt, wodurch umfangreiche Datenmengen analysiert werden können. Allerdings ist die Modellierung von Systemabhängigkeiten hier weniger detailliert und nicht zukunftsgerichtet im Vergleich zu Digitalen Zwillingen. Dies befähigt Digitale Schatten besonders für die Untersuchung spezifischer Systemsegmente (beispielsweise einzelner Maschinen). Ihre umfangreiche Datenbasis unterstützt Prognosen, etwa im Hinblick auf den Einsatz von Ressourcen. Digitale Zwillinge in der letzten Ebene zeichnen sich ebenfalls durch eine große Datenmenge aus. Digitale Zwillinge verfügen häufig über ein Simulations- oder Prognosemodell, die die erforderlichen Daten automatisch verarbeiten und somit präskriptive Analysen ermöglicht [KKTH18]. Ein Merkmal Digitaler Zwillinge ist der bidirektionale Informationsfluss zwischen dem physischen System und seiner digitalen Abbildung. Die Frequenz, mit der diese Datenübertragungen stattfinden, ist dabei nicht festgelegt [HaMG19]. Der gegenseitige Datenaustausch simulationsbasierter Digitaler Zwilling ermöglicht es, den aktuellen Zustand des Systems abzubilden und die Umsetzbarkeit von Entscheidungen vorab durch Simulationen im digitalen Modell zu prüfen. [TCQZ18].

Zuhr et al. beschreiben einen Digitalen Prozess-Zwilling (DPZ) als ein funktionales, virtuelle und dynamisches Modell eines realen Prozesses. Der DPZ besteht aus einer Kombination von Diensten und Modellen und wird kontinuierlich mit aktuellen Betriebsdaten versorgt. Der DPZ dient als Überwachungs- und Steuerungssystem, das in drei mögliche Hauptdienste gegliedert werden kann: Zustandsüberwachung, Prozessanalyse und Prozessverbesserung. Bei der Zustandsüberwachung kann der Zustand von Hardware- und Softwarekomponenten überwacht werden, um die Verfügbarkeit und Leistung sicherzustellen. Die Prozessanalyse identifiziert Schwachstellen und gibt Warnungen aus, während sie gleichzeitig zukünftige Zustände vorhersagt, um präventive Maßnahmen zu erleichtern. Die Prozessverbesserung nutzt Simulationswerkzeuge und optional künstliche Intelligenz, um die Produktivität zu steigern und ermöglicht die Selbststeuerung durch regelbasierte Algorithmen.

Darüber hinaus wird eine zugehörige Architektur für den Digitalen Prozess-Zwilling definiert, welche in Abbildung 2 vereinfacht dargestellt ist.

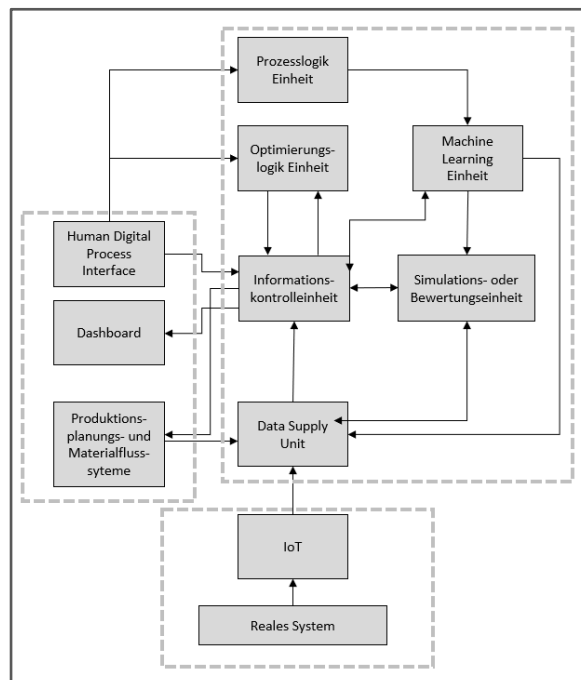


Abb. 2: Rahmenwerk Digitaler Prozess-Zwilling (eigene Darstellung nach Zuhr [ZuRM22])

Der DPZ ist an vier Stellen mit seiner Umgebung verbunden. Es erfasst dabei Daten direkt aus der physischen Umgebung (reales System) und ebenso mit bestehenden Produktionsplanungs- und Materialflusssystemen wie dem Enterprise Resource Planning (ERP) und dem Warehouse Management System (WMS) verbunden, um Operations-, Planungs- und Materialflussinformationen zu integrieren. Weiter ermöglicht ein Human

Digital Process Twin Interface (HDPTI) Nutzern die Interaktion mit dem System und visualisiert wichtige Informationen über ein Dashboard. Die Architektur des DPZ setzt sich aus der Datenversorgungseinheit und der Informationssteuerungseinheit zusammen. Die Datenversorgungseinheit ist für das Sammeln, Aufbereiten, Speichern und Bereitstellen von Daten zuständig, während die Informationssteuerungseinheit die Daten analysiert und auf Basis dieser Analysen Handlungsempfehlungen präsentiert. Diese Entscheidungsfindung basiert auf Algorithmen und Leistungsindikatoren (KPIs). Das Rahmenwerk erlaubt durch seine modulare Struktur die Integration zusätzlicher Dienste wie Simulationen und maschinelles Lernen, je nach spezifischem Bedarf. Durch die bidirektionale Verbindung zwischen Datenversorgungs- und Informationssteuerungseinheit wird ein kontinuierlicher Datenfluss sichergestellt, der eine dynamische Anpassung und Optimierung der Prozesse ermöglicht [ZuRM22].

Simulationsbasierter Digitaler Zwilling als Entscheidungsunterstützungssystem

Als Simulation wird die Durchführung von Experimenten an einem Modell zur Bewertung und Optimierung realer oder hypothetischer Systeme verstanden. Ein Simulationsmodell ist dabei eine abstrahierte Darstellung des zu untersuchenden Systems, in der dessen Struktur oder Verhalten mit reduziertem Detaillierungsgrad abgebildet wird. Dies ermöglicht es, durch Experimente am Modell Rückschlüsse auf das reale System zu ziehen und unterstützt Entscheidungsprozesse durch die Analyse verschiedener Systemvarianten. [Eley12]

Aufbauend auf den Definitionen von Simulation und Simulationsmodellen, zeigen sich simulationsbasierte Digitale Zwillinge als spezialisierte Anwendungen dieser Konzepte, die in der Praxis zur Unterstützung von operativen Entscheidungen eingesetzt werden.

Simulationsbasierte Digitale Zwillinge sind effektive Werkzeuge die vorab genutzt werden können, um Entscheidungen zu bewerten [KaGF21]. Diese digitalen Modelle können operative Abläufe unterstützen, wie beispielsweise die Planung des Ressourceneinsatzes zum Beispiel von Material und Mitarbeiter. Bei ihrer Entwicklung ist es wichtig, den richtigen Detaillierungsgrad für die Abbildung des realen Systems festzulegen [SGMM23]. Es gilt, die wesentlichen Entscheidungsparameter zu identifizieren, damit der Digitale Zwilling effektiv operative Entscheidungen unterstützen kann [VaDe11].

Die jeweilige Komplexität des Systems erfordert eine sorgfältige Auswahl des Detaillierungsgrads für die digitale Abbildung, damit die erforderliche Qualität der Ergebnisse erreicht werden kann. Brooks in Robinson et al. heben hervor, dass die Beziehung zwischen der Leistung des Modells und der Detailgenauigkeit noch wenig erforscht ist und weitere Untersuchungen erfordert. Die Autoren nehmen an, dass ein Modell mit geringerem Detailgrad genauso wirkungsvoll sein kann, da der Nutzen weiterer Einzelheiten über einen bestimmten Punkt hinaus nicht mehr zur Genauigkeit der Ergebnisse beiträgt [RBKV10].

3 Realisierung des Digitalen Prozess-Zwillings in der Lern- und Musterfabrik des TZ PULS

Im Rahmen des Projekts Process InnovationCenter (PRINCE) werden simulationsbasierte Digitale Prozess-Zwillinge als Entscheidungsunterstützungssystem erforscht, wobei verschiedene Simulationsmodelle mit unterschiedlichen Detailierungsgraden erstellt wurden. Der folgende Abschnitt bietet einen Überblick über den Entwicklungsprozess eines Digitalen Prozess-Zwillings. Es wird dargestellt, wie digitale Modelle schrittweise zu simulationsbasierten Digitalen Zwillingen erweitert werden können. Zudem wird die Visualisierung des Digitalen Zwillings mithilfe eines Dashboards beschrieben, das die Erkenntnisse des Zwillings für die Anwender darstellt. Der Anwendungsfall wird in der U-Zelle der Lern und Musterfabrik des TZ-PULS umgesetzt. Hier werden Industrie 4.0-Technologien wie Robotik, IoT-Komponenten und intelligente Sensorik, um Produktionslogistikprozesse zu simulieren und zu optimieren, erforscht und angewendet. Auf einer Fläche von 900 m² wird eine Wertschöpfungskette abgebildet, die von Wareneingang über verschiedene Fertigungsbereiche bis zum Warenausgang reicht. Unter den Fertigungsbereichen befindet sich eine U-Zelle zur Fertigung von Bodenrollern. Neben der Forschung dient das Technologiezentrum auch als Ort der Weiterbildung und für Fachvorträge, um die Potenziale von Digitalisierung und Automatisierung in der Produktionslogistik zu erforschen und zu demonstrieren. Im Rahmen dieser Forschung werden die Simulationsmodelle mit der ereignisdiskreten Simulationssoftware Plant Simulation von Tecomatix [Htp00].

3.1 Vorgehen zur Entwicklung eines simulationsbasierten Digitalen Zwillings

Die Entwicklung des Verfahrens erfolgt in Anlehnung an den Vorgehensplan für Ablaufsimulationen von Mayer et al. [MPSW20]. Der Prozess beginnt mit einer Zieldefinition, gefolgt von einer Anforderungsanalyse. In dieser Phase erfolgt die Identifikation der erforderlichen Datenquellen und notwendigen technischen Ressourcen, einschließlich geeigneter Simulationssoftware. Im Anschluss erfolgt die Entwicklung eines digitalen Modells, das die physischen und funktionalen Merkmale eines Produkts oder Prozesses wiedergibt. Dieser Schritt erfordert zunächst die Datenerhebung, -analyse und -aufbereitung, insbesondere in Bezug auf Layout, Struktur und Prozessabläufe des realen Systems. Anschließend wird durch die Integration von Echtzeitdaten aus Sensoren sowie Wartungs- und Zustandsdaten von Anlagen das digitale Modell zu einem Digitalen Schatten weiterentwickelt. Die Implementierung einer Verbindung zwischen dem Digitalen Schatten, der Simulation und dem realen System ermöglicht die Entwicklung zum Digitalen Zwillings. Bei der Realisierung des Digitalen Zwillings wurde das von Zühr et al. entwickelte Framework verwendet. Im nächsten Schritt werden Funktionen beschrieben, die auf Basis aktueller Daten und Simulationsergebnisse automatische Anpassungen der Simulationsparameter oder Entscheidungsvorschläge ermöglichen. Das Dashboard wird während der Phase „Datensammlung“ initiiert und kontinuierlich mit jeder Ebene in seinen Funktionen und Anwendungen weiterentwickelt. Der wesentliche

Unterschied zwischen Ebene 1 und 2 liegt in der Aktualisierungsfrequenz der integrierten Daten, während ab Ebene 3 das Dashboard dynamisch mit aktuellen Daten aktualisiert wird. Abschließend erfolgt die Erweiterung des Dashboards um eine bidirektionale Verbindung zum Digitalen Zwilling. Dadurch ermöglicht das Dashboard nicht nur die Visualisierung relevanter Kennzahlen, sondern bietet auch Interaktivität, die es Anwendern erlaubt, Simulationen zu steuern und Parameter anzupassen. Es dient als Schnittstelle für Entscheidungsträger, die auf Basis der gesammelten Daten Entscheidungen treffen und Prozesse optimieren können.

3.2 Anwendungsbeispiel: Implementierung in der U-Zelle

In diesem Abschnitt wird das zuvor beschriebene Vorgehen und Framework anhand eines praktischen Anwendungsfalls beschrieben und validiert. Diese Umsetzung konzentriert sich vorrangig auf die Ausarbeitung eines Digitalen Prozess-Zwilling als Entscheidungsunterstützungssystem.

Ziel ist es, einen Digitalen Prozess-Zwilling zu erforschen, der Entscheider bei der Kapazitätsplanung unterstützt. Als Demonstrationsplattform dienen Prozesse und Aufbau einer realitätsnahen Produktion in der Lern- und Musterfabrik.

Der betrachtete Fertigungsprozess umfasst sieben Montageschritte. In diesen sieben Schritten wird ein Bodenroller montiert. Ein U-Zellen-Montage-Layout wird verwendet. Die Montage besteht aus sieben Einzelarbeitsplätzen (OP 1-7), die von jeweils einem Mitarbeiter bedient werden, der Bearbeitungstakt beträgt 75 Sekunden. Die Produktion erfolgt an 240 Tagen pro Jahr im Zwei-Schichtbetrieb zu jeweils 7,5 Stunden pro Schicht. Auf Grundlage der vorhandenen Daten der Lern- und Musterfabrik wird zu Beginn ein digitales Modell der U-Zelle erstellt. Hierfür wird die U-Zelle mit den Einzelarbeitsstationen und ihrer zugehörigen internen Logistik abgebildet. Der Detaillierungsgrad wird dabei so gewählt, dass alle nötigen Bereiche abgebildet sind, um ein reales Abbild der Montage zu gewährleisten. In Abbildung 3 ist die physische U-Zelle und ihre digitale Abbildung in Plant Simulation dargestellt.

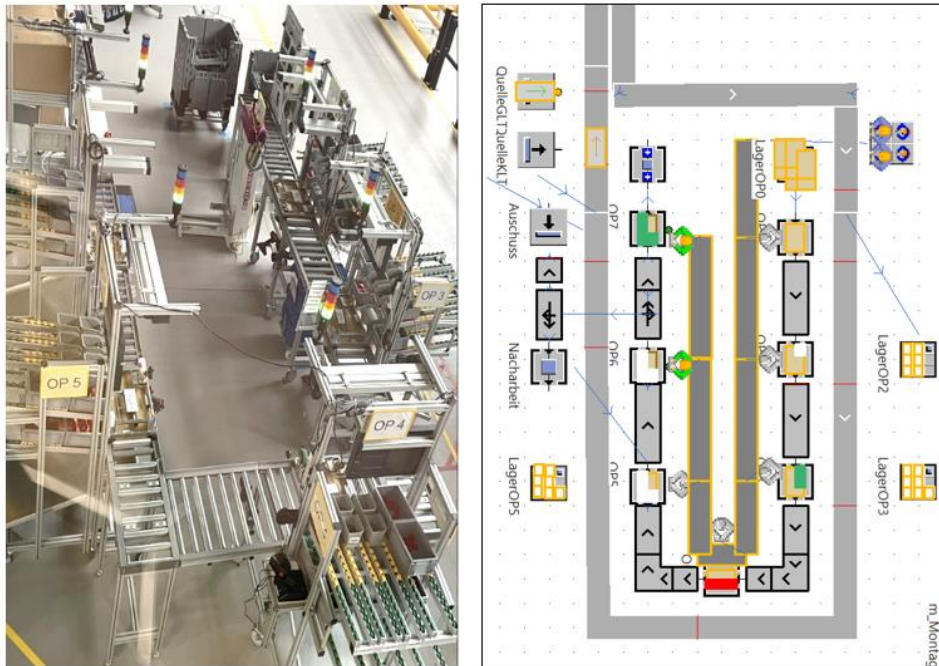


Abb. 3: Die physische U-Zelle und ihre digitale Abbildung

Im nächsten Schritt werden aktuelle Daten aus dem U-Zellen-Montageszenario in die Simulationsumgebung integriert. Dazu gehören Kundenaufträge, Teilelisten, Produktionszeit, Lagerbestände, Mitarbeiter sowie Schichtpläne. Die Daten werden über eine SQLite-Datenbank in die Simulation eingespeist. Im Anschluss wird ein Arbeitstag simuliert. Zu diesem Zweck werden mehrere Simulationsläufe durchgeführt wobei hier unterschiedliche Produktionsreihenfolgen angenommen werden. Zweck der Simulationsdurchläufe ist es, die tatsächliche Bearbeitungszeit an den Stationen und die Gesamtdurchlaufzeit zu evaluieren, um dadurch die Korrektheit der Kapazitätsplanung des Digitalen Zwillinges zu gewährleisten. Ein Ziel ist die Sicherstellung der täglichen Ausbringungsmenge von 770 Bodenrollern. Das Ergebnis der Simulation ist die Ausbringungsmenge. Abbildung 5 zeigt die digitale Abbildung der U-Zelle mit ihren sieben Einzelstationen im Bearbeitungsprozess.

Nachdem die Produktionsdaten erfolgreich in das digitale Modell der U-Zelle erfasst wurden, erfolgt die praktische Umsetzung des Frameworks für den Digitalen Zwilling. Die Simulationseinheit übernimmt dabei die Aufgabe die Simulation mit den notwendigen Daten zu versorgen. Im Anwendungsfall findet die Datenübertragung mithilfe einer Datenbank statt. Die Auswahl der Konfigurationen die simuliert werden sollen, erfolgt in der Informationskontrollereinheit. Hierzu kann auf die Datenversorgungseinheit (Data

Supply Unit) zugegriffen werden. Die Data Supply Unit ist als Datenbank (SQLite) realisiert, in der sich alle verfügbaren Daten befinden (u.a. historische Simulationsläufe sowie deren Ergebnisse und Daten aus dem System).

Die Ergebnisse der Simulation werden im Anschluss mit Hilfe der Simulationseinheit übertragen und in der Data Supply Unit gespeichert. Die Informationskontrolleinheit kann daraufhin die Simulationsergebnisse und Daten im Dashboard darstellen. Das Dashboard ist so gestaltet, dass es nicht nur Handlungsempfehlungen und aktuelle KPIs darstellt, sondern auch Entscheidungsträgern die Möglichkeit bietet, zusätzliche Simulationen vorzunehmen. Durch Änderungen im Dashboard können Simulationen mit geänderten Parametern gestartet werden. Dies wird durch die Informationskontrolleinheit ermöglicht, die die notwendigen Anpassungen an die Simulationskontrolleinheit weiterleitet. Die aus den veränderten Parametern resultierende Handlungsempfehlung und ihren Auswirkungen werden wiederum direkt durch die Informationskontrolleinheit im Dashboard visualisiert.

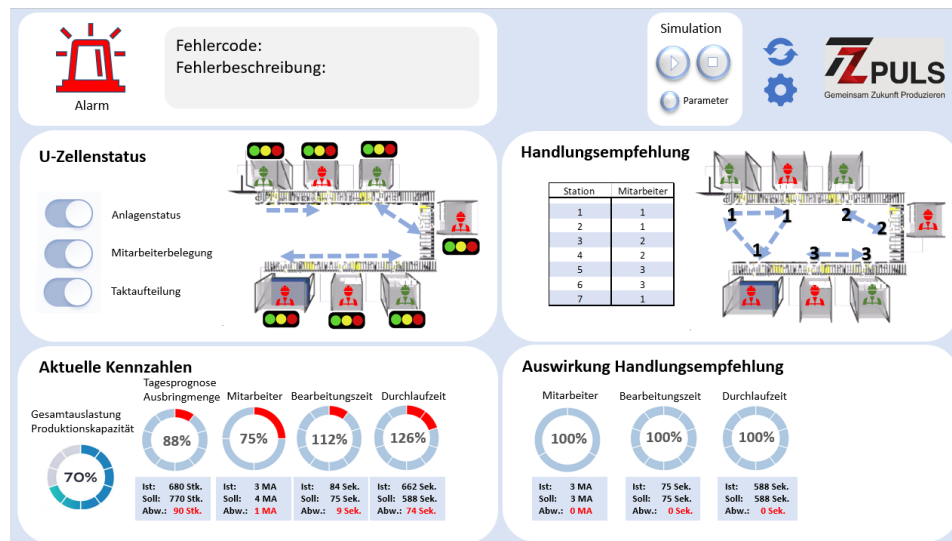


Abb. 4: Dashboard als Visualisierung des Digitalen Prozess Zwillings

Abbildung 4 zeigt ein Dashboard, das in zwei Hauptbereiche aufgeteilt ist. Im linken Bereich ist der Fokus auf der Zustandsüberwachung, die mit den Echtzeitdaten aus der U-Zelle eine kontinuierliche Überwachung ermöglicht. Der U-Zellenstatus zeigt den aktuellen Fertigungsstatus der Kundenbestellungen, Mitarbeiterbelegung und Taktzuweisung sowie den Status der Anlagen. Darunter werden relevante Kennzahlen und deren Erreichungsgrade visualisiert. Besonders kritische Abweichungen der Kennzahlen werden im Rahmen der Prozessanalyse angezeigt, was eine rasche Identifikation und

Reaktion auf mögliche Betriebsstörungen ermöglicht. Die rechte Seite des Dashboards zeigt mögliche Prozessverbesserungen, und visualisiert die aus der Simulation ermittelten Handlungsempfehlungen und deren Auswirkungen sowie die Steuerungselemente der Simulation.

Im abgebildeten Anwendungsfall wird die U-Zelle als visuelle Darstellung abgebildet. An den Einzelarbeitsstationen zeigen grün hinterlegte Symbole von Mitarbeitern die Anwesenheit und rote die Abwesenheit. Die Pfeile bilden die Taktzuweisung des jeweiligen Mitarbeiters ab. Abweichungen in den KPIs werden rot hervorgehoben. Durch die gleiche visuelle Abbildung der U-Zelle auf der Prozessverbesserungsseite (rechts) werden Änderungen der Mitarbeiterbelegung und deren Taktzuweisung übersichtlich dargestellt. Zudem werden die direkten Auswirkungen der Handlungsempfehlung auf die betreffenden Kennzahlen ersichtlich. Somit kann der Entscheiderkreis die Handlungsempfehlung bei Bedarf übernehmen und im System umsetzen.

4 Zusammenfassung und Ausblick

In diesem Artikel wird die Implementierung eines simulationsbasierten Digitalen Prozess-Zwillings auf Basis den von Zuhr et al. definierten Frameworks beschrieben und dessen Umsetzbarkeit verifiziert. Ziel des Frameworks ist es, Entscheidungsunterstützung und Prozessoptimierung durch simulationsbasierte Digitale Prozess-Zwillinge zu ermöglichen und dabei die strukturierte Vorgehensweise aufzuzeigen [Zuhr21]. Die erfolgreiche Umsetzung dieses Frameworks wurde in der Lern- und Musterfabrik des Technologiezentrums Produktions- und Logistiksysteme (TZ PULS) demonstriert.

Das Framework beschreibt die Datenaufnahme und -analyse durch eine zweiteilige Architektur aus Datenversorgungs- und Informationssteuerungseinheit, was die Integration von Betriebsdaten und den Einsatz von simulationsbasierten Methoden zur Entscheidungsfindung unterstützt.

Das vorgestellte Dashboard visualisiert die Ergebnisse der Simulation und stellt Echtzeitdaten des realen Systems dar. Es setzt folgende im Framework definierten Dienste um: Zustandsüberwachung, Prozessanalyse und Prozessverbesserung. Durch die Darstellung von Handlungsempfehlungen und deren Auswirkungen unterstützt das Dashboard Entscheidungsträger bei der Entscheidungsfindung. Darüber hinaus fördert die Integration von Simulationsergebnissen mit Echtzeitbetriebsdaten im Dashboard ein besseres Verständnis der Betriebsprozesse und deren Potenzialen bei den Anwendern.

Durch die Implementierung und Anwendung des Digitalen Prozess-Zwillings kann die Effektivität der Entscheidungsfindung verbessert werden. Jedoch ergeben sich bei der Implementierung Digitaler Prozess-Zwillinge in Unternehmen auch Herausforderungen. Eine Schwierigkeit besteht in der Notwendigkeit, die Aktualität des Digitalen Prozess-

Zwillings stets zu gewährleisten. Jede Veränderung im realen System muss rechtzeitig im digitalen Abbild aktualisiert werden, um die Funktionsfähigkeit des Digitalen Prozess-Zwillings zu sichern. Ebenso muss der Datenfluss zwischen den beiden Systemen jederzeit aufrechterhalten werden. Eine zusätzliche Herausforderung ergibt sich aus dem Design des Dashboards. Dieses muss in der Lage sein, alle relevanten Informationen so darzustellen, dass sie den Entscheidungsträgern eine zuverlässige Entscheidungsgrundlage bieten. Dabei müssen diverse Perspektiven integriert werden, um die Anforderungen aller Beteiligten zu erfüllen.

In der weiterführenden Forschungsarbeit soll die Geeignetheit des gezeigten Vorgehens evaluiert werden. Dafür sollte ein Kriterienkatalog erarbeitet werden, anhand dessen die Umsetzung des Framework bewertet werden kann. Dieser Katalog soll insbesondere darauf eingehen, inwiefern das Framework allgemeingültig umgesetzt werden kann oder ob gegebenenfalls noch Änderungen am Modell vorgenommen werden müssen. Ebenso sollen die Einsatzvoraussetzungen und Grenzen des Frameworks bei praktischen Umsetzungsstudien erforscht werden. Auf Grundlage der daraus resultierenden Erkenntnisse soll das Framework weiter angepasst und zu einem allgemeingültigen Referenzmodell weiterentwickelt werden.

Darüber hinaus sollten zukünftige Forschungsarbeiten untersuchen, wie die technischen Fähigkeiten von Entscheidungsunterstützungssystemen weiterentwickelt werden können. Eine Möglichkeit ist, den Digitalen Prozess-Zwilling um maschinelles Lernen zu erweitern, um zu untersuchen, ob sich dadurch die Effizienz des Systems weiter verbessern lässt. Durch den vermehrten Einsatz von Modellen des maschinellen Lernens lässt sich der Simulationsaufwand reduzieren und damit die Reaktionsgeschwindigkeit des Digitalen Zwillings erhöhen, ohne auf die Prognosegenauigkeit der Simulation zu verzichten. Dies kann die Entscheidungsfindung der Anwender weiter optimieren [MuRM22].

Literatur

- [BCKM09] BLUTNER, DORIS ; CRAMER, STEPHAN ; KRAUSE, SVEN ; MÖNKS, TYCHO ; NAGEL, LARS ; REINHOLZ, ANDREAS ; WITTHAUT, MARKUS: Assistenzsysteme für die Entscheidungsunterstützung. In: BUCHHOLZ, P. ; CLAUSEN, U. (Hrsg.): *Große Netze der Logistik*. Berlin, Heidelberg : Springer Berlin Heidelberg, 2009 — ISBN 978-3-540-71047-9, S. 241–270
- [Eley12] ELEY, MICHAEL: *Simulation in der Logistik: eine Einführung in die Erstellung ereignisdiskreter Modelle unter Verwendung des Werkzeuges „Plant Simulation“*, Springer-Lehrbuch. Berlin Heidelberg : Springer Gabler, 2012 — ISBN 978-3-642-27372-8

- [FRBM17] FELLMANN, MICHAEL ; ROBERT, SEBASTIAN ; BÜTTNER, SEBASTIAN ; MUCHA, HENRIK ; RÖCKER, CARSTEN: Towards a Framework for Assistance Systems to Support Work Processes in Smart Factories. In: HOLZINGER, A. ; KIESEBERG, P. ; TJOA, A. M. ; WEIPPL, E. (Hrsg.): *Machine Learning and Knowledge Extraction, Lecture Notes in Computer Science*. Bd. 10410. Cham : Springer International Publishing, 2017 — ISBN 978-3-319-66807-9, S. 59–68
- [Grie16] GRIEVES, MICHAEL: Origins of the Digital Twin Concept, [object Object] (2016)
- [HaMG19] HARPER, K. ERIC ; MALAKUTI, SOMAYEH ; GANZ, CHRISTOPHER: Digital Twin Architecture and Standards., [object Object] (2019)
- [Http00] <https://www.haw-landshut.de/forschungseinrichtungen/technologiezentren/tz-puls>.
- [JSNY20] JONES, DAVID ; SNIDER, CHRIS ; NASSEHI, AYDIN ; YON, JASON ; HICKS, BEN: Characterising the Digital Twin: A systematic literature review. In: *CIRP Journal of Manufacturing Science and Technology* Bd. 29 (2020), S. 36–52
- [KaGF21] KAUKE, DIRK ; GALKA, STEFAN ; FOTTNER, JOHANNES: Digital Twins in Order Picking Systems for Operational Decision Support. In: , 2021
- [KKTH18] KRITZINGER, WERNER ; KARNER, MATTHIAS ; TRAAR, GEORG ; HENJES, JAN ; SIHN, WILFRIED: Digital Twin in manufacturing: A categorical literature review and classification. In: *IFAC-PapersOnLine* Bd. 51 (2018), Nr. 11, S. 1016–1022
- [KoSc23] KORNBLUM, WOLFGANG ; SCHUMANN, RENÉ: Digitale Zwillinge ermöglichen fundierte Entscheidungen durch vernetzte Daten. In: *Bautechnik* Bd. 100 (2023), Nr. 4, S. 198–205
- [KoSZ18] KORTH, BENJAMIN ; SCHWEDE, CHRISTIAN ; ZAJAC, MARKUS: Simulation-ready digital twin for realtime management of logistics systems. In: *2018 IEEE International Conference on Big Data (Big Data)*. Seattle, WA, USA : IEEE, 2018 — ISBN 978-1-5386-5035-6, S. 4194–4201
- [MPSW20] MAYER, GOTTFRIED ; PÖGE, CARSTEN ; SPIECKERMANN, SVEN ; WENZEL, SIGRID: Einleitung. In: MAYER, G. ; PÖGE, C. ; SPIECKERMANN, S. ; WENZEL, S. (Hrsg.): *Ablaufsimulation in der Automobilindustrie*. Berlin, Heidelberg : Springer Berlin Heidelberg, 2020 — ISBN 978-3-662-59387-5, S. 1–19
- [MuRM22] MUEHLBAUER, KONSTANTIN ; RISSMANN, LUKAS ; MEISSNER, SEBASTIAN: Decision Support for Production Control based on Machine Learning by Simulation-generated Data: In: *Proceedings of the 14th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*. Valletta, Malta : SCITEPRESS - Science and Technology Publications, 2022 — ISBN 978-989-758-614-9, S. 54–62

- [RBKV10] ROBINSON, STEWART; BROOKS, ROGER; KOTIADIS, KATHY; VAN DER ZEE, DURK-JOUKE (Hrsg.): *Conceptual Modeling for Discrete-Event Simulation*. 0. Aufl. : CRC Press, 2010 — ISBN 978-1-4398-1038-5
- [SFHQ21] SINGH, MAULSHREE ; FUENMAYOR, EVERT ; HINCHY, EOIN ; QIAO, YUANSONG ; MURRAY, NIALL ; DEVINE, DECLAN: Digital Twin: Origin to Future. In: *Applied System Innovation* Bd. 4 (2021), Nr. 2, S. 36
- [SGMM23] SCHMID, FLORIAN ; GALKA, STEFAN ; MARINKOVIC, MINJA ; MEIBNER, SEBASTIAN: Simulationsbasierte Untersuchung von Abstraktionsgraden im Kontext Digitaler Zwillinge für innerbetriebliche Transportsysteme. In: *20. ASIM Fachtagung Simulation in Produktion und Logistik 2023*, [object Object] (2023), S. 135 — ISBN 9783863602765
- [TCQZ18] TAO, FEI ; CHENG, JIANGFENG ; QI, QINGLIN ; ZHANG, MENG ; ZHANG, HE ; SUI, FANGYUAN: Digital twin-driven product design, manufacturing and service with big data. In: *The International Journal of Advanced Manufacturing Technology* Bd. 94 (2018), Nr. 9–12, S. 3563–3576
- [TiBa02] TIMPE, KLAUS-PETER; BAGGEN, ROBERT. (Hrsg.): *Mensch-Maschine-Systemtechnik: Konzepte, Modellierung, Gestaltung, Evaluation*. 2. Aufl., Stand: Februar 2002. Düsseldorf : Symposion Publ, 2002 — ISBN 978-3-933814-83-8
- [VaDe11] VASUDEVAN, KARTHIK; DEVIKAR, ASHISH: Selecting simulation abstraction levels in simulation models of complex manufacturing systems. In: *Proceedings of the 2011 Winter Simulation Conference (WSC)*. Phoenix, AZ, USA : IEEE, 2011 — ISBN 978-1-4577-2108-3, S. 2268–2277
- [Zuhr21] ZUHR, PASCAL; MUEHLBAUER, KONSTANTIN; BAEUML, STEPHANIE; MEIBNER, SEBASTIAN: An approach for target-oriented process analysis for the implementation of Digital Process Optimization Twins in the field of intralogistics. In: *Proceedings of the 33rd European Modeling & Simulation Symposium* : CAL-TEK srl, 2021 — ISBN 978-88-85741-57-7, S. 183–191
- [ZuRM22] ZUHR, PASCAL; RISSMANN, LUKAS ; MEIBNER, SEBASTIAN: Framework for planning and implementation of Digital Process Twins in the field of internal logistics. In: *IFAC-PapersOnLine, 10th IFAC Conference on Manufacturing Modelling, Management and Control MIM 2022*. Bd. 55 (2022), Nr. 10, S. 2221–2227

Teil B

Kurzbeiträge und Prototypen

Einfluss von Gamification auf die Akzeptanz von digitalen Gesundheitsanwendungen (DiGA) - Ergebnisse einer qualitativen Studie

Kurzbeitrag

Lena Ulrich¹, Birte Malzahn²

Abstract: Der vorliegende Beitrag untersucht, welche Faktoren die Nutzerakzeptanz von DiGA beeinflussen, und insbesondere auch, welchen Einfluss Gamification in diesem Zusammenhang ausübt. Hierfür wurden Hypothesen aufgestellt, die mittels einer stichpunktbezogenen qualitativen Datenerhebung überprüft wurden. Die Ergebnisse zeigen, dass Gamification zwar den Spaß an der Nutzung erhöht, dieser aber keinen starken Einfluss auf die Absicht zur Nutzung der untersuchten DiGA ausübt. Der Beitrag zeigt mögliche Konsequenzen der Ergebnisse für Forschung und Praxis auf.

Keywords: Gamification, Digitale Gesundheitsanwendungen (DiGA), Nutzerakzeptanz

1 Einleitung

DiGA unterstützen u. a. die Therapie von Krankheiten oder Verletzungen [Bund23]. Sie ermöglichen eine Versorgung von Patient*innen „remote“ und ohne notwendige menschliche Interaktion [FrLü21]. Der Erfolg des Konzepts ist jedoch maßgeblich davon abhängig, ob die DiGA verschrieben werden können, denn nur dann ist ein ausreichender Absatz wahrscheinlich [FrLü21]. Hierfür ist die Aufnahme einer DiGA in das DiGA-Verzeichnis notwendig, was in Deutschland seit 2020 erfolgen kann [FrLü21]. Essentiell für den Erfolg ist des Weiteren die Akzeptanz der Anwendungen bei Patient*innen und auch bei der Ärzteschaft, die die DiGA verordnen muss [FrLü21]. *Gamification* bezeichnet den Einsatz von Designelementen, die charakteristisch für Spiele sind, in Anwendungen, die nicht Teil eines Unterhaltungsspiels sind [DDKN11]. Gängige Elemente sind u. a. das Erwerben von Punkten, Bestenlisten, Auszeichnungen, das Erreichen von Leveln oder das Anzeigen des Fortschritts [HaKS14]. Der vorliegende Beitrag untersucht, welche Faktoren die Nutzerakzeptanz von DiGA beeinflussen. Dabei liegt ein Fokus auf der Untersuchung des Einflusses von Gamification.

¹ HTW Berlin, Studiengang Wirtschaftsinformatik, Treskowallee 8, 10318 Berlin

² HTW Berlin, Studiengang Wirtschaftsinformatik, Treskowallee 8, 10318 Berlin,
birte.malzahn@htw-berlin.de, ORCID ID 0000-0001-5361-1794

2 Gamification in E-Health-Anwendungen

Der Einsatz von Gamification im Gesundheitsbereich wird seit einigen Jahren erforscht [SaIF17]. [HJHV22] untersuchen beispielsweise die Nutzung einer gamifizierten E-Health-Anwendung bei älteren Personen auf Basis des Technology Acceptance Models und zeigen u. a., dass der Spaß an der Nutzung die wahrgenommene Nützlichkeit beeinflusst und dass die wahrgenommene Nützlichkeit die Intention beeinflusst, die Anwendung weiterhin zu benutzen. [SaIF17] zeigen in ihrer Metastudie, dass Forschungsarbeiten v. a. kurzfristige Effekte von Gamification untersucht haben. [DaMB23] weisen in ihrer Metastudie darauf hin, dass einige Studien positive Effekte von Gamification / Serious Games feststellen, andere dagegen keine Effekte nachweisen können. Sie fordern daher, die Wirkungsweisen weiter zu untersuchen, insbesondere auch bei unterschiedlichen Bevölkerungsgruppen.

3 Aufbau und Durchführung der Studie

Folgende Hypothesen wurden in Anlehnung an das Technology Acceptance Models [Davi86] und dessen Weiterentwicklungen aufgestellt:

H1: Die wahrgenommene Nützlichkeit hat einen positiven Einfluss auf die Nutzungsabsicht [Venk00], [VeDa00].

H2: Die Leichtigkeit der Nutzung hat einen positiven Einfluss auf die Nutzungsabsicht [Venk00], [VeDa00].

H3: Die Leichtigkeit der Nutzung hat einen positiven Einfluss auf die wahrgenommene Nützlichkeit [Davi86], [VeDa00].

H4: Die subjektive Norm hat einen positiven Einfluss auf die Nutzungsabsicht [VeDa00].

H5a: Der Einsatz von Gamification erhöht den wahrgenommenen Spaß an der Nutzung einer DiGA [BoMM19] [LoPT14].

H5b: Der wahrgenommene Spaß hat einen positiven Einfluss auf die Nutzungsabsicht [DaBW92] [BoMM19].

In Zusammenarbeit mit der akquinet tech@spree GmbH und der G. Pohl-Boskamp GmbH & Co. KG wurde eine DiGA für Belastungsinkontinenzbeschwerden für die Zielgruppe biologisch anatomischer Frauen ab 50 Jahren entwickelt [Ulri23]. Für die Studie wurden zwei Prototypen dieser DiGA mittels der Software Figma [Figm24] erstellt: **Prototyp A** beinhaltet vier Gamification-Elemente: Punkte, Trophäen, Status und Fortschrittsanzeige. **Prototyp B** enthielt keine Gamification-Elemente.

Da es sich bei Gamification von DiGA um ein relativ neues Forschungsgebiet handelt, wurde eine qualitative Forschungsmethodik gewählt [Ulri23]. Die Datenerhebung wurde zwischen dem 28.07.2023 und dem 19.08.2023 einzeln mit insgesamt 14 Frauen im Alter von 48 bis 72 Jahren durchgeführt. Die Probandinnen wurden vorab per Zufallsprinzip in

zwei Testgruppen eingeteilt: Gruppe A arbeitete mit Prototyp A, Gruppe B mit Prototyp B. Zu Beginn erhielten die Teilnehmerinnen eine kurze Einführung in das Thema und es wurden Fragen über Erwartungen und Befürchtungen bezüglich der DiGA sowie zur allgemeinen Nutzungsabsicht der Anwendung gestellt. Die Probandinnen führten dann mehrere Aufgaben mit dem zugeteilten Prototyp aus. Anschließend wurden ihnen mittels Leitfadens offene Fragen zu den Konstrukten gestellt. Da es sich um eine qualitative Studie handelt, wurden Fragen aus früheren quantitativen Forschungsarbeiten angepasst. Beispielsweise wurde die auf einer Skala zu bewertende Aussage „Es macht mir Spaß, das System zu benutzen.“ in die offene Frage „Inwieweit hat die Nutzung der App dir Spaß bereitet?“ umformuliert [Ulri23]. Alle Gespräche wurden aufgezeichnet.

4 Auswertung und Ergebnisse

Zur qualitativen Analyse der Interviews wurde die Software MAXQDA [Verb24] genutzt. Es wurde die skalierende Strukturierungstechnik [Mayr15] verwendet, um die Aussagen der Probandinnen in deduktive Kategorien zu codieren, die aus dem Modell dieser Studie abgeleitet wurden. Anschließend wurden die Inhalte bezüglich der untersuchten Konstrukte auf einer Ordinalskala bewertet [Ulri23]. Es wurden die folgenden drei Ausprägungen verwendet: "+" für positive Aussagen, "o" für neutrale Aussagen und "-" für negative Aussagen. Tabelle 1 zeigt die Bewertung der Konstrukte.

Kategorie	Gruppe A: mit Gamification							Gruppe B: ohne Gamification						
	P2	P3	P4	P7	P9	P10	P14	P1	P5	P6	P8	P11	P12	P13
Wahrgenommene Nützlichkeit	+	+	+	+	o	+	+	+	+	+	+	-	+	+
Leichtigkeit der Nutzung	+	+	o	+	+	+	+	+	o	+	o	+	+	+
Wahrgenommener Spaß	+	+	o	+	o	+	+	+	o	o	-	o	o	-
Subjektive Norm	+	+	+	+	o	+	+	+	+	+	+	+	+	+
Nutzungsabsicht (ex post)	o	+	+	+	o	+	+	+	+	+	o	o	+	+

Tabelle 1 Bewertung der Konstrukte nach Durchführung des Tests (Eigene Darstellung)

H	Beschreibung	Prozent	Bestätigung
1	Nützlichkeit → Nutzungsabsicht	79%	stark bestätigt
2	Leichtigkeit der Nutzung → Nutzungsabsicht	64%	bestätigt
3	Leichtigkeit der Nutzung → Nützlichkeit	64%	bestätigt
4	Subjektive Norm → Nutzungsabsicht	79%	stark bestätigt
5a	Gamification → Spaß an der Nutzung	---	stark bestätigt
5b	Spaß → Nutzungsabsicht	50%	<i>schwach</i> bestätigt

Tabelle 2 Auswertung der Hypothesen (Eigene Darstellung)

Die Überprüfung aller Hypothesen außer H5a erfolgte für beide Gruppen zusammen. Es wurde für jede dieser Hypothese ermittelt, zu welchem Anteil die Bewertung der beiden Konstrukte der jeweiligen Hypothese übereinstimmen (z. B. positiv-positiv, neutral-neutral, negativ-negativ). Bei einem Anteil von über 75 Prozent an Übereinstimmung wurde die Hypothese als stark bestätigt angesehen. Ein Anteil von 51 bis 75 Prozent wurde als Bestätigung der Hypothese angesehen. Bei einem Anteil von 26 bis 50 Prozent wurde die Hypothese als schwach bestätigt gewertet. Auf eine Darstellung der Ergebnisse mittels Kreuztabellen wird in diesem Beitrag aus Platzgründen verzichtet. Zur Auswertung von Hypothese 5a wurden alle positiven Bewertungen des *wahrgenommenen Spaßes* mit 1 bewertet, neutrale Bewertungen mit 0 und negative Bewertungen mit -1. Für Gruppe A (mit Gamification) ergibt sich dabei ein Durchschnittswert von 0,71, für Gruppe B (ohne Gamification) ein Durchschnittswert von -0,14. Die sich ergebende Abweichung zwischen diesen Gruppen wurde als starke Bestätigung von H5a gewertet. Tabelle 2 fasst die Ergebnisse zusammen.

5 Diskussion der Ergebnisse

Die Hypothesen des Modells können anhand der qualitativen Daten mit unterschiedlicher Stärke als bestätigt angesehen werden. In Übereinstimmung mit früheren Studien zeigt die *wahrgenommene Nützlichkeit* einen starken Einfluss auf die *Nutzungsabsicht (ex post)*. Die vorliegenden Daten weisen zudem darauf hin, dass Gamification den *wahrgenommenen Spaß* durchschnittlich deutlich erhöht. Jedoch konnte nur ein schwacher direkter Einfluss des *wahrgenommenen Spaßes* auf die *Nutzungsabsicht (ex post)* festgestellt werden. Die Interviewaussagen deuten darauf hin, dass ein Spaßfaktor bei der Nutzung einer DiGA teilweise als unangemessen empfunden wird. Als Konsequenz ließe sich ableiten, dass der Einsatz von Gamification in diesem Kontext verzichtbar wäre.

Es ist jedoch denkbar, dass sich bei einer Untersuchung der Nutzungsabsicht nach verschiedenen Nutzungszeiträumen [VeDa00] andere Ergebnisse ergeben würden. So kann nach einiger Zeit ein Gewöhnungseffekt an Gamificationelemente auftreten, der zu positiven Auswirkungen führen kann [RPTP22]. Die Langzeitwirkung des Einsatzes von Gamification bei dieser DiGA konnte aufgrund der gegebenen Rahmenbedingungen nicht untersucht werden. Dies sollte aufbauend auf der vorliegenden Studie erfolgen, um die Wirkungsweisen weiter zu erforschen. Die Ergebnisse könnten zudem mit dem Anwendungskontext der DiGA variieren. Auch ist möglich, dass das Alter oder das Geschlecht der Probandinnen einen Einfluss auf die Ergebnisse haben [KoHa14]. Auch das Forschungsdesign führt zu Limitationen: Aufgrund der geringen Teilnehmerzahl sind keine weitergehenden statistischen Auswertungen möglich. Die qualitative Studie kann jedoch als Vorstudie wichtige Hinweise darauf geben, welche Aspekte in einer quantitativen Studie weiter untersucht werden sollten. Für Anbieter von DiGA ist eine kontinuierliche Nutzungsabsicht bzw. Nutzung durch die Anwender*innen unerlässlich, um dauerhaft am Markt bestehen können. Sie sollten daher weiter die Einflussfaktoren auf die Nutzung von DiGA untersuchen.

Literaturverzeichnis

- [BoMM19] BORRÁS-GENÉ ; MARTÍNEZ-NÚÑEZ ; MARTÍN-FERNÁNDEZ: Enhancing Fun Through Gamification to Improve Engagement in MOOC. In: *Informatics* Bd. 6 (2019), Nr. 3, S. 28
- [Bund23] BUNDESINSTITUT FÜR ARZNEIMITTEL UND MEDIZINPRODUKTE: *Das Fast-Track-Verfahren für digitale Gesundheitsanwendungen (DiGA) nach § 139e SGB V - Ein Leit-faden für Hersteller, Leistungserbringer und Anwender*, 2023
- [DaBW92] DAVIS, FRED. D. ; BAGOZZI, RICHARD P. ; WARSHAW, PAUL R.: Extrinsic and Intrinsic Motivation to Use Computers in the Workplace. In: *Journal of Applied Social Psy-chology* Bd. 22 (1992), Nr. 14
- [DaMB23] DAMAŠEVIČIUS, ROBERTAS ; MASKELIŪNAS, RYTIS ; BLAŽAUSKAS, TOMAS: Serious Games and Gamification in Healthcare: A Meta-Review. In: *Information* Bd. 14 (2023), Nr. 2, S. 105
- [Davi86] DAVIS, FRED. D.: *A Technology Acceptance Model for Empirically Testing New End-User Information Systems: Theory and Result*. Cambridge, MIT Sloan School of Management, 1986
- [DDKN11] DETERDING, SEBASTIAN ; DIXON, DAN ; KHALED, RILLA ; NACKE, LENNART: From Game Design Elements to Gamefulness: Defining „Gamification“. In: *Proceedings of the 15th International Academic MindTrek Conference*. Tampere Finland : Asso-ciation for Computing Machinery New York NY United States, 2011 — ISBN 978-1-4503-0816-8, S. 9–15
- [Figm24] FIGMA GMBH: *Figma*. URL <https://www.figma.com/de/>. - abgerufen am 2024-02-27
- [FrLü21] FRIESENDORF, CORDELIA ; LÜTTSCHWAGER, SABRINA: *Digitale Gesundheitsanwen-dungen: Assessment der Ärzteschaft zu Apps auf Rezept, essentials*. Wiesbaden [Hei-delberg] : Springer, 2021 — ISBN 978-3-658-33983-8
- [HaKS14] HAMARI, JUHO ; KOIVISTO, JONNA ; SARSA, HARRI: Does Gamification Work? — A Literature Review of Empirical Studies on Gamification. In: *Proceedings of the 47th Hawaii International Conference on System Sciences*. Hawaii, USA, 2014, S. 3025–3034
- [HJHV22] HURMUZ, MARIAN ZM ; JANSEN-KOSTERINK, STEPHANIE M ; HERMENS, HERMIE J ; VAN VELSEN, LEX: Game not over: Explaining older adults' use and intention to con-tinue using a gamified eHealth service. In: *Health Informatics Journal* Bd. 28 (2022), Nr. 2, S. 146045822211060

- [KoHa14] KOIVISTO, JONNA ; HAMARI, JUHO: Demographic Differences in Perceived Benefits from Gamification. In: *Computers in Human Behavior* Bd. 35 (2014), S. 179–188
- [LoPT14] LOUNIS, STAVROS ; PRAMATARI, KATERINA ; THEOTOKIS, ARISTEIDIS: Gamification is all about Fun: The Role of Incentive Type and Community Collaboration. In: *ECIS 2014 Proceedings*. Tel Aviv, 2014, S. 1–15
- [Mayr15] MAYRING, P.: *Qualitative Inhaltsanalyse: Grundlagen und Techniken*. 12. Auflage. : Beltz Verlag, 2015 — ISBN ISBN: 978-3-407-25730-7
- [RPTP22] RODRIGUES, LUIZ ; PEREIRA, FILIPE D. ; TODA, ARMANDO M. ; PALOMINO, PAULA T. ; PESSOA, MARCELA ; CARVALHO, LEANDRO SILVA GALVÃO ; FERNANDES, DAVID ; OLIVEIRA, ELAINE H. T. ; U. A.: Gamification Suffers from the Novelty Effect but Benefits from the Familiarization Effect: Findings from a Longitudinal Study. In: *International Journal of Educational Technology in Higher Education* Bd. 19 (2022), Nr. 1
- [SaIF17] SARDI, LAMYAE ; IDRI, ALI ; FERNÁNDEZ-ALEMÁN, JOSÉ LUIS: A systematic review of gamification in e-Health. In: *Journal of Biomedical Informatics* Bd. 71 (2017), S. 31–48
- [Ulri23] ULRICH, LENA: *Untersuchung des Einsatzes von Gamification zur Steigerung der Wirksamkeit einer digitalen Gesundheitsanwendung*, HTW Berlin, Masterarbeit, 2023
- [VeDa00] VENKATESH, VISWANATH ; DAVIS, FRED. D.: A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. In: *Management Science* Bd. 46 (2000), Nr. 2
- [Venk00] VENKATESH, VISWANATH: Determinants of Perceived Ease of Use: Integrating Control, Intrinsic Motivation, and Emotion into the Technology Acceptance Model. In: *Information Systems Research* Bd. 11 (2000), Nr. 4
- [Verb24] VERBI – SOFTWARE. CONSULT. SOZIALFORSCHUNG. GMBH: *MAXQDA*. URL <https://www.maxqda.com/de/>. - abgerufen am 2024-02-27

Mobile Applikation zur Selbsttherapieunterstützung chronisch Erkrankter mittels Bewegungsmotivation

Prototyp

Umsetzung des ersten Anwendungsfall für das Krankheitsbild Parkinson

Martin Przewloka¹, Nassim Abuhaltam², Michael Lüttebrandt³, Ingrid Profitlich, Markus Maria Profitlich

Abstract: In Deutschland sind ca. 200.000 Menschen von Parkinson betroffen, verbunden mit einer deutlichen, maßgeblich demografisch verursachten jährlichen Steigerung. Bewegungstherapeutischen Maßnahmen, wie sie häufig verordnet werden, steht der Mensch entgegen: Trägheit, Verzweiflung, Resignation, offensichtliche Ausweglosigkeit sind nur einige Motive, die dazu führen, dass nach einem initialen Aufbruch das eigenständige Festhalten an regelmäßiger Bewegung ausbleibt und damit unweigerlich die Spirale der gefühlten und nachweisbaren Verschlechterung des Gesundheitszustands in Gang gebracht wird. Die hier vorgestellte Anwendung ist für das Zielkrankheitsbild Parkinson entwickelt. Kontinuierlich werden hiermit Motivationsimpulse zur Bewegung mittels kurzweiliger Videosequenzen gegeben. Die Videos wurden von dem Parkinsonerkranktem Schauspieler und Komiker Markus Maria Profitlich und der Diplomsportlehrerin, Gymnastik- und Tanzpädagogin Ingrid Profitlich in Szene gesetzt. Erweitert ist die Anwendung um ein neuartiges, einfaches Spiels, um zusätzliche Anreize zur Bewegungsausführung zu schaffen. Die Applikation ist seit Oktober 2022 auf den Plattformen App Store (Apple) und seit Mai 2023 auf Google Play frei verfügbar.

Keywords: Mobile Applikation, Cross-Platform-Apps, Flutter, Parkinson, Therapieunterstützung, Bewegungsmotivation, Gamifizierung, DiGA, Akzeptanz, Virtual Reality

1 Einleitung

Wenngleich eine allgemeine Definition des Begriffs Krankheit nicht gelingt, so ist vielen Krankheitsbildern gemeinsam, dass Betroffene objektiv Einschränkungen hinsichtlich ihres Wohlbefindens wahrnehmen und Bereitschaft zeigen, diesen entgegen zu wirken. Nicht nur bei zivilisationsbedingten Erkrankungen, sondern im Zusammenhang mit einer Großzahl chronischer Erkrankungen, werden Bewegung und Sport sowohl zur Prophylaxe wie auch als wesentliche Therapieunterstützung angesehen. Dies betrifft bspw. Herz-Kreislauferkrankungen, Diabetes, Osteoporose sowie neurologische Erkrankungen, zu

¹ Institut für Digitale Assistenzsysteme e.V., Technologiezentrum Gießen, Kerkrader Straße 9, 35394 Gießen

² Technische Hochschule Mittelhessen, Wiesenstraße 14, 35390 Gießen

³ Brainsport GmbH & Co. KG, Friemer Straße 12, 37284 Waldkappel

denen Parkinson als zweithäufigste neurodegenerative Erkrankung zählt. Dies betrifft bspw. Herz-Kreislaufkrankungen, Diabetes, Osteoporose sowie neurologische Erkrankungen, zu denen Parkinson als zweithäufigste neurodegenerative Erkrankung zählt.

Studien und Metaanalysen zeigen, dass die Inzidenz von Herz-Kreislauf-, Stoffwechsel-, Krebs- und psychischen Erkrankungen invers mit dem körperlichen Aktivitäts- oder Fitnessniveau assoziiert ist (für einen Überblick siehe [Wi15]). Umgekehrt stehen physische Inaktivität und chronische Erkrankung in einem positiv korrelierten Zusammenhang, wennngleich eine Vielzahl an Drittvariablen in diesem komplexen System existieren (für einen Überblick siehe [Bu15] und für eine weltweite Studie [Lee12]). Ebenso sind die Zusammenhänge von Bewegung und Sport im Umfeld der neurologischen Erkrankungen ausgiebig untersucht worden. Dabei werden deren positiven Effekte signifikant nachgewiesen. Dies gilt nicht nur für die in diesem Anwendungsfall zugrundeliegende Erkrankung Parkinson (vgl. u.a. [Fe93], [Ta09], [Ts22]), sondern erstreckt sich über nahezu alle neurologischen Krankheitsbilder hinweg (siehe u.a. bspw. für die aktuell vielfach diskutierte Erkrankung Demenz [Is15]).

In der Praxis stehen Ärztinnen und Ärzte, Therapeutinn:en und das Gesundheitssystem allerdings vor der Herausforderung physischer Inaktivität der Betroffenen. Punktuelle Aufforderungen und Motivationsimpulse, wie diese regelmäßig beim Arzttermin ausgesprochen werden, sind häufig wenig nachhaltig. Die sogenannte Compliance und damit verbundene Adhärenz (Motivation) erreicht hierbei nicht die notwendigen Zielgrößen und gilt es zu verbessern. Die Gründe hierfür sind vielschichtig (eine ausführliche Anlyase hierzu findet sich bei [Co23]). Für die von uns entwickelte Anwendung sind insbesondere die folgenden Kriterien, welche mittel- und unmittelbar die Adhärenz von Bewegungstherapien beeinflussen, von besonderem Interesse:

- Charakteristiken der Anwendungen in Form von Übungen (Art der Übung, Dauer, Häufigkeit, Intensität u.a.) sowie potenzielle Risiken
- Eigene Fähigkeiten, die Übungen umzusetzen, Komplexität, Hilfsmiteinsatz
- Eigene Grundeinstellung
- Kommunikation, Involvierung und Rückmeldung der Expertinn:en (Ärztin/Arzt, Therapeut:in) und weiterer Beteiligter (Angehörige, Bekannte)
- Beobachtung, Kontrolle
- Integration in das tägliche Leben
- Messung und Darstellung des Fortschritts
- Zieldefinition, Belohnung
- Spaß, Freude, Unterhaltung bei der Durchführung der Übungen

Es ist daher naheliegend, mit Hilfe digitaler Ansätze eine Lösung für das Grundproblem der Adhärenz im Zusammenhang mit Selbsttherapien zu entwickeln. Die Kernidee dabei ist, dass Betroffene einem durch Expertinn:en oder selbst bestimmten Bewegungsplan treu bleiben, indem diese kontinuierlich zur Durchführung der Übungen motiviert werden. Spaß, Freude und Unterhaltung wie auch Belohnungssysteme sollen intrinsisch Teil der

Applikation sein. Die hier vorgelegte Anwendung für das Krankheitsfeld Parkinson zeichnet sich durch die folgenden Kernelemente aus:

- Kurzweilige Bewegungsübungen in Form von Vormachvideos gemäß Empfehlungen fachspezifischer Therapeut:innen. Die Übungen werden von einem bundesweit sehr bekanntem Schauspieler und Komiker, Markus Maria Profitlich, vorgeführt.
- Universalität der Übungen, indem diese auch bei sehr fortgeschrittenem Krankheitsbild immer noch ausgeführt werden können.
- Einfache Bedienung und Erinnerungsfunktion der Applikation
- Belohnungssystem über eine Sammelalbum-Funktion (Gamifizierung)

Die Anwendung soll auf den mobilen Betriebssystemen Android und iOS gleichermaßen lauffähig sein und idealerweise nur in einer Code-Strecke implementiert werden. Dies reduziert Wartungsaufwendungen signifikant und erlaubt schnelle Releasezyklen bzw. die effiziente Anpassung der Applikation an weitere Krankheitsbilder. Weiterhin soll die Anwendung vollends den Anforderungen an den Datenschutz genügen und keine persönlichen Daten offenlegen (privacy by design).

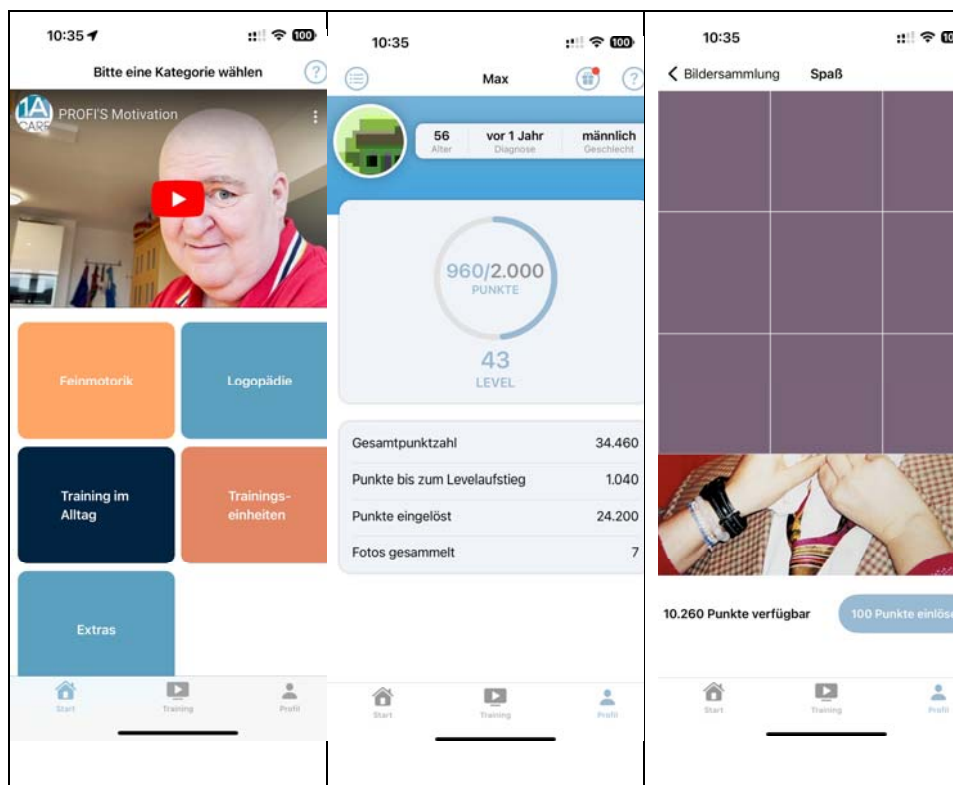
Die mobile Anwendung ist kostenfrei und uneingeschränkt verfügbar. Limitierungen und/oder ergänzende Geschäftsmodelle wie bspw. Werbung wurden nicht implementiert.

2 Inhaltlicher Aufbau der MotivationsApp für Parkinsonerkrankte

Die mobile Anwendung (App) für Parkinson-Betroffene wurde so konzipiert, dass mittels unterhaltsamer Trainings- und Übungsvideos diese regelmäßig zur Bewegung animiert werden sollen. Die Grundfunktionen bilden hierbei Anmeldung bzw. Registrierung, das Anlegen eines Profils und das beliebige Favorisieren und Abspielen der Trainingsvideos. Letzteres erlaubt damit auch die Erstellung eines persönlichen Trainingsplans. Ein einfaches Benachrichtigungs- bzw. Erinnerungssystem ergänzt diese Funktionen.

Völlig neuartig gestaltet sich die Hinzunahme eines Spielsystems (Gamifizierung) zum Nachweis durchgeführter Übungen. Das Grundprinzip hierbei ist, dass ein in Kacheln zerlegtes, exklusives Bild zunächst unsichtbar ist, schrittweise aber aufgedeckt werden kann, indem Übungen absolviert werden. Statt nur Punkte zu sammeln, wie es viele spielerische Anwendungen nutzen, werden so Neugier und Anreize weiter zu machen gleichermaßen geweckt. In der ersten Ausbaustufe handelt es sich um besondere und hochwertige Photographien von Markus Maria Profitlich. In der nächsten Version soll es möglich sein, dass Angehörige Bilder den Betroffenen zukommen lassen können, wie bspw. Familienbilder, Urlaubsbilder usw., die aber nur über das Erreichen von Punkten zur Aufdeckung der zunächst verdeckten Bilder angeschaut werden können. So entsteht kontinuierlich ein Anreiz Übungen durchzuführen, ohne dass wie bei vielen Gamifizierungslösungen das Problem des Verlassens des sogenannten Flow-Korridors auftritt (vgl. [Ha14]).

Abbildung 1a) zeigt die Kategorisierung der Bewegungsübungen in Form von Videos, Abbildung 1b) den aktuellen Punktestand und weitere Erfolgsparameter und Abbildung 1c) die Umsetzung des Spiels zum Aufdecken von Bildern. Das Grundprinzip der einfachen Menüführung wird stringent eingehalten, zumal zu beachten ist, dass je nach Krankheitsfortschritt die Feinmotorik, wie diese zum Betätigen von Buttons oder Listen notwendig sein kann, der Betroffenen stark eingeschränkt sein kann. Auf feine Strukturen wird daher weitestgehend verzichtet. Die Videos bzw. Übungen werden zusätzlich beschrieben und in Textform dokumentiert. Grundsätzlich werden aber alle Videos beim Abspielen von Markus Maria Profitlich sehr unterhaltsam erklärt. Final wird auf eine passende Gesamtspieldauer der Videos geachtet, indem diese im Regelfall eine Spieldauer von 2 bis 3 Minuten aufweisen. Sämtliche Übungen sind eng mit fachkundigen Therapeut:innen abgestimmt worden und decken damit die aktuellsten wissenschaftlichen Erkenntnisse ab.



Abbildungen 1: a) Auswahl der Trainingsvideos aus unterschiedlichen Kategorien; b) Aktueller Punktestand und „gewonnene Bilder“; c) Aufdecken von Bildern mittels erzielter Punkte über das Durchführen von Übungen

3 Technischer Aufbau der Mobilen Applikation

Die Motivationsapplikation für Parkinson-Erkrankte ist zunächst nativ mittels iOS Swift für Apple-Geräte entwickelt und Oktober 2022 produktiv gesetzt worden. Die Gründe hierfür waren vielschichtig und wurden sowohl von Anforderungen aus der Reihe Betroffener wie auch aus dem Vorhandensein von Erfahrungen im Umfeld der iOS Swift Entwicklungsumgebung heraus getragen. Mit Veröffentlichung der App und einer hieraus entstandenen medialen Präsenz (siehe bspw. [Bi22], [nt22]) stieg der Druck, diese auch auf der zweiten, aber meistverbreiteten mobilen Plattform Android, verfügbar zu machen. Eine parallele, native Entwicklung mit Android Studio wurde aus wirtschaftlichen Gründen heraus verworfen, da es praktisch zu einem doppelten Aufwand in der Wartung und Weiterentwicklung der Anwendung führt. Die Entscheidung fiel für einen Cross-Plattform-Ansatz, der beide Plattformmodelle aus einer Code-Linie heraus bedienen sollte. Dabei ist in Kauf zu nehmen, dass dieser Ansatz Performance- und User-Experience-Nachteile gegenüber nativen Anwendungen mit sich bringt. Diese Aspekte wurden im Vorfeld ausgiebig untersucht und konnten schlussendlich im Zusammenhang mit den nichtfunktionalen Anforderungen der Parkinson-App verworfen werden. Dabei war es von Vorteil, dass die Applikation im Rahmen des Entscheidungsprozesses bereits auf der iOS-Plattform produktiv im Einsatz war und hierdurch direkt relevante Parameter (Benchmarks) ermittelt werden konnten.

Zur Entscheidung standen 3 Plattformkonzepte, die nach der Methode der Nutzwertanalyse miteinander verglichen wurden:

- React Native, entwickelt von Facebook/Meta unter Verwendung von React und Javascript
- Xamarin, entwickelt von Microsoft unter Verwendung von .NET und C#
- Flutter, entwickelt von Google und das jüngste Produkt dieser Techniken unter Verwendung der Programmiersprache Dart

In den Entwicklercommunities werden regelmäßig Vergleiche von Cross-Plattform-Konzepten durchgeführt, die ebenfalls für den Entscheidungsprozess recherchiert wurden (siehe bspw. [Sc22]).

Die Entscheidung fiel schlussendlich auf das Framework Flutter, da es u.a. die besten UI-Fähigkeiten besitzt und die so erzeugten Implementierungen nur geringste Unterschiede zu den nativen Anwendungen aufweisen. Letzteres gilt insbesondere auch für den Aspekt der Performance, wenngleich dieser Punkt weniger Relevanz für die umgesetzte Parkinsonanwendung besitzt. Nachteilig ist die bis dato noch nicht weit verbreitete Programmiersprache, die eine anfängliche Eintrittshürde für dieses Framework darstellt.

Der Applikation liegt das MVVM-Entwurfsmuster (Model View ViewModel) zugrunde. View und ViewModels kommunizieren über Datenbindung (data binding) und Kommandos (commands). Damit wird die Benutzerschnittstelle von der Geschäfts- bzw. Anwendungslogik abstrahiert. Eine effizientere Wartung, Weiterentwicklung und

einfacheres Testen ist hierdurch möglich geworden. Die somit vorhandene lose Kopplung und Modularität ermöglicht zudem eine vereinfachte Anpassung an plattformspezifische Funktionen. Dies ist zudem wesentlich für die Portierung der initialen iOS-Anwendung in das Flutter-Framework gewesen.

Die Motivationsvideos wurden professionell und hochauflösend aufgezeichnet, so dass sich selbst mit den gängigen Komprimierungstechniken Dateigrößen von 50MB und mehr per Video ergaben. Im nächsten Schritt wurden diese auflösungsseitig um den Faktor 10 in der Dateigröße herunterskaliert und ein privater youtube-Videokanal zur zentralen Speicherung der Videos genutzt. Eine bereits vorhandene Bibliothek `youtube_player_flutter` [Fl18] konnte direkt Einsatz finden und reduzierte den Entwicklungsaufwand signifikant. Zusätzlich kommt das Cloud-Framework Firebase zum Einsatz, um die zusätzlichen Kernfunktionen der App, wie die erzielten Punkte und deren Verwendung für das Bildkachelaufdecken, zu verwalten. Die Applikation muss damit in der jetzigen Version zwingend über eine Internetverbindung verfügen, welche aber bandbreitenoptimiert ist und zudem keine datenschutzbedenklichen Aspekte beinhaltet.

In einer weiteren Ausbaustufe ist es verhältnismäßig leicht möglich, die Applikation auch offline betreibbar zu machen, indem favorisierte oder auch die kompletten Videodateien auf dem mobilen Endgerät persistiert werden. Dies erzwingt damit aber einen höheren lokalen Speicherressourcenbedarf. Bis dato und basierend auf einer über einjährigen Nutzung der Anwendung im realen Umfeld ist dieser Aspekt niemals von Relevanz geworden, so dass diese Anforderung aktuell nicht weiter verfolgt werden soll.

4 Ergebnisse, Zusammenfassung und Ausblick

Die Motivations-Anwendung ist seit Oktober 2022 für Apple-Geräte verfügbar [Ap22], für Android-basierte Modelle seit Mai 2023 [Go23]. Auf beiden Plattformen läuft die App ausgesprochen stabil. Weder signifikante Fehler in der Software noch Absprungraten der Anwender sind zu verzeichnen. Dies demonstriert, dass die Nutzung der Plattform Flutter als Cross-Plattform-Lösung eine richtige Entscheidung war. Mit Beginn der Veröffentlichung stieg die Anzahl der Nutzer kontinuierlich und organisch auf ca. 5.000 mit Stand April 2024. Werbung oder anderweitige Mechanismen zur beschleunigten Verbreitung werden aktuell nicht eingesetzt. Das Nutzerfeedback auf beiden Plattformen liegt im Höchstbereich und dokumentiert zusätzlich Nutzen und Erfolg dieser Anwendung. Die App wird darüber hinaus in unterschiedlichen Informationskanälen für Parkinson-Erkrankte platziert und empfohlen (siehe z.B. [Pa24]).

Im nächsten Schritt soll nun wissenschaftlich fundiert die Evidenz, d.h. die Wirksamkeit dieser Therapiergänzung aufgezeigt werden. Hierzu ist es notwendig, eine entsprechende kontrollierte Studie an Betroffenen umzusetzen. Im Rahmen einer randomisierten Doppelblindstudie werden App-Nutzer und Nicht-App-Nutzer im Zeitverlauf miteinander verglichen, um statistisch validierte Aussagen zur Evidenz der Anwendung treffen zu können. Dieser Schritt wird es final erlauben, die App einem Zulassungsprozess zur

Digitalen Gesundheitslösung (sogenannte DiGA) zuzuführen [Bf24]. Parallel ist es angedacht, das Konzept der Applikation auf weitere Krankheitsfelder zu übertragen. Im Fokus stehen hierbei zunächst die Krankheitsbilder Demenz, Alzheimer, Diabetes und post-operative Therapiefelder. Schlussendlich ist es geplant, die Migration auf ein nationales/europäisches Cloud-Framework durchzuführen, um den Zulassungsanforderungen der DiGA gerecht zu werden..

Mit Bezug zur Wirtschaftsinformatik erfordern, wie in zahlreichen Untersuchungen aufgezeigt, nicht nur das Krankheitsbild Parkinson, sondern generell digitale, medizinische Anwendungen bewusste Maßnahmen zur Steigerung der Nutzerakzeptanz (vgl. für einen Überblick [Su19]). Die von uns entwickelte Applikation verzichtete daher bewusst auf erweiterte Funktionen, wie diese bspw. im Rahmen der Integration einer SmartWatch möglich sind. Zudem wurde sowohl im Design wie auch in der Bedienung der Anwendung jedwede Komplexität vermieden. Die kaum messbare Absprungrate der aktuellen Nutzerbasis ist sicherlich ein Indiz dafür, dass die von uns umgesetzten Ansätze zu einer hohen Nutzerakzeptanz beitragen.

In einem zusätzlichen Prototypen wurde parallel eine weitere Technologiedomäne erschlossen. In einer virtuellen Umgebung (erzeugt durch das Tragen einer VR-Brille) müssen im Rahmen eines Spiels fiktive Türen geöffnet werden. Hierbei wird ein Sollöffnungsablauf vorgegeben, der therapeutisch festgelegt wurde (wie bspw. ein weiter, ausholender Öffnungsvorgang in einem vergrößerten Abstand zur simulierten Tür), und im Rahmen der Ausführung mittels Inertialsensorik erfasst sowie beurteilt. Erste Experimente mit Parkinson-Erkrankten verliefen vielsprechend und sollen ebenfalls weiter verfolgt werden. Auch hier erfolgt eine direkte Bewegungsmotivation, eingebettet in ein einfaches, kurzweiliges Spiel.

Literaturverzeichnis

- [Ap22] App Store; PROFIS Motivation; <https://apps.apple.com/de/app/profis-motivation/id1631284796>; Version 1.1. (2022)
- [Bf24] Bundesinstitut für Arzneimittel und Medizinprodukte; DiGA Digitale Gesundheitsanwendungen; https://www.bfarm.de/DE/Medizinprodukte/Aufgaben/DiGA-und-DiPA/DiGA/_node.html (2024)
- [Bi22] Bild online; Seine App soll Alles besser LACHEN. Markus Maria Profitlich hilft Parkinson-Erkrankten; <https://www.bild.de/regional/koeln/koeln-aktuell/markus-maria-profitlich-hilft-parkinson-kranken-mit-eigener-app-81940642.bild.html> 15.11.2022
- [Bu15] Burtscher, M.; Körperliche Inaktivität und Erkrankungen; in: Therapeutische Umschau, Sport- und Bewegungsmedizin; Volume 72, Issue 5, <https://doi.org/10.1024/0040-5930/a000679>, May 2015.
- [Co23] Collado-Mateo, D. et al.; Key Factors Associated with Adherence to Physical Exercise in Patients with Chronic Diseases and Older Adults: An Umbrella Review; in: Int. J.

- Environ. Res. Public Health 2021, 18(4); <https://doi.org/10.3390/ijerph18042023>; (2023)
- [Fe93] Fertl, E. et al.; Physical activity and sports in patients suffering from Parkinson's disease in comparison with healthy seniors. J. Neural Transm. Park. Dis. Dement. Sect. **5**, p. 157–161 (1993).
- [Fl18] Flutter – Google; Play and Pause a Video; in Official Flutter Documentation; <https://docs.flutter.dev/cookbook/plugins/play-video> (2018)
- [Go22] Google Play; PROFIS'S Motivation; https://play.google.com/store/apps/details?id=de.einsacare.motivation&hl=en_US (2023)
- [Ha14] Hamari, J., Koivisto, J.; Measuring flow in gamification: Dispositional Flow Scale-2. Computers in Human Behavior. DOI: 10.1016/j.chb.2014.07.048 (2014).
- [Is15] Iso-Markku, P., et al.; Physical activity and dementia: Long-term follow-up study of adult twins. Ann Med, 2015: p. 1–7. (2015)
- [Lee12] Lee, I.M., et al.; Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. Lancet, 2012. 380(9838): p. 219–29.
- [Pa24] PARKOUR; „Profi's Motivation“ bringt in Bewegung; PARKOUR, das Parkinson Magazin, Ausgabe 16, 2024
- [Nt22] NTV online Der Tag; Mit einer Prise Humor: Comedian hilft Parkinson-Betroffenen mit App; https://www.n-tv.de/der_tag/Mit-einer-Prise-Humor-Comedian-hilft-Parkinson-Betroffenen-mit-App-article23718738.html 15.11.2022
- [Sc22] Scanbot SDK; Flutter vs. React Native vs. Xamarin: Pros and cons of the leading cross-platform frameworks; <https://scanbot.io/developer/techblog/flutter-vs-react-native-vs-xamarin/> 13.10.2022
- [Su19] Suman, R., Sahibuddin, Sh.; User Acceptance Testing in Mobile Health Applications; in: Proceedings of the 2nd International Conference on Information Science and Systems; <https://dl.acm.org/doi/abs/10.1145/3322645.3322670>; p. 145-149, (2019).
- [Ta09] Tanaka, K. et al. Benefits of physical exercise on executive functions in older people with Parkinson's disease. Brain Cogn. 69, p. 435–441 (2009)
- [Ts22] Tsukita, K.; Long-Term Effect of Regular Physical Activity and Exercise Habits in Patients With Early Parkinson Disease; in: Neurology Volume 98, Issue 8, <https://doi.org/10.1212/WNL.0000000000013218>; February 2022; p. 859-871.
- [Wi15] Wiskemann, J.; Friedmann-Bette, B.: Potenzial körperlicher Aktivität in Prävention und Therapie chronischer Erkrankungen; In: Aktuelle Ernährungsmedizin 2015; 40(06); Georg Thieme Verlag, DOI: 10.1055/s-0035-1552782, S. 389-408.

Einsatz von Generativer KI im Self-service Micro-retailing

Kurzbeitrag und Prototyp

Erster Anwendungsfall für einen regionalen Lebensmittel-Shop und Aufbau eines Living-labs zur Untersuchung des phygitalen Einkaufserlebnisses

Martin Przewloka¹, Markus Häfele²

Abstract:

Personallose Selbstbedienungssshops im Einzelhandel sind im Vormarsch. Die Gründe hierfür sind vielfältig und aller Voraussicht nach auch im Zusammenhang mit Veränderungen, die durch die Pandemie ausgelöst wurden, zu sehen. Personalmangel, Schaffung von Möglichkeiten einer flexiblen, dezentralen und effizienten Versorgung wie auch Aspekte der Nachhaltigkeit können hierbei aufgezählt werden. Einen besonderen Stellenwert nimmt die Vermarktung regionaler Lebensmittelprodukte ein. Beinahe zwingend ergibt sich hieraus, dass dies mittels kleinerer, dezentraler Verkaufseinheiten umgesetzt werden kann. Dem entgegen stehen Kosten der zu schaffenden Infrastruktur und für den Betrieb, die schlussendlich aus Effizienzgründen den Einsatz digitaler Technologien erfordern. Im Rahmen dieser Arbeit soll ein Versuchsfeld (Living-lab) geschaffen werden, um Chancen, Risiken und Handlungsempfehlungen für den Einsatz generativer KI-Techniken innerhalb sogenannter Micro-shops zu untersuchen. Primär sollen hierbei einfache bis hin zu multimediale ChatBot-Systeme zum Einsatz kommen, mit dem primären Ziel, das Einkaufserlebnis der Kund:innen zu steigern. Das sekundäre Ziel besteht darin, einen effizienteren Betrieb dieser Verkaufsformen sicherzustellen.

Keywords: Micro-shops, Micro-retailing, Künstliche Intelligenz, phygital, Generative KI, Akzeptanz, ChatBots, Customer Experience, Einkaufserlebnis, Self-service, Digitale Transformation, Nachhaltigkeit, ChatGPT, LangChain, RAG

1 Einleitung

Wenngleich sich der Begriff des Micro-retailing (siehe bspw. [Fo92], [Br93], [Zi04]), welches den Trend zu immer kleiner werdenden Verkaufseinheiten beschreibt, im bundesdeutschen Umfeld noch vergleichsweise wenig Verbreitung gefunden hat, so sind die hiermit verbundenen Artefakte schon sehr lange eingeläutet. Als wesentliche Gründe hierfür sind zu sehen:

- Flexiblere Öffnungszeiten
- Effizienteres Warenmanagement durch eingeschränktes Sortiment
- Etablieren von Teststores für Innovationen und Neuprodukte sowie Premiumstores, Store-Experimente mit verringertem Risiko (vgl. [Br13])

¹ Institut für Digitale Assistenzsysteme e.V., Technologiezentrum Gießen, Kerkrader Straße 9, 35394 Gießen

² EMMI'S 4.0, Markus Häfele, Taubenäcker 25, 35435 Wettenberg

- Ergänzung des eCommerce Angebots durch physische (Kleinst-)Verkaufsstellen
- Shops für ländliche und wenig entwickelte Regionen (vgl. [Eu23])
- Rückbesinnung auf die lokale Versorgung der Konsument:innen (vgl. [Fe15])

Eine Besonderheit stellt hierbei der Lebensmittelmarkt dar, der von zusätzlichen, wesentlichen Trends geprägt wird (siehe u.a. [Ha24]):

- Gesunde, nachhaltige Ernährung und Produktangebote
- Das vermehrte Angebot von lokalen Produkten aus der Landwirtschaft, dem Bäckerei-, Fisch- und Fleischereigewerbe
- Die Experience Economy, d.h. die Schaffung eines Einkaufserlebnisses angereichert durch Zusatzangebote wie Rezeptkarten, Kochkurse sowie eine verstärkte Interaktion zwischen Kund:innen und Anbieter bzw. Erzeuger

Je kleiner sich die Verkaufseinheiten darstellen, desto mehr gilt es, den Herausforderungen (Fläche, Warenumsatz, Skalierbarkeit etc.), die diese mit sich bringen, zu entgegen. Dies gelingt genau dann, wenn sich das Produktangebot automatisiert und mit geringstmöglichem Personaleinsatz vermarkten lässt. Hersteller bzw. Erzeuger von Lebensmittelprodukten können dann in die Lage versetzt werden, ihr Angebot direkt vor Ort zu verkaufen und erzielen damit einen zusätzlichen Ertrag, da Zwischenhändlermargen entfallen. Zusätzlich ist es möglich, den Kund:innen ein völlig neues Einkaufserlebnis zu schaffen, sei es, indem ein erhebliches Maß an Vertrauen in die Produkte und deren Erzeugung transportiert wird, ein sich ständig neues, überraschendes und saisonales Sortiment angeboten wird oder Zusatzangebote platziert werden (z.B. die Selbsternte von landwirtschaftlichen Produkten).

Der Entfall des Personals stellt eine der größten Herausforderungen des Micro-retailing dar. Die Gefahren des Diebstahls, Vandalismus und ähnliche Tatbestände werden häufig hervorgehoben und sind nicht vollends vermeidbar. Ab einer gewissen Missbrauchsschwelle, die es zu akzeptieren gilt, übersteigen die Aufwendungen erweiterter technischer Maßnahmen zur Schadensreduktion den potenziellen Schaden. Von viel höherer Relevanz ist die Fokussierung auf die Schaffung eines Einkaufserlebnisses, welches Verkäufer:innen und Berater:innen nicht vermissen lässt. Fragen, die Kund:innen zu Produkten, zum Bezahlvorgang und darüber hinaus haben, dürfen nicht unbeantwortet bleiben, da dies zu Unsicherheit, Unzufriedenheit und schlussendlich auch zu einem nicht nachhaltig positivem Einkaufserlebnis führt.

Die Einführung eines direkt zugänglichen ChatBots-Systems, welches kontinuierlich über aktuelle Informationen zu den Produkten, deren Verwendungsmöglichkeiten, Erfahrungen uvm. führt, kann genau diese Lücke schließen. Dabei kann der Zugang über Voice oder Texteingabe ermöglicht werden, verbunden mit einer variierbaren Dynamik, mit der das virtuelle System und Kund:innen in den Dialog treten.. Das entwickelte System soll so gestaltet sein, dass es technisch beliebig skaliert werden kann. Es bildet so die Basis für detaillierte Untersuchungen zum Verbraucherverhalten im personallosen Micro-retailing.

2 Der Micro-shop EMMI'S 4.0

Die erste Implementierung einer ChatBot-Applikation zur Steigerung des Einkaufserlebnisses sollte in-situ (Living-lab) umgesetzt werden. Hierfür konnte der Micro-shop EMMI'S 4.0 [EM40] identifiziert werden, der alle wesentlichen Elemente einer modernen, hochflexiblen Kleinstvermarktungseinheit für Lebensmittelprodukte aufweist:

- Angebot und Direktverkauf von saisonalen Vor-Ort-Regionalprodukten; klare Sortimentfokussierung und Distanzierung vom Kernsortiment-retailing
- Mobiles, autarkes und preisgünstiges Komplettsystem, welches schnell und einfach installiert und in Betrieb genommen werden kann
- Vollständig ausgelegt für den sicheren, personallosen self-service-Betrieb
- Erprobt, aber auch ständig offen für Innovationen

Der Micro-shop EMMI'S 4.0 ist ein Container-System mit einer vergleichsweise geringen Verkaufsfläche, die aber genau den Anforderungen von Erzeugern und Anbietern für eine flexible Vor-Ort-Vermarktung entspricht. Besonderer Wert wurde hierbei auf die Entwicklung eines preisgünstigen, leicht transportierbaren Systems gelegt, um die Eintrittsschwelle zur Anschaffung oder Anmietung so gering wie möglich zu halten.

Gleichermaßen wurden die Anforderungen an einen sicheren und für die Kund:innen angenehmen Einkauf, wie Klimatisierung und adäquate Beleuchtung, berücksichtigt. Die wesentlichen Daten einer Verkaufseinheit sind (siehe auch Abbildungen 1):

- Nettoverkaufsfläche 2,0m x 2,3m, Platzierung von 2 bis 3 Kühleinheiten plus zusätzliche Verkaufsregale problemlos möglich
- Raumhöhe innen mindestens 2,4m
- Elektronischer Zugang mit Vorabauthentifizierung (Kredit-/Bankkarte), Self-service Kassensystem und Kamera-Überwachung des Innenraums
- Handhabbares Containergewicht von netto 1.000kg, bewegbar mit einem Stapler
- Erfüllung sämtlicher regulatorischer Auflagen incl. Abnahmefähigkeit durch die Lebensmittelbehörden
- Micro-Solaranlage mit Batteriepuffersystem (Leistung im Regelfall mit 1kWp völlig ausreichend)



Abbildungen 1: Außen- und Innenansicht des Micro-Verkaufscontainers EMMI'S 4.0 und Produktplatzierung in einer Kühleinheit

EMMI'S 4.0 distanziert sich bewusst von den Ansätzen der Anbieter von reduzierten Grundversorgungssortimenten. Hierzu zählen bspw. die Mini-Shops an Tankstellen und Bahnhöfen wie auch die vermehrt sichtbaren Automatenangebote. Ebenso ist als Abgrenzung die Markteinführung des Shopsystems Teo ([Te24]) als Ableger der tegut... gute Lebensmittel GmbH & Co. KG zu nennen, welches den Ansatz einer angebotsreduzierten Grundversorgung (Kernsortiment-retailing) mit beinahe 1.000 Artikeln verfolgt.

3 Einsatzszenario für eine Generative ChatBot KI zur Steigerung des Einkaufserlebnisses im EMMI'S 4.0 Micro-shop

Die Einsatzmöglichkeiten, Chancen und Risiken des Einsatz von ChatBot-Lösungen im Handel und hierbei maßgeblich im eCommerce, sind ein kontinuierlicher Untersuchungsgegenstand (siehe bspw. [Ch21], [CH20]). Der technologische Fortschritt, wie jüngst die breite Positionierung und Verfügbarmachung von generativen KI-Systemen, hat zu erweiterten Lösungen geführt und sich deutlich dem natürlichen Dialog, wie er zwischen Menschen geführt wird, angenähert. Der maßgebliche Fokus liegt hierbei in der Steigerung der Customer Experience im ausschließlich digitalen Kontext, wie bspw. im Rahmen des online-Shoppings (vgl. [Ch21]).

Hingegen sind Untersuchungen, insbesondere mit wissenschaftlicher Methodik, in einem kombiniert physischen und digitalen/virtuellen („phygital“) Umfeld kaum zu finden. Aus Sicht der Kund:innen ist es leicht nachvollziehbar, dass gerade im Zusammenhang mit personallosen, physischen Verkaufseinheiten ein Bedarf an ad-hoc-Informationen und -Kommunikation besteht, die es darzustellen gilt.

Mögliche Fragestellungen sind:

- „Wie ist das Produkt angebaut worden, welche Bodenaufbereitungsmittel kamen zum Einsatz, wie wurde das Produkt geerntet?“
- „Über welchen Zeitraum wird dieses Produkt in der kommenden Zeit verfügbar sein?“
- „Wie lange kann ich dieses Produkt zu Hause im Kühlschrank aufbewahren? Eignet es sich zum Einfrieren und wenn ja, was sollte ich dann beachten?“
- „Gibt es eine besondere, regionale Zubereitungsempfehlung für dieses Produkt?“
- „Welches der angebotenen Produkte ist der Top-Seller in diesem Monat?“
- „Kann ich auch mit einer Debit-Karte bezahlen?“
- Uvm.

Die Bereitstellung dieser Informationen über vorgefertigtes Prospektmaterial, Plakate oder gar den Verweis auf online-Seiten, gelinkt mittels gedruckter QR-Codes, ist verwirrend und medienbrechend. Schlussendlich kann dies zu einem deutlich verminderten Einkaufserlebnis führen. Die neue Lösung besteht darin, dass ein virtueller Vorort-Berater, zunächst implementiert in Form eines in natürlicher Sprache kommunizierenden ChatBots (Tablet-System, vgl. Abbildungen 2), die sich so ergebende Lücke schließen und zum Einsatz kommen soll. Die breite Verfügbarkeit vortrainierter, generativer KI-Modelle soll die Basis für das erste System bilden (MVP). Hierauf aufbauend sollen die Living-lab-Untersuchungen zum Kaufverhalten aufgesetzt werden.

Das System wurde so universell konzipiert, dass ein Ausrollen einer sprachfähigen Variante ebenfalls sofort möglich ist. Im Rahmen des Living-lab-Ansatzes sollen nun folgend unterschiedliche Ausprägungen des Bots (Sprache vs. Texteingabe, robot-push³, human-pull⁴, Mischformen) sowie das Akzeptanzverhalten im Zusammenhang mit unterschiedlichen Kommunikationsformen während des Einkaufs untersucht werden.

Die Kommunikation der Verkaufseinheit erfolgt über ein im Container verbautes Mobilfunkgateway. Eine 4G-Verbindung, wie diese heute in Deutschland grundsätzlich verfügbar ist⁵, erfüllt sämtliche Anforderungen an Bandbreite und Latenz der verbauten Anwendungen.

4 Technische Umsetzung des Demonstrators und Minimum Viable Products als Basis für künftige Living-lab-Untersuchungen

Das Ziel des Vorhabens bestand in der Entwicklung einer flexiblen und beliebig skalierbaren ChatBot-Anwendung, die in der Lage ist, in natürlicher Sprache für sämtliche Fragen und Probleme im Kontext eines Einkaufs in einem personallosen Micro-shop zur

³ Der ChatBot = Roboter ist proaktiv in der Kommunikation, d.h. Kund:innen werden aktiv angesprochen

⁴ Der ChatBot = Roboter ist reaktiv in der Kommunikation, d.h. er reagiert nur auf Fragen der Kund:innen

⁵ In bis dato nicht aufgetretenen Ausnahmesituationen sind alternative Internetverbindungen umsetzbar

Verfügung zu stehen. Wesentliche nichtfunktionale Kriterien sind eine komplexitätsfreie und einfache Bedienung sowie die Möglichkeit einer datenschutzkonformen Erfassung von Daten zum Einkaufsverhalten. Hierzu bot es sich an, zunächst allgemein verfügbare, vortrainierte KI-Sprachmodelle zu verwenden, mit der Einschränkung, dass diese eine ausreichende Abdeckung im Bereich der deutschen Sprache aufweisen. Die erste Wahl fiel zunächst auf die kommerziellen Sprachmodelle ChatGPT 3.5 und ChatGPT 4.0, da diese bereits in Vorprojekten erfolgreich Einsatz fanden und das technische know-how zur Einbindung in Individualanwendungen vorhanden war. Aus den gleichen Überlegungen heraus, wurde eine Python-Umgebung gewählt, um die Anwendung zu erstellen.



Abbildungen 2: Point of Sale and Interaction – ChatBot Kommunikation über Tablet und bargeldloses Kassen- bzw. Bezahlssystem. Direktkommunikation einer Kundin mit dem ChatBot im Rahmen ihres Verkaufsbesuchs im EMMI'S 4.0 - Container

Parallel wurden Dokumentationen zu den Produkten, Rezepten, den Anbaumethoden wie auch zum Ablauf des Einkaufs wie insbesondere dem Bezahlvorgang erstellt und so aufbereitet, dass diese als erweiterte Lerndaten der ChatBot-Anwendung zugeführt werden können. Hierbei kam im ersten Schritt das LlamaIndex Framework (vgl. [LI24]) zum Einsatz, da es sämtliche benötigte Funktionen zum Laden, Indizieren und Querying der Daten in geeigneten Bibliotheken zur Verfügung stellt. Auf die Implementierung einer RAG-Anwendung (Retrieval Augmented Generation, [Le20]) wurde im ersten Schritt ebenfalls verzichtet, ist aber in einer Folgeausbaustufe explizit vorgesehen. Die anonymen user-Interaktionen (statistische Kennzahlen, Sentimente, Fragedomänen, Käufe) werden von der Anwendung klassifiziert und für spätere analytische Auswertungen persistiert.

In ersten Untersuchungen konnte gezeigt werden, dass das vortrainierte Modell GPT 3.5 bereits eine ausreichende Präzision erreicht. Zudem besitzt es einen Geschwindigkeitsvorteil im Antwortverhalten gegenüber dem deutlich erweiterten Modell GPT 4.0, welcher damit der user experience zugute kommt (siehe bspw. in einfacher Gegenüberstellung [Bö23]). In allen Fällen muss die von OpenAI kostenpflichtig zur Verfügung gestellte API genutzt werden. Nachteilig für den Anwendungsfall ist das damit von OpenAI aktuell verfolgte Geschäftsmodell, da dieses seine Erlöse maßgeblich aus der Menge der verarbeiteten Daten erzielt (vgl. [Re24] und [OA24]). Dies kann aus Anwender:innensicht ein erhebliches Risiko darstellen, da der Datenverkehr nur schwer

kalkulierbar ist. Mit dem hier realisierten MVP-Anwendung (minimum viable product) lagen die Betriebskosten im Bereich von 2 – 3 Cent pro user-Interaktion, welches noch akzeptabel erscheint. Allerdings ist schon jetzt die Umstellung auf alternative Sprachmodelle (incl. open source) angedacht, um Alternativen insbesondere hinsichtlich einer absicherbaren betriebswirtschaftlichen Kalkulation anbieten zu können.

Abbildung 2 (rechts) zeigt einen Dialog einer Kundin mit dem ChatBot-System. Basierend auf den zusätzlich antrainierten Daten ist der ChatBot in der Lage, die Frage der Kundin stichhaltig zu beantworten. Der Dialog ist in dieser Variante effizient und präzise gehalten. Eine Ausweitung auf einen stärker werbenden Dialog ist leicht umsetzbar und soll ebenfalls Grundlage der nun folgenden Untersuchungen des Verbraucherverhaltens werden.

5 Ergebnisse, Zusammenfassung und Ausblick

Das Ziel, die Entwicklung eines ersten Systems zur interaktiven Kommunikation zwischen Kund:innen und einer mobilen Verkaufseinheit, konnte vollends umgesetzt werden. Hierbei wurde zunächst das kommerzielle und generische Sprachmodell ChatGPT 3.5, eingebettet in einer dem Anwendungsfall entsprechenden mobilen Applikation, eingesetzt. Es wurde mit spezifischen Datenquellen aus dem Umfeld der Verkaufsprodukte, Verkaufsumgebung uvm. angelernt. Zum Einsatz kamen ausschließlich deutschsprachige Dokumente, um auch diesbezüglich die Robustheit des Systems zu validieren. Parallel wurden erfolgreich lokal installierbare Sprachmodelle getestet wie bspw. OpenChat-3.5-0106 [OC24].

Die Besonderheit dieses Projekts liegt in der Mehrdimensionalität der digitalen Anwendung, da diese einen fundamentalen Mehrwert für den Betrieb kleiner, mobiler Verkaufseinheiten darstellt. Es gibt keinerlei Beschränkungen der Anwendungsszenarien und oder der Verkaufsprodukte. Immer dann, wenn Fachpersonal nicht oder nur teilweise zur Verfügung steht, können diese Systeme beinahe barrierefrei und zudem sehr kostengünstig installiert werden. Zudem kann die Anwendung unmittelbar mehrsprachig arbeiten, da vortrainierte Sprachmodelle in den Kernsprachen mit guter Qualität zur Verfügung stehen und sich darüber hinaus außergewöhnlich schnell weiter entwickeln.

Es wurden diverse wissenschaftliche Ansätze zur Akzeptanzmessung des Einsatz von ChatBot-Systemen beschrieben (z.B.[Re20]). Mit dieser ersten assistierenden Implementierung sind nun die Grundlagen für das völlig neue Feld der in-situ-Untersuchungen zum Verhalten der Verbraucher im Micro-retailing mit digitaler Unterstützung geschaffen worden. Unter Berücksichtigung sämtlicher Aspekte des Datenschutzes können anonymisiert Akzeptanzuntersuchungen im Umgang mit dem ChatBot-System unter Variation der Parameter Text vs. Sprache vs. Bilder/Videos, passiver vs. aktiver ChatBot, Relevanz des Contents, uvm. umgesetzt werden. Gleichmaßen können Korrelationsuntersuchungen zur Veränderung des Einkaufserlebnisses sowie Szenarien neuer Geschäftsvarianten, wie bspw. die Möglichkeit

der Schaffung personenbezogener Angebote sowie Verkaufsförderungsmaßnahmen durchgeführt werden. In einer fortgeschrittenen Ausbaustufe sollen Untersuchungen zum Einfluss der Entwicklung einer ChatBot-Persönlichkeit auf die Customer Experience durchgeführt werden. Die Fragestellung besteht hierbei, ob die Schaffung von Wiedererkennungseffekten eines benannten ChatBots „EMMI“ (analog zu Siri oder Cortana) bis hin zur Platzierung von sozialen Signalen („social cues“)⁶ Einflüsse auf das Kaufverhalten der Kund:innen haben.

Literaturverzeichnis

- [Bö23] Böhl, L.; ChatGPT 3.5 vs.4: Wesentliche Unterschiede; in: Stuttgarter Nachrichten 13.11.2023
- [Br13] Bressoud, E.; Testing FMCG innovations: experimental real store versus virtual; In: Journal of Product & Brand Management, 22(4), pp. 286-292 (2013)
- [Br93] Brown, S.; Micro-scale retail location: Cinderella or ugly sister?. International Journal of Retail & Distribution Management, 21(7) (1993)
- [Ch20] Chui, M., Hazan, E., Roberts, R., Singla, A., Smaje, K.; The economic potential of generative AI. (2020)
- [Ch21] Chen, J. S., Le, T. T. Y., Florence, D.; Usability and responsiveness of artificial intelligence chatbot on online customer experience in e-retailing. International Journal of Retail & Distribution Management, 49(11), pp. 1512-1531 (2021)
- [EM40] EMMI'S 4.0; Produkte wie früher, Verkaufen wie heute; www.emmis4.com (2024); letzter Zugriff 25.4.2024
- [Eu23] Eustis, E. N., & Sonnenberg, H. J.; Transforming Micro-Retailing in Emerging Markets; (2023)
- [Fe15] Feldmann, C., Hamm, U; Consumers' perceptions and preferences for local food: A review. Food quality and preference; 40, pp. 152-164 (2015)
- [Fo92] Foxall, G. R., Hackett, P. M.; Consumers' perceptions of micro-retail location: wayfinding and cognitive mapping in planned and organic shopping environments. International Review of Retail, Distribution and Consumer Research, 2(3), pp. 309-327 (1992)
- [Ha24] Hassoun, A., et al; The fourth industrial revolution in the food industry—part II: Emerging food trends; In: Critical Reviews in Food Science and Nutrition, 64(2), pp. 407-437. (2024)
- [Le20] Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., ... & Kiela, D.; Retrieval-augmented generation for knowledge-intensive nlp tasks. Advances in Neural Information Processing Systems, 33, pp. 9459-9474. (2020)

⁶ Hierzu zählen verbale, visuelle, auditive und unsichtbare ‚cues‘ wie Geschlecht, Tonalität, Sprechtempo, spezifische Stimmmerkmale etc. des ChatBots

- [LI24] Llamaindex; Turn your enterprise data into production-ready LLM applications; <https://www.llamaindex.ai> ; (2024) letzter Abruf 25.4.2024
- [OA24] OpenAI; Introducing ChatGPT; <https://openai.com/blog/chatgpt/> ; letzter Zugriff 25.4.2024
- [OC24] OPENCHAT; Advancing Open-source Language Models with Mixed Quality Data; <https://huggingface.co/openchat/openchat-3.5-0106> (2024); letzter Zugriff 25.4.2024
- [Re20] Rese, A., Ganster, L., & Baier, D.; Chatbots in retailers' customer communication: How to measure their acceptance?; Journal of Retailing and Consumer Services, 56, 102176 (2020)
- [Re24] Reiff, N., Velasquez, V.; What Is ChatGPT, and How Does it Make Money?; <https://investopedia.com/what-is-chatgpt-7094342> 20.2.2024; letzter Zugriff 25.4.2024
- [Te24] tegut... teo; So kauft man heute ein; <https://www.teegut.com/teo.html> ; (2024) letzter Zugriff 25.4.2024
- [Zi04] Ziliani, C., & Bellini, S.; Retail micro-marketing strategies and competition. The International Review of Retail, Distribution and Consumer Research; 14(1), pp. 7-18. (2004)

Researching Financial Market Dynamics through Algorithmic Trading Agents

Short Paper

Paul Helstab¹, Christof Mosler², Daniel Kaiser³

Abstract: The evolving landscape of financial markets demands the integration of advanced technologies and quantamental paradigms like market simulations and artificial intelligence. This study explores the use of agent-based interactive discrete event trading simulations to analyze financial market scenarios and detect anomalies. The research involves various market fees and synthetic competition between exchanges, utilizing an inter-market spread arbitrage machine for price stability. This approach enables the observation of anomalies related to spreads, execution speed, traded volumes, liquidity, order execution probability, and agent decision-making influence. This study could serve as a valuable tool for financial institutions and regulatory authorities in strategic decision-making concerning financial market challenges.

Keywords: autonomous agents, simulations, financial markets, DEMAS

1 Introduction

Our research explores the use of Agent-based Interactive Discrete Event Simulations (ABIDES) for the analysis of complex financial market scenarios and the identification of potential anomalies. The goal is to experimentally simulate financial market issues instead of observing them, for example, in controversial pilot programs on real reference markets. Our work encompasses various market fees and a synthetic competition between two exchanges, employing an inter-market spread arbitrage machine to enable price stability between the exchanges. In other experiments various market scenarios were simulated, and key metrics such as the impact on execution quality, generated turnaround, and market share of exchanges were analyzed under different marketplace fee scenarios. This approach allows the observation of anomalies related to spreads, execution speed, traded volumes, liquidity, probability of order execution, and the impact on agents' decision-making.

¹ Boerse Stuttgart Group, Data Engineering, Börsenstraße 4, 70174 Stuttgart, Paul.Helstab@boerse-stuttgart.de

² Hochschule für Technik Stuttgart, Schellingstr. 24, 70174 Stuttgart, christof.mosler@hft-stuttgart.de

³ Boerse Stuttgart Group, Data Engineering, Börsenstraße 4, 70174 Stuttgart,
Daniel.Kaiser@boerse-stuttgart.de

2 ABIDES and Related Work

The application of scientific computing models from information and game theory with discrete event multi-agent simulations (DEMAs) in the study of financial markets has garnered significant attention among economic researchers [Ha12]. This is particularly relevant in light of the 2010 flash crash, in which Navinder Singh Sarao, acting from his bedroom, was able to trigger a significant disruption in the U.S. economy through the use of high-frequency trading (HFT) of the e-mini S&P 500 [VL20]. This incident highlighted the need for a deeper understanding of the underlying dynamics of such anomalies in order to ensure effective regulation within HFT.

The discrete event multi-agent simulation (DEMAs) experiments, developed and customized for our research, are implemented using the open-source framework ABIDES, introduced by Byrd et al. [BHB19] in collaboration with researchers from J.P. Morgan AI Research [Am21]. The framework is utilized to evaluate the defined reference market problems. ABIDES is a multipurpose discrete event multi-agent simulation environment designed based on the NASDAQ equity trading protocols ITCH and OUCH [Am21b]. ABIDES offers access to complex financial market problems, enables nanosecond precision, and repeatable double-auction trading simulations [BHB20].

There is vast literature on financial market simulations in general: A good overview is given e.g. in [AF22]. Discrete event multi-agent simulations (DEMAs) combine the advantages of parallel discrete event simulations (PDES), for fast and efficient execution, and agent-based models (ABM), for flexibility and logical processes. In [BHB19] all major research related to DEMAs and ABIDES is presented, including former research projects and frameworks.

Hedge funds and investment banks discovered the opportunities of experimental simulations like DEMAs as support for backtesting on investment strategies [Pr18, Am21]. As such, researchers utilized such test-bed environments in experiments to gain a more comprehensive understanding of the events that led to the flash crash [Ba19].

In order to establish the capability to infer deterministic causal relationships in such financial market scenarios, we extended the existing model and introduced a competing exchange to the original simulation model. This approach follows a similar methodology as a study from the Japan exchange group (JPX) [HMY22]. To achieve price stability and high liquidity between these exchanges, we implemented a so-called arbitrage machine between markets with the possibility of arbitrage transactions. O'Hara et al. describe such a strategy as inter-market spread arbitrage by HFT market makers [Oh13].

The development and injection of reinforcement learning (RL) agents into the presented ABIDES models was introduced in another study, addressing similar research goals [DVB22]. In the meantime, J.P. Morgan AI Research introduced Phantom, another Multi-agent reinforcement learning (MARL) framework, giving also its comparison to other RL frameworks in [Ar22].

3 Approach

For our research we use the Agent-based interactive discrete event simulations (ABIDES) framework described above and add some new developments to investigate the impact of various market fee structures on trading agents with two competing exchanges.

All experiments are implemented in Python. Logged data from the ABIDES framework is used to feed our newly developed post-experiment analysis dashboard. Due to its broad functionality and possibilities, the dashboard is implemented as a web-based application.

Our goal was to extend the base model experiments for market turnover and execution quality by a secondary competing exchange with separate market fee models, and provide specific agents the ability to make a price and fee-effective order placement decision. We wanted to realistically constitute a synthetic competition between two exchanges and obtain evaluable results for market share, turnover and execution quality. To achieve price stability and high liquidity between exchanges, a so-called arbitrage machine between markets with the possibility of arbitrage transactions was implemented. The primary objective of this inter-market spread arbitrage machine is to achieve price stability, akin to that of high-frequency traders, by utilizing an arbitrage strategy.

4 Results and Outlook

The primary objective of our experiments was to simulate a security's trading activity with a fundamental value of \$1,000 on a U.S. stock exchange during a usual trading day, which spans from 9:30 AM to 4:00 PM. In our experiments, we have implemented three different types of trading agents. The simulation aimed to replicate the dynamic interactions of market participants and the price movements in the security. The experiment was run five times, each utilizing a different distinct pseudorandom number generator (PRNG) seed.

The preliminary results are promising and demonstrate the potential utility of this framework. We could observe interesting behaviors between agents in different market scenarios and also conclude significant differences in the turnover and the speed of execution between the exchanges with different fee strategies. Our newly introduced arbitrage machine showed a realistic behavior in all simulations.

It should be emphasized that our experiments are still very simplified and that real-world financial markets are complex, but the objective should be to strive for comprehensive and highly realistic simulations. Further efforts are required to enhance these methods and optimize them into a fully operational system. There are plenty of options to extend the experiments presented in this research. E.g. the implementation of more extensive experimentation is planned for the future, including more market participants and a wider variety of trading strategies, in addition to more detailed parameterization. Also, the Machine Learning-based ABIDES-Gym with RL agents could be used to make decision even more realistic. These extensions could allow a more detailed analysis of HFT agents

between two exchanges and the observation of manipulation strategies such as price layering, spoofing, and quote stuffing [Oh13].

References

- [AF22] Axtell, R.L., Farmer, J. D.: Agent-Based Modeling in Economics and Finance: Past, Present, and Future, INET Oxford Working Paper No. 2022-10, available online: <https://www.inet.ox.ac.uk/files/JEL-v2.0.pdf>, 2022.
- [Am21] Amrouni, S. et.al.: Abidesgym: Gym environments for multi-agent discrete event simulation and application to financial markets, 2021.
- [Am21b] Amrouni, S. et.al.: ABIDESgym, in Proceedings of the Second ACM International Conference on AI in Finance, ACM, 11/2021, DOI: 10.1145/3490354.3494433, available online: <https://dl.acm.org/doi/10.1145/3490354.3494433>, 2011.
- [Ar22] Ardon, L. et.al.: Phantom – A RL-driven multi-agent framework to model complex systems, in Proceedings of the 2022 ACM International Conference on AI in Finance, 10/2022, available online: <https://arxiv.org/abs/2210.06012>, 2022.
- [Ba19] Balch, T.H. et.al.: How to evaluate trading strategies: Single agent market replay or multiple agent interactive simulation?, DOI: 10.48550/ARXIV.1906.12010, available online: <https://arxiv.org/abs/1906.12010>, 2019.
- [BHB19] Byrd, D.; Hybinette, M.; Balch, T. H.: Abides: Towards high-fidelity market simulation for ai research, available online: <https://arxiv.org/abs/1904.12066>, 2019.
- [BHB20] Byrd, D.; Hybinette, M.; Balch, T. H.: Abides: Towards high-fidelity multiagent market simulation, in Proceedings of the 2020 ACM SIGSIM Conference on Principles of Advanced Discrete Simulation, pp. 11–22, ISBN: 9781450375924, available online: <https://doi.org/10.1145/3384441.3395986>, 2020.
- [DVB22] Dwarakanath, K., Vyetenko, S., Balch, T.: Equitable marketplace mechanism design, in Proceedings of the Third ACM International Conference on AI in Finance, pp. 232–239, available online: <https://doi.org/10.1145/3533271.3561673>, 2022.
- [Ha12] Hayes, R. et.al.: Agent based model of the e-mini future: Application for policy making, in Proceedings of the 2012 Winter Simulation Conference (WSC), 12/2012, pp. 1–12. DOI: 10.1109/ WSC.2012.6465037, 2012.
- [HMY22] Hoshino, M.; Mizuta, T.; Yagi, I.: Investigation into effectiveness of maker-taker fees in stock markets using artificial market, JPX working paper, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4145195, 2022.
- [Oh13] O’Hara; de Prado; Easley: High-frequency Trading. London: Incisive Media Investments Ltd, 2013.
- [Pr18] de Prado, M. L.: Advances in Financial Machine Learning, 1st. Wiley Publishing, ISBN: 1119482089, 2018.
- [VL20] Verity, A.; Lawrie, E.: Hound of hounslow: Who is navinder sarao, the ’flash crash trader’? - bbc news, <https://www.bbc.com/news/explainers-51265169>, 2020, accessed 16/03/2024.

Simulationsgestützte Analyse der neuen dynamischen Strompreise und Alternativen zur Anreiz-Optimierung für Endkunden

Prototyp

Thomas Wiedemann ¹

Abstract: Die seit 1.1.2024 auch für private Endkunden verfügbaren dynamischen Strompreise werden durch die gleichzeitig gültige Gesetzgebung zu den Strommarktabgaben wie Netzentgelte, Konzessionsabgaben und Steuern weitgehend negiert, da fast immer 20 oder mehr €-Cent an konstanten Abgaben anfallen. Auf der Basis von stündlichen dynamischen Strompreisen seit Anfang 2023 werden verschiedene Simulationsszenarien zu Marktpreis- und Abgaben-Verteilungs-Alternativen modelliert und berechnet, welche deutlich höhere Anreize für Endkunden bezüglich des Speicherausbaus bewirken könnten. Eine reale Umsetzung würde eine Novellierung der Gesetze und Abgabenordnungen für Energieversorger erfordern. Ob dies im gegenwärtigen politischen und Markt-technischen Umfeld möglich ist, wird aktuell gemeinsam mit Energiemarkt-Experten diskutiert und evaluiert.

Keywords: Simulationsgestützte Analyse, dynamische Strompreise, Anreiz-Optimierung

1 Einleitung

Seit dem 1.1.2024 sind große Energieversorger gesetzlich verpflichtet, Endkunden optional dynamische Strompreise anzubieten [En23]. Ab 2025 betrifft dies alle Energieversorger. Diese dynamischen Strompreise bilden die Leipziger Strombörse ab. Je nach Marktnachfrage und Angebot an alternativen Energien schwanken die dynamischen Netto-Preise zwischen negativen Werten von bis zu -10 Cent/KWh und bis zu 40 Cent/KWh (vgl. Abb. 1). Leider wird dieser sehr interessante neue Ansatz der Politik durch die gleichzeitig gültige Gesetzgebung zu den Strommarktabgaben wie Netzentgelte, Konzessionsabgaben und Steuern wieder weitgehend negiert, da fast immer 20 oder mehr €-Cent an konstanten Abgaben anfallen. Bei 1 Cent Strombörsen-Netto-Preis liegt der Bruttopreis bei 21 Cent und die Abgabenlast folglich bei über 2000%! (vgl. Abb. 1)

In der Folge ist zu vermuten, dass dynamische Strompreise für Kunden ohne eigenen Stromspeicher im Haus kaum attraktiv sind. Fraglich ist auch, ob bei Vorhandensein eines Stromspeichers die Einsparungen über das Jahr ausreichen, um nach 10...20 Jahren den dann technisch am Lebensende befindlichen Speicher zu ersetzen oder gar zu erweitern. Weiterhin zeigen auch die zunehmend häufiger werdenden negativen Strompreise eine

¹ HTW Dresden, Fak. Informatik/Math., F.-List-Platz 1, 01069 Dresden, wiedem@informatik.htw-dresden.de;
<https://orcid.org/0009-0004-1994-6901>

klare Dysfunktionalität der Strom-Markt-Mechanismen an. Eine genaue Datenanalyse und darauf aufsetzende Simulationen von Lösungsoptionen sollen dies genauer evaluieren. In der Literatur gibt es Überlegungen zu sogenannten variablen Netzentgelten [Dena24]. Leider enthalten die Ausführungen (auch in [IEG24]) keine Informationen zur Umsetzbarkeit und zur Akzeptanz derartiger Konzepte bei den Entscheidungsträgern in Politik und Energiewirtschaft. Zu befürchten sind Einbußen für den Staat und damit angesichts der Haushaltslage eine generelle Ablehnung. Der vorliegende Artikel versucht daher eine aufkommensneutrale Optimierung der Strompreis-Abgaben.



Abb. 1 Dynamische Strompreise (Quelle: www.tibber.de am 28.3.24 [Ti24])

2 Datenerfassung und Basisanalyse

2.1 Datenquellen für den Abruf historischer dynamischer Strompreise

Dank der OpenData-Initiative existieren im Internet sehr viele Datenquellen zum Strompreismarkt. Die vom Fraunhofer-Institut für Solare Energiesysteme ISE [Fh24] gepflegte Website energy-charts.info bietet auf einer Unterseite [Ise24a] verschiedene Datenschnittstellen im OAS3-Format an (für Details [OAS3]). Über die Schnittstelle zu den Börsenpreisdaten [Ise24b] können historische Daten zum Börsen-Strommarktpreis abgerufen werden. Der „Day-Ahead“-Strompreis der Leipziger Börse wird jeweils 24 Stunden vorher festgelegt und ist online abrufbar. Damit entfallen für den Endabnehmer komplexe Prognosemodelle oder Wetteranalysen, da diese in den 24h-Preis-Vorabdaten inkludiert sind. Mit einer Minimalwert-Suche über den Preisen der nächsten 24 Stunden kann jeder Endkunde oder dessen (KI-) Steuerung leicht den besten Zeitpunkt zum Abruf von günstigem Strom für den nächsten Tag ermitteln.

Nach dem Export aus [Ise24b] wurden die Daten im JSON-Format in eine Datenbank-Tabelle eingelesen und analog zu Verfahren des Data-Mining auf Tages-, Wochen und

Monatswerte konsolidiert. Zusätzlich wurden günstige (grün) und ungünstige (rot) Perioden anhand eines frei definierbaren Preisgrenzwertes (im Beispiel 0,05 €) berechnet und visualisiert (vgl. Abb. 3).

Datum	Netto-Preis	Brutto-Preis	Dauer	Wartezeit bis
20211119	0,04 €	0,10 €	25	9
20211120	0,11 €	0,15 €	0	10
20211121	0,12 €	0,15 €	0	10
20211122	0,11 €	0,15 €	0	10
20211123	0,06 €	0,12 €	0	10
20211124	0,07 €	0,11 €	0	10
20211125	0,10 €	0,13 €	0	10
20211126	0,12 €	0,15 €	0	10
20211127	0,11 €	0,15 €	0	10
20211128	0,12 €	0,15 €	0	10
20211129	0,14 €	0,18 €	0	10
20211130	0,17 €	0,20 €	0	10
20211201	0,16 €	0,20 €	0	10
20211202	0,12 €	0,15 €	0	10
20211203	0,10 €	0,13 €	0	10
20211204	0,11 €	0,15 €	0	10
20211205	0,11 €	0,15 €	0	10
20211206	0,13 €	0,18 €	0	10
20211207	0,12 €	0,15 €	0	10
20211208	0,10 €	0,13 €	0	10
20211209	0,08 €	0,10 €	0	10
20211210	0,06 €	0,08 €	0	10
20211211	0,08 €	0,10 €	0	10
20211212	0,11 €	0,15 €	0	10
20211213	0,09 €	0,13 €	0	10
20211214	0,11 €	0,15 €	0	10
20211215	0,09 €	0,13 €	0	10
20211216	0,07 €	0,09 €	0	10
20211217	0,05 €	0,08 €	0	10
20211218	0,08 €	0,10 €	0	10
20211219	0,08 €	0,10 €	0	10
20211220	0,06 €	0,08 €	0	10
20211221	0,02 €	0,05 €	0	10
20211222	0,02 €	0,05 €	0	10
20211223	0,03 €	0,07 €	0	10
20211224	0,06 €	0,10 €	0	10
20211225	0,06 €	0,10 €	0	10
20211226	0,02 €	0,05 €	0	10
20211227	0,06 €	0,10 €	0	10
20211228	0,02 €	0,05 €	0	10
20211229	0,01 €	0,03 €	0	10
20211230	0,04 €	0,08 €	0	10
20211231	0,01 €	0,03 €	0	10
20211231	0,02 €	0,06 €	0	10
20211231	0,05 €	0,08 €	0	10
20211231	0,05 €	0,08 €	0	10

Abb. 3: Strompreismix

Die rot markierten Wartezeiten auf günstigen Strom sind kritisch für kleinere Speicher, da dieser dann leerläuft und mit teurem Strom (zum Bruttopreis) nachgeladen werden muss. Da die Werte auf reinen Preisdaten beruhen, soll dieser Bereich zur eindeutigen begrifflichen Abgrenzung als **Preisdarkflaute** bezeichnet werden. Im Beispiel vom November 2023 dauerte diese Preisdarkflaute 15 Tage, der Maximalwert wurde erreicht mit 17 Tagen im Januar 2024.

Die rot markierten Tage in Abb. 3 sind eine anschauliche Visualisierung des Kernproblems der alternativen Energien und der gesamten Energiewende: Um einer solchen **Preisdarkflaute** bei einem kompletten Abschalten herkömmlicher Kraftwerke zu begegnen, müsste eine **Stromspeicherkapazität von mindestens 20 Tagen** bereitgestellt werden, besser noch mehr!

3 Simulation eines adaptiven dynamischen Strompreises

Zur genaueren Analyse wurde ein flexibel parametrierbares Simulationsmodell auf der Basis der realen stündlichen Börsenstrompreise entwickelt. Im Simulationsmodell wird durch die modellierte Steuerung des Energiespeichers eine einfache Minimalwert-Analyse der jeweils nächsten 24h dynamischer Strompreise durchgeführt. Bei mindestens mittlerer Speicherladung wird genau zum 24h-Minimum der Speicher vollgeladen. Falls die Speicherladung weniger als 25% beträgt, wird auch zu nicht ganz optimalen Zeitpunkten der Speicher auf 50% aufgeladen. Bei ungünstigen Strompreisen wird der gesamte Haushalt aus dem Speicher versorgt. Wenn die Speicherkapazität erschöpft ist, wird automatisch auf das öffentliche Stromnetz umgeschaltet. Bei der Berechnung des Modells werden Preisdaten ab dem 1.1.2023 verwendet, d.h. das Modell spielt alle Szenarien auf den realen Strombörsen-Werten von Anfang 2023 bis Mitte 2024 durch. Um die Sinnhaftigkeit dynamischer Strompreise (DynStromPreis) zu bewerten, wird mit einem Standard-Stromtarif von 0,35 €/ KWh verglichen.

In Abb. 5 ist die Bedienmaske des entwickelten Energiemarkt-Simulators dargestellt. Aufgrund der starken Datenbezüge zu realen Strompreisdaten wurde eine Implementierung mit einer Microsoft-Access-Datenbank durchgeführt. Auch die Algorithmen zur Simulation sind mit dem darin verfügbaren Visual Basic von Microsoft realisiert. Portierungen zu .NET- oder webbasierten Versionen sind geplant.

Die auf ein Jahr normierte Differenz zwischen Standard-Tarif und DynStromPreis von 332 € zeigt den geringen Effekt der dynamischen Strompreise unter realen Bedingungen. Mit den eingesparten 333 € / Jahr ist eine Refinanzierung des Stromspeichers mit aktuell ca. 10.000 € Investkosten selbst nach 20 Jahren nicht möglich. Stromspeicher sind ohne zugehörige PV-Anlage ein sehr teures Hobby und wirtschaftlich ein Totalschaden!

The screenshot shows the user interface of the simulation model. At the top, there are input fields for 'SzenarioID' (2), 'Algorithmus' (100), and 'Optionen'. Below this is a green bar with the scenario name: 'Szenario mit konstantem Strompreis und Brutto-DynStrom variabel MIT Speicher'. To the right of this bar is a blue button labeled 'Simulation starten'.

The main area is divided into two columns. The left column, titled 'Vorgaben (editierbar zum Testen von Konfigurationen)', contains input fields for:

- Standard-Kunde : Preis pro kWh: 0,35 €
- DynStrom: nur mit GÜNSTIGEM Strom ab: 0,05 €
- Dauer-Leistung im Haus [KW]: 0,5
- Speichergöße in KWh: 12

 The right column, titled 'Simulationsergebnisse für die Vorgabewerte', displays the following results:

- == > Gesamtpreis: 1.990,28 €
- == > Gesamtpreis: 1.558,26 €
- Differenz: 432,01 €
- Differenz normiert auf 1 Jahr : 332,67 €
- Verlust an Abgaben /DynKunde: 0,00 €
- Verlust an Abgaben gesamt: 0,00 €
- Einkünfte DynKunde real: 0,28 €

On the far right, there are date and duration settings:

- DatumStart: 2023010101
- DatumStart: 2024041821
- Dauer h: 11373
- Dauer Tage: 474
- Dauer Jahre: 1,30

 Below these are checkboxes for 'debugg' and 'SaveLog', and buttons for 'Teste Aufschläge' and 'Zeige Sim-LOG'.

At the bottom left, there is a 'Zusatzinfos:' field showing 'SpeicherLeer=3788 h (157 Tage)=33%'.

Abb. 5: Die Bedienmaske des Simulationsmodells (mit Standard-Konfiguration)

4 Schlussfolgerungen aus den Simulationsuntersuchungen

4.1 Best-Case-Szenario „Ohne Abgaben“

Bei einer Befreiung dynamischer Strompreise von allen Abgaben ergibt sich eine Differenz von ca. +2.500 € pro Jahr. Damit wäre eine Refinanzierung und sogar Erweiterung der Stromspeicher leicht möglich. In der Summe ergeben sich nur leider Abgabenverluste in gleicher Höhe. Diese können für Deutschland insgesamt bis zu 10 Mrd. €/Jahr betragen. Eine derartige Befreiung ist bei der aktuellen Haushaltssituation nicht vorstellbar, da neben dem Staat (MwSt. usw.) auch die Kommunen und Energieunternehmen diese Summe anteilig pro Jahr verlieren würden.

Das aktuelle Ist-Verhältnis von Netto- zu Brutto-Strompreisen zeigt Abb. 6.

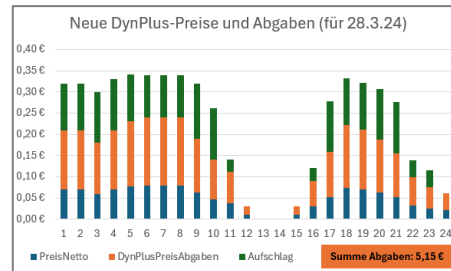
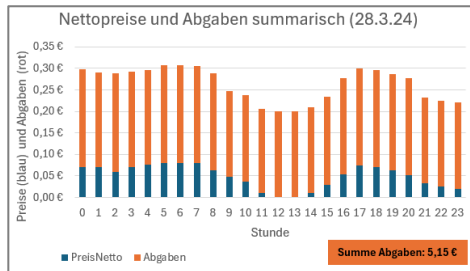


Abb. 6: Aktuelle Netto- und Brutto-Strompreise Abb. 7: Preisverlauf mit DynPlus-Strompreisen

4.2 Modellszenario „DynPlus-Strompreise“

Es erscheint sinnvoll, eine Verschiebung der Abgaben über der Zeit anzustreben. Bei geringen Nettopreisen sollten nur geringe Abgaben fällig werden und bei mittleren Preisen diese mit einem Aufschlag gegenfinanziert werden. Hohe Netto-Preise sollten uneinflusst bleiben. Im aktuellen Entwurf wird bei geringen Nettopreisen nur noch eine prozentuale Abgabenlast von maximal 200% angenommen. (vgl. Abb. 8). Der resultierende Gesamtverlauf kann Abb. 7 entnommen werden.

Stunde	PreisNetto	DynPlusPreis	Aufschlag	BruttoPlus	Kommentar
7	0,08 €	0,16 €	0,10 €	0,34 €	Aufschlag
8	0,06 €	0,13 €	0,13 €	0,32 €	Aufschlag
9	0,05 €	0,09 €	0,12 €	0,26 €	Aufschlag
10	0,04 €	0,07 €	0,03 €	0,14 €	Aufschlag
11	0,01 €	0,02 €	0,00 €	0,03 €	ProzMax 200%
12	0,00 €	0,00 €	0,00 €	0,00 €	ProzMax 200%
13	0,00 €	0,00 €	0,00 €	0,00 €	ProzMax 200%
14	0,01 €	0,02 €	0,00 €	0,03 €	ProzMax 200%
15	0,03 €	0,06 €	0,03 €	0,12 €	Aufschlag
16	0,05 €	0,11 €	0,12 €	0,28 €	Aufschlag
17	0,07 €	0,15 €	0,11 €	0,33 €	Aufschlag

Abb. 8: Neue Preis- und Abgaben-Verteilungs-Tabelle (Auszug)

Zur Mittagszeit ergibt sich das angestrebte Brutto-Preisminimum zum Laden der Stromspeicher. Durch die Zusatzabgaben ergeben sich gleiche Abgaben (in Abb. 7 rechts unten: 5,15 €) in der Summe über den Tag, d.h. der neue Preis-Ansatz ist für den Staat und die Energiewirtschaft aufkommensneutral! Simulationen dieses Ansatzes zeigen eine prinzipielle Umsetzbarkeit dieser Idee durch eine Neugestaltung der dynamischen Strompreise, nachfolgend als DynPlus-Preise bezeichnet.

5 Eine Empfehlung an die Gesetzgebung

Eine reale Umsetzung des DynPlus-Preismodells würde entsprechende Novellierungen der Strommarkt-Gesetzgebung und/oder der Ausführungsbestimmungen für die Abgabeberechnung erfordern. Dazu wird gegenwärtig die Akzeptanz dieses Ansatzes bei Wissenschaftlern im Energiepreismarkt, den Vertretern der Energiewirtschaft und final der Politik geprüft. Da die Freiheitsgrade bei der neuen Abgabenverteilung sehr groß sind und das letzte Wort die Politik und die Lobby der Energieversorger haben werden, wird folgender, sehr offener Vorschlag zur Bestimmung neuer adaptiver dynamischer Strompreise gemacht, wobei der enthaltene Prozentwert beliebig änderbar ist: *„Die Abgabenlast bei dynamischen Strompreisen darf 200% des Netto-Strompreises nicht übersteigen. Die summarische Differenz zur bisherigen Berechnung dynamischer Brutto-Strompreise kann zu anderen Zeiten durch Aufschläge ausgeglichen werden.“*

Die Verwendung einer prozentualen Begrenzung der Strompreisabgaben führt zu einer sanften und nicht sprunghaften Reduzierung der Abgabenlast. Damit erfolgen im Ergebnis wiederum auch keine sprunghaften Lastwechsel durch die Endabnehmer und das Stromnetz kann insgesamt stabiler und besser ausgeregelt werden. Beim obigen Prozentwert von maximal 200% Abgaben würde sich bei Netto-Stromkosten von einem Cent eine Abgabenlast von 2 Cent ergeben und damit ein finaler Bruttopreis von 3 Cent gegenüber bisher 23 Cent (vgl. Abb. 8). Ab 0.00 € Nettostrompreis und darunter gehen dann auch diese Abgaben wirklich auf 0,00 €! Ob diese temporären „Nichtabgaben“ (auch im EU-Kontext) gesetzlich zulässig sind, wird ebenfalls mit Experten geklärt. Alternativ könnten auch bei negativen Strompreisen die gesetzlich unbedingt notwendigen Abgaben erhoben werden, möglichst jedoch auch begrenzt auf einen prozentualen Anteil und einem sehr geringen konstanten Abgabenbetrag. Der bewusst offen gehaltene Textvorschlag kann einen Wettbewerb der Energieversorger um das attraktivste Preismodell für zukünftige adaptive dynamische Strompreise bewirken. Zu erwartende Folgen sind eine starke Motivation der Endkunden und Energieversorger zum Ausbau der Stromspeicher und ein nachfolgendes besseres Ausbalancieren von Angebot und Nachfrage.

6 Ausblick

Die vorgestellten Untersuchungen werden aktuell gemeinsam mit Fachkollegen aus den Bereichen Energiewirtschaft und der wirtschaftswissenschaftlichen Investitionsplanung diskutiert. Weiterhin ist eine umfangreiche Validierung der Modelle, dabei auch ein weiteres Modell zur Langzeitsimulation des Ansatzes bis 2050, mit Fachkolleginnen aus dem Simulationsbereich in Abstimmung und Arbeit.

Wenn bei der technischen Validierung und juristischen Prüfung keine grundsätzlichen Restriktionen oder Probleme bekannt werden, sollen anschließend gemeinsam mit KollegInnen aus dem VDI (Verein Deutscher Ingenieure), VDE (Verband der Elektrotechnik Elektronik Informationstechnik e.V.) und anderen Fachverbänden die entsprechenden Fachexperten in den politischen Gremien angesprochen werden.

Literatur- und Linkverzeichnis

- [En23] ENWG-Gesetz §41a: https://www.gesetze-im-internet.de/enwg_2005/_41a.html
Gesetz über die Elektrizitäts- und Gasversorgung (Energiewirtschaftsgesetz - EnWG)
§ 41a Lastvariable, tageszeitabhängige oder dynamische und sonstige Stromtarife
- [Echart24] Energy-Chart: Strom-Markt-Daten:
www.energy-charts.info/charts/installed_power/chart.htm
- [Dena24] Webinar am 16. Juli 2024 (mit Überblick über die aktuellen Veröffentlichungen zu dynamischen Tarifen und Netzentgelten) : <https://www.set-hub.de/newsroom/veranstaltungen-1/2024/dynamische-tarife-netzentgelte-in-der-aktuellen-debatte/>
- [Dynplus] Website zu diesem Projekt mit aktuellen Ständen: <https://dynplus.de/>
- [IEG24] Fraunhofer-Einrichtung für Energieinfrastrukturen und Geothermie IEG:Kapazitätsbasierte Netzentgelte
<https://www.ieg.fraunhofer.de/de/presse/pressemitteilungen/2024/kapazitaetsbasierte-netzentgelte.html>
- [Ti24] Tibber-Website zu dynamischen Strompreisen: <https://tibber.com/de>
- [Fh24] Fraunhofer-Institut für Solare Energiesysteme ISE,
Website: <https://www.ise.fraunhofer.de/>
- [Ise24a] ISE-API-Überblick: <https://api.energy-charts.info/>
- [Ise24b] ISE-API für "Day-ahead-price"
https://api.energy-charts.info/#/prices/day_ahead_price_price_get
- [OAS3] OpenAPI Specification v3: <https://spec.openapis.org/oas/v3.0.3>

(Alle Webseiten wurden erfolgreich abgerufen am 1. August 2024 23:00 Uhr)

Modellierung von Process Mining im agilen Projektmanagement am Beispiel Scrum

Kurzbeitrag

Pia Schwägerl¹, Anna Schidek ², Holger Timinger ³ und Maren Martens⁴

Abstract: Process Mining als Schnittstelle zwischen Data Mining und Prozessmanagement nutzt digital aufgezeichnete Ereignisinformationen, um so Prozesse mit Hilfe von Algorithmen zu analysieren. Während die Technik bisher vorwiegend in operativen Prozessen zum Einsatz kommt, zeigt diese Arbeit Anwendungsfelder von Process Mining im agilen Projektmanagement auf. Anhand des Rahmenwerks Scrum werden drei Möglichkeiten zur Integration von Process Mining präsentiert und entsprechend modelliert.

Keywords: Process Mining, agiles Projektmanagement, Scrum, Modellierung

1 Einleitung

Process Mining bietet durch die Nutzung und Auswertung digital aufgezeichneter Ereignisinformationen von Arbeits- bzw. Prozessabläufen viele Möglichkeiten zur umfassenden Analyse von Prozessen. Je nach Zielsetzung der Analyse können so Prozessabläufe identifiziert, mit bestehenden Prozessen verglichen oder optimiert werden. [Val16] In vielen Branchen und Disziplinen bereits erfolgreich zur Prozessanalyse genutzt [Re20] ist die Anwendung von Process Mining in Hinblick auf Projekte noch wenig erforscht, obwohl es auch hier große Chancen bietet bestehende Prozesse besser zu verstehen und die Projektarbeit zu verbessern. Auch die zunehmende Bedeutung von Projekten in der Wirtschaft [GP23] unterstreicht die Wichtigkeit der Fragestellung, wie Process Mining speziell im agilen Projektmanagement genutzt werden kann. Zur Beantwortung dieser Frage, wurde im Rahmen eines laufenden Studienprojektes ein Modell erarbeitet, welches drei verschiedene Anwendungsmöglichkeiten von Process Mining im agilen Projektmanagement nach Scrum aufgreift.

¹ Institute for Data and Process Science, Hochschule Landshut, Am Lurzenhof 1, 84036 Landshut, pia-schwaegerl@web.de

² Institute for Data and Process Science, Hochschule Landshut, Am Lurzenhof 1, 84036 Landshut, anna.schidek@haw-landshut.de, <https://orcid.org/0000-0003-3724-4626>

³ Institute for Data and Process Science, Hochschule Landshut, Am Lurzenhof 1, 84036 Landshut, holger.timinger@haw-landshut.de, <https://orcid.org/0000-0001-7992-0392>

⁴ Institute for Data and Process Science, Hochschule Landshut, Am Lurzenhof 1, 84036 Landshut, maren.martens@haw-landshut.de

2 Modellierung der verschiedenen Anwendungsmöglichkeiten

In systematischen Literaturrecherchen zum Thema Process Mining im Projektmanagement [SSM24] sowie speziell im agilen Projektmanagement [Sc24] konnten essenzielle Informationen extrahiert und in das in Abb. 1 dargestellte Gesamtmodell überführt werden. Die Literatur wurde dabei unter verschiedenen Gesichtspunkten, wie der Anwendungsdomäne und der angewandten Projektmanagementmethode sowie der behandelten Arten und Perspektiven des Process Mining, betrachtet. Ebenfalls Bestandteil der Literaturanalyse war die Auswertung der jeweiligen Datenquellen, genutzten Tools und Algorithmen. Bisherige Forschungen im gewählten Themenfeld des agilen Projektmanagements beschränken sich hauptsächlich auf die Softwareentwicklung mit Scrum. Im unten gezeigten Modell (Abb.1) soll deshalb eine erste Idee zur Integration von Process Mining in einen typischen Ablauf von Scrum [SS20] umgesetzt werden. Visualisiert werden neben dem im Zentrum stehenden zeitlichen Ablauf der Scrum Events zudem die Scrum Artefakte (blau dargestellt), zugrundeliegende Prozessmodelle (grün=Referenzmodelle, orange=Ist-Prozessmodelle) sowie Datenquellen und daraus erstellte Ereignislogs. Als Modellierungssprache wurde BPMN 2.0 gewählt [Ob14].

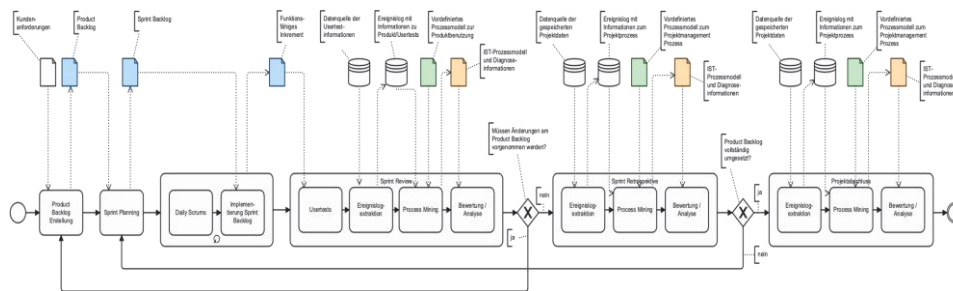


Abb. 1: Modell zur Integration von Process Mining in das agile Rahmenwerk Scrum [Sc24] (Visualisierung unter Verwendung von BPMN 2.0 [Ob14] und BPMN.io [We23])

2.1 Modellvarianten

Das entwickelte Modell sieht dabei drei verschiedene Möglichkeiten zur Integration von Process Mining in den agilen Projektablauf nach Scrum vor. So kann Process Mining bereits entwicklungsbegleitend im Zuge der Sprint Reviews oder der Sprint Retrospektive eingesetzt werden, aber auch nachgelagert in der Projektabschlussphase. Die Idee der verschiedenen Varianten, die im Folgenden kurz beschrieben werden, ist es, diese sowohl gezielt einzeln als auch als ganzheitliche Betrachtung einsetzen zu können. Tab. 1 gibt zudem einen Überblick über die durch die verschiedenen Optionen abgedeckten Process-Mining-Arten und -Perspektiven [Va16]. Der Fokus des Modells liegt dabei auf der Erkennung von Abweichungen sowie der Übereinstimmungsprüfung eines Ist- und Soll-Prozesses unter Berücksichtigung mehrerer Perspektiven.

Option A – Process Mining im Rahmen des Sprint Reviews: Im Laufe des Sprint Reviews, also der Begutachtung des während des Sprints entstandenen Inkrements, können Usertests durchgeführt werden, die wiederum die Datenbasis für eine

Übereinstimmungsprüfung mit den gewünschten Vorstellungen der Stakeholder bilden [Ru14]. Gemäß der Definition of Done können dann Referenzmodelle mit dem aktuellen IST-Zustand verglichen werden und bei Bedarf Änderungen in den Product Backlog eingearbeitet werden. [Sc24]

Option B – Process Mining im Rahmen der Sprint Retrospektive: Angewendet in der Sprint Retrospektive kann Process Mining zur Reflexion des Entwicklungsteams und seiner Vorgehensweise während des aktuellen Sprints genutzt werden [MMR18]. Aus Datenquellen, wie beispielsweise Jira oder GitHub, können Informationen über die konforme Durchführung des Sprints und seiner Events extrahiert und für eine Analyse durch Process Mining herangezogen werden. So kann beispielsweise durch die User-ID der Einträge geprüft werden, welche Rollen beteiligt sind, oder es können durch den Zeitstempel der Projektaktivitäten Aussagen zum Entwicklungsablauf innerhalb des Rahmenwerks Scrum getroffen werden. [Sc24]

Option C – Process Mining im Rahmen des Projektabschlusses: Um zukünftige Projekte optimierter gestalten zu können, ist es sinnvoll den agilen Projektmanagementprozess in Scrum um eine Projektabschlussphase zu ergänzen. Auch hier ist die Anwendung von Process Mining als „Gesamtretrospektive“ des gesamten Projektes möglich [SC14]. Ablauf und Zielsetzung folgen demselben Schema wie Option B. [Sc24]

Variante	Arten Process Mining			Perspektiven Process Mining			
	Erkennung	Übereinstimmung	Modellverbesserung	Kontrollfluss	Organisation	Zeit	Fall
Option A (Sprint Review)	x	x		x		x	
Option B (Sprint Retrospektive)	x	x		x	x	x	
Option C (Projektabschluss)	x	x		x	x	x	x

Tab. 1 Durch das Modell unterstützte Arten und Perspektiven des Process Mining [Sc24]

2.2 Evaluierung

Anhand einer ersten qualitativen Expertenbefragung mithilfe eines Fragebogens in Anlehnung an Porst [Po14] konnte das Modell in einer initialen Iteration evaluiert werden. Der Fragebogen ist dabei in zwei Teile untergliedert und wurde qualitativ ausgewertet. [Sc24] Die ersten Fragen stellen den Expertenstatus [MN91] der Personen sicher. Befragt wurden drei Personen mit mehrjähriger Erfahrung im Bereich des Process Mining und des agilen Projektmanagements. Im zweiten Teil wurden Fragen zum Modell in offenen Fragestellungen beantwortet. Das Modell wurde dabei als vielversprechend bewertet und die Experten betonten, dass die Anwendung von Process Mining im agilen Projektmanagement wichtige Erkenntnisse liefern kann, konnten aber auch Limitierungen identifizieren. Insbesondere wurden die Verfügbarkeit und Extraktion der Daten als problematisch beachtet. Da die Qualität und Verfügbarkeit von Daten allgemein eine

Herausforderung in der Nutzung von Process Mining darstellt bzw. als Voraussetzung gilt, sollte darauf auch im Bereich des agilen Projektmanagements ein gesonderter Blick geworfen werden, besonders aufgrund der Einzigartigkeit von Projekten. Auch die Verfügbarkeit von zusätzlichen personellen und monetären Ressourcen für die Anwendung wurde von den Experten als Herausforderung angeführt.

3 Diskussion

Das entwickelte Integrationsmodell verdeutlicht, dass Process Mining im Projektmanagement nach Scrum durchaus angewandt werden kann und Potenziale zur Prozessoptimierung besitzt. Jedoch ist der Nutzen, der aus den Erkenntnissen der Process-Mining-Anwendung gezogen werden kann, mit dem zeitlichen Aufwand und den Kosten, welche dadurch entstehen, abzuwägen – auch im Zuge einer weiteren Evaluierungssiteration, z.B. durch die Anwendung von Case Studies. Weiterhin ist zu beachten, dass sich Option A auf solche Projekte beschränkt, deren Projektgegenstand getestet werden kann und dabei elektronisch aufgezeichnete Daten hinterlässt. Auch bei der Anwendung nach Option B und Option C ist Voraussetzung, dass alle Aktivitäten digital dokumentiert sind und somit ein vollständiger und qualitativ hochwertiger Ereignislog zur Analyse vorliegt. Laut Expertenaussagen, die sich aus der Evaluierung ergaben, ist dies in der Praxis oft nicht der Fall, was entweder zu einer verminderten Qualität der Process-Mining-Ergebnisse oder zu Mehraufwand führen würde, sofern speziell dafür sämtliche Projektschritte elektronisch erfasst werden müssten.

4 Fazit und Ausblick

Das Modell zeigt verschiedene Anwendungsmöglichkeiten von Process Mining im agilen Projektmanagement innerhalb des Rahmenwerks Scrum und somit auch das Potenzial Process Mining in weiteren Bereichen neben der Analyse von operativen Betriebsabläufen einzubinden. Process Mining bietet durch die Betrachtung gespeicherter Daten einen objektiven Blick auf Prozessabläufe und somit eine gute Ergänzung zur Reflexion in Sprint Review und Sprint Retrospektive. Durch eine erste Evaluierung konnte einerseits die Sinnhaftigkeit der Anwendung bestätigt, aber auch schon in diesem Stadium Limitierungen des Modells gefunden werden, welche einen Ausgangspunkt für weitere Forschungsthemen in diesem Feld bieten. So könnten in zukünftigen Forschungsarbeiten zum einen die Datenextraktion und -aufbereitung der Ereignislogs genauer betrachtet werden, um so das Modell auch in der Praxis anwenden zu können. Zum anderen ist es auf Basis dessen auch denkbar, das Modell um einen weiteren Anwendungsfall während der Planungsphase zu ergänzen.

Literaturverzeichnis

- [GP23] GPM Deutsche Gesellschaft für Projektmanagement e. V.: Makroökonomische Vermessung der Projektstätigkeit in Deutschland. URL <https://www.gpm-ipma.de/wissen/studien/makrooekonomische-vermessung-der-projekttaetigkeit-in-deutschland>. - abgerufen am 04.01.2024.
- [MMR18] Marques, R.; Mira da Silva, M.; R. Ferreira, D.: Assessing Agile Software Development Processes with Process Mining: A Case Study. IEEE 20th Conference on Business Informatics, 2018.
- [MN91] Meuser, M.; Nagel, U.: ExpertInneninterviews - vielfach erprobt, wenig bedacht: ein Beitrag zur qualitativen Methodendiskussion. Westdt. Verlag, Opladen, 1991.
- [Ob14] Object Management Group: Business Process Model and Notation (BPMN), Version 2.0. URL <https://www.omg.org/spec/BPMN/2.0.2/PDF> - abgerufen am 17.01.2024.
- [Po14] Porst, R.: Fragebogen: ein Arbeitsbuch. VS-Verlag, Wiesbaden, 2014.
- [Re20] Reinkemeyer, L. (Hrsg.): Process mining in action: principles, use cases and outlook. Springer Nature Switzerland AG, Cham Schweiz, 2020.
- [Ru14] Rubin, V. et.al.: Agile development with software process mining. Proceedings of the 2014 International Conference on Software and System Process, Nanjing China, 2014.
- [SC14] Sebu, M.; Ciocârlie, H.: Applied process mining in software development. 9th IEEE International Symposium on Applied Computational Intelligence and Informatics (SACI), Timișoara Romania, 2014.
- [Sc24] Schwägerl, P.: Anwendungsmöglichkeiten von Process Mining im agilen Projektmanagement, Bachelorarbeit Hochschule Landshut, 2024.
- [SSM24] Schidek, A.; Schwägerl, P.; Martens, M.: Mapping the Landscape of Process Mining in Project Management: A Comprehensive Literature Review, Research Notes on Data and Process Science, Issue 6, DOI: 10.57688/397. 2024.
- [SS20] Schwaber, K.; Sutherland, J.: The Scrum Guide. URL <https://scrumguides.org/docs/scrumguide/v2020/2020-Scrum-Guide-German-m%C3%A4nnlich-male-version.pdf> – aufgerufen am 01.11.2023.
- [Va16] Van der Aalst, W.: Process Mining: Data Science in Action. 2. Auflage. Springer-Verlag Berlin Heidelberg, New York, u.a., 2016.
- [We23] Web-based tooling for BPMN, DMN and Forms. URL <https://bpmn.io/> - abgerufen am 06.01.2024.

A Use Case Analysis assessing Process Mining Model Quality for Simulation

Short Paper

Marianne Schmolke ¹, Muhammad Jawad Javaid²

Abstract: Ever-evolving business processes can be re-engineered with process mining and simulation as a technology, as the approach shapes future outcomes, analyzing their present and past. Developing effective models can make it easier and more cost-effective, as there are potential benefits such as working with what-if scenarios and eliminating bottlenecks, but their quality check is always an issue. Different criteria have been applied to tackle this quality issue, disclosing expected benefits and weaknesses. This can help build suitable ground for analyzing and implementing relevant use cases. The criteria extracted from process model quality research demonstrate the effectiveness of model assessment and have practical implications for improving use case modeling. Results display the quality of particular models and indicate how use case modeling can be improved.

Keywords: Comprehensive Process Model Quality Framework (CPMQF), optimization, process mining simulation

1 Introduction

Our research findings testify to the significant benefits that process mining and simulation approaches offer [Aa18]. A special focus is testing the quality of modeled process mining cases for simulation. For this, the work of [Sc22] can be applied and enhanced. We will show which criteria from literature are applicable and whether the use cases conform to the criteria. As Wagner et al. [WSM16] discovered, simulation modeling is a powerful tool for comprehending the complexity of business processes, a challenge that lies in the creation of realistic models. Mesabbah et al. [MHM19] argue that the complexity of processes, uncertain parameters, time-critical activities, budget, and resource constraints typify an environment suitable for scenario testing. As Norambuena [NN18] aptly points out, the key distinction between process mining and simulation is that the former is used

¹ Fakultät für Wirtschaftswissenschaften/Hochschule Wismar, Softwarearchitekturen und Geschäftsprozesse, Philipp-Müller-Str. 14, 23966 Wismar, Marianne.Schmolke@hs-wismar.de, <https://orcid.org/0000-0002-4293-603X>

² Fakultät für Ingenieurwissenschaften/Hochschule Wismar, Operations and Management of Maritime Systems, Philipp-Müller-Str. 14, 23966 Wismar, m.javaid@stud.hs-wismar.de

to understand a process's current or past behavior, while the latter can be applied to predict future behavior.

2 Related Work

Considerable research has been conducted in the area of process mining simulation. Whereas van der Aalst [Aa18] defines the formal prerequisites for the match of process mining and simulation, and Aguirre et al. [APA12] focus on Business Process Reengineering, this paper focuses on the practical application of robust models. Norambuena [APA12] and Martin et al. [APA12] conduct systematic (literature) reviews on the integration of process mining and (Business Process) simulation. While Norambuena and Martin et al. focus more on challenges or open issues, this paper focuses on clarifying the current state of Process Mining model quality for simulation. It is useful to extract applicable analysis criteria from existing works to utilize the potentials by tackling the weaknesses starting from the conceptualization of studies.

3 Search Criteria and Analysis

We reviewed English articles (more than 4 pages) containing “process mining” AND “simulation” in the title from “2007 on” from the following databases: ACM Digital Library (dl.acm.org), IEEE Xplore (ieeexplore.ieee.org), Proquest (proquest.com), Science Direct (sciencedirect.com), and Springer (link.springer.com). Papers with process mining models were selected from the findings. 31 articles were found of which 9 contained a visualized process mining model. This set of 9 papers was considered for the analysis.

The motivation was to find literature with a strong focus on process mining and simulation in the sense that both approaches appear in the document title. This selection can be seen as exemplary, i.e. this is an exploratory study.

[Sc22] extracted guidelines for good Process Mining Case Study Models to enrich their fit-for-use level by applying the relevant criteria of the Comprehensive Process Model Quality Framework (CPMQF) [MC18]. This focus is applicable to the context at hand.

The summary of applicable criteria is [Sc22] (continued on next page):

1. Use no more than 31 nodes
2. Use verb-object activity labels
3. Use clearly defined domains and scopes
4. Use internal and external feedback loops
5. Try to capture the ‘soft’ process-related issues

The first criterion is important to keep the model manageable. Verb-object activity labels ensure better understandability, whereas clearly defined domains and scopes help ensure the modeling result's validity.

Internal and external feedback loops (at least one of these two needs to be fulfilled) ensure the reliability of the results. It can be assumed that all papers embrace this criterion but it is possible that the methods are applied with a clear exploratory focus (discovery in process mining) and with a separate simulation approach. ‘Soft’ process-related issues can be interpreted as less quantifiable, i.e., perceptual and social aspects, mostly addressed by change management initiatives in [...] projects ([Sc22], p.141))

All papers fulfilled the first criterion, „Use no more than 31 nodes. " Therefore, it is not shown in Tab 1. The first criterion („Use verb-object activity labels“) in the table will not be discussed, so it is only shown for reasons of completeness. The criteria have been applied thoroughly to the use cases. These are the results:

4 Results

Criterion Paper	(a) ³	(b) ⁴	(c) ⁵	(d) ⁶	(e) ⁷	(f) ⁸	(g) ⁹	(h) ¹⁰	(i) ¹¹
Use verb-object activity labels	N	Y	N	Y	N	N	N	N	N
Use clearly defined domains and scopes	Y	N	Y	Y	N	N	N	Y	Y
Use internal and external feedback loops	Y	Y	Y	Y	Y	Y	Y	Y	Y
Try to capture the ‘soft’ process-related issues	N	Y	N	N	N	N	N	N	Y

Tab. 1 Criteria fulfilled in studies with Process Mining Models

³ [MB09]

⁴ [ARA17]

⁵ [DD19]

⁶ [Ja20]

⁷ [ZSB20]

⁸ [FS21]

⁹ [HMK2]

¹⁰ [RL22]

¹¹ [La22]

5 Conclusion

Limitations of the study are a small selection of cases and the exploratory character.

Nevertheless, the analysis shows that none of the papers fulfilled all quality criteria for their modeled use cases. Nevertheless, the results are valuable and show that process mining and simulation are valid instruments for gaining relevant insights.

The analysis also showed that the main application fields of process mining and simulation are creating what-if scenarios and eliminating bottlenecks. Additionally, the process optimization focus is strongly present. The implementation category is of minor relevance as process mining and simulation are not automatically connected to Workflow Management logic.

A criterion fulfilled in all cases was „External/Internal feedback loops“. It could be considered not to test this criterion for process mining/simulation studies. This criterion is generally fulfilled by the method because a simulation can be seen as model feedback.

An important aspect rarely considered is the „soft“ process-related one. Its relevance is even increasing due to technological advances such as in Artificial Intelligence. AI systems are less capable of dealing with ambiguous information [HP24] than the experts dealing with the interpretation of process model results.

The size of the model (max. 31 nodes) is still of relevance and verb-object activity labels make models more understandable. Therefore, these criteria still need to be considered in the future.

The final set of criteria is therefore for process mining simulation:

- Use no more than 31 nodes (not necessarily respected by models)
- Use verb-object activity labels
- Use clearly defined domains and scopes
- Try to capture the ‘soft’ process-related issues

Applying the criteria found here the outputs of process mining simulation analysis can be improved. Further research should include how the implementation of process mining model simulation can be supported by embracing the approach at hand.

References

- [APA12] Aguirre, S., Parra, C., and Alvarado, J., “Combination of Process Mining and Simulation Techniques for Business Process Redesign: A Methodological Approach,” International Symposium on Data-Driven Process Discovery and Analysis, 2012, 24–43, Springer, Berlin, Heidelberg. doi: 10.1007/978-3-642-40919-6_2.

- [ARA17] Abo-Hamad, W., Ramy, A., Arisha, A., “A hybrid process-mining approach for simulation modeling,” Online World Conference on Soft Computing in Industrial Applications, 2017, pp. 1527–1538. doi: 10.1109/wsc.2017.8247894.
- [Aa18] van der Aalst, W. M. P., “Process mining and simulation: a match made in heaven!,” Summer Simulation Multiconference,; Bordeaux, France, 2018. doi: 10.22360/summersim.2018.ssc.005.
- [DD19] Dorrer, M., Dorrer, A., “Generation of agent simulation models by using process mining methods on the example of E-learning process,” Journal of Physics: Conference Series; Vol. 1399, No. 3, 2019, p. 33077. doi: 10.1088/1742-6596/1399/3/033077.
- [FS21] Ferronato, J. J., and Scalabrin, E. E., “PM2Sim: The Automated Creation of a Simulation Model from Process Mining,” IEEE International Conference on Systems, Man and Cybernetics; Melbourne, Australia, 2021. doi: 10.1109/smc52423.2021.9659211.
- [HMK21] Halawa, F., Madathil, S. C., Khasawneh, M. T., “Integrated framework of process mining and simulation–optimization for pod structured clinical layout design,” Expert Systems With Applications; Vol. 185, 2021, p. 115696. doi: 10.1016/j.eswa.2021.115696.
- [HP24] Hutson, J, Plate, D, “Disrupting Algorithmic Culture: Redefining the Human(ities),” Generative AI in Teaching and Learning, IGI Global, 2024, pp. 1–30.
- [Ja20] Jadrić, M., Pašalić, I. N., and Ćukušić, M., “Process Mining Contributions to Discrete-event Simulation Modelling,” Business Systems Research; Vol. 11, No. 2, 2020, 51–72, Sciendo, Warsaw, Poland. doi: 10.2478/bsrj-2020-0015.
- [La22] Le Lay, J., Neveu, J., Dalmás, B., Augusto, V., “Automated Generation of Patient Population for Discrete-Event Simulation Using Process Mining,” Annual Modeling and Simulation Conference; San Diego, USA, 2022. doi: 10.23919/annsim55834.2022.9859406.
- [MB09] Mărușter, L., van Beest, N. R. T. P., “Redesigning business processes: a methodology based on simulation and process mining techniques,” Knowledge and Information Systems [online], Vol. 21, No. 3, 2009, pp. 267–297, <https://link.springer.com/article/10.1007/s10115-009-0224-0>.
- [MC18] Meyer, P. de, Claes, J., An overview of process model quality literature - The Comprehensive Process Model Quality Framework, 2018.
- [MHM19] Mesabbah, M., Abo-Hamad, W., and McKeever, S., “A hybrid process mining framework for automated simulation modelling for healthcare,” Online World Conference on Soft Computing in Industrial Applications, 2019, 1094–1102. National Harbor, MD, USA. doi: 10.1109/wsc40007.2019.9004800.
- [NN18] Brian Keith Norambuena, and Norambuena, B. K., “Integration of Process Mining and Simulation: A Survey of Applications and Current Research,” WorldCIST, 2018, 287–294, Cham. doi: 10.1007/978-3-319-77703-0_29.
- [Ph19] Phan, R., Vincent Augusto, Augusto, V., Martin, D., Marianne Sarazin, and Sarazin, M., “Clinical pathway analysis using process mining and discrete-event simulation: an application to incisional hernia,” Online World Conference on Soft Computing in Industrial Applications, 2019, pp. 1172–1183. doi: 10.1109/wsc40007.2019.9004944.

- [RL22] Rashid, Khandakar M., Louis, J. “Integrating Process Mining with Discrete-Event Simulation for Dynamic Productivity Estimation in Heavy Civil Construction Operations,” *Algorithms*; Vol. 15, No. 5, 2022, p. 173. doi: 10.3390/a15050173.
- [Sc22] Schmolke, M., “Process Mining Model Quality in Software Development Case Studies: An Analysis,” Tagungsband zur 35. Jahrestagung des Arbeitskreises Wirtschaftsinformatik an Hochschulen für Angewandte Wissenschaften im deutschsprachigen Raum (AKWI) vom 11.09. bis 13.09.2022, Berlin (2022).
- [WSM16] Wagner, G., Seck, M., and McKenzie, F. D., “Process modeling for simulation: observations and open issues,” *Online World Conference on Soft Computing in Industrial Applications*, 2016, 1072–1083, Washington, DC, USA. doi: 10.1109/wsc.2016.7822166.
- [ZSB20] Zakarija, I., Skopljanac-Macina, F., Blašković, B., “Automated simulation and verification of process models discovered by process mining,” *arXiv: Artificial Intelligence*, 2020. doi: 10.1080/00051144.2020.1734716.

Applying Predictive Process Monitoring to Predict User Behavior with Click Data

Short Paper

Jana Feiler¹ and Tobias Walter¹

Abstract: Predictive Process Monitoring (PPM) is used to predict the future behavior of running process instances using machine learning techniques. While traditional PPM focuses and works well on highly structured business processes, the planned research explores the application of and challenges related to PPM on rather unstructured processes, such as those discovered through user click traces on websites. Therefore, this research's overall objective is to develop and evaluate appropriate AI-supported PPM approaches for unstructured processes based on click traces provided by Germany's public broadcaster ARD with its audio and video streaming services ARD Mediathek & ARD Audiothek.

Keywords: Process Mining, Predictive Process Monitoring, Unstructured Business Processes, Click Data, User Behavior Analysis

1 Introduction

Process Mining is characterized as the intersection of data science and process science by providing a well-established collection of techniques to utilize process-based event data in order to discover, compare, improve and predict business processes [Va22]. While discovery, conformance checking and the enhancement of processes are considered as backward-looking process mining techniques, relying solely on historic event data, *Predictive Process Monitoring* (PPM) represents a forward-looking approach to provide predictions on running process instances [Di18]. Those predictions on real-time event data may be operationalized e. g., through warning and anomaly detection systems, providing the opportunity to intervene and apply preventive countermeasures [Br16; DG22].

Traditional process mining and likewise PPM performs well when drawing on structured data extracted from process-aware information systems, such as enterprise resource planning software (e. g., SAP) and being applied to event data of highly structured processes [Ko24; La22; VG07]. On the contrary, applying process mining to unstructured data or loosely structured processes remains challenging and oftentimes leads to rarely interpretable so-called “spaghetti-models” [DGP16; Ko24; St20]. Such *Unstructured Business Processes* (UBP) do not build-upon a pre-defined set and sequence of activities, leading to increased

¹ Hochschule Mainz – University of Applied Sciences, Lucy-Hillebrand-Straße 2, 55128 Mainz,
jana.feiler@hs-mainz.de; tobias.walter@hs-mainz.de

process flexibility, complexity and limited predictability as often found in the customer-centric service sector or the healthcare sector [St20]. Hence, new process mining methods tailored specifically to UBP are necessary [Ko24]. Currently, there are no universal PPM approaches suitable for predictions on UBP [Di18; St20].

Process Mining based on user behavior data, such as click traces extracted e. g. , from web logs, like the BPI Challenge 2016 data set [Dv16], represents a concise example of UBP. The planned research, outlined within this paper, will introduce a new real-life data set by using click traces provided by Germany's public broadcaster ARD with its audio and video streaming services *ARD Mediathek* and *ARD Audiothek*. Predicting user click behavior, for example, may help recognise user preferences, proactively offer content to or guide users, thus making desired content easier to find. It is expected that this will lead to increased user satisfaction and in turn positively impact the business success of media companies.

The remainder of the paper is structured as follows. Sect. 2 provides a compact overview on the current state of research with regard to PPM on structured and less structured processes as well as application domains for process mining of UBP. The research methodology is described in Sect. 3 while research objectives are outlined in Sect. 4. The paper concludes in Sect. 5 with expected results and contributions of the planned research.

2 Background

Literature and research on PPM, mainly with regard to structured processes, increased considerably within recent years providing a comprehensive overview on data types used for predictions, prediction techniques and different prediction outcomes [DG22; Di18; Sp20]. As traditional process mining methods perform rather poorly on UBP due to the high variability and heterogeneity inherent to loosely structured processes, applying those methods generates either highly complex and barely comprehensible or over-generalized process models [DGP16; St20]. Therefore, pre-processing of event data during the event log preparation is required, for example through aggregation of low-level activities to high level ones, abstraction or discarding of less relevant activities or clustering and partitioning of the event log into several less complex sub logs [St20]. With regard to process discovery, the Fuzzy Miner [GV07] has been proven suitable for less structured processes though other common discovery algorithms and mixed methods have been likewise analyzed and evaluated, e. g. by Stefanini et al. [La22; St20].

A large interest in applying process mining to UBP can be observed in the healthcare sector, where numerous case studies, evidenced by multiple literature reviews, have been conducted [Da22; DM22] and prevailing challenges are addressed for example by Munoz-Gama et al. [Mu22]. Other application domains such as in e-commerce [AH24; Pe22], online education [Ha21; IKT18; Mu15], online news [SI21], social networks [Di17] and social collaboration [BI21] exist but appear less refined and aggregated. Especially within the services sector, customer journey analysis or mapping is recently combined with process mining to analyze

customer behavior, identify customer needs and to increase customer satisfaction [BA17]. A customer journey represents the sequence of touch points with a customer when delivering a service, which may include click data but potentially spans across different IT systems, communication channels and service providers [HH21; WH23]. In this context, several studies went even further by making predictions and consecutive recommendations on running customer journeys [AH24; HH21; TH18; WH23]. Thereof the study of Terragni & Hassani [TH18] poses to be the most comparable one to the planned research, as the authors use click traces of an advertising web portal to discover process models and to provide individual recommendations, wherefore it may serve as a initial structure for the first iteration within the planned PM^2 approach.

3 Research Methodology

The proposed research based on audio and video streaming click traces, will be realized following the PM^2 methodology by van Eck et al. [Ec15], as this approach is tailored specifically to process mining projects. Therefore, Fig. 1 outlines the PM^2 phases, adapted to the planned research, while visualizing major artifacts and involved parties, such as the researchers in the role of the process experts and the data partner as the domain expert.

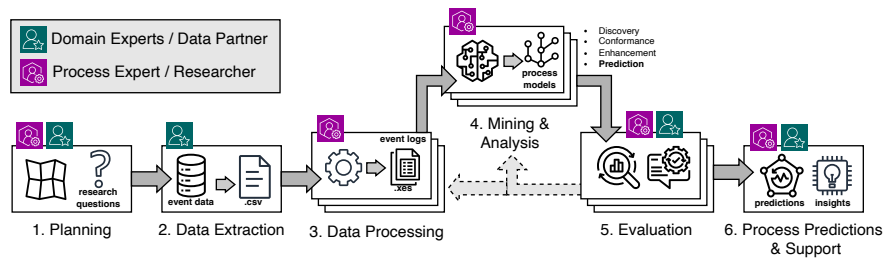


Fig. 1: Adapted PM^2 research approach based on [Ec15]

The research is currently situated in the initial planning phase where process and domain experts jointly explore challenges and characteristics of the video and audio streaming domain to derive consequential research questions. In a next step, the event data will be extracted by the data partner and provided to the researchers, which will in turn convert the data into event logs and take necessary pre-processing measures. The subsequent mining & analysis phase will initially feature backward-looking process mining techniques such as process discovery and enhancement in order to provide predictions on previously discovered process models. Ultimately, the PPM methods are to be comprehensively evaluated to verify and ensure their scientific and practical value. It should be noted that data processing, mining & analysis as well as the evaluation phase may be run iteratively before predictions and respective insights can be operationalized in the final phase.

4 Research Objective

The objective of the research outlined within this paper is to develop appropriate AI-supported PPM approaches for UBP and to subsequently evaluate them. Real click data from users of the German video and audio streaming services, namely ARD Audiothek and ARD Mediathek, are used for the analysis and evaluation.

Research Objective 1: Identify and/or develop methods or algorithms for event log preprocessing in the context of PPM for unstructured processes

In the discipline of process mining for UBP, it has been proven to pre-process log data so that the derived process models become interpretable [Bl21; Ze21]. It is expected that with abstracted models derived from aggregated logs, more plausible predictions can be made than with “spaghetti models”.

Research Objective 2: Identify and/or develop methods or algorithms for predicting the behavior of unstructured processes

In the field of process mining for UBP, it has been shown that focusing on entire process models is no longer effective. Instead, the aim is to identify process patterns. Predictions are not made based on a single event trace, but rather on process patterns visited by the trace.

Research Objective 3: Apply and evaluate the performance of developed PPM methods

Real click data is delivered by Germany’s public broadcaster ARD. This data is used for method and outcome evaluation of objectives 2 and 3. Besides, real data for further research may be provided by additional partners in the area of media and health.

5 Expected Results and Contribution

In accordance with the aforementioned research objectives the planned research is expected to deliver different kinds of predictions based on the provided click data, such as user-specific recommendations based on user types or other forms of behavior clustering. Predictions on potential premature process terminations could be utilized to identify process weaknesses or UX deficiencies and to intervene in order to ensure a satisfactory user experience.

Furthermore, the research will apply PPM to unstructured processes in a new domain, namely the media domain and enrich state-of-the-art research by conducting a case study based on a new real-world data set. Beyond that, the applied methods and received results may serve as a basis for further research with other real data sets provided by additional data partners from the media and healthcare domain.

Acknowledgments This research is funded by the Ministry of Science and Health Rhineland-Palatinate, Germany, research program AI-DPA (<http://ai-dpa.hs-mainz.de>)

References

- [AH24] Akhavan, F.; Hassannayebi, E.: A Hybrid Machine Learning with Process Analytics for Predicting Customer Experience in Online Insurance Services Industry. *Decision Analytics Journal* 11, p. 100452, 2024, doi: 10.1016/j.dajour.2024.100452.
- [BA17] Bernard, G.; Andritsos, P.: A Process Mining Based Model for Customer Journey Mapping. In: CAiSE-Forum-DC. 2017, URL: <https://api.semanticscholar.org/CorpusID:1695038>.
- [Bl21] Blatt, J.; Delfmann, P.; Schwade, F.; Schubert, P.: Log Shifting for Enterprise Collaboration Systems: A Supervised Event Abstraction Approach (Extended Abstract). In (Jans, M.; Janssenswillen, G.; Kalenkova, A.; Maggi, F. M., eds.): *Proceedings of the ICPM Doctoral Consortium and Demo Track 2021 Eindhoven, The Netherlands, November, 2021. Workshop Proceedings, 3rd International Conference on Process Mining ICPM 2021, CEUR/RWTH, Aachen*, pp. 21–22, 2021, URL: https://ceur-ws.org/Vol-3098/demo_190.pdf.
- [Br16] Breuker, D.; Matzner, M.; Delfmann, P.; Becker, J.: Comprehensible predictive models for business processes. *MIS Quarterly* 40 (4), pp. 1009–1034, 2016, doi: 10.25593/opus4-fau-18267.
- [Da22] Dallagassa, M. R.; Dos Santos Garcia, C.; Scalabrin, E. E.; Ioshii, S. O.; Carvalho, D. R.: Opportunities and Challenges for Applying Process Mining in Healthcare: A Systematic Mapping Study. *Journal of Ambient Intelligence and Humanized Computing* 13 (1), pp. 165–182, 2022, doi: 10.1007/s12652-021-02894-7.
- [DG22] Di Francescomarino, C.; Ghidini, C.: Predictive Process Monitoring. In (van der Aalst, W. M. P.; Carmona, J., eds.): *Process Mining Handbook*. Springer International Publishing, Cham, pp. 320–346, 2022, doi: 10.1007/978-3-031-08848-3_10.
- [DGP16] Diamantini, C.; Genga, L.; Potena, D.: Behavioral Process Mining for Unstructured Processes. *Journal of Intelligent Information Systems* 47 (1), pp. 5–32, 2016, doi: 10.1007/s10844-016-0394-7.
- [Di17] Diamantini, C.; Genga, L.; Marozzo, F.; Potena, D.; Trunfio, P.: Discovering Mobility Patterns of Instagram Users through Process Mining Techniques. In: *2017 IEEE International Conference on Information Reuse and Integration (IRI)*. IEEE, San Diego, CA, USA, pp. 485–492, 2017, doi: 10.1109/IRI.2017.69.
- [Di18] Di Francescomarino, C.; Ghidini, C.; Maggi, F. M.; Milani, F.: Predictive Process Monitoring Methods: Which One Suits Me Best? In (Weske, M.; Montali, M.; Weber, I.; vom Brocke, J., eds.): *Business Process Management*. Springer International Publishing, Cham, pp. 462–479, 2018, doi: 10.1007/978-3-319-98648-7_27.
- [DM22] De Roock, E.; Martin, N.: Process Mining in Healthcare – An Updated Perspective on the State of the Art. *Journal of Biomedical Informatics* 127, p. 103995, 2022, doi: 10.1016/j.jbi.2022.103995.
- [Dv16] Dees, M.; van Dongen, B.: BPI Challenge 2016, 2016, doi: 10.4121/UUID:360795C8-1DD6-4A5B-A443-185001076EAB, URL: https://data.4tu.nl/collections/_/5065538/1.
- [Ec15] van Eck, M. L.; Lu, X.; Leemans, S. J. J.; van der Aalst, W. M. P.: PM^2 : *A Process Mining Project Methodology*. In (Zdravkovic, J.; Kirikova, M.; Johansson, P., eds.): *Advanced Information Systems Engineering*. Springer International Publishing, Cham, pp. 297–313, 2015, doi: 10.1007/978-3-319-19069-3_19.

- [GV07] Günther, C. W.; Van Der Aalst, W. M. P.: Fuzzy Mining – Adaptive Process Simplification Based on Multi-perspective Metrics. 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.; Alonso, G.; Dadam, P.; Rosemann, M., eds.): *Business Process Management*. Vol. 4714, Springer, Berlin, Heidelberg, pp. 328–343, 2007, DOI: 10.1007/978-3-540-75183-0_24.
- [Ha21] Hachicha, W.; Ghorbel, L.; Champagnat, R.; Zayani, C. A.; Amous, I.: Using Process Mining for Learning Resource Recommendation: A Moodle Case Study. *Procedia Computer Science* 192, pp. 853–862, 2021, DOI: 10.1016/j.procs.2021.08.088.
- [HH21] Hassani, M.; Habets, S.: Predicting Next Touch Point In A Customer Journey: A Use Case In Telecommunication. In (Al-Begain, K.; Iacono, M.; Campanile, L.; Bargiela, A., eds.): *Proceedings of the 35th International Conference on Modelling and Simulation (ECMS)*. European Council for Modeling and Simulation, pp. 48–54, 2021, DOI: 10.7148/2021-0048, URL: <https://doi.org/10.7148/2021-0048>.
- [IKT18] Intayoad, W.; Kamyod, C.; Temdee, P.: Process Mining Application for Discovering Student Learning Paths. In: 2018 International Conference on Digital Arts, Media and Technology (ICDAMT). IEEE, Phayao, pp. 220–224, 2018, DOI: 10.1109/ICDAMT.2018.8376527.
- [Ko24] Koschmider, A.; Aleknonytė-Resch, M.; Fonger, F.; Imenkamp, C.; Lepsien, A.; Apaydin, K.; Janssen, D.; Langhammer, D.; Ziolkowski, T.; Zisgen, Y.: Process Mining for Unstructured Data: Challenges and Research Directions. In: *Modellierung 2024*. Gesellschaft für Informatik e.V., Bonn, pp. 119–136, 2024, DOI: 10.18420/modellierung2024_012.
- [La22] Lamghari, Z.: Unstructured Business Processes Improvement Using Process Mining Techniques. *ASM Science Journal* 17, pp. 1–13, 2022, DOI: 10.32802/asmscj.2022.965.
- [Mu15] Mukala, P.; Buijs, J.; Leemans, M.; Van der Aalst, Wil M. P.: Learning Analytics on Coursera Event Data: A Process Mining Approach. In: *Proceedings of the 5th International Symposium on Data-driven Process Discovery and Analysis (SIMPDA)*. Vol. 1527, CEUR, Vienna, Austria, pp. 18–32, 2015, URL: <https://api.semanticscholar.org/CorpusID:563734>.
- [Mu22] Munoz-Gama, J.; Martin, N.; Fernandez-Llatas, C.; Johnson, O. A.; Sepúlveda, M.; Helm, E.; Galvez-Yanjari, V.; Rojas, E.; Martinez-Millana, A.; Aloini, D.; Amantea, I. A.; Andrews, R.; Arias, M.; Beerepoot, I.; Benevento, E.; Burattin, A.; Capurro, D.; Carmona, J.; Comuzzi, M.; Dalmas, B.; De La Fuente, R.; Di Francescomarino, C.; Di Ciccio, C.; Gatta, R.; Ghidini, C.; Gonzalez-Lopez, F.; Ibanez-Sanchez, G.; Klasky, H. B.; Prima Kurniati, A.; Lu, X.; Mannhardt, F.; Mans, R.; Marcos, M.; Medeiros De Carvalho, R.; Pegoraro, M.; Poon, S. K.; Pufahl, L.; Reijers, H. A.; Remy, S.; Rinderle-Ma, S.; Sacchi, L.; Seoane, F.; Song, M.; Stefanini, A.; Sulis, E.; Ter Hofstede, A. H.; Toussaint, P. J.; Traver, V.; Valero-Ramon, Z.; Weerd, I. V. D.; Van Der Aalst, W. M.; Vanwersch, R.; Weske, M.; Wynn, M. T.; Zerbato, F.: *Process Mining for Healthcare: Characteristics and Challenges*. *Journal of Biomedical Informatics* 127, p. 103994, 2022, DOI: 10.1016/j.jbi.2022.103994.
- [Pe22] Pegoraro, M.; Uysal, M. S.; Hülsmann, T.-H.; Van Der Aalst, W. M. P.: Uncertain Case Identifiers in Process Mining: A User Study of the Event-Case Correlation Problem on Click Data. In (Augusto, A.; Gill, A.; Bork, D.; Nurcan, S.; Reinhartz-Berger, I.; Schmidt, R., eds.): *Enterprise, Business-Process and Information Systems Modeling*. Vol. 450, Springer International Publishing, Cham, pp. 173–187, 2022, DOI: 10.1007/978-3-031-07475-2_12.

- [SI21] Sarirah Husin, H.; Ismail, S.: Process Mining Approach to Analyze User Navigation Behavior of a News Website. In: Proceedings of the 4th International Conference on Information Science and Systems. ICISS '21, ACM, New York, NY, USA, pp. 7–12, 2021, doi: 10.1145/3459955.3460593.
- [Sp20] Spree, F.: Predictive Process Monitoring - A Use-Case-Driven Literature Review. EMISA Forum 40 (1), pp. 43–54, 2020, URL: <https://dl.gi.de/items/ba4e5170-e09e-47f1-bbab-23a9c831256c>.
- [St20] Stefanini, A.; Aloini, D.; Benevento, E.; Dulmin, R.; Mininno, V.: A Process Mining Methodology for Modeling Unstructured Processes. Knowledge and Process Management 27 (4), pp. 294–310, 2020, doi: 10.1002/kpm.1649.
- [TH18] Terragni, A.; Hassani, M.: Analyzing Customer Journey with Process Mining: From Discovery to Recommendations. In: 2018 IEEE 6th International Conference on Future Internet of Things and Cloud (FiCloud). IEEE, Barcelona, Spain, pp. 224–229, 2018, doi: 10.1109/FiCloud.2018.00040.
- [Va22] Van Der Aalst, W.M.P.: Process Mining: A 360 Degree Overview. In (Van Der Aalst, W.M.P.; Carmona, J., eds.): Process Mining Handbook. Vol. 448, Springer International Publishing, Cham, pp. 3–34, 2022, doi: 10.1007/978-3-031-08848-3_1.
- [VG07] Van Der Aalst, W.; Gunther, C.: Finding Structure in Unstructured Processes: The Case for Process Mining. In: Seventh International Conference on Application of Concurrency to System Design (ACSD 2007). IEEE, Bratislava, pp. 3–12, 2007, doi: 10.1109/ACSD.2007.50.
- [WH23] Wolters, L.; Hassani, M.: Predicting Activities of Interest in the Remainder of Customer Journeys Under Online Settings. In (Montali, M.; Senderovich, A.; Weidlich, M., eds.): Process Mining Workshops. Vol. 468, Springer Nature Switzerland, Cham, pp. 145–157, 2023, doi: 10.1007/978-3-031-27815-0_11.
- [Ze21] van Zelst, S. J.; Mannhardt, F.; de Leoni, M.; Koschmider, A.: Event abstraction in process mining: literature review and taxonomy. Granular Computing 6, pp. 719–736, 2021, doi: 10.1007/s41066-020-00226-2.

A 3 phases model towards an industrialization of (generative) artificial intelligence developed for corporate environments

short paper

Martin Przewloka¹

Abstract:

Companies face major challenges when it comes to developing a strategic approach related to artificial intelligence (AI). The massive growth of generative models is another indication that technology adaptation is associated with very large uncertainties. As a result, undirected activities in corporates occur in which so-called use cases, prototypes and minimum viable products are implemented. A targeted and strategic approach is hardly recognizable. The aim of the work carried out here was to develop a practical model that meets all parameters: moving forward quickly, learning from mistakes, developing unique selling points and thus competitive advantages ultimately resulting in an AI strategy. This is exactly what can be achieved in 3 steps, the efficiency phase, the competitive advantage phase and the transformation phase, all of which build on one another sequentially. The model presented here also makes a significant contribution to the development and adaptation of a responsible AI.

Keywords: digital transformation, generative AI, competitive advantages, technology-push, clinical trials, clinical studies, key performance indicators, change management, responsible AI, efficiency

1 Introduction

Almost 1,5 years have passed: The hype - if that's what it is - surrounding Generative Artificial Intelligence, represented in particular by OpenAI, Aleph Alpha, Google Bard and others, is still present, although a little calm seems to have returned. In any case, one thing has happened in this relatively short period of time: Artificial intelligence (AI) is now no longer just a technology. It has become tangible and imaginable (cf. [Da24]). We all currently recognize what machines called digital computers are obviously capable of, even if they work extremely inefficiently, but can and will ultimately take over valuable tasks from humans.

Current research work is intensively concerned with the introduction of artificial intelligence applications in companies (for a systematic to date overview, please refer to [Ki21 and [Bo21]). The focus here is on selecting the right AI tools, knowledge management, change management aspects and decision-making processes. It is also

¹ Institut für Digitale Assistenzsysteme e.V., Technologiezentrum Gießen, Kerkrader Straße 9, 35394 Gießen
und Technische Hochschule Mittelhessen, Wiesenstraße 14, 35390 Gießen

assumed that the business strategies must lead to the selection of AI applications, but the reverse path is not taken into account. The model proposed, tested and unique to date in this work allows significantly more degrees of freedom: in particular the targeted adjustment of the corporate strategy considering current and future AI technologies.

2 The dilemma: short-term AI efforts versus the need to develop a sustainable and stable AI strategy

There is no doubt that artificial intelligence will penetrate all industrial areas (cf. [Mc18], [Co18]). Strategic actions are necessary, and this is the only way to achieve lasting success. However, it is noticeable that the need to develop an AI strategy (cf. [Mi20]) is overshadowed by general questions:

- What level of maturity does the technology have and how trustworthy are the AI-generated results? [Sa21]
- How is the use of AI regulated externally and internally (e.g. works council)? What influence does this have on your own fields of application? [Bu18]
- What legal aspects arise when using AI, such as liability, protection of intellectual property, protection of personal rights, etc.? [Ca18]
- Do I have the necessary data, or how do I ensure the required data foundation? cf. [Se23]

These questions cannot currently be answered, or at least only partially. The level of uncertainty in companies is high (cf. [EF24]) and means that AI does not receive the necessary strategic attention² or even is advanced in a comparatively unstructured manner.

The result to overcome the dilemma described is a solution scenario that initially seems surprising: the industrialization of AI requires an agile approach from which the strategy gradually develops and not the other way around. This means that from initially specific use cases, the clues and later the critical success factors of the AI strategy to be developed are formed.

3 Towards AI industrialization: an approach in 3 steps

We have seen this time and again in the past. Especially in the context of technologies, the risk of overvaluation is very high (cf. [Co23]). The first to do so will only rarely benefit from the time advantages that may exist, especially if the goals and expectations are too

² while the awareness in context with AI in companies is very high, this does not necessary lead to a strategic attention aligned with the vision and goals of the enterprise

high and ambitious. However, short-term successes are essential for building trust and security in this domain and therefore also form the basis for strategic investments.

Step 1: Generating efficiencies

A quick start requires quick successes. If these do not materialize, it will not be possible to develop a suitable AI vision for the company, as there is a lack of positive experience. Visionary promises, which are also disseminated by the media via all digital channels, can lead to ill-considered action and thus to disappointment. In this respect, it makes sense to align the first projects closely with the current business area and to optimize the associated processes and activities. Replacing human activities with automated systems is a good idea in precisely those areas where time-consuming, often routine-driven processes predominate (cf. [Ho00] and [Cu14]). Corresponding use cases that are barely coupled or even separated from each other will be developed, implemented and deployed (figure 1, step 1).

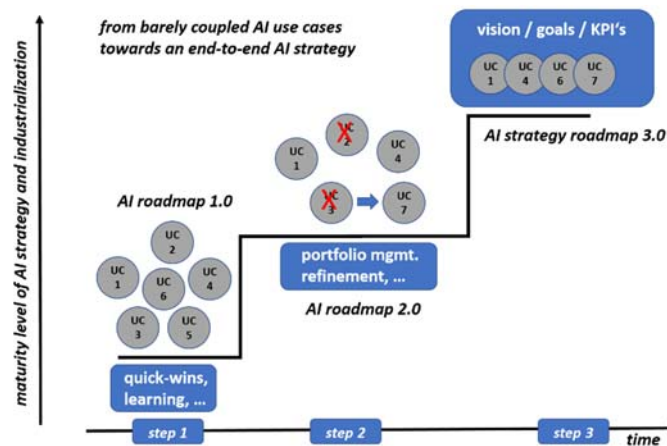


Figure 1: in three steps, use cases form evolutionary the AI strategy and are finally leading to an AI industrialization approach

Step 2: Creating competitive advantages through the targeted use of AI

The next stage involves a much more strategic approach, i.e. the formation of a perspective that goes beyond short-term successes (cf. [Mc23], [Co22]). This also involves making initial investment decisions with a medium-term horizon. The competition is sounded out here, be it through benchmarking or to identify unique selling points for future entrepreneurial activity. The question is not about replacing humans with a machine, but rather about optimal implementation. Areas will emerge that are primarily processed by machines, i.e. automated, while others will absolutely require humans to carry out the

corresponding tasks. As a consequence, uses cases already identified and tested in the 1st phase will disappear, some have to be adjusted and finally, new ones will come up (refer to figure 1, step 2).

Step 3: The AI-based transformation of the company through to leapfrog innovation

The aim of an AI-based transformation will be to expand the existing business model, whether by opening up new markets, developing new products and/or services or even diversifying completely (cf. [Ma19]). Innovations are now more tangible than ever and "creating something new" with the help of generative technologies is becoming a lived part of entrepreneurial activity. This can only succeed if people and machines work together, which means that the efficiency debate will initially take a back seat and give way to a strategic roadmap (figure 1, step 3).

4 Summary and Conclusion

A step-by-step approach to implementing (generative) AI solutions in the corporate environment is necessary in order to successfully apply this currently highly valued technology. Initially, the rapid creation of experience with an appropriate short-term perspective is preferable to long-term, strategically plannable action. In this way, areas of success and problems can be identified at an early stage and transformational steps can be initiated on the basis of the relevant experience. Paradoxically, rudimentary human labor will initially be replaced by machine labor before the actual potential can be leveraged in the next step by merging humans and machines. It is only in this final stage, i.e. step 3, that an awareness of the responsible use of AI will develop. In the current hype phase, this is obviously still falling by the wayside.

References

- [Bo21] Borges, A. F., Laurindo, F. J., Spínola, M. M., Gonçalves, R. F., Mattos, C. A.; The strategic use of artificial intelligence in the digital era: Systematic literature review and future research directions; In: International journal of information management, 57, 102225 (2021).
- [Bu18] Buiten, M. C.; Towards intelligent regulation of artificial intelligence; in: European Journal of Risk Regulation, 10(1), pp. 41-59 (2019)
- [Ca18] Cath, C.; Governing artificial intelligence: ethical, legal and technical opportunities and challenges; in: Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 376(2133), 20180080 (2018).
- [Co18] Cockburn, I. M., Henderson, R., Stern, S.; The impact of artificial intelligence on innovation (Vol. 24449). Cambridge, MA, USA: National bureau of economic research (2018).

- [Co23] Cooban, A.; AI investment is booming. How much is hype? In: CNN News; <https://edition.cnn.com/2023/07/23/business/ai-vc-investment-dot-com-bubble/index.html> (23rd of July, 2023) (last access 29th of April, 2024).
- [Cu14] Cummings, M. M.; Man versus machine or man + machine?; in: IEEE Intelligent Systems, 29(5), pp. 62-69 (2014).
- [Da24] databricks; The great acceleration: CIO perspectives on generative AI; MIT report (2024).
- [EF24] EFI report; 2024 report on research and innovation; commission of experts for research and innovation (EFI); (2024).
- [Ha23] Hardman, T.C., Aitchison, R., Scaife, R., Edwards, J., Slater G.; The future of clinical trials and drug development in 2050; in Drugs Context 2023, 12;2023-2-2; doi: 10.7573/dic.2023-2-2; (2023).
- [Ho00] Hoc, J. M.; From human-machine interaction to human-machine cooperation; in: Ergonomics, 43(7), pp. 833-843 (2000).
- [Ho22] Horowitz, M. C., Allen, G. C., Kania, E. B., Scharre, P.; Strategic competition in an era of artificial intelligence. Center for a New American Security. (2022).
- [Ki21] Kitsios, F., Kamariotou, M.; Artificial intelligence and business strategy towards digital transformation: A research agenda; in: Sustainability, 13(4), 2025 (2021).
- [Ma19] Magistretti, S., Dell’Era, C., Petruzzelli, A. M. How intelligent is Watson? Enabling digital transformation through artificial intelligence; in: Business Horizons, 62(6), pp. 819-829 (2019).
- [Mc18] McKinsey; The promise and challenge of artificial intelligence; McKinsey Executive Briefing; <https://www.mckinsey.com/featured-insights/artificial-intelligence/the-promise-and-challenge-of-the-age-of-artificial-intelligence> (2018) (last access Apr. 29th, 2024)
- [Mc23] McKinsey; Rewired to Outcompete; <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/rewired-to-outcompete> (2013) (last access Apr. 29th, 2024)
- [Mi20] S. Mithas, S. Murugesan, P. Seetharaman; What is Your Artificial Intelligence Strategy?, in: IT Professional, vol. 22, no. 2, pp. 4-9, 1 March-April 2020, doi: 10.1109/MITP.2019.2957620.
- [Sa21] Sadiq, R. B., Safie, N., Abd Rahman, A. H., Goudarzi, S.. Artificial intelligence maturity model: a systematic literature review. In: PeerJ Computer Science, 7, e661 (2021).
- [Se23] Segal, R.; AI needs data more than data needs AI; Forbes Technology council; <https://www.forbes.com/sites/forbestechcouncil/2023/10/05/ai-needs-data-more-than-data-needs-ai/> (2023) (last access Apr. 29th, 2024)

Sicherheitsmaßnahmen für den ERP-Betrieb

Kurzbeitrag

Dennis Buckenauer¹ und Sandy Eggert²

Abstract: In dieser Forschungsarbeit wurden zunächst allgemeine IT-Sicherheitsbedrohungen im Bereich kleiner und mittlerer Unternehmen betrachtet. Anschließend erfolgte die Identifikation von Sicherheitslücken im IT-Betrieb, mit dem Fokus auf der Nutzung von ERP-Systemen. Eine Umfrage zu Sicherheitsrisiken und Schutzmaßnahmen untermauerte die Notwendigkeit für Unternehmen, einen proaktiven und ganzheitlichen Ansatz zur Gewährleistung der Sicherheit im ERP-Betrieb zu verfolgen. Die Identifikation von Schwachstellen und die Implementierung von Schutzmaßnahmen sind entscheidend, um die Sicherheit und Integrität von Unternehmensdaten zu gewährleisten. Die Ergebnisse der Arbeit mündeten in die Entwicklung eines Sicherheitsleitfadens, welcher speziell auf die Sicherheit des Einsatzes von ERP-Systemen in kleinen und mittleren Unternehmen abzielt.

Keywords: Sicherheitsrisiken, Sicherheitslücken, Schutzmaßnahmen, Sicherheitsleitfaden, ERP-Betrieb.

1 Einleitung

Kleine und mittlere Unternehmen sind heutzutage einer Vielzahl von IT-Sicherheitsbedrohungen ausgesetzt [BJ18]. Dies zeigt u.a. eine Untersuchung der Bitkom, in der die Arten digitaler IT-Angriffe erhoben wurden, die in Unternehmen entsprechende Schäden verursacht haben [Bi23]. Am häufigsten wurden Phishing-Angriffe mit 31% genannt, gefolgt von Angriffen auf Passwörter (29%) und Infektionen mit Schadsoftware bzw. Malware (28%). Ransomware-Angriffe (23%) sind nach wie vor ein großes Problem, da sie zu erheblichen Datenverlusten und finanziellen Schäden führen können. SQL-Injection-Angriffe (21%) stellen ein Risiko für die Datensicherheit dar, während Spoofing-Angriffe (14%) und DDoS-Attacken (12%) Betriebsunterbrechungen verursachen können. Weniger häufig, aber dennoch ernstzunehmende Bedrohungen sind Cross-Site-Scripting (XSS) (10%), Man-in-the-Middle-Angriffe (7%) und CEO Fraud (6%) [Bi23]. Diese hohen Angriffszahlen deuten drauf hin, dass das Bewusstsein für die Notwendigkeit einer systematischen Herangehensweise an die Informationssicherheit noch nicht ausreichend verbreitet ist und Sicherheitsmaßnahmen in Unternehmen unzureichend implementiert sind. Ziel ist es, Maßnahmen zu identifizieren, die die Sicherheit im IT-Betrieb von KMU erhöhen. Da mehr als die Hälfte der deutschen KMU ein ERP-System nutzen [Bi23], liegt der Fokus der Betrachtung auf den ERP-Betrieb.

¹ HWR Berlin, Badensche Str. 52, 10825 Berlin, dennis.buckenauer@gmail.com,

² HWR Berlin, Badensche Str. 52, 10825 Berlin, sandy.eggert@hwr-berlin.de

2 Identifikation von Sicherheitslücken

Um technische und organisatorische Schutzmaßnahmen ableiten zu können, erfolgte zunächst die Identifikation und Einordnung von Sicherheitslücken. Dazu wurde neben einer Literaturuntersuchung eine Onlinebefragung von Unternehmen aus dem Bereich KMU im Zeitraum von Januar bis Februar 2024 herangezogen. Die Befragung fand im Rahmen einer ERP-Marktstudie [Eg24] statt, an der sich insgesamt 64 ERP-Anbieter- und -Dienstleistungsunternehmen beteiligt haben. Folgende Fragen wurden im Bereich der IT-Sicherheit beantwortet:

- Was waren aus Ihrer Sicht die häufigsten Sicherheitsrisiken im Jahr 2023?
- Welche Schutzmaßnahmen setzen Sie für Ihr ERP-System ein?

Abbildung 1 (links) zeigt die prozentuale Verteilung verschiedener Angriffsarten. Die Ergebnisse zeigen, dass Phishing-E-Mails mit 53% den häufigsten Angriffstyp darstellen, gefolgt von Credential Sniffing (13%) und Cross-Site Scripting (13%). Diese Ergebnisse betonen die fortwährende Relevanz von Phishing als eine der vorherrschenden Bedrohungen im Bereich der IT-Sicherheit. Des Weiteren wird ersichtlich, dass DDoS-Angriffe mit 9% der gemeldeten Vorfälle signifikant vertreten sind, was auf die kontinuierliche Bedrohung durch Verfügbarkeitsangriffe hinweist. Ebenso sind Angriffe wie SQL-Injection (8%) und Man-in-the-Middle (5%) relevante Bedrohungen, die auf Schwachstellen in der Software und Netzwerkinfrastruktur abzielen. Die Präsenz von physischen Angriffen (11%) weist daraufhin, dass die Sicherheit der physischen Infrastruktur von Unternehmen ebenso wichtig ist, wie die Sicherheit der digitalen Systeme.

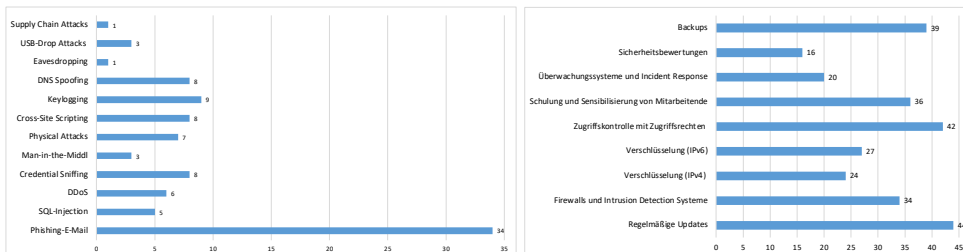


Abb. 1: Häufigkeit von Sicherheitsrisiken nach Angriffsarten (links) und Implementierungen von Sicherheitsmaßnahmen (rechts), n=64

Eine Übersicht, welche Sicherheitsmaßnahmen ERP-Anbieter und -Dienstleister bereits in ihren Systemen implementiert haben, zeigt die Abbildung 1 (rechts). Aus den Daten geht hervor, dass mehr als 60 % der befragten Unternehmen „regelmäßige Updates“, „Zugangskontrolle mit Zugriffsrechten“ und „Backups“ systemseitig implementiert haben. Deutlich mehr als die Hälfte der Unternehmen gaben an „Schulung und Sensibilisierung der Mitarbeiter“ in ihren Projekten einzugliedern. Weiterhin werden Firewalls und Intrusion Detection Systeme von etwas mehr als der Hälfte der Systemanbieter implementiert.

3 Schutzmaßnahmen

Im Rahmen der Literaturstudie wurden unterschiedliche Schutzmaßnahmen identifiziert und in technische und organisatorische Maßnahmen gegliedert.

3.1 Technische Schutzmaßnahmen

Die Sicherung von Unternehmensdaten in ERP-Systemen stellt insbesondere für kleine und mittlere Unternehmen aufgrund der damit einhergehenden Kosten und Personaleinsatzes eine große Herausforderung dar. Effektive Sicherheit in ERP-Systemen erfordert eine Kombination verschiedener technischer Schutzmaßnahmen, die speziell darauf ausgelegt sind, die Integrität, Vertraulichkeit und Verfügbarkeit sensibler Informationen zu gewährleisten. Folgende Schutzmaßnahmen wurden identifiziert:

Firewall- und Virenschutzsysteme: Die Kombination von Firewall- und Virenschutzsystemen bildet eine grundlegende Verteidigungslinie für ERP-Systeme in KMU. Durch die Implementierung von Richtlinien und regelmäßigen Updates können sie dazu beitragen, das Risiko von Angriffen zu minimieren und die Integrität der Unternehmensdaten zu schützen [Bu21].

Verschlüsselungstechnologien: Grundsätzlich sind im Rahmen der Verschlüsselungstechnologien in der Informationssicherheit die symmetrische und die asymmetrische Verschlüsselung zu unterscheiden. Hinzu kommen hybride Verschlüsselungsverfahren, die eine Kombination aus symmetrischer und asymmetrischer Verschlüsselung verwenden. Diese kombinierten Methoden bieten eine ausgewogene Lösung zwischen Sicherheit und Effizienz und werden oft in komplexen Sicherheitsumgebungen eingesetzt. Ein effektives Schlüsselmanagement ist entscheidend, um die Sicherheit der verschlüsselten Daten zu gewährleisten [BS24].

System- und Netzwerküberwachung: Die Überwachung von ERP-Systemen und Netzwerken ermöglicht die frühzeitige Identifikation von Sicherheitsvorfällen. Überwachungssysteme bieten Einblicke in Geräte und Verbindungen und unterstützen IT-Experten bei der Identifizierung und Meldung von Fehlern, der Planung von Wartungsmaßnahmen und der Einhaltung von Compliance- und Sicherheitsstandards. Sie tragen dazu bei, Netzausfälle zu vermeiden und die Ressourcen für IT-Anpassungen zu optimieren [Sh23]. Identifizierte Faktoren im Rahmen der System- und Netzwerküberwachung sind: Früherkennung von Sicherheitsvorfällen, Tools und Methoden für die Überwachung sowie Intrusion Detection Systems (IDS) und Intrusion Prevention Systems (IPS). IDS und IPS bieten eine eingehende Überprüfung und Identifikation potenzieller Sicherheitsbedrohungen im Vergleich zu herkömmlichen Firewalls und Virenschutzsystemen und sind in der Lage, sowohl bekannte als auch unbekannte Bedrohungen zu identifizieren und genauere Informationen über die Art und Herkunft des Angriffs bereitzustellen. Durch die Echtzeitüberwachung des Netzwerkverkehrs können sie auf verdächtige Aktivitäten reagieren und potenzielle Angriffe abwehren [Ma23].

3.2 Organisatorische Sicherheitsmaßnahmen

Schulungen und Sensibilisierung: Schulungen und Sensibilisierungsmaßnahmen spielen eine maßgebliche Rolle, um die Mitarbeitenden für Sicherheitsrisiken zu sensibilisieren und bewusstes Verhalten zu fördern [Is23].

Erstellung von Sicherheitsrichtlinien: Sicherheitsrichtlinien dienen als Regeln, die konfiguriert werden können, um Ressourcen auf einem Gerät oder Netzwerk zu schützen. Sie sollten als integraler Bestandteil der allgemeinen Sicherheitsimplementierung verwendet werden, um Domänencontroller, Server, Clientgeräte und andere Ressourcen in einer Organisation zu schützen [CY24].

Incident Response Plan: Die Entwicklung und Implementierung eines Incident Response Plans (IRP) für ERP-Systeme ist ein entscheidender Schritt, um auf Sicherheitsvorfälle schnell und effektiv reagieren zu können. ERP-Systeme sind zentrale Komponenten der IT-Infrastruktur in Unternehmen, verarbeiten sensible Daten und stellen dadurch ein attraktives Ziel für Sicherheitsangriffe dar [BS21].

4 Erste Ergebnisse und Ausblick

Im ersten Schritt wurde ein Sicherheitsleitfaden konzipiert, der KMU dabei unterstützt, systematisch sicherheitsrelevante Aspekte zu überprüfen und geeignete Maßnahmen zu ergreifen und aktuelle Lücken im Sicherheitsmanagement zu identifizieren.

	ja/nein	entsprechende Maßnahme umsetzung	nicht erfüllt
1. Entwicklung von Sicherheitsrichtlinien:			
Festlegung klarer Anforderungen an das Passwortmanagement			
Verbot der Installation von nicht autorisierter Software			
Beschränkung der Nutzung von externen Speichermedien			
2. Benutzerschulung und Sensibilisierung:			
Schulung der Mitarbeitenden zur sicheren Nutzung von ERP-Systemen und den Umgang mit sensiblen Daten.			
Sensibilisierung für Phishing-Angriffe und andere Cyber-Bedrohungen sowie Maßnahmen zur Erkennung und Vermeidung.			
Integration von regelmäßigen Schulungen und Schulungsmaterialien in die Onboarding-Prozesse neuer Mitarbeitenden.			
Durchführung regelmäßiger Phishing-Tests, um die Reaktionsfähigkeit der Mitarbeitenden zu verbessern.			
3. Zugriffskontrolle:			
Ermittlung und Dokumentation aller Benutzernden, die Zugriff auf das ERP-System haben.			
Implementierung von strengen Zugriffskontrollen und Berechtigungsstufen basierend auf den Rollen und Verantwortlichkeiten der Benutzernden.			
Regelmäßige Überprüfung und Aktualisierung von Benutzerberechtigungen.			
Implementierung einer Zwei-Faktor-Authentifizierung zur weiteren Sicherung der Zugriffskontrollen.			
4. Systemintegrität:			
Regelmäßige Überprüfung der Systemintegrität durch Sicherheits- und Penetrationstests.			
Implementierung von Sicherheitspatches und -aktualisierungen für das ERP-System und alle verbundenen Softwarekomponenten.			
Überwachung und Erkennung von Anomalien im Systemverhalten.			
Implementierung eines Systems zur Überwachung von Sicherheitsupdates und Patches.			
5. Compliance und Datenschutz:			
Einhaltung geltender Datenschutzgesetze und -verordnungen, insbesondere im Hinblick auf den Umgang mit personenbezogenen Daten.			
Regelmäßige Überprüfung der Compliance mit branchenspezifischen Standards und Best Practices im Bereich der Informationssicherheit.			
Dokumentation über die Einhaltung von Datenschutzgesetzen.			
Implementierung regelmäßiger interner Überprüfungen und Audits zur Sicherstellung der Einhaltung von Standards und Best Practices.			
6. Datensicherheit:			
Implementierung von Verschlüsselungstechnologien für sensible Daten sowohl während der Speicherung als auch der Übertragung.			
Regelmäßige Sicherung und Archivierung von Daten, einschließlich Offsite-Backups, um im Falle von Datenverlusten wiederherstellen zu können.			
Überwachung und Schutz vor Datenverlust oder -diebstahl durch Zugriffskontrollen und Firewalls.			
Regelmäßige Überprüfung der Effektivität der implementierten Verschlüsselungstechnologien.			
7. Notfallvorsorge und Incident Management:			
Entwicklung eines Notfallplans für den Umgang mit Systemausfällen, Datenverlust oder Sicherheitsverletzungen.			
Einrichtung eines Incident Response-Teams zur schnellen Reaktion auf Sicherheitsvorfälle.			
Regelmäßige Überprüfung und Aktualisierung des Notfallplans basierend auf aktuellen Bedrohungen und Risiken.			
Durchführung von Simulationen von Sicherheitsvorfällen zur Überprüfung der Wirksamkeit des Notfallplans.			
8. Externe Überwachung und Auditierung:			
Analyse der Ergebnisse externer Audits und Implementierung von Maßnahmen zur Verbesserung der Sicherheit.			
Einbindung externer Sicherheitsexperten in regelmäßige Sicherheitsüberprüfungen und Audits zur unabhängigen Bewertung der Sicherheitsmaßnahmen.			

Abb. 2: Sicherheitsleitfaden für KMU

Durch die Kombination von technischen und organisatorischen Ansätzen verfolgt der Leitfaden einen ganzheitlichen Sicherheitsansatz. Der Leitfaden ermöglicht es den Unternehmen, ihre Sicherheitsmaßnahmen kontinuierlich zu verbessern und sich den sich ständig verändernden Bedrohungslandschaften anzupassen.

In einer weiteren Entwicklungsstufe soll der Leitfaden um einen Maßnahmenkatalog ergänzt werden, der technischen und organisatorische Schutzmaßnahmen mit konkreten Umsetzungsvorschlägen ausstattet.

Literaturverzeichnis

- [Bi23] Bitkom; Statista: Welche der folgenden Arten von digitalen IT-Angriffen haben innerhalb der letzten 12 Monate in Ihrem Unternehmen einen Schaden verursacht?, <https://de-statista-com.ezproxy.hwr-berlin.de/statistik/daten/studie/928943/umfrage/von--digitalen-angriffen-betroffene-unternehmen-nach-art-des-angriffs/>, Stand: 13.03.2024.
- [BJ18] Bollhöfer, E.; Jäger, A.: Wirtschaftsspionage und Konkurrenzausspähung. Vorfälle und Prävention bei KMU im Zeitalter der Digitalisierung, Max-Planck-Institut für ausländisches und internationales Strafrecht eV., Freiburg i.Br., 2018.
- [BS21] BSI, Bundesamt für Sicherheit in der Informationstechnik, APP.4.2: SAP-ERP-System. https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Grundschutz/Kompendium_Einzel_PDFs_2021/06_APP_Anwendungen/APP_4_2_SAP_ERP_System_Edition_2021.html, Stand: 03.03.2024.
- [BS24] BSI, Bundesamt für Sicherheit in der Informationstechnik, Arten der Verschlüsselung. <https://www.bsi.bund.de/dok/6597616>, Stand: 06.03.2024.
- [Bu21] van der Burgt, R., 2021. Antiviren-Software vs. Firewall: Was ist der Unterschied? <https://vpnoverview.com/de/antiviren-information/unterschied-antivirus-firewall/>, Stand: 06.02.2024.
- [CY24] Cyberpilot, Kostenlose Vorlage für IT-Sicherheitsrichtlinie – Eine Schritt-für-Schritt-Anleitung. <https://www.cyberpilot.io/de/cyberpilot-blog/it-sicherheitsrichtlinie-eine-schritt-fuer-schritt-anleitung>, Stand: 06.02.2024.
- [Eg24] Eggert, S.: ERP-Marktanalyse 2024. In: ERP Information 1/2024, S. 74-102, DPI Verlag Berlin, 2024.
- [Is23] Isler, K., 2023. Sicherheitsbewusstsein schaffen: Wie Unternehmen ihre Mitarbeiter mit Trainings sensibilisieren können. <https://www.hagel-it.de/it-insights/sicherheitsbewusstsein-schaffen-wie-unternehmen-ihre-mitarbeiter-mit-trainings-sensibilisieren-koennen.html>, Stand: 06.02.2024.
- [Ma23] Mahmood, S., 2023. Der Unterschied zwischen Firewall und IDS, IPS: Die Grundlagen verstehen. <https://nextdoorsec.com/de/der-unterschied-zwischen-firewall-und-ids-ips-die-grundlagen-verstehen/>, Stand: 06.02.2024.
- [Sh23] Sharif, A., 2023. Was ist Netzwerküberwachung? <https://www.crowdstrike.de/cyber-security-101/observability/network-monitoring/>, Stand: 06.02.24.

