Grußworte

Liebe Fachgruppenmitglieder,

Ich freue mich, Sie als Empfänger der neusten Ausgabe des EMISA Forums der Fachgruppe EMISA begrüßen zu dürfen. Leider können wir an dieser Stelle nicht an unser alljährliches Fachgruppentreffen erinnern, da uns Covid dieses Jahr einen Strich durch die Rechnung gemacht hat. Die Begutachtung der Beiträge fand trotzdem statt und wir danken Agnes Koschmider, Judith Michael und Bernhard Thalheim für die Erstellung der Proceedings. Diese sind unter

http://ceur-ws.org/Vol-2628/

verfügbar. Zudem freue ich mich, dass Agnes Koschmider sich von diesen Umständen nicht hat einschüchtern lassen, im Jahr 2021 einen erneuten Anlauf zu starten. Sie wird unterstützt von Judith Michael das nächste Fachgruppentreffen in Kiel zu organisieren. Der Termin steht mit dem 20.-21. Mai 2021 bereits fest und es gibt sogar schon eine Website:

https://www.emisa2021.de/

Wir bitten alle, sich diesen Termin bereits einzutragen und die Reise nach Kiel zu planen. Selbstverständlich müssen wir aber auch mit weiteren Einschränkungen im kommenden Jahr rechnen. Daher planen wir aktuell mit einem Hybridformat. Sofern möglich freuen wir uns darauf, wieder einmal zu persönlichen Gesprächen zusammen zu kommen. Einreichungen von Extended Abstracts und Short Papers in den Kategorien PhD Research Proposal, Current Research Talk und Novel Directions Talk werden erbeten. Zudem steht wieder einmal eine Wahl an: es gibt das Leitungsgremium für die Periode von 2022-24 zu bestimmen. Wir freuen uns auf eine weitere schöne EMISA-Tagung in Kiel!

Mit herzlichen Grüßen,

Jan Mendling
(EMISA-Sprecher)

11th International Workshop on Enterprise Modeling and Information Systems Architectures

https://www.emisa2021.de

Call for Papers

EMISA 2021 is the eleventh international workshop in a series that provides a key forum for researchers and practitioners in the field on design methods for information systems. The workshop series emphasizes a holistic view on this field, fostering integrated approaches that address and relate business processes, business people and information technology.

The workshop is open for a broad range of subjects. Possible topics include, but are not limited to:

- Patterns for information systems architectures
- Process modeling and process-aware information systems for IoT
- Complex event processing and event-driven architectures
- Information systems for IoT
- Modelling social information and innovation networks
- Domain-specific modeling methods and languages for IoT
- Metamodeling
- Method and model engineering
- Quality of modeling methods, models, architectures and languages
- Learning and teaching design methods for information systems

Organization

The workshop is organized by the *GI Special Interest Group on Design Methods for Information Systems* (GI-SIG EMISA), which provides a forum for researchers from various disciplines who develop and apply methods to support the analysis and design of information systems.

Submission Types

EMISA 2021 calls for submissions in the following categories:

1. PhD Research Proposals:
EMISA 2021 invites PhD students to submit research proposals. There will be a dedicated slot in the program to discuss PhD research proposals including the current status and the further plan of the research work. PhD research proposals shall be submitted as a short paper of 5 pages.

2. Current Research Talk Proposals:
EMISA 2021 invites proposals for scientific talks of international excellence. Eligible are proposal submissions that are based on published or accepted papers from international conferences or journals. Proposals for research talks shall be submitted as an extended abstract of up to 2 pages.

3. Novel Directions Talk Proposals:
EMISA 2021 invites proposals for talks that motivate a novel research direction, outline the research gaps to address, and carve out major challenges. These talks shall serve as a stimulus for discussions as part of a dedicated slot in the workshop program. Novel directions talk proposals shall be submitted as a short paper of 5 pages.

All accepted submissions (PhD Research, Current Research Talk, Novel Directions Talk) will be published in the next print edition of EMISA Forum. The short papers proposing PhD Research or a Novel Directions Talk will also be published as an electronic CEUR proceedings volume. All submissions have to strictly follow the formatting guidelines of LNI. Template and explanations can be found at the GI website. Submissions have to be made via easychair.

Important Dates

Submission Deadline	February 27, 2021
Notification	March 20, 2021
Final papers	April 01, 2021

Chairs

Agnes Koschmider, Kiel University
Judith Michael, RWTH Aachen

ZEUS2021

13th Central-European Workshop on Services and their Composition (ZEUS 2020)

https://zeus2021.pi.uni-bamberg.de/

Objectives

ZEUS focuses on the discussion of fresh ideas, the presentation of work in progress, and the establishment of a scientific network between young researchers in the region.

1. Discuss fresh ideas

We offer a forum to discuss ideas at a level that is more work-in-progress than in a traditional conference. We thereby want to attract especially PhD students in the early phases of their work. Participants can get feedback from outside their group before a submission to a reviewed conference. This makes ZEUS a great opportunity to discuss ideas.

2. Practice scientific work

We see the ZEUS workshop as an opportunity to practice the whole range of scientific work. We do not put the sole focus on the submitted papers themselves, but also on the presentations and the discussions during the workshop. To this end, we hand out a Best Presentation Award since 2010 at the end of the workshop to appreciate high quality presentations.

3. Establish contacts between young researchers in the region

We aim at bringing together young researchers who work in the same geographic and scientific region. This way, we would like to provide an opportunity for people to establish a scientific network that can be intensely used, including mutual visits at affordable costs. The workshop will serve as a platform to present current research ideas and research directions.

Topics

The topics of the ZEUS workshop are centered around service technologies, which include a rich set of facets. The purpose of analysis, synthesis, or simulation of service technologies are as welcome as practical evaluations, use case-driven feasibility studies, or technology adoption models. ZEUS also calls for contributions in the field of Cloud Computing, RESTful services, and microservices.

Topics include, but are not limited to:

- Service lifecycle: analysis, specification, modelling, testing, deployment,
- execution, monitoring, adaptation
- Patterns, languages, reference models, and model extensions
- Multi-view and multi-perspective engineering (SOA, choreographies, collaborations, conversations, artifact-centric systems)
- Formal methods, models, simulation, and verification
- System architectures for service composition
- RESTful Web services (design aspects, hypermedia, linked data, mashups, conversations)
- Microservices and Nanoservices (architecture, lifecycle, deployment, composition)
- Workflows, business processes, and business decisions (modelling, execution, analysis, mining, as well as papers on blockchains and BPM)
- Complex event processing (correlation, aggregation, transformation, monitoring, extraction)
- Security, compliance, and non-functional requirements and properties
- Cloud-enabled applications, migration to/from the Cloud, Cloud Integration, Serverless Computing
- Composable Big Data Analytics Pipelines
- Applications, frameworks, methods, tool demonstrations, and case studies

Submission

We are looking forward to three types of contributions for ZEUS. All papers must be submitted following the instructions at the ZEUS submission site handled by EasyChair: https://easychair.org/conferences/?conf=zeus2021

Results can be presented in talks or tool demonstrations. Submissions will be reviewed by at least three reviewers each in order to assure general fitness regarding content, readability and scope and to give first feedback to the authors. Depending on innovation, technical soundness and presentation clarity, papers may be rejected or accepted as position or workshop papers.

Workshop papers:

Workshop papers are "regular" contributions that describe original solutions in field of ZEUS. These papers must not exceed 6 pages (LNCS style). The 6 pages does not include references, so there is more space for your work. Workshop papers are reviewed according to the call for papers. Accepted papers shall be included in the proceedings and presented at the workshop.

Positions papers:

Position papers should draft a new idea and put it up for discussion at the workshop. Position papers should only be an extended abstract and must not exceed 3 pages (LNCS style) without references. Position papers are briefly reviewed according to the call for papers. The main idea and the relation to existing work should be contained. Accepted papers shall be included in the proceedings. Position papers allow authors to get early feedback during the workshop, but should not disallow extending the paper to a full paper submitted to a first class conference – even if the position paper is referenced and the delta is explained properly.

Tools demonstrations:

ZEUS also offers a forum to demonstrate implementations of techniques and algorithms in the area of the aforementioned topics to get early feedback and provide interesting insights for the audience. Tool demonstrators are asked to submit a demo script of no more than 3 pages (LNCS style) without references which states how the tool is linked to the call for papers and what to expect during the demonstration.

Important Dates

Submission	January 19, 2021
Notification	February 14, 2021
Camera-ready (pre-proceedings) version	February 21, 2021
Registration	February 19, 2021
Workshop	February 25-26, 2021

Submission Guidelines

Template: LNCS style - https://github.com/latextemplates/LNCS
Workshop Paper: 6 pages excluding references
Position Paper: 3 pages excluding references
Tool Demonstration: 3 pages excluding references

Program Chair

Johannes Manner, University of Bamberg, Germany

Local Organizer

Robin Lichtenthäler, University of Bamberg, Germany

Contact

E-mail: zeus2021@easychair.org

A Case-Study to Teach Process-Aware Information Systems[1]

Carlo Simon[2], Stefan Haag[2]

The digitalization of products, processes, and services - where IoT-technologies facilitate the communication of real-world elements with a virtual instance - pose a challenge in teaching and learning. Modelling of these systems – the real world, its digital twin and their interaction – is hardly done in a theoretical way. Rather gaining and exchanging experiences throughout the learning process should be a significant learning output for the students.

Concrete challenges for teaching how to develop digital twins lie in the following aspects

- Static and dynamic behaviour must be understood and modelled, i.e. the components of the system, their interactions and the interfaces to the outside.
- A wide variety of application domains from end-to-end order business processes to processes on the shop-floor must be considered.
- A number of formalisms is used for the different purposes like scheduling or control.
- Different academic disciplines are involved from information systems to automation.

In order to convey such a comprehensive learning experience, the authors apply problem-based and research-oriented learning in form of a case-study. Students learn how to recognize complex tasks in teams, structure them, weigh up solutions against each other, implement their chosen one, evaluate the implementation and present the results. The case study is applied to 3rd grade bachelor students and 1st grade master students. The students' central task reads as follows: *Develop a holistic, integrated model of a company according to the automation pyramid, link it to reality and control reality via the model*!

The concrete elaboration of this tasks differs on the students' previous knowledge. Bachelors follow a problem-based approach, masters a research-oriented one.

Case setting is a fabrication plant given as a fischertechnik model upgraded with a Raspberry Pi. Sensors and actuators are used to control the operation of work pieces in accordance to given orders. Though the production cell is kept as simple as possible, product variants are viable as to simulate an IoT-setting. Further variations occur from secondary conditions like order priority or limited time budget, all of which need to be considered.

[1] The original article is published as C. Simon, S. Haag: Digitale Zwillinge modellieren und verstehen. In: J. Michael, D. Bork, J.-R. Rehse, M. Striewe, M. Ullrich (Eds.) Joint Proceedings of Modellierung 2020: Short, Workshop and Tools & Demo Papers. Wien, Austria, 2020, CEUR-WS, S. 101 – 112.

[2] HS Worms, FB Informatik, Erenburgerstr. 19, 67549 Worms, Deutschland, {simon,haag}@hs-worms.de

The problem-based approach begins with the dissemination of theoretical basics, after which the task gets announced. The students now need to consider which of the principles given contribute to a solution and how to apply these. Afterwards, their theoretical solution gets implemented while being guided by the lecturers. Presenting the chosen solution is part of the teaching experience where students both give and receive feedback in plenary. At the end of term, the students reflect their learning process and search for ways to improve it.

The higher degree of freedom research-oriented learning entails requires a more open approach. Students get an overview of the case-study and are tasked themselves with understanding the challenges of modelling holistically. Afterwards, they need to formulate properties modelling languages should possess to be usable in the given context and examine known languages for applicability. As the students don't have expertise in production or automation technology, they have to acquire the relevant knowledge. Contrary to the bachelor students, the master students evaluate possible languages rather than learning one. Now they search for criteria to choose their course of action and implement their solution. Lastly, presentations including the reasoning for the chosen solution are held, feedback is given and taken in plenary and the learning process is reflected and assessed.

While not being able to formally evaluate the success of the presented teaching approach due to varying cohort sizes and a continuing development of the used tool which influences the possibilities to elaborate the task, a series of observations can be made:

- As the students need to present interim results in the course of the term and receive feedback on them, the presentations on term's end are of high quality.
- Even in early semesters, students experience the university as a place of research and teamwork. Thus, they become aware of the difference to school-based learning.
- Due to the received feedback, students develop scrutiny and a more self-aware attitude.

The courses make use of a new web-based tool for higher Petri-nets which reduces the number of needed process modelling languages to one. The choice is due to the following:

- All of the mentioned diverse application domains can be modelled using Petri nets.
- Petri nets are able to represent every level of abstraction.
- The inherent semantics allows for simulatable process models.

The web app is novel as it implements aspects not yet found in other software. It can be used distributed, even on mobile devices. Time data types are included and usage of higher Petri nets enables modelling of complex systems because users may define own data types. The tool supports organizational and data modelling as well as process maps. All of these models can be connected and integrated – even with the real twin – making it possible to impart different aspects of modelling and the competencies needed to create and evaluate the results, making it possible to both teach and utilize process-aware information systems.

Business process representation and performance of different task types (Extended Abstract)

Hamzah Ritchi[1], Mieke Jans[2], Jan Mendling[3], Hajo A. Reijers[4]

Extended Abstract

The analysis of business processes is an integral part of risk assessment procedures and audit methodology. In both auditing research and process modeling research, there is an ongoing debate on which representation format might be best suited to support the analysis task. Most important in this context is the question whether business process models as visual representation might be superior to textual narratives. In this research, we refer to our recent article [RJMR20] that investigates the affinity of different tasks with two process representational formats: textual narratives and visual diagrams (BPMN process models). Our findings demonstrate that the representation format has an impact on task performance and that the direction of this impact depends upon the affinity of the tasks type with the representation format. This implies that auditors are best provided with different process representations, depending on the task they are performing at that moment. These findings have important implications for research on auditing tasks, and more broadly also for software engineering and information systems research.

Bibliography

[RJMR20] Ritchi, H., Jans, M. J., Mendling, J., & Reijers, H. (2020). The influence of business process representation on performance of different task types. *Journal of Information Systems*. In press.

[1] Padjadjaran University, Jl. Dipati Ukur No.35,Bandung 40132, Indonesia, hamzah.ritchi@unpad.ac.id
[2] Hasselt University, Martelarenlaan 42, 3500 Hasselt, Belgium, mieke.jans@uhasselt.be
[3] Wirtschaftsuniversität Wien, Department of Information Systems and Operations, Welthandelsplatz 1, 1020 Wien, Austria, jan.mendling@wu.ac.at
[4] Utrecht University, Department of Information and Computing Sciences, Buys Ballot Gebouw, Princetonplein 5, 3584 CC Utrecht, The Netherlands, h.a.reijers@uu.nl

Current Research Talk: Conceptual Modelling and Artificial Intelligence*

Overview and research challenges from the perspective of predictive business process management

Peter Fettke[1,2]

[1] German Research Center for Artificial Intelligence (DFKI), Saarbrücken, Germany
[2] Saarland University, Saarbrücken, Germany
peter.fettke@dfki.de

Abstract. Currently, the visibility of Artificial Intelligence (AI) and society's expectations of AI are very high, particularly compared to other research topics, namely Modelling. However, between Conceptual Modelling (short: Modelling) and AI exist many interesting and important interrelationships. This position paper overviews possible applications of AI for Modelling and Modelling for AI. After this general discussion, the field of predictive business process management is focused as a particular application case of AI and Modelling. Predictive process management uses machine learning for predicting the future state of a running process instance. The paper closes with some general remarks and research challenges.

Keywords: Artificial Intelligence, Modelling, Business Process Modelling, Deep Learning, Explainability

1 Motivation

The field of Artificial Intelligence (AI) receives tremendous public visibility and expectations of the society regarding the transformational potential of AI are extremely high. Although it is not the first time that AI receives so much attention in society, it is safe to say that the field has made some important and remarkable progress, e.g. machine translation, speech recognition, image classification, or playing board games archives results and quality levels which were not foreseen a decade before.

On the other hand, the field of Conceptual Modelling (short: Modelling) does not receive similarly high attention from the general audience. Moreover, from the tremendous success of using data for machine learning often the conclusion is drawn that the explicitly, hand-crafted making of a model which represents a domain is not necessary or useful during system development anymore. Such a negative conclusion about the importance of Modelling is false and dangerous because it is well-known

* This extended abstract is based on Peter Fettke: Conceptual Modelling and Artificial Intelligence Overview and research challenges from the perspective of predictive business process management. In: Judith Michael, Dominik Bork: Joint Proceedings of Modellierung 2020 Short, Workshop and Tools & Demo Papers Workshop on Models in AI, CEUR Workshop Proceedings, Vol. 2542, urn:nbn:de:0074-2542-0, pp. 157-164

that AI in general and machine learning in particular has important application prerequisites and severe limitations under particular application characteristics [1].

Hence, it is much more fruitful to explore and to elaborate the various and rich intellectual interrelationships between AI and Modelling. At the moment, no clear understanding exists in how Modelling and AI fit together. Against this background, the main objective of this position paper is to elaborate on interrelationships between AI and Modelling. As such, this short position paper does not aim to make a final statement on this topic, but it stimulates further discourse.

The paper unfolds as follows: After this introduction, Section 2 frames and position the fields of AI and Modelling. General application potentials of AI and Modelling are overviewed by Section 3. Section 4 focusses on the case of predictive process management. The paper closes with some remarks and research challenges.

To elaborate more deeply on particular challenges, the case of predictive business process management is presented [3]. Business process monitoring is a phase of the business process management life cycle. Typical examples of business processes are order-to-cash, purchase-to-pay, and complaint-to-resolution. Running process instances, also known as cases, are monitored and managed during the process execution, also known as process run-time. Typically monitored parameters and process characteristics are its current status, the executed process steps, the time taken to execute particular steps, or the throughput time (see Fig. 1). The objective of predictive process management is to gain insights about the future of a case. Based on the current case status, the future of the case is predicted. Typical questions are: What will be the next action to be taken for this case? When will the next event occur? When will this case terminate? Will the case be completed on time?

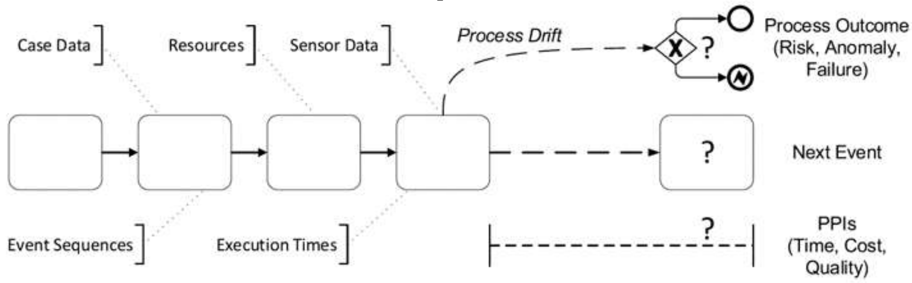


Fig. 1. Input data and predictands of process prediction (source: [2])

References

1. Ng, A.: What Artificial Intelligence Can and Can't Do Right Now. Harvard Business Review Digital Articles, pp. 2-4. Havard Business Review (2016)
2. Rehse, J.-R., Dadashnia, S., Fettke, P.: Business process management for Industry 4.0 – Three application cases in the DFKI-Smart-Lego-Factory. it – Information Technology 60, 133-141 (2018)
3. Evermann, J., Rehse, J.-R., Fettke, P.: Predicting process behaviour using deep learning. Decision Support Systems 100, 129-140 (2017)

Monotone Precision and Recall Measures for Comparing Executions and Specifications of Dynamic Systems (Extended Abstract)[1]

Artem Polyvyanyy[2], Andreas Solti[3], Matthias Weidlich[4], Claudio Di Ciccio[5], Jan Mendling[6]

1 Motivation

The behavioural comparison of systems is an important concern of software engineering research. For example, the areas of specification discovery and specification mining are concerned with measuring the consistency between a collection of execution traces and a program specification. This problem is also tackled in process mining with the help of measures that describe the quality of a process specification automatically discovered from execution logs.

Technically, behavioural comparisons are often formulated in a *relative* manner, defining a *quotient* of some aspect of one behaviour over the same aspect of another behaviour. For instance, the quotients of the behaviours of a system at different points in time reveal how the system has changed. In process mining, in turn, the quotient of the behaviour of a system as recorded in a log over the behaviour as specified can be used to analyse the trustworthiness of the latter [Ca18]. Yet, defining such quotients is challenging: A recent commentary on measures in process mining identifies a set of intuitive properties and shows that none of the available measures fulfils them [Ta18]. That is, measures lack monotonicity or cannot handle infinite behaviour.

[1] The original article appeared in the ACM Transactions on Software Engineering and Methodology (TOSEM) [Po20]

[2] The University of Melbourne, Australia, artem.polyvyanyy@unimelb.edu.au
[3] Vienna University of Economics and Business, Austria, solti@ai.wu.ac.at
[4] Humboldt-Universität zu Berlin, Germany, matthias.weidlich@hu-berlin.de
[5] Sapienza University of Rome, Italy, diciccio@di.uniroma1.it
[6] Vienna University of Economics and Business, Austria, mendling@ai.wu.ac.at

2 Approach

In this work, we address the problem of behavioural comparison by introducing a new framework for the definition of behavioural quotients. To this end, we approach the problem of behavioural comparison based on the notion of a formal language, which is a suitable starting point to capture the sequential (state-based) behaviour of a dynamic system. Specifically, we address the problem of *how to define meaningful quotients for behavioural comparison of finite and infinite languages.*

Our contributions include a framework for the definition of behavioural quotients that guarantee desired properties. Our framework is then used to instantiate two quotients that are grounded in the cardinality of a language (for finite languages) and the entropy of an automaton (for finite and infinite languages). It was recently shown that our entropy-based measures satisfy all the properties put-forward in the literature for recall and precision measures [STvdA19]. We further show how the proposed quotients can be used to define monotone precision and recall measures between the behaviour as recorded in an execution log of a system and the behaviour captured in a specification of the system.

3 Evaluation

Based on the jBPT library [PW13], we implemented the proposed precision and recall quotients in a tool that is publicly available.[7] Comprehensive experiments with execution logs of real IT systems highlight that existing measures indeed violate the monotonicity property. Moreover, the results of controlled experiments underpin that our quotients enable meaningful conclusions on the relation of the behaviours of two systems. Finally, we explore the scalability of the computation of our measures. The results illustrate the importance of convergence properties on the runtime of the method.

References

[Ca18] Carmona, Josep; Van Dongen, Boudewijn F.; Solti, Adreas; Weidlich, Matthias: Conformance Checking – Relating Processes and Models. Springer, 2018.

[Po20] Polyvyanyy, Artem; Solti, Andreas; Weidlich, Matthias; Ciccio, Claudio Di; Mendling, Jan: Monotone Precision and Recall Measures for Comparing Executions and Specifications of Dynamic Systems. ACM Trans. Softw. Eng. Methodol., 2020. In press.

[PW13] Polyvyanyy, Artem; Weidlich, Matthias: Towards a Compendium of Process Technologies - The jBPT Library for Process Model Analysis. CAiSE'13 Forum. Vol. 998 of CEUR Workshop Proceedings. CEUR-WS.org, pp. 106–113, 2013.

[STvdA19] Syring, Anja F.; Tax, Niek; van der Aalst, Wil M. P.: Evaluating Conformance Measures in Process Mining Using Conformance Propositions. TOPNOC 14:192–221, 2019.

[Ta18] Tax, Niek; Lu, Xixi; Sidorova, Natalia; Fahland, Dirk; van der Aalst, Wil M.P.: The imprecisions of precision measures in process mining. IPL, 135:1–8, 2018.

[7] https://github.com/jbpt/codebase/tree/master/jbpt-pm

Estimating Process Conformance by Trace Sampling and Result Approximation (Extended Abstract)[1]

Martin Bauer, Han van der Aa, Matthias Weidlich[2]

1 Motivation

Process-oriented information systems allow for the coordination of process models, describing the interplay between activities drawn out to reach a certain business goal, and recorded data, which captures the execution of these activities in real life. In particular, conformance checking methods, allow for the analysis and the comparison of these two views on a business process at hand, thus enabling a process analyst to comprehensively answer the question of how envisioned and executed behavior relate to each other [Ca18]. As the volume and frequency with which data is recorded increases, event logs comprise billions of events [VdA14]. Due to the exponential run time complexity of alignment-based conformance checking techniques [VD18], which are considered the de facto standard, the need for run time improvements is apparent. While various techniques for efficient alignment computation have been proposed, these, fundamentally, still require the consideration of *all*, possibly billions, of recorded events in a log.

Often, however, conformance checking aims to provide a general understanding of the overall conformance of process execution, rather then identifying all individual deviations. Recognizing this, we argue that for the calculation of general conformance insights, only a representative fraction of the traces in a log is required. Therefore, we are able to improve the run time performance of conformance checking in two ways. First, we reduce the number of traces required to obtain a conformance result through a sampling procedure. Second, we further reduce the required number of alignment calculations by using a worst-case approximation of the conformance result, for traces that are sufficiently similar to previously seen ones. In our work, we instantiate both the sampling and the approximation methods for two types of conformance results: a numerical fitness measure and a distribution of conformance issues over all activities.

[1] The original article was presented at the 17th International Conference on Business Process Management (2019) in Vienna, Austria [BVdAW19]

[2] All authors are with the Humboldt-Universität zu Berlin, Department of Computer Science, Berlin, Germany, {martin.bauer | han.van.der.aa | matthias.weilich}@hu-berlin.de

2 Approach

For trace sampling, we formulate the sampling of a new trace, calculating its conformance result, and updating the aggregated conformance result as a series of binomial experiments. In particular, by quantifying change induced in the aggregated conformance result ϵ, we can view this procedure as a binomial trial with two outcomes: either the trace introduces new information or it does not. Furthermore, based on a statistically grounded minimal sample size N, we determine when enough traces have been seen to consider the conformance result as sufficiently representative of the complete log. Formally, N represents the number of consecutive traces without new information that are required to conclude that the unknown probability p of the next trace containing new information is less than a chosen parameter δ.

To approximate conformance results, we compute the maximal impact that a newly sampled trace ξ can have on an aggregated conformance result. This approximation is based on the most similar observed trace ξ', as well as its edit distance to ξ. The worst-case impact of ξ on the overall conformance result is approximated based on the alignment of ξ' plus all differences between ξ and ξ'. The potential impact of ξ on the aggregated conformance result is checked against a significance threshold ϵ. Only if this check signals a potential significant change, an alignment is computed for ξ. Otherwise, we instead use the estimation as the conformance result of ξ, thus avoiding the need to compute an alignment for ξ.

3 Evaluation Results

We evaluated our techniques in comprehensive experiments with real-world and synthetic datasets for conformance results resembling the global fitness of the log and the relative distributions of non-conformant activities. Our results highlight dramatic improvements in terms of conformance checking efficiency compared to the baseline approach [VD18]. With samples as small as 0.1% to 1% of a log, we obtain conformance results that are virtually equivalent to those obtained when considering the complete log. This translates into respective reductions of the run times of conformance checking algorithms.

References

[BVdAW19] Bauer, Martin; Van der Aa, Han; Weidlich, Matthias: Estimating Process Conformance by Trace Sampling and Result Approximation. In: BPM. pp. 179–197, 2019.

[Ca18] Carmona, Josep; Van Dongen, Boudewijn F.; Solti, Adreas; Weidlich, Matthias: Conformance Checking – Relating Processes and Models. Springer, 2018.

[VD18] Van Dongen, Boudewijn F.: Efficiently Computing Alignments - Using the Extended Marking Equation. In: BPM. pp. 197–214, 2018.

[VdA14] Van der Aalst, Wil M. P.: Data Scientist: The Engineer of the Future. Springer International Publishing, Cham, pp. 13–26, 2014.

Integrated Simulation of Domain-Specific Modeling Languages with Petri Net-based Transformational Semantics

David Mosteller, Michael Haustermann, Daniel Moldt, Dennis Schmitz[1]

Abstract: The development of domain specific models requires appropriate tool support for modeling and execution. Meta-modeling facilitates solutions for the generation of modeling tools from abstract language specifications. The Rmt approach (Renew Meta-Modeling and Transformation) applies transformational semantics using Petri net formalisms as target languages in order to produce quick results for the development of modeling techniques. The problem with transformational approaches is that the inspection of the system during execution is not possible in the original representation. We present a concept for providing simulation feedback for domain specific modeling languages (DSML) that are developed with the Rmt approach on the basis of meta-models and transformational semantics using Petri nets. Details of the application of this new approach are illustrated by some well-known constructs of the Business Process Model and Notation (BPMN). The results summarized in this extended abstract have been published in [Mo19].

Keywords: Meta-Modeling; Petri Nets; Reference Nets; Simulation; Graphical Feedback

1 Challenge of DSMLs: Animation and Simulation

As Meta-Modeling is used to provide new domain-specific modeling languages (DSML) several tools provide tool support to build corresponding models. Bryant et al. identified “the mapping of execution results (e.g., error messages, debugging traces) back into the DSML in a meaningful manner, such that the domain expert using the modeling language understands the result” [Br11, p. 228] as one of the challenges for the translation semantics approach. Concerning the user experience, meaningful visual representation of the domain concepts is vital for the communication between different stakeholders, especially for the domain experts that are often non-software engineers [Ab17, p. 233]. The representation of DSML in execution is still considered a challenge in tool generation in general [MC18, p. 196].

2 The Rmt-Approach

As solution we extend the Renew Meta-Modeling and Transformation (Rmt) framework with a direct simulation of the DSML’s original representation and discuss the integration of the approach. The presented concept for simulation visualization is based on the highlighting

[1] Universität Hamburg, Fachbereich Informatik, `https://www.inf.uni-hamburg.de/inst/ab/art`

of model constructs as graphical feedback. This is achieved by reflecting simulation events of the underlying executed Petri nets (target language) into the DSML (source language). As a special dialect of Petri nets we use reference nets (see [Ku02]) which can be executed by our tool Renew. Several types of mappings are evaluated regarding their expressiveness and features for modeling. A major challenge for the provision of direct simulation support is the integration into model-driven approaches in the sense that the DSML developer can specify the desired representation of the executed models in a model-driven fashion. The concept presented includes tools that enable DSML developers to create the necessary artifacts and configurations to manage this task. We describe the current implementation that processes these artifacts and configurations to initialize a simulation of the DSML model with graphical feedback. Based on these implementations we discuss multiple alternatives to provide support for DSML developers to specify the desired representation of the executed models in the Rmt approach. As a part of our contribution, a generic compiler is implemented in Renew. This is used in the processing of artifacts and configurations. On this basis, the generated technique may be executed within Renew's simulation engine in its original representation. The applicability is demonstrated by providing a solution of simulation / animation for a selected subset of BPMN concepts via our approach.

Based on our concept for providing simulation feedback for DSML that are developed with the Rmt approach on the basis of meta-models and transformational semantics using Petri nets we plan to extend the transformation and variations of the target language in the direction of traditional Petri net formalisms for which analysis can be integrated into the Rmt-approach.

Literaturverzeichnis

[Ab17] Abrahão, Silvia; Bourdeleau, Francis; Cheng, Betty H. C.; Kokaly, Sahar; Paige, Richard F.; Störrle, Harald; Whittle, Jon: User Experience for Model-Driven Engineering: Challenges and Future Directions. In: 20th ACM/IEEE International Conference on Model Driven Engineering Languages and Systems, MODELS 2017, Austin, TX, USA, September 17-22, 2017. IEEE Computer Society, S. 229–236, 2017.

[Br11] Bryant, Barrett R.; Gray, Jeff; Mernik, Marjan; Clarke, Peter J.; France, Robert B.; Karsai, Gabor: Challenges and directions in formalizing the semantics of modeling languages. Comput. Sci. Inf. Syst., 8(2):225–253, 2011.

[Ku02] Kummer, Olaf: Referenznetze. Logos Verlag, Berlin, 2002.

[MC18] Mayerhofer, Tanja; Combemale, Benoit: The Tool Generation Challenge for Executable Domain-Specific Modeling Languages. In (Seidl, Martina; Zschaler, Steffen, Hrsg.): Software Technologies: Applications and Foundations. Springer International Publishing, Cham, S. 193–199, 2018.

[Mo19] Mosteller, David; Haustermann, Michael; Moldt, Daniel; Schmitz, Dennis: Integrated Simulation of Domain-Specific Modeling Languages with Petri Net-Based Transformational Semantics. Transactions on Petri Nets and Other Models of Concurrency, 14:101–125, 2019.

Monotone Precision and Recall Measures for Comparing Executions and Specifications of Dynamic Systems (Extended Abstract)[1]

Artem Polyvyanyy[2], Andreas Solti[3], Matthias Weidlich[4], Claudio Di Ciccio[5], Jan Mendling[6]

1 Motivation

The behavioural comparison of systems is an important concern of software engineering research. For example, the areas of specification discovery and specification mining are concerned with measuring the consistency between a collection of execution traces and a program specification. This problem is also tackled in process mining with the help of measures that describe the quality of a process specification automatically discovered from execution logs.

Technically, behavioural comparisons are often formulated in a *relative* manner, defining a *quotient* of some aspect of one behaviour over the same aspect of another behaviour. For instance, the quotients of the behaviours of a system at different points in time reveal how the system has changed. In process mining, in turn, the quotient of the behaviour of a system as recorded in a log over the behaviour as specified can be used to analyse the trustworthiness of the latter [Ca18]. Yet, defining such quotients is challenging: A recent commentary on measures in process mining identifies a set of intuitive properties and shows that none of the available measures fulfils them [Ta18]. That is, measures lack monotonicity or cannot handle infinite behaviour.

[1] The original article appeared in the ACM Transactions on Software Engineering and Methodology (TOSEM) [Po20]

[2] The University of Melbourne, Australia, artem.polyvyanyy@unimelb.edu.au
[3] Vienna University of Economics and Business, Austria, solti@ai.wu.ac.at
[4] Humboldt-Universität zu Berlin, Germany, matthias.weidlich@hu-berlin.de
[5] Sapienza University of Rome, Italy, diciccio@di.uniroma1.it
[6] Vienna University of Economics and Business, Austria, mendling@ai.wu.ac.at

2 Approach

In this work, we address the problem of behavioural comparison by introducing a new framework for the definition of behavioural quotients. To this end, we approach the problem of behavioural comparison based on the notion of a formal language, which is a suitable starting point to capture the sequential (state-based) behaviour of a dynamic system. Specifically, we address the problem of *how to define meaningful quotients for behavioural comparison of finite and infinite languages.*

Our contributions include a framework for the definition of behavioural quotients that guarantee desired properties. Our framework is then used to instantiate two quotients that are grounded in the cardinality of a language (for finite languages) and the entropy of an automaton (for finite and infinite languages). It was recently shown that our entropy-based measures satisfy all the properties put-forward in the literature for recall and precision measures [STvdA19]. We further show how the proposed quotients can be used to define monotone precision and recall measures between the behaviour as recorded in an execution log of a system and the behaviour captured in a specification of the system.

3 Evaluation

Based on the jBPT library [PW13], we implemented the proposed precision and recall quotients in a tool that is publicly available.[7] Comprehensive experiments with execution logs of real IT systems highlight that existing measures indeed violate the monotonicity property. Moreover, the results of controlled experiments underpin that our quotients enable meaningful conclusions on the relation of the behaviours of two systems. Finally, we explore the scalability of the computation of our measures. The results illustrate the importance of convergence properties on the runtime of the method.

References

[Ca18] Carmona, Josep; Van Dongen, Boudewijn F.; Solti, Adreas; Weidlich, Matthias: Conformance Checking – Relating Processes and Models. Springer, 2018.

[Po20] Polyvyanyy, Artem; Solti, Andreas; Weidlich, Matthias; Ciccio, Claudio Di; Mendling, Jan: Monotone Precision and Recall Measures for Comparing Executions and Specifications of Dynamic Systems. ACM Trans. Softw. Eng. Methodol., 2020. In press.

[PW13] Polyvyanyy, Artem; Weidlich, Matthias: Towards a Compendium of Process Technologies - The jBPT Library for Process Model Analysis. CAiSE'13 Forum. Vol. 998 of CEUR Workshop Proceedings. CEUR-WS.org, pp. 106–113, 2013.

[STvdA19] Syring, Anja F.; Tax, Niek; van der Aalst, Wil M. P.: Evaluating Conformance Measures in Process Mining Using Conformance Propositions. TOPNOC 14:192–221, 2019.

[Ta18] Tax, Niek; Lu, Xixi; Sidorova, Natalia; Fahland, Dirk; van der Aalst, Wil M.P.: The imprecisions of precision measures in process mining. IPL, 135:1–8, 2018.

[7] `https://github.com/jbpt/codebase/tree/master/jbpt-pm`

Quantifying the Re-identication Risk of Event Logs for Process Mining (Extended Abstract)*

Saskia Nuñez von Voigt[1], Stephan A. Fahrenkrog-Petersen[2], Dominik Janssen[3], Agnes Koschmider[3], Florian Tschorsch[1], Felix Mannhardt[4,5], Olaf Landsiedel[3], Matthias Weidlich[2]

Event logs for process mining often contain sensitive information that could be linked to individual process stakeholders by cross-correlating background information, e.g., in an emergency room process, certain events can indicate that a patient is in a certain condition. In general, case attributes can contain various kinds of sensitive data, revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, as well as financial or health information. Likewise, an event log can reveal information about the productivity [PLWF17] or the work schedule of hospital staff. Such kind of staff surveillance is a critical privacy threat. We therefore argue that, when publishing event logs, the risk of such re-identification attacks must be considered. The intention of this extended abstract is to raise awareness to the re-identification risk of event logs and to highlight the importance of privacy-preserving techniques in process mining [MKBW19]. We therefore provide measures to quantify this risk. Specifically, we provide an approach to express the uniqueness of data, which is derived from models that are commonly adopted by process mining techniques. The higher the uniqueness of an event log, the higher an adversary's chances to identify the target. In our case, a targeted re-identification is assumed, i.e., an adversary has information about specific individuals, which includes a subset of the attribute values. Given this background information, the adversary's goal is to reveal sensitive information, e.g., a diagnosis.

Our approach therefore explores the number of cases that are uniquely identifiable by the set of case attributes or the set of event attributes. We use this information to derive a measure of uniqueness for an event log, which serves as a basis for estimating how likely a case can be re-identified. We evaluate a number of so-called projections that can be considered as a kind of data minimization, effectively reducing the potential risk of re-identifying an individual in an event log. Projections refer to a subset of attributes in the event log. For instance, one projection might contain the sequence of all executed activities with their timestamps, while another projection only contains the case attributes. It has been shown that even sparse projections of event logs hold privacy risks [FaAW19].

* The original article is accepted for the 32nd International Conference on Advanced Information Systems Engineering (CAiSE 2020), Grenoble, France, June 8-12, 2020.
[1] Technische Universität Berlin, Germany, firstname.lastname@tu-berlin.de
[2] Humboldt-Universität zu Berlin, Germany, firstname.lastname@hu-berlin.de
[3] Kiel University, Germany, doj|ak|ol@informatik.uni-kiel.de
[4] SINTEF Digital, Trondheim, Norway, felix.mannhardt@sintef.no
[5] NTNU Norwegian University of Science and Technology, Trondheim, Norway

Therefore, in our evaluation, we will consider the re-identification risk for various projections.

To demonstrate the importance of uniqueness considerations for event logs, we conducted a large-scale study with 12 publicly available event logs from the 4TU.Centre for Research Data repository[1]. We categorized the records and assessed the uniqueness where cases refer to a natural person. Our results suggest that an adversary can potentially re-identify up to all of the cases, depending on prior knowledge. We show that an adversary needs only a few attributes of a trace to successfully mount such an attack.

In conclusion, generalization of attributes certainly helps to reduce the re-identification risk [ZBBC17]. Our results, however, show that combining several attributes, such as case attributes and activities, still yields many unique cases. In combination with lowering the resolution of values, e.g., publishing only the year of birth instead of the full birthday, we are able to reduce the re-identification risk. Such generalization techniques can also be applied to timestamps, activities, or case attributes. Along the lines of the data minimization principle, i.e., limiting the amount of personal data, omitting data is simply the most profound way to reduce the risk, which we clearly see when taking our projections into account. Consequently, the projections can be used to reduce the re-identification risk.

This paper shows that we as a community have to act more carefully, though, when releasing event logs, while also highlighting the need to develop privacy-preserving techniques for event logs. We believe that this work will foster the trust and increases the willingness for sharing event logs while providing privacy guarantees.

Literaturverzeichnis

[FaAW19] Fahrenkrog-Petersen, S.A., van der Aa, H., Weidlich, M.: PRETSA: Event Log Sanitization for Privacy-Aware Process Discovery. In: ICPM '19: Proceedings of the International Conference on Process Mining. (2019) 1-8

[MKBW19] Mannhardt, F., Koschmider, A., Baracaldo, N., Weidlich, M., Michael, J.: Privacy-Preserving Process Mining. Business & Information Systems Engineering 61(5) (2019) 595-614

[PLWF17] Pika, A.; Leyer, M.; Wynn, M.T.; Fidge, C.J.; Ter Hofstede, A.H.; van der Aalst, W.M.: Mining Resource Profiles from Event Logs. TMIS '17: Proceedings of ACM Transactions on Management Information Systems 8(1) (2017)

[ZBBC17] Zook, M.; Barocas, S.; Boyd, D.; Crawford, K.; Keller, E.; Gangadharan, S.P.; Pasquale, F.: Ten Simple Rules for Responsible Big Data Research (2017)

[1] https://data.4tu.nl/repository/collection:event_logs_real

Research Challenges for the Modelling of the Resource and Organizational View

Ralf Laue[1], Thomas Bauer[2]

Abstract: One reason for the success of automating business processes by means of business process management systems (BPMS) is that with BPMN, a widely accepted industry standard exists. The availability of a standard together with powerful visual process diagramming tools allows (up to a certain degree) to model processes independently of the runtime environment. However, a closer look reveals that not all relevant information is equally well supported by BPMN. The ISO standard 19439 [In06] differentiates between four views: The function view (process steps and decisions) and the information view (flow of data) are what can be called the core function of BPMN. On the other hand, the resource view (describing human as well as technical resources) and the organizational view (describing responsibilities and authorities) are to a large degree outside the scope of the BPMN standard.

Nevertheless, for describing the execution of business processes and for automating them, the latter two views have to be considered as well. In our work, we investigated the state of research on modelling the resource and organizational view. We found that while there is quite a lot of research on this topic, some requirements of practical relevance are still not yet fully covered. Furthermore, we examined how information on resource and organizational view has to be modelled in four current BPMS. We found that all four tools provided basic support for this purpose, but did not cover all requirements, for example with respect to more complex substitution plans.

Keywords: business process management system, workflow system, resource perspective, resource modelling

1 Motivation

When the execution of business processes shall be supported at run-time by business process management systems, it is necessary to define to whom manual activities have to be assigned. In exceptional cases, it must be possible to delegate the execution of a task. Furthermore, BPMS should be able to monitor the execution of tasks and to escalate them if necessary (e. g., when they have not been completed after a given time). In case of an absence of a responsible person, substitution plans should be applied.

[1] University of Applied Sciences Zwickau, Germany, Dr.-Friedrichs-Ring 2a, 08056 Zwickau, Germany ralf.laue@fh-zwickau.de

[2] University of Applied Sciences Neu-Ulm, Wileystr. 1, 89231 Neu-Ulm, Germany thomas.bauer@hs-neu-ulm.de

While thanks to the BPMN standard, no programming knowledge is necessary for defining the flow of control for a business process, we observed that this is not the case for defining the mentioned information on resources. In current BPMS, inclusion of such resource-related information is tool-specific. For instance, to define sophisticated actor assignments, as required in many scenarios from practice, languages as JavaScript or XPath have to be used. It is the aim of our ongoing work to get an overview on existing work on the topic, to compare it with requirements from practice and to identify open research challenges.

2 Findings and Further Questions

It has to be stated that there is no lack of work on the resource view. Remarkable work include the seminal work on workflow resource patterns [Ru05], suggestions for adding the resource perspective to BPMN [SCV15] and even a graphical language for defining resource requirements and constraints [Ca15]. Anyway, certain details such as substitution plans can turn out as more complicated as it might seem at a first glance.

On the other hand, an examination of four current BPMS (Bizagi Studio, IBM Business Process Designer, K2 Cloud and Signavio Workflow Accelerator) revealed that none of them was powerful enough to fulfil all our requirements on flexibility and ease for defining resource assignments. In addition, there is currently no standard that would allow the exchange of business process models with resource information between different tools.

In a next step, we want to compare the existing work on meta-models for the resource and organizational view. In particular, it is an open question whether the requirements on resource allocation is domain-independent. Further outstanding issues include the understandability of different formalisms for specifying resource-related requirements, a concept to map such formalisms to directory services such as LDAP, and concepts that are not yet sufficiently covered (in particular, complex escalation rules and substitution plans).

References

[Ca15] Cabanillas, Cristina; Knuplesch, David; Resinas, Manuel; Reichert, Manfred; Mendling, Jan; Cortés, Antonio Ruiz: RALph: A Graphical Notation for Resource Assignments in Business Processes. In: Advanced Information Systems Engineering, 2015. pp. 53–68, 2015.

[In06] International Organization for Standardization: ISO 19439:2006: Enterprise integration - Framework for enterprise modelling. 2006.

[Ru05] Russell, Nick; van der Aalst, Wil M. P.; ter Hofstede, Arthur H. M.; Edmond, David: Workflow Resource Patterns: Identification, Representation and Tool Support. In: Advanced Information Systems Engineering, 2005. pp. 216–232, 2005.

[SCV15] Stroppi, Luis Jess Ramhiotti, Omar; Villarreal, Pablo David: Defining the resource perspective in the development of processes-aware information systems. Information and Software Technology, 59:86 – 108, 2015.

The Diverse Ends-in-View of Research on Business Processes and Organizational Routines (Extended Abstract)

Jan Mendling[1], Nicholas Berente[2], Stefan Seidel[3] and Thomas Grisold[3]

Extended Abstract

A recent debate in the field of information systems research has raised the question whether the arguments against method articulated by Paul Feyerabend [Fey93] offers a basis for reflecting the methodological foundations of the field [Tre18]. In this research, we build on our recent commentary [MBSG20] that contributes to this debate. More specifically, we reflect on Habermas's pragmatist perspective of social science. We argue that research on business processes and organizational routines exemplifies a pluralism that goes beyond the relativistic conclusions of Feyerabend. Research into business processes and organizational routines exhibits a healthy diversity of epistemological and methodological approaches. Accompanying this diversity is an openness to novelty and change. Yet, at the same time, this does not necessitate the abandonment of rigor and a cumulative tradition implied by Feyerabend's "anything goes." There is not a singular, hegemonic approach to what constitutes strong business process and organizational routines research, but neither have we devolved into anarchy.

Bibliography

[Fey93] Feyerabend, P. (1993): Against method: Verso.

[Men20] Mendling, J., Berente, N., Seidel, S., Grisold, T. (2020): Pluralism and Pragmatism in the Information Systems Field: The Case of Research on Business Processes and Organizational Routines. In *ACM SIGMIS Database: the DATABASE for Advances in Information Systems*. In press.

[Tre18] Treiblmaier, H. (2018): The Philosopher's Corner: Paul Feyerabend and the Art of Epistemological Anarchy-A Discussion of the Basic Tenets of Against Method and an Assessment of Their Potential Usefulness for the Information Systems Field. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems, 49*(1), 93-101

[1] Wirtschaftsuniversität Wien, Department of Information Systems and Operations, Welthandelsplatz 1, 1020 Wien, Austria, jan.mendling@wu.ac.at

[2] University of Notre Dame, Mendoza College of Business, Notre Dame, IN 46556, United States, nberente@hd.edu

[3] Universität Liechtenstein, Institut für Wirtschaftsinformatik, Fürst-Franz-Josef-Strasse, 9490 Vaduz, Liechtenstein, firstname.lastname@uni.li

Towards Privacy-Preserving IoT Systems Using Model Driven Engineering (Extended Abstract)

Judith Michael[1], Lukas Netz[1], Bernhard Rumpe[1], Simon Varga[1]

Abstract: Collecting, storing and processing data from machines, processes and workers in production processes is increasing with technological possibilities and the availability of sensors. This raises the challenge of ensuring privacy for personal data within this context. For MDE approaches it is important to consider privacy already at the model level. This paper discusses a way to create privacy-preserving IoT systems using an MDE approach to support privacy and data transparency. We show the relevance and application on a use case from industrial production processes. Additionally, we discuss abilities for practical realization and its limitations. The work summarized in this extended abstract has been presented at the MDE4IoT workshop (MODELS 2019) and is published by CEUR-WS.org [Mi19b].

Keywords: Domain-Specific Languages, Generated Enterprise Information Systems, Information Portals, Internet of Things, Model-Based Software Engineering, Privacy-By-Design, Privacy Modeling

1 Main Findings

Problem Description. The rise of wearable technologies makes it possible to equip workers, products and machines in production processes with miniaturized sensors. This development goes hand in hand with questions about the informational self-determination, the security of data collected as well as data protection and transparency. This paper shows how it is possible to include privacy considerations in the MDE development process already on model level. The foundation for this work lies in the privacy-design strategies [CHH16] as well as an idea to include privacy checkpoints into system architectures. In preceding publications [Ma19; Mi19a], we have already taken these ideas into account. This paper goes a step further and discusses them in relation with model- driven engineering (MDE).

Use Case. The use case shows a part of a production process, so one station in a manufacturing area with several operators and robots collaborating with each other. Operators are wearing smart glasses and watches and are using smartphones and tablets which collect sensory information. Further sensory information is collected on products and related machines and robots. Assistive systems, which use such data ,e.g., to do ergonomic analyses to improve the ergonomic intervention processes have to follow security and privacy regulations and should provide informational self-determination facilities to inform operators what happens with

[1] Software Engineering, RWTH Aachen, Germany, `www.se-rwth.de`, {michael,netz,rumpe,varga}@se-rwth.de

the data. Thus, besides data collection, storage, removal as well as primary and secondary use, we have included an information portal into our approach which ensures privacy control and provides information. Operators can give their consent for data usage, withdraw it and get options to delete the data at any point. Privacy checkpoints within the components of the system architecture can help to ensure a privacy-preserving system design.

Results. The approach presented in this paper uses MDE tools and frameworks together with a set of domain-specific languages (DSLs) to create an Enterprise Information System (EIS). The EIS is considering privacy-preservation and makes the relevant information available for users and data providers to allow for informed decisions about their data use. Starting with structural information in the domain model, which is needed for MDE approaches, we add privacy concepts (privacy model) to support the execution of the privacy checkpoints. At the instance level, a purpose tree has to be defined which is used in concrete rules attached to a privacy policy instance. The purpose tree defines hierarchical relations between purposes as well as connected attributes and classes, which are needed for this purpose.

Employees define their privacy policy rules and can allow the data controller to give access to their data based on these rules. They can see decisions for data access based on the automatic comparison of privacy policies between data providers and data consumers. The data controller can add and change concrete purposes for data collection and storage. Employees are informed about changes and get possibilities to define new privacy policy rules or change existing ones. The traceability of data is ensured in the system architecture by considering all checkpoints. In our approach we demonstrate how to enable engineers to define software with model-driven approaches which meet the growing requirements for data protection and transparency.

References

[CHH16] Colesky, M.; Hoepman, J.; Hillen, C.: A Critical Analysis of Privacy Design Strategies. In: IEEE Security and Privacy Workshops (SPW). Pp. 33–40, 2016.

[Ma19] Mannhardt, F.; Koschmider, A.; Baracaldo, N.; Weidlich, M.; Michael, J.: Privacy-Preserving Process Mining: Differential Privacy for Event Logs. Business & Information Systems Engineering (BISE) 61/5, pp. 595–614, 2019, ISSN: 1867-0202.

[Mi19a] Michael, J.; Koschmider, A.; Mannhardt, F.; Baracaldo, N.; Rumpe, B.: User-Centered and Privacy-Driven Process Mining System Design for IoT. In: Information Systems Engineering in Responsible Information Systems. Vol. 350, LNBIP, Springer, pp. 194–206, 2019, ISBN: 978-3-030-21296-4.

[Mi19b] Michael, J.; Netz, L.; Rumpe, B.; Varga, S.: Towards Privacy-Preserving IoT Systems Using Model Driven Engineering. In: MDE4IoT & ModComp Workshops 2019. Vol. 2442, CEUR-WS.org, pp. 15–22, Sept. 2019.

Using Action Design Research for Co-creating Service-Dominant Business Artifacts between Academia and Industry

Egon Lüftenegger[0000−0003−0998−5814]
egon@businessmodelradar.com

BusinessModelRadar.com, Salzburg, Austria

Abstract. Our project started as an innovation initiative for shaping the future of an international Dutch conglomerate in the financial services industry. This endeavor took the shape of a collaboration between academics and practitioners that evolved from unique goals for industry as fast results and real scientific results towards a collaborative approach. We formalized this collaboration by using action design research (ADR) for achieving impact within the company while generating new knowledge. The use of ADR resulted in the co-creation of artifacts that brought mutual benefits, resulting in a win-win situation for the academia and industry. On the one hand, academics were able to develop a framework and its underlying artifacts for service-dominant business design and engineering. On the other hand, the framework helped to achieve organizational change for shifting from an asset-dominant business towards a service-dominant one.

Keywords: Action Design Research · Service-Dominant Logic · Service-Dominant Business Model Radar · Service-Dominant Business Engineering

1 Introduction

Collaboration between academics and practitioners is becoming increasingly important because this interaction can generate reciprocal benefits for all parties involved [25]. Practitioners from industry obtain access to scientific competence, and scientists, in exchange, learn about the industry's needs and interests [8]. From our experience, in our industry-academy collaboration, there are expectation challenges between practitioners and researchers: On the practitioner's side, there is an urgency of rapid results for achieving a return on investment within the company. On the academic side, there is a goal of delivering academic results by performing research that requires time and effort. Therefore, there is a challenge to balance both sides of the project collaboration: Achieving impact within the company, with the generation of new academic knowledge. However, balance is usually not well established from the beginning of the project, specifically, for a long-term PhD project like ours. In particular, for satisfying the needs of a Dutch conglomerate in financial services for researching about growing opportunities by the design of new business models driven by information systems [9].

Our project, named originally as CoProFind (contract-based process outsourcing in the Financial Services industry), was initiated by a former chief executive officer for facing the future challenges driven by the rapid market changes. A committee conformed by academic researchers of the university and executives of the Dutch conglomerate recruited a PhD candidate with professional experience with the ability to think outside-the-box. Hence, the project team was conformed and started the journey of our project.

In this paper, we further explore our collaboration challenges, experience, and gains. In particular, we discuss our shift from researcher-driven design research towards a collaborative action design research approach. With this research approach, we were able to produce a framework that achieved organizational change while generating academic knowledge.

This paper is structured as follows: In Section 2, we explain how the collaboration took form. In Section 3, we discuss our collaboration challenges and learned lessons. In Section 4, we describe the criteria for selecting the research method. In Section 5, we describe good practices of the selected collaborative research method: Action Design Research. In Section 6, we discuss the result of applying the collaborative research approach in the Dutch conglomerate. Finally, we end this paper with conclusions.

2 How the Collaboration took Form

During the first year, the PhD candidate researched the solution space: Innovations within financial services. In this first step, the PhD candidate identified the importance of business model innovation and brought the company the use of visual representations for communicating business models [10].

As the next step, we developed a framework for establishing business innovation directions (BID). As presented in [17], the BID framework has four dimensions: Logic, openness, competitiveness, and newness. From these four dimensions, we focus our attention in this paper on the logic dimension. This logic dimension has two possible values: Goods-dominant or service-dominant. A goods-dominant logic implies a traditional focus on products and value chains with a manufacturer mindset. In this goods-dominant logic [24], each link of the value chain adds value for producing products and offering adding value services for such products. This logic has a focus on value-in-ownership, and the consumer plays a passive role by being a receptor of a product produced by the firm. The service-dominant logic focuses on solutions: the value-in-use instead of value-in-ownership. This value-in-use is co-created within a value network or ecosystem. In this logic, all the business actors participate actively in the co-creation process [24].

In the information systems (IS) group from the Dutch University, we designed business models by following a goods-dominant perspective that is heavily asset oriented like FLAME. We also designed business models by following a service-dominant perspective, like Servestment. The PhD candidate led the development of the Servestment business concept, a crowd-funding finance platform for service providers by selling tokens for service delivery [11]. The practitioners from the company decided that business concepts that followed the service-logic perspective were the most innovative ones. Furthermore, they manifested interest in designing a business by following this way of thinking: The practice inspired problem.

3 Challenges and Learned Lessons

During our project, we faced challenges in the co-operation between academics and practitioners. We conceptualize them as follows:

Consultant trap. There is a conflict between rapid results and methodological academic rigors required for academic contributions. Practitioners want rapid results like a consultancy, resulting in a challenging environment.

High revenue trap. Financial success for a high-income earner conglomerate can slow the progress down due to the lack of urgency on developing new concepts as a result of increasing financial performance.

Artifact acceptance. Balancing usable and understandable artifacts driven by novel theory, new jargon, and new ways of doing things.

Due to the challenging collaboration, we were able to learn valuable lessons. We describe them as follows:

Identify the practitioner's needs. Do not start with solutions before knowing the problem. We started the project with e-contracting technologies. However, after we starting working together with practitioners, we were able to find the real needs for bringing benefits to the company and the university.

Inclusive instead of exclusive. Include the industry side in the artifact development process. The collaboration was successful because the interaction between academics and practitioners helped to reduce the knowledge gap between them. Academics gained insights from practitioners, and practitioners gained state-of-the-art knowledge from academics.

Associate with industry partners. Middle management can offer access to resources on the industry's side. An innovator manager was essential in the company side for facilitating workshops, giving feedback, and spreading the new artifacts within the organization. Usually, senior managers have less time and energy to focus on non-urgent tasks like a research project.

Artifact co-creation. The project meetings and workshops enabled practitioners to collaborate. They brought valuable insights and to establish a reality check of the tool in the real business environment.

Convince with evidence. Proofs for following a specific research direction are essential for practitioners and academics. Proofs are particularly useful when the PhD candidate is new to the organizations involved in the project or even new in the country. The attendance to a workshop at the University of Cambridge helped to convince the practitioners and then academics to solve the practice-inspired business design problem by developing the service-dominant business framework.

Keep it simple but complete. Artifacts can not be too complicated or too simple. Finding the right balance is achieved with the interaction between academic practitioners and end-users.

Inside-out. The PhD candidate worked as a staff member within the university and also spent time working within the conglomerate. This configuration helped to bring academics and practitioners closer and also influenced on reducing the knowledge gap by constant interaction and exchange.

4 Finding the Right Balance between Action and Design Science Research

In our project, we started with design science research (DSR): A well-known research method for developing artifacts grounded in academic theory. However, in DSR, the involvement of end-users occurs only during the evaluation, once the artifact is already wholly developed [7].

We experienced that by using DSR alone, our artifacts were not well received. Then, the PhD candidate looked at Action Research (AR)[3] for making an impact within the company [12]. AR [4] is an iterative research method where researchers intervene in the real world to solve practitioners' problems and to gain scientific knowledge [2]. This research method is usually performed as an iterative process and combines theory generation with researcher intervention for solving an immediate organization problem [27]. However, it lacks the artifact development process from DSR.

Therefore, we ask ourselves the following question: How can we construct artifacts with organizational impact and academic quality? Our collaboration shifted from researcher only driven artifact development towards a collaborative approach between academics and practitioners. The collaboration between academics and researchers required the adoption of a new research approach within the IS group. Our research aimed to achieve the dual goal of creating academic knowledge and solving practitioners' problems. AR was a method taught in the Netherlands research school for information and knowledge systems (SIKS), where the PhD candidate attended. However, the PhD candidate found in the IS literature a new kind of research method not taught at SIKS called to combine DS with AR: Action Design Research (ADR) [27]. ADR [27] has emerged as a new design research method that combines DSR with AR to focus clearly on artifact development while taking into consideration user participation and feedback during the experimentation.

We applied ADR successfully in our research project because the method combines design research with action research for achieving a dual goal of organizational impact and knowledge generation: Mutual benefits for practitioners and researchers. This co-creative approach to performing research balanced the development of new artifacts with an organizational impact: A shift towards a service-dominant business, from a goods-dominant one.

5 Action Design Research Best Practices for Co-creating Artifacts between Academia and Industry

In this section, we present good practices for performing an action design research (ADR) process. We describe the good practices of the ADR process as the following stages [27]:

ADR Stage 1: Problem formulation. In this stage, we identify and conceptualize a research opportunity based on existing theories. Two principles drive this stage: Principle 1, practice-inspired research; and, principle 2, Theory-ingrained artifact. The former emphasizes viewing problems as knowledge-creation opportunities. The later, emphasize that theories inform artifacts that are created and evaluated within ADR.

ADR Stage 2: Building, intervention, and evaluation (BIE). In this stage, we use the problem and theoretical foundation for artifact development. Three principles drive this stage: Principle 3, reciprocal shaping: states that an ADR team formed by academics and practitioners engage in the artifact iterative process. Principle 4, mutually influential roles: Stress the importance of mutual learning from the participants within the ADR process. Finally, principle 5, authentic and concurrent evaluation: Emphasizes that evaluation is not a separate stage of the research process that follows building. In ADR, the artifact development process is iterative: First, we present a researcher-driven version to the practitioners: The alpha version. This one or more alpha versions are formative for refining the artifact. Then, the practitioners contribute to feedback. The captured feedback is processed by academics, resulting in one or more beta versions of the artifact. We use the beta versions with end-users in workshop settings. In this beta version, we asses the value and utility of the outcomes.

ADR Stage 3: Reflection and learning. In this stage, we reflect on the development process from building a particular solution to a broader class of problems. The resulting artifact, also known as the ensemble, will reflect not only the original design but also the practitioner's perspectives within the organizational use. This stage works in parallel with Stage 1 and Stage 2.

ADR Stage 4: Formalization of learning. In this stage, we formalize the outcome as a tool for solving a class of problems. Generalized outcomes drive this stage: Principle 7. The resulting artifact or ensemble is, by definition, a solution to address a problem that can be generalized.

6 Using ADR for Developing the Framework and its Underlying Artifacts

In this section, we describe how we applied ADR for developing the framework artifact and two underlying artifacts: The service-dominant strategy canvas and service-dominant business model radar (In a few words, the business model radar).

6.1 Service-dominant Business Framework as the Overall View Artifact

By following ADR stage 1, we started the development of the service-dominant business framework for solving a practical problem: How to design businesses by following a service-dominant mindset. By following the ADR stage 1, the PhD candidate identified the theory behind: The link between strategy, business models, and business processes [26] [1].

The PhD candidate proposed the development of a service-logic driven framework as a foundation for his PhD thesis. However, the practitioners required confirmation of the service trend for going in this direction. The PhD candidate proposed to attend a service design and innovation workshop at the Institute for Manufacturing (IfM) at the University of Cambridge. The company showed interest in the research direction by sending an innovation manager to join the workshop with the PhD candidate. The innovation manager confirmed in practice the research direction on the service-dominant logic established by the PhD candidate, and the executives were eager to continue in this direction. As a result, the innovation manager supported the PhD candidate for developing a framework following a service-logic: The innovation manager and the PhD candidate sent a memo to the upper management for applying this framework in the company [5]. Then, due to the interest of the company in this direction, the IS group allowed the development of the framework as PhD thesis [14].

In Figure 2, we present the evolution of the iterative construction process that shaped the artifact. By following ADR stage 2, the framework evolved from an alpha version proposed by the PhD candidate [12] to a beta version (first presented in [20]) that included the feedback from academics and practitioners [6], [14]. The original alpha framework shown in Figure 1, includes from top

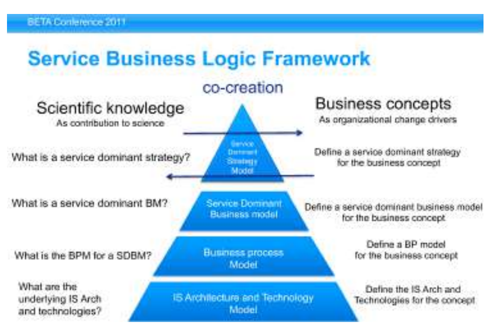


Fig. 1. The service business Logic framework [12]: The alpha service-dominant business. Proposed to the Dutch conglomerate in [5].

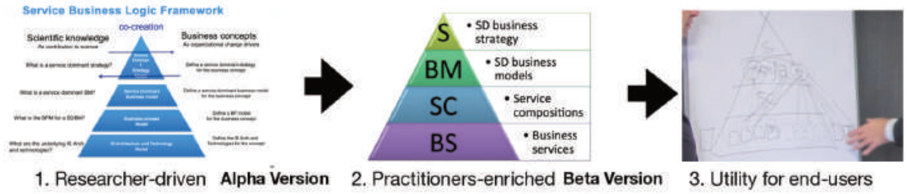


Fig. 2. Service-dominant business framework development with action design research

to bottom: Service-dominant strategy model, service-dominant business model, business process model, and IS architecture and technology model. The PhD candidate defined the last layer at that time concerning the company's desire to implement business models. As shown in Figure 3, the artifact evolved into a business-only because our users were business executives: The beta version. This version was a result of ADR stage 3 by reflecting on the evolving practitioners' business needs. Hence, this updated the design principles and resulted in the service-dominant business framework with strategy formulation, business model design, and business process compositions with business services. In Figure 2, in the end, we show how business-oriented practitioners interacted with the framework during one of our workshops within the Dutch conglomerate.

As suggested by ADR stage 4, the framework can solve a class of problems in designing service-dominant business models. For instance, for modelling business adopting Industry 4.0 [15]. Furthermore, the framework guided the development of the underlying artifacts. In the service-dominant strategy layer, we have the service-dominant strategy canvas. In the service-dominant business model layer, we have the business model radar. At the bottom two layers, the service composition and business services, we used concepts with less novelty. At the business services layer, we used the concept of a business service catalog inspired in the service-oriented architecture and applied at the business level. At the business service composition layer, we use the concepts of service blueprints and business processes for illustrating the idea of using a business service catalog.

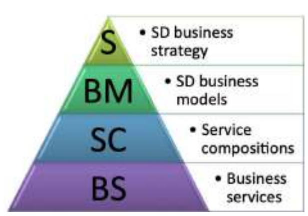


Fig. 3. Service-Dominant Business Framework [20]

6.2 Service-dominant Strategy Artifact

The first practical problem was how to formulate a service-dominant strategy. The PhD candidate reviewed the literature regarding the service-dominant logic theory and its strategic development [13]. This theory ingrained the design of an academic version of the strategy canvas: The alpha artifact [21].

As stated by ADR stage 2, we used the alpha version with the practitioners [22]. However, the jargon was too complicated. For instance, we used categories such as exogenous and endogenous. Then, we constructed the practitioner a user-focused version: The beta version [19]. In Figure 6, we show the service-dominant strategy canvas artifact: from the alpha version to the utility for the users in the last version. We tested the beta version in a workshop setting with executives from the conglomerate. During the workshop, we used an interactive approach with sticky notes and a poster version of the tool for enabling the collaboration. At first, when we started to use the elements, the executives were not too collaborative. However, after ten minutes with the tool, we were able to interact with the practitioners. By following ADR stage 3, we reflected on the workshop experience. As a lesson, we learned that the interactive poster with sticky notes approaches worked well with end-users, and we decided to use this approach with the remaining artifacts.

By following ADR stage 4, we have the service-dominant strategy canvas with four categories as a generalized outcome: The value-in-use, the service ecosystem, and collaboration management. From the first category, the executives shifted their thinking from car leasing towards mobility solutions. From the second category, the executives defined their role within an ecosystem of service partners: The orchestrator. Finally, they identified the kind of partnership for playing this role. By following ADR stage 4, we can generalize the outcomes of the tool. Practitioners can not only design an orchestrating strategy but also identify other roles that could lead to different types of business models

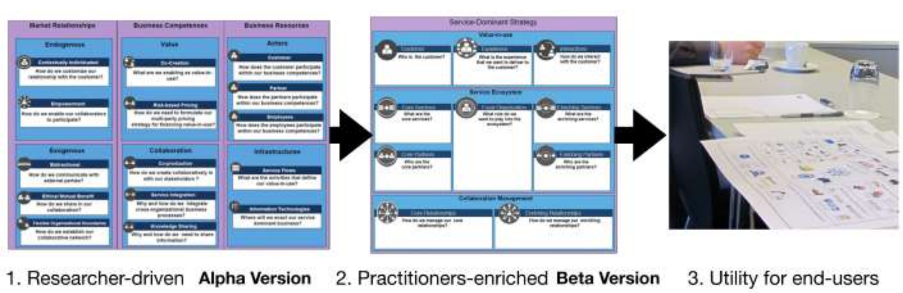


Fig. 4. service-dominant strategy canvas artifact evolution: From the researchers driven Alpha version to the utility for end-users in a workshop session from the practitioners-enriched Beta version

6.3 The Service-dominant Business Model Artifact: The Business Model Radar

As stated by ADR stage 1, a practical problem drove the research: The need for designing solution-based business models derived from the adoption of a service-dominant strategy. In particular, as we identified in the strategic workshop session: The design of a mobility business model by playing the role of orchestrator. The PhD candidate identified the theory for developing the service-dominant business model artifact: The business model canvas that contains the elements of a traditional business model and the service-dominant strategy that contains the elements of a service-dominant mindset [20], [14].

By following ADR stage 2, the PhD candidate developed the first business model radar (BMR) artifact by confronting the elements of the business model canvas with the elements of the service-dominant logic [20], [14]: the alpha artifact. We discussed the version in working meetings with the ADR team. However, the first goal was to test the circular shape of the service-dominant business model tool. The BMR has a circular shape for emphasizing the co-creation process due to the adoption of a value network structure. Once the ADR team accepted the PhD candidate argument of doing a circular-based representation on the tool, he produced a second alpha version by improving the confrontation process between the theoretical elements [18].

During our work on the BMR by following ADR Stage 2 (BIE), the practitioners influenced on the practicality and usability of the business model tool and the academics with the theory and the artifact: Achieving mutually influential roles (principle 4). The BMR alpha artifact versions were evaluated internally within the ADR team by following the principle 3 (reciprocal shaping) and then with the beta version we tested with a broader audience in a workshop setting [14]. In Figure 5, we present the evolution of BMR as an iterative development process: From the alpha version to the utility for the users in the last version.

As shown in Figure 5, the complexity increased as the BMR evolves. The increase in complexity is explained by discussing with practitioners and by updating our design principles. At first, there was not an explicit separation between costs and benefits because one senior manager insisted that he does not care about the benefits of the other parties involved in the business model. However,

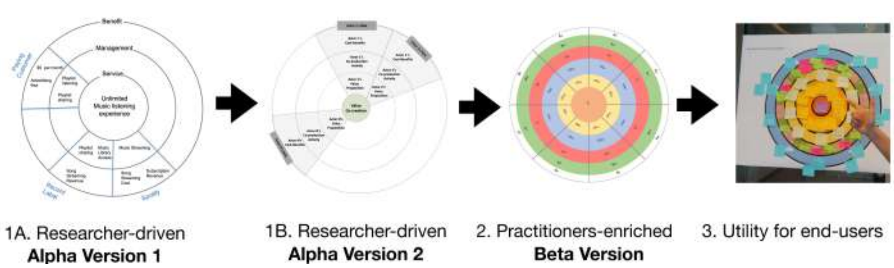


Fig. 5. ADR development process for the service-dominant business model radar

this was a particular aspect that hinders the artifact as a solution for a class of problems.

By following ADR stage 3 on reflection and learning, the PhD candidate reflected on this issue and established a beta version of the BMR by including an explicit separation of cost-benefits to emphasize the business model aspect of the BMR. In this way, the end-users have to think about the costs and the benefits for each party involved. During the workshops, the BMR tool achieved the goal of designing a business model within the conglomerate: A mobility orchestrator business model. For instance, during a workshop with end-users from the conglomerate, a participant expressed "it provides a bird's eye view on the designed business model".

As stated by ADR stage 4, we formalize the outcome: A service-dominant business model is the reframing of the business model concept by following a service-dominant strategy. This concept takes shape as a conceptual modeling tool for business model design: The business model radar. The business model radar takes a value network organizational structure where each co-creation actor contributes to the overall solution with value propositions. For delivering a value proposition, each actor must perform a co-creation activity. By participating in the business model, a co-creation actor can incur in costs and gain benefits.

Practitioners and academics can use the BMR for designing business models as ecosystems beyond mobility business models [16]. The former PhD candidate has tested this in business model innovation workshops in Austria and international lectures about business models with students from universities from Finland, Germany, Indonesia, and Austria.

7 Conclusions

The main result of the project is the service-dominant business design framework. This framework acted as the first dimension of the BASE/X (Business Agility through Service Engineering in an eXtended enterprise) framework [6]. The BASE/X is currently active research in the IS group. The framework covers the spectrum of the formulation of service-dominant business strategies to the process-based execution of business services.

Our project produced academic impact by the use of the framework and the tools. For instance, the service-dominant business model radar, developed in [14], [20] and [20] and applied in mobility solutions [20], has been used in a mobility project conducted by the IS group at the School of Industrial Engineering in Eindhoven University of Technology: C-MobILE (Accelerating C-ITS Mobility Innovation and depLoyment in Europe). The BMR has been also applied in regional projects in Austria [23] and interregional projects between Austria and Slovenia [15], [16]. Regarding the business impact, the organization shifted towards financial solutions by including consultancy services on servitization and a mobility business model.

The ADR method proved in our project to be the right research approach for collaborating with the industry by co-creating solutions rather than just

Fig. 6. BASE/X (Business Agility through Service Engineering in an eXtended enterprise) [6]

delivering them. By including the practitioner in the research, we were able to minimize the gaps in domain-specific knowledge. Furthermore, the company executives accepted and used the resulting artifact for defining a new direction of doing business.

Acknowledgements. ADR Team that participated in the framework and underlying artifacts development process: Egon Lüftenegger (former PhD candidate), Marco Comuzzi (former co-promotor), Paul Grefen (former promotor), and, Caren Weisleder (former innovation manager within the conglomerate).

References

1. Al-Debei, M., Avison, D.: Developing a unified framework of the business model concept. European Journal of Information Systems **19**(3), 359–376 (2010)
2. Avison, D., Lau, F., Myers, M., Nielsen, P.: Action research. Communications of the ACM (1999)
3. Baskerville, R., Wood-Harper, A.T.: Diversity in information systems action research methods. European Journal of information systems **7**(2), 90–107 (1998)
4. Coughlan, P., Coghlan, D.: Action research for operations management. International Journal of Operations & Production Management **22**(2), 220–240 (2002)
5. C.Weisleder, Lüftenegger, E.: Internal memo: Service-dominant framework proposal. Tech. rep., DLL Financial Solutions (2010)
6. Grefen, P., Lüftenegger, E., Linden, van der, E., Weisleder, C.: BASE/X business agility through cross-organizational service engineering : the business and service design approach developed in the CoProFind project. BETA publicatie : working papers, Technische Universiteit Eindhoven (2013)
7. Hevner, A., March, S., Park, J., Ram, S.: Design science in information systems research. MIS Quarterly **28**(1), pp. 75–105 (2004)
8. Kelli, A., Mets, T., Jonsson, L., Pisuke, H., Adamsoo, R.: The changing approach in academia-industry collaboration: from profit orientation to innovation support. Trames: A Journal of the Humanities and Social Sciences **17**(3), 215 (2013)
9. Lüftenegger, E., Angelov, S., van der Linden, E., Grefen, P.: Coprofind: Contract based process outsourcing in the financial industry new business structures. Poster presented at 4th SIKS/BENAIS Conference on Enterprise Information Systems, The Netherlands, Nijmegen (2009). https://doi.org/2010.13140/RG.2.2.13414.98880

10. Lüftenegger, E., Angelov, S., van der Linden, E., Grefen, P.: The state of the art of innovation-driven business models in the financial services industry. BETA publicatie : working papers; Vol. 310, Technische Universiteit Eindhoven (2010)
11. Lüftenegger, E., Zhao, X., Angelov, S., Grefen, P.: Servestment: A service oriented business model for the financial service industry. CoProFind Deliverable 005 (2010)
12. Lüftenegger, E.: Service logic framework. Presentation at the Beta Research Conference (2011). https://doi.org/10.13140/RG.2.2.20388.01923
13. Lüftenegger, E.: Service strategy. Tech. rep., Eindhoven University of Technology and De Lage Landen (2011). https://doi.org/10.13140/RG.2.2.13257.70245/1
14. Lüftenegger, E.: Service-dominant business design. Ph.D. thesis, Eindhoven: Technische Universiteit Eindhoven (2014). https://doi.org/10.6100/IR774591
15. Lüftenegger, E.: Management-tools in smart service engineering für industrie 4.0. In: Intelligente Produktion: intelligente Produktion - Management- und Mitarbeiteraspekte, pp. 208–230. Harlow: Pearson ; Maribor : Univerza v Mariboru, Ekonomsko-poslovna fakulteta (2019)
16. Lüftenegger, E.: Servitization und industrie 4.0: Neue geschäftsmodelle in der service-dominant industrie 4.0. In: Intelligente Produktion: intelligente Produktion - Management- und Mitarbeiteraspekte, pp. 255–277. Harlow: Pearson ; Maribor : Univerza v Mariboru, Ekonomsko-poslovna fakulteta (2019)
17. Lüftenegger, E., Angelov, S., Grefen, P.: A framework for business innovation directions. BETA Working Paper Series 351 (2011)
18. Lüftenegger, E., Comuzzi, M., Grefen, P.: The service-dominant ecosystem: Mapping a service dominant strategy to a product-service ecosystem. In: Collaborative Systems for Reindustrialization, pp. 22–30. Springer Berlin Heidelberg (2013). https://doi.org/10.1007/978-3-642-40543-3_3
19. Lüftenegger, E., Comuzzi, M., Grefen, P.: Designing a tool for service-dominant strategies using action design research. Service Business **11**(1), 161–189 (2017). https://doi.org/10.1007/s11628-015-0297-7
20. Lüftenegger, E., Comuzzi, M., Grefen, P., Weisleder, C.: The Service Dominant Business Model: A Service Focused Conceptualization. BETA publicatie : working papers; vol. 402, Technische Universiteit Eindhoven (2013)
21. Lüftenegger, E., Grefen, P., Weisleder, C.: The service dominant strategy canvas: defining and visualizing a service dominant strategy through the traditional strategic lens, BETA publicatie : working papers; vol. 383, vol. 383. Technische Universiteit Eindhoven (2012)
22. Lüftenegger, E., Grefen, P., Weisleder, C.: The service dominant strategy canvas: Towards networked business models. In: Collaborative Networks in the Internet of Services, pp. 207–215. Springer Berlin Heidelberg (2012). https://doi.org/10.1007/978-3-642-32775-9_21
23. Lüftenegger, E., Softic, S.: Service-dominant business model financial validation: Cost-benefit analysis with business processes and service-dominant business models. In: Proceedings of 30th Central European Conference on Information and Intelligent Systems (CECIIS 2019). pp. 161–172. University of Zagreb (2019)
24. Lusch, R., Vargo, S.: Service-Dominant Logic: Premises, Perspectives, Possibilities. Cambridge University Press, 1st edn. (2014)
25. Muscio, A.: What drives the university use of technology transfer offices? evidence from italy. The Journal of Technology Transfer **35**(2), 181–202 (2010)
26. Osterwalder, A.: The Business Model Ontology: A Proposition in a Design Science Aproach. Ph.D. thesis, Universite de Lausanne (2004)
27. Sein, M.K., Henfridsson, O., Purao, S., Rossi, M., Lindgren, R.: Action design research. MIS quarterly pp. 37–56 (2011)

Predictive Process Monitoring: A Use-Case-Driven Literature Review

Florian Spree

Technical University - Sofia Bulgaria
Florian.spree@t-online.de
https://orcid.org/0000-0003-1071-7948

Abstract. Predictive process monitoring is the subject of a growing interest in academic research and industry. As a result, an increased number of papers on this topic have been published. Due to the high complexity in this research area, a poor comparability is given. Several researchers have already tackled this issue by providing different academic categorizations. However, it seems that the needs of the industry are not considered. The situation makes it difficult to identify relevant papers and thus possibilities for collaboration. Therefore, this paper contributes to the knowledge domain by developing a taxonomy of three identified business use cases. First, a wide-ranging systematic literature review identifies research papers in this area. Then a use-case-driven taxonomy is proposed to establish an efficient and fast framework. Finally, the identified papers get categorized based on the meta-data and by reading the full text of each paper. Hence, the papers data will support practitioners and researchers in identifying relevant papers based on use cases.

Keywords: Business Process Management, Predictive Process Monitoring, Process Mining, Systematic Literature Review

1 Introduction

In recent years, organizations have tried to exploit historical process data to get data-driven insights from the day-to-day business operations. One opportunity to improve process performance by exploiting historical data is to train models based on different types of machine learning. Predictive process monitoring (PPM) takes historical process data (a set of completed business process executions) as input and uses machine learning techniques to predict a user specified need during the runtime of a selected business process. In the past, different setups have been applied because of the high complexity. That stems from the fact that researchers have used different algorithms, datasets, domains or prediction goals. Because of the high complexity and poor comparability, a variety of taxonomies for different scenarios have been developed.

In 2017 [28] and 2018 [53], the most representative time prediction setups of business processes were summarized. Even though both papers had the same intention, the methodology differs as shown in chapter 4. Another review in 2018 by [17] tackled the issue of the high variety of techniques and developed a value-driven framework based on prediction type. Finally, [48] presented a categorized collection of outcome-oriented PPM methods to enable researchers to compare methods in a unified setting.

The aim of this paper is to summarize, evaluate and categorize relevant literature based on business use cases (UC). The motivation of this is to provide a simple and easy to understand framework that promotes communication and collaboration between the industry and academia. This is achieved by a systematic literature review (SLR) that identifies published papers and by devising a taxonomy to classify the observations in UC. Further goals are to evaluate the results by means of different dimensions and to provide an overview with references to relevant literature to support practitioners and researchers in their future work.

The paper is structured as follows: the second chapter describes the main terms connected to the PPM area. Section three proceeds the SLR methodology and the review protocol. In sections four, a taxonomy gets developed that categorizes the result of section three. Section five discusses the results of this paper. The final section summarizes the academic and industrial contribution of this paper and identifies topics for future work.

2 Background

2.1 Business Process Management

Business Process Management (BPM) is a set of methods, tools and techniques to see how work is performed in an organization [10]. As a central element of contemporary organizations, BPM can support and monitor processes that are e.g. subject to policies, regulations and laws. The capability to optimize or support business decisions while running on an enterprise resource planning or workflow system is known as business activity monitoring [40]. However, BPM does not provide predictive solutions for a specific running process. That is where PPM comes into play. PPM focuses on exploiting generated process data and provides business insights that allow business users to take countermeasures during runtime.

2.2 Predictive Process Monitoring

PPM aims to predict the future of quantifiable values during a running process execution [25, 50] whereas for example business process intelligence focus on long term predictions such as key performance indicators [32]. To predict the outcome of running processes, PPM exploits historical data of already executed processes of the same type [25]. The set of historical data consists of events that correspond to the execution of activities of each process instance. Based on the prediction of these traces, the idea is to enable the business to proactively improve process performance and mitigate risks [39]. There are many scenarios where it is useful to have reliable process predictions, such as predicting compliance violations [8], the remaining sequence of activities [11, 43] or the remaining execution time of a case [9, 38].

2.3 Process Mining

In recent decades, process mining has emerged as a research field that focuses on analysing the execution of processes. Process mining is a collection of techniques to extract valuable process data [3]. Depending on the BPM lifecycle, different approaches such as process model discovering, monitoring or improving can be accomplished. In this paper the target is to support making decisions during runtime by using logfiles [2]. PPM makes use of process mining by retrieving the information from Process Aware Information Systems (PAIS) that are stored in logfiles for example in order to make time [1] or cost predictions [49].

3 Research goals and method

This paper applies an SLR in order to review a specific area in a thorough and unbiased manner [21]. The review ensures a rigorous and complete documentation and will be used to create a use-case-driven taxonomy that categorizes research papers. Furthermore, in chapter 4 the results will be evaluated and analysed by different dimensions to identify the relevance and to provide an easy access for the industry.

3.1 Systematic Review Protocol

The systematic review protocol specifies the research questions, the search protocol, and the selection criteria. Below, the research questions (RQ) are formulated, electronic databases are identified and inclusion as well as exclusion criteria are defined. The paper aims at answering the following research question:

1. RQ (Existing published papers): “Which types of academic publications exist in the field of predictive process monitoring?”

In line with the main research question, the paper also answers the following sub-research questions:

1. Sub-RQ (Taxonomy): “How should the published papers be categorized?”
2. Sub-RQ (Ranking): “What are the most relevant published papers?”

The first step was to develop search strings that are used to query electronic databases with the goal of producing a broad outcome of academic papers in the area of PPM. The following search strings derive from the terms introduced in chapter 2 and are used as keywords:

- “(business) process” – the domain of the paper is in the area of business processes
- “prediction” or “predictive” – a relevant paper needs to discuss the area of prediction
- “(business) process monitoring” – a paper targets the area of process monitoring
- “process mining” – a relevant paper targets the area of process mining

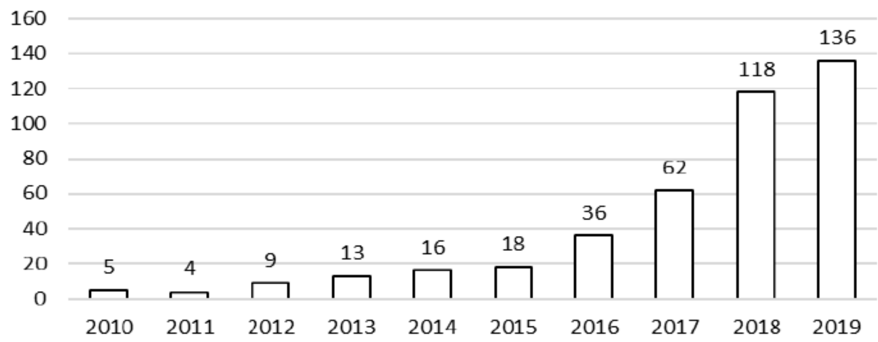
Because the research areas of the keywords "monitoring" and "prediction" or "predictive" are vast and encompass areas outside of the BPM domain, the search strings are always combined with "(business) process" to limit the results in a manageable way. BPM itself can include different angles and different domains. To ensure a broad field of literature and to not limit the search, the keywords "algorithm", "techniques" and "prediction type" were left out. Presuming that different authors might use a variety of wordings to refer to PPM, the following datable shows all constructed phrases. The generated phrases were then executed by a software that is designed to present cases for research in a structured and manageable way [19].

Table 1. Applied search strings for the SLR to academic databases

Search Strings	Results	Date of Search
"business process" AND "prediction"	51 Papers	07.02.2020
"prediction" AND "business process"	33 Papers	07.02.2020
"prediction" AND "process mining"	17 Papers	07.02.2020
"predictive" AND "business process monitoring"	289 Papers	07.02.2020
"predictive" AND "business process"	316 Papers	07.02.2020
"predictive" AND "process mining"	37 Papers	07.02.2020
"predictive" AND "process monitoring"	248 Papers	07.02.2020

The phrases were applied to the Google Scholar academic database, a well-known electronic literature database in the field of computer science, that encompasses other electronic databases such as ResearchGate, arXiv, Elsevier, IEEE Xplore and Springer. The search was conducted in February 2020 and retrieved all studies that contained at least one of the constructed phrases in the title, keywords, abstract, or full text of the paper. The results were exported and merged into one Excel sheet for further processing. It returned in total 991 papers, 507 excluding duplicates. Duplicate are identified as papers that appeared in more than one search result of a phrase that have the identical title and author(s) [22]. The following figure shows how the studies are distributed from 2010 to 2019. Figure 1 shows that the number of publications on PPM is constantly increasing. Thus, a significant growth of published papers from the beginning of 2016 can be observed.

Fig. 1. Number of published predictive process monitoring papers from 2010 to 2019



In order to be considered the results of the SLR were matched against inclusion and exclusion criteria that are based on the research question. To assess the study's relevance, it then must satisfy all inclusion and exclusion criteria. In this paper the three inclusion criteria were applied in the following chronological order:

1. Inclusion criteria: The paper is cited at least five times (An exception was made for papers published in 2019. Due to the lack of time to get the necessary number of citations it was lowered to three)
2. Inclusion criteria: The paper is written in English
3. Inclusion criteria: The paper was published in conferences proceedings or journals

After applying the inclusion criteria, the number of papers was reduced to 106. The remaining papers were further assessed with respect to the exclusion criteria by viewing the abstracts of each paper.

1. Exclusion criteria: The paper is not related to the computer science field
2. Exclusion criteria: The paper is not concerned with prediction in the context of BPM
3. Exclusion criteria: The paper is not accessible on the web

After proceeding the exclusion criteria, 39 unique papers were kept in accordance with the main RQ. However, literature reviews come with its limitations. Consequently, to have a benchmark the data of the paper gets compared with the findings of other reviews from the same research field. Four different papers that also applied an SLR were identified. Table 2 shows the review methodology and results of each paper.

Table 2. Comparison of papers review methodology sorted by years covered

	Keywords	**Search scope**	**Min. number of citat.**	**Years covered**	**Papers after filtering**
Method in [53]	"predictive process monitoring" "predictive business process monitoring" "predict (the) remaining time" "remaining time prediction" "predict (the) remaining * time"	Title, full text	5 (except 2017 papers)	2005-2017	53
Method in [48]	"predictive process monitoring" "predictive business process monitoring" "business process prediction"	Title, abstract, keywords, full text	5 (no exception)	2005-2017	14
Method in [17]	"predictive" AND "business process" "predictive" AND "process mining" "prediction" AND "business process" "prediction" AND "process mining"	Titel, abstract	10 (if published before 2016)	2005-2018	51
Method in [28]	"business process" AND "prediction" "predictive monitoring" AND "business process"	Title, abstract, keywords	5 (except 2016-2017 paper)	2010-2016	41

	Keywords	Search scope	Min. number of citat.	Years covered	Papers after filtering
Papers method	“business process” AND “prediction” “prediction” AND “business process” “prediction” AND “process mining” “predictive” AND “business process monitoring” “predictive” AND “business process” “predictive” AND “process mining” “predictive” AND “process monitoring”	Title, abstract, keywords, full text	5 (except 2019 papers)	2011-2019	39

It can be observed that the methodology does not vary noticeable in the use of keywords and their combinations. Since different authors might use different terms for the same meaning, it is difficult to identify how the number of search phrases and combinations affect the result of papers after filtering. Besides that, the used inclusion and exclusion criteria and the RQ of each paper can have a strong impact on the result in terms of quantity and quality. That can explain why the numbers of [48] differ considerably from the overall result. Moreover it can be observed that the papers method has a lower result than [28] [17] [53] although more keywords were used.

From the papers final list, standard meta-data such as authors, year of publication and number of citations were extracted. In addition, the type of publication, type of domain (identified by the origin of the data set or log) and type of prediction were extracted by reading the full text of each paper.

4 Taxonomy and data analysis

Usually the literature on PPM can be classified by two main dimensions: type of prediction and algorithm [17]. In this section, a taxonomy is proposed to classify the identified papers by business use cases and thus to answer Sub-RQ1. At the beginning a parent use case describes the need of the paper which correlates to the RQ: “*As the industry I want to have an understanding what relevant academic research exists to identify collaborations in the area of PPM*”. The starting point is further broken down in business use cases to identify and organize the need of the industry more in detail. This step is necessary to establish an easy to understand communication basis between the industry needs and academic papers. Based on each use case a subcategory is introduced to provide an efficient and fast overview. For a better understanding, each business use case is elucidated by a brief explanation and further examples. Finally, different academic papers are assigned to the use-case-driven framework and the chapter concludes by presenting results in three data tables.

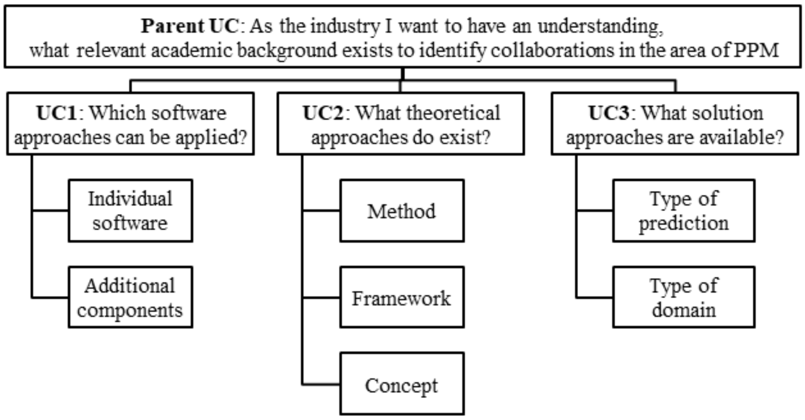
UC1: Which software approaches can be applied? Academic papers that contribute a technical implementation are assigned to this UC. The software solution can include frameworks [13, 16], improvements [12] or fill the gap between research and practice [20]. To ensure an easy understanding on the industry side, the type of implementation is categorized as individual software [20] or as additional component for instance for a mining toolset [13]. Commercial solutions are not taken into consideration.

UC2: *What theoretical approaches do exist?* The use case focuses mainly on a theoretical level by introducing methods, frameworks and concepts. Methods or techniques focus on a particular procedure to fulfil a more general [18, 25] or specified goal [14, 24]. Then again frameworks [17, 28] help to accomplish a goal that can include different methods [15]. At last concepts have a more abstract level of detail and therefore provide a rough overview of thoughts and approaches [37]. The academic contribution often gets confirmed by a case study or review.

UC3: *What solution approaches are available?* The use case aims to tackle a particular problem and proposes solutions. Each solution approach is related to one of the three macro-categories of prediction as outlined in [17]. The solution can provide a theoretical approach that gets confirmed by a set of generated [18, 34] or real life data [31, 35]. Based on the used data set the domain of each paper was specified.

In figure 2 the use-case-driven framework is visualized to provide a better understanding.

Fig. 2. Taxonomy of use cases for predictive process monitoring



Next, the SLR results get categorized by use case. Each paper gets assigned once and is shown in the use-case-driven framework below. The tables are structured from the left to the right as follows: The first column references the academic paper to provide an easy access for the industry. Followed by use case specific subcategories, the number of citations, type of publication and year of publication. All tables are sorted by number of citations to provide a ranking by relevance and to answer Sub-RQ2. As a result, 4 software, 22 theoretical and 13 problem-solution papers are identified.

Table 3. Overview of four software approaches (UC1: Which software approaches can be applied?)

Ref.	Type of implementation	Number of citations	Type of publication	Year of publication
[16]	Addit. component	69	Journal	2017
[20]	Indiv. software	7	Conference	2017
[12]	Indiv. software	7	Conference	2017
[13]	Addit. component	6	Conference	2015

Table 4. Overview of 22 theoretical approaches (UC2: What theoretical approaches do exist?)

Ref.	Type of theory	Domain	Number of citat.	Type of pub.	Year of pub.
[25]	Framework	Healthcare	141	Conference	2014
[31]	Framework	Logistic	100	Journal	2014
[18]	Method	-	99	Journal	2014
[24]	Method	Financial, healthcare	87	Conference	2016
[42]	Method	Customer supp., financial	72	Journal	2015
[47]	Framework	Financial	42	Conference	2016
[28]	Time	-	40	Journal	2017
[43]	Framework	Healthcare	34	Conference	2012
[51]	Method	Healthcare, insurance	32	Conference	2016
[48]	Framework	-	32	Journal	2019
[14]	Framework	Healthcare, public admin.	29	Conference	2016
[27]	Method	Automotive, healthcare	28	Journal	2017
[17]	Framework	-	26	Conference	2018
[15]	Framework	Healthcare, public admin.	22	Journal	2018
[52]	Method	Financial, public admin.	15	Conference	2016
[53]	Framework	-	13	Journal	2019
[26]	Concept	Insurance	10	Conference	2017
[46]	Framework	Financial, healthcare, manufac-turing, public admin.	9	Journal	2018
[44]	Method	Customer supp., financial, healthcare	9	Journal	2018
[4]	Concept	Logistic	7	Conference	2016
[37]	Concept	-	5	Conference	2018
[23]	Method	Manufacturing	5	Journal	2018

Table 5. Overview of 13 solution approaches (UC3: What solution approaches are available?)

Ref.	Type of prediction	Domain	Number of citat.	Type of pub.	Year of pub.
[45]	Sequence of activities	Customer supp., financial, public admin.	137	Conference	2017
[7]	Outcome	-	90	Journal	2016
[35]	Time	Customer supp., financial, public admin.	54	Journal	2018
[11]	Sequence of activities	Automotive, financial	49	Conference	2016
[36]	Time	Public admin.	48	Conference	2014
[41]	Time	Healthcare, manufacturing	30	Conference	2017
[33]	Time	Customer supp., financial	20	Conference	2017
[29]	Sequence of activities	Automotive, customer supp., financial	17	Conference	2017
[34]	Time	Generated Data Set	16	Conference	2011
[6]	Next activity	Financial, manufacturing	14	Journal	2019
[54]	Time	Financial	10	Conference	2017
[5]	Time	Logistic	9	Conference	2013
[30]	Next activity	Customer supp., financial	5	Journal	2018

Following the papers proposed use-case-driven framework, practitioners and researchers can easily navigate through the field of interest with the aim to identify academic papers for further collaboration. To provide an easy and efficient framework each use case provides its own subcategories. Furthermore, papers can get selected by the number of citations and the year or type of publication.

5 Discussion

To review the above contribution, an SLR was conducted to identify relevant research in the area of PPM. In contrast, the study was not limited to a specific type of prediction [48] or methods [28]. The papers taxonomy followed a more generic approach by providing the industry a relevant overview of different business use cases. A similar approach has been done by [17] where the taxonomy guides companies through a selection of prediction goals to find the best technique matching their needs. In comparison, the scope of this paper also includes the software and theoretical perspectives.

Concerning the number of papers after filtering, it is difficult to state if all relevant academic papers were identified. The search result depends heavily on the methodology as introduced in table 2 and the databases that were used to conduct the search. To have a benchmark, the result of papers after filtering for each identified SLR in the field of PPM gets compared by a cross table. An overlap is identified when papers have an identical title and the same author(s). To simplify the result, the overlapping is shown in percentage.

Table 6. Overlapping of literature review results in the search area of PPM

	[53]	[48]	[17]	[28]	**Paper**
Results of [53] in	-	21%	31%	28%	21%
Results of [48] in	12%	-	10%	13%	15%
Results of [17] in	64%	36%	-	56%	44%
Results of [28] in	44%	36%	43%	-	36%
Results of this paper in	32%	43%	33%	36%	-

The comparison shows that the performed review includes at least 30% but always less than 45% of all other review results. The identified gap can be an indicator that not all relevant work is identified and therefore seen as a threat of incompleteness.

6 Conclusion and future work

The research area of PPM has been growing significantly in recent years. However, the high degree of complexity makes it difficult for the industry to find a suitable overview that promotes collaborations. The novelty of this research is, on the one hand, to provide by the means of an SLR a relevant academic background. On the other hand, this paper introduces a new communication approach by providing a use-case-driven framework to guide the industry in an understandable and easy way through the academic domain. Therefore, this paper contributes to the knowledge domain by proposing a novel taxonomy that helps to overcome the lack of connection between academy and industry. The framework helps to navigate among the three different use cases of software, theory and problem-solving to support the communication and collaboration between industry and academia.

In future work, this taxonomy can be further developed and used to identify mutual benefits in the area of PPM or can be applied in other research areas because of its generic nature. This presumption makes the author believe that the framework may also be applicable outside of the PPM area. Additionally, the conducted SLR provides a qualified background to identify research gaps or promote further investigations.

References

1. van der Aalst, W.M.P., Schonenberg, M.H., Song, M.: Time prediction based on process mining. Inf. Syst. 36, 450–475 (2011). https://doi.org/10.1016/j.is.2010.09.001
2. van der Aalst, W.: Process Mining - Data Science in Action. Springer Berlin Heidelberg, Berlin, Heidelberg (2016)
3. van der Aalst, W., Adriansyah, A., de Medeiros, A.K.A., Arcieri, F., Baier, T., Blickle, T., Bose, J.C., van den Brand, P., Brandtjen, R., Buijs, J., Burattin, A., Carmona, J., Castellanos, M., Claes, J., Cook, J., Costantini, N., Curbera, F., Damiani, E., de Leoni, M., Delias, P., van Dongen, B.F., Dumas, M., Dustdar, S., Fahland, D., Ferreira, D.R., Gaaloul, W., van Geffen, F., Goel, S., Günther, C., Guzzo, A., Harmon, P., ter Hofstede, A., Hoogland, J., Ingvaldsen, J.E., Kato, K., Kuhn, R., Kumar, A., La Rosa, M., Maggi, F.M., Malerba, D., Mans, R.S., Manuel, A., McCreesh, M., Mello, P., Mendling, J., Montali, M., Motahari-Nezhad, H.R., zur Muehlen, M., Munoz-Gama, J., Pontieri, L., Ribeiro, J., Rozinat, A., Seguel Pérez, H., Seguel Pérez, R., Sepúlveda, M., Sinur, J., Soffer, P., Song, M., Sperduti, A., Stilo, G., Stoel, C., Swenson, K., Talamo, M., Tan, W., Turner, C., Vanthienen, J., Varvaressos, G., Verbeek, E., Verdonk, M., Vigo, R., Wang, J., Weber, B., Weidlich, M., Weijters, T., Wen, L., Westergaard, M., Wynn, M.: Process Mining Manifesto. In: BPM 2011: Business Process Management Workshops. pp. 169–194 (2012)
4. Alias, C., Rawet, V.L., Neto, H.X.R., Reymão, J. do E.N.: Investigating into the Prevalence of Complex Event Processing and Predictive Analytics in the Transportation and Logistics Sector: Initial Findings From Scientific Literature. In: MCIS 2016 Proceedings. pp. 1–17 (2016)
5. Bevacqua, A., Carnuccio, M., Folino, F., Guarascio, M., Pontieri, L.: A data-adaptive trace abstraction approach to the prediction of business process performances. In: ICEIS 2013 - Proceedings of the 15th International Conference on Enterprise Information Systems. pp. 56–65 (2013)
6. Borkowski, M., Fdhila, W., Nardelli, M., Rinderle-Ma, S., Schulte, S.: Event-based failure prediction in distributed business processes. Inf. Syst. 81, 220–235 (2019).
7. Breuker, D., Matzner, M., Delfmann, P., Becker, J.: Comprehensible Predictive Models for Business Processes. MIS Q. 40, 1009–1034 (2016). https://doi.org/10.25300/misq/2016/40.4.10
8. Castellanos, M., Casati, F., Dayal, U., Shan, M.-C.: A Comprehensive and Automated Approach to Intelligent Business Processes Execution Analysis. Distrib. Parallel Databases. 16, 239–273 (2004). https://doi.org/10.1023/B:DAPD.0000031635.88567.65
9. van Dongen, B.F., Crooy, R.A., van der Aalst, W.M.P.: Cycle Time Prediction: When Will This Case Finally Be Finished? In: OTM 2008: On the Move to Meaningful Internet Systems: OTM 2008. pp. 319–336 (2008)
10. Dumas, M., La Rosa, M., Mendling, J., Reijers, H.A.: Fundamentals of Business Process Management. Springer Berlin Heidelberg, Berlin, Heidelberg (2018)
11. Evermann, J., Rehse, J.-R., Fettke, P.: A Deep Learning Approach for Predicting Process Behaviour at Runtime. In: International Conference on Business Process Management 2016. pp. 327–338 (2017)

12. Evermann, J., Rehse, J.R., Fettke, P.: XES tensorflow - Process prediction using the tensorflow deep-learning framework. CEUR Workshop Proc. 1848, 41–48 (2017)
13. Federici, M., Rizzi, W., Di Francescomarino, C., Dumas, M., Ghidini, C., Maggi, F.M., Teinemaa, I.: A ProM operational support provider for predictive monitoring of business processes. In: CEUR Workshop Proceedings. pp. 1–5 (2015)
14. Di Francescomarino, C., Dumas, M., Federici, M., Ghidini, C., Maggi, F.M., Rizzi, W.: Predictive business process monitoring framework with hyperparameter optimization. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). pp. 361–376 (2016)
15. Di Francescomarino, C., Dumas, M., Federici, M., Ghidini, C., Maggi, F.M., Rizzi, W., Simonetto, L.: Genetic algorithms for hyperparameter optimization in predictive business process monitoring. Inf. Syst. 74, 67–83 (2018). https://doi.org/10.1016/j.is.2018.01.003
16. Francescomarino, C. Di, Dumas, M., Maggi, F.M., Teinemaa, I.: Clustering-Based Predictive Process Monitoring. IEEE Trans. Serv. Comput. 12, 896–909 (2019).
17. Di Francescomarino, C., Ghidini, C., Maggi, F.M., Milani, F.: Predictive Process Monitoring Methods: Which One Suits Me Best? In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). pp. 462–479 (2018)
18. Ghattas, J., Soffer, P., Peleg, M.: Improving business process decision making based on past experience. Decis. Support Syst. 59, 93–107 (2014). https://doi.org/10.1016/j.dss.2013.10.009
19. Harzing, A.W.: Publish or Perish
20. Jorbina, K., Rozumnyi, A., Verenich, I., Di Francescomarino, C., Dumas, M., Ghidini, C., Maggi, F.M., La Rosa, M., Raboczi, S.: Nirdizati: A web-based tool for predictive process monitoring. In: CEUR Workshop Proceedings (2017)
21. Kitchenham, B.: Procedures for performing systematic reviews. Keele, UK, Keele Univ. 33. 1–26 (2004)
22. Kofod-petersen, A.: How to do a structured literature review in computer science. Researchgate. 1–7 (2014)
23. Lee, W.L.J., Parra, D., Munoz-Gama, J., Sepúlveda, M.: Predicting process behavior meets factorization machines. Expert Syst. Appl. 112, 87–98 (2018). https://doi.org/10.1016/j.eswa.2018.05.035
24. Leontjeva, A., Conforti, R., Di Francescomarino, C., Dumas, M., Maggi, F.M.: Complex symbolic sequence encodings for predictive monitoring of business processes. Int. Conf. Bus. Process Manag. 2016. 9253, 297–313 (2015). https://doi.org/10.1007/978-3-319-23063-4_21
25. Maggi, F.M., Di Francescomarino, C., Dumas, M., Ghidini, C.: Predictive monitoring of business processes. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). 8484 LNCS, 457–472 (2014). https://doi.org/10.1007/978-3-319-07881-6_31
26. Maisenbacher, M., Weidlich, M.: Handling Concept Drift in Predictive Process Monitoring. In: Proceedings - 2017 IEEE 14th International Conference on Services Computing, SCC 2017. pp. 1–8 (2017)
27. Márquez-Chamorro, A.E., Resinas, M., Ruiz-Cortés, A., Toro, M.: Run-time prediction of business process indicators using evolutionary decision rules. Expert Syst. Appl. 87, 1–14 (2017). https://doi.org/10.1016/j.eswa.2017.05.069
28. Marquez-Chamorro, Eduardo, A., Resinas, M., Ruiz-Cortes, A.: Predictive monitoring of business processes: A survey. IEEE Trans. Serv. Comput. 11, 962–977 (2018). https://doi.org/10.1109/TSC.2017.2772256
29. Mehdiyev, N., Evermann, J., Fettke, P.: A multi-stage deep learning approach for business process event prediction. 2017 IEEE 19th Conf. Bus. Informatics. 1, 119–128 (2017). https://doi.org/10.1109/CBI.2017.46
30. Mehdiyev, N., Evermann, J., Fettke, P.: A Novel Business Process Prediction Model Using a Deep Learning Method. Bus. Inf. Syst. Eng. 62, 143–157 (2020). https://doi.org/10.1007/s12599-018-0551-3
31. Metzger, A., Leitner, P., Ivanovic, D., Schmieders, E., Franklin, R., Carro, M., Dustdar, S., Pohl, K.: Comparing and Combining Predictive Business Process Monitoring Techniques. IEEE Trans. Syst. Man, Cybern. Syst. 45, 276–290 (2015). https://doi.org/10.1109/TSMC.2014.2347265
32. zur Mühlen, M., Shapiro, R.: Business Process Analytics. In: Handbook on Business Process Management 2. pp. 137–157. Springer Berlin Heidelberg, Berlin, Heidelberg (2010)
33. Navarin, N., Vincenzi, B., Polato, M., Sperduti, A.: LSTM networks for data-aware remaining time prediction of business process instances. In: 2017 IEEE Symposium Series on Computational Intelligence, SSCI 2017 - Proceedings. pp. 1–7 (2018)

34. Pandey, S., Nepal, S., Chen, S.: A test-bed for the evaluation of business process prediction techniques. 7th Int. Conf. Collab. Comput. Networking, Appl. Work. 382–391 (2011). https://doi.org/10.4108/icst.collaboratecom.2011.247129
35. Polato, M., Sperduti, A., Burattin, A., Leoni, M. de: Time and activity sequence prediction of business process instances. Computing. 100, 1005–1031 (2018).
36. Polato, M., Sperduti, A., Burattin, A., De Leoni, M.: Data-aware remaining time prediction of business process instances. In: Proceedings of the International Joint Conference on Neural Networks. pp. 816–823 (2014)
37. Poll, R., Polyvyanyy, A., Rosemann, M., Röglinger, M., Rupprecht, L.: Process forecasting: Towards proactive business process management. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). pp. 496–512 (2018)
38. Rogge-Solti, A., Weske, M.: Prediction of Remaining Service Execution Time Using Stochastic Petri Nets with Arbitrary Firing Delays. In: ICSOC 2013: Service-Oriented Computing. pp. 389–403 (2013)
39. Salfner, F., Lenk, M., Malek, M.: A survey of online failure prediction methods. ACM Comput. Surv. 42, 1–42 (2010). https://doi.org/10.1145/1670679.1670680
40. Schmidt, W.: Business Activity Monitoring (BAM). In: Rausch, P., Sheta, A.F., and Ayesh, A. (eds.) Business Intelligence and Performance Management. pp. 229–242. Springer London, London (2013)
41. Senderovich, A., Di Francescomarino, C., Ghidini, C., Jorbina, K., Maggi, F.M.: Intra and inter-case features in predictive process monitoring: A tale of two dimensions. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). pp. 306–323 (2017)
42. Senderovich, A., Weidlich, M., Gal, A., Mandelbaum, A.: Queue mining for delay prediction in multi-class service processes. Inf. Syst. 53, 278–295 (2015).
43. van der Spoel, S., van Keulen, M., Amrit, C.: Process Prediction in Noisy Data Sets: A Case Study in a Dutch Hospital. In: International Symposium on Data-Driven Process Discovery and Analysis. pp. 60–83 (2013)
44. Tax, N., Teinemaa, I., van Zelst, S.J.: An interdisciplinary comparison of sequence modeling methods for next-element prediction. Softw. Syst. Model. (2020). https://doi.org/10.1007/s10270-020-00789-3
45. Tax, N., Verenich, I., La Rosa, M., Dumas, M.: Predictive business process monitoring with LSTM neural networks. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). pp. 477–492 (2017)
46. Teinemaa, I., Dumas, M., Leontjeva, A., Maggi, F.M.: Temporal stability in predictive process monitoring. Data Min. Knowl. Discov. 32, 1306–1338 (2018). https://doi.org/10.1007/s10618-018-0575-9
47. Teinemaa, I., Dumas, M., Maggi, F.M., Di Francescomarino, C.: Predictive Business Process Monitoring with Structured and Unstructured Data. In: BPM 2016: Business Process Management. pp. 401–417 (2016)
48. Teinemaa, I., Dumas, M., Rosa, M. La, Maggi, F.M.: Outcome-oriented predictive process monitoring: Review and benchmark. ACM Trans. Knowl. Discov. Data. 13, (2019). https://doi.org/10.1145/3301300
49. Tu, T.B.H., Song, M.: Analysis and Prediction Cost of Manufacturing Process Based on Process Mining. In: 2016 International Conference on Industrial Engineering, Management Science and Application (ICIMSA). pp. 1–5. IEEE (2016)
50. Verenich, I.: A general framework for predictive business process monitoring. (2016)
51. Verenich, I., Dumas, M., La Rosa, M., Maggi, F.M., Di Francescomarino, C.: Complex symbolic sequence clustering and multiple classifiers for predictive process monitoring. In: International Conference on Business Process Management 2016. pp. 218–229 (2016)
52. Verenich, I., Dumas, M., La Rosa, M., Maggi, F.M., Di Francescomarino, C.: Minimizing overprocessing waste in business processes via predictive activity ordering. Int. Conf. Adv. Inf. Syst. Eng. 9694, 186–202 (2016). https://doi.org/10.1007/978-3-319-39696-5_12
53. Verenich, I., Dumas, M., Rosa, M. La, Maggi, F.M., Teinemaa, I.: Survey and Cross-benchmark Comparison of Remaining Time Prediction Methods in Business Process Monitoring. ACM Trans. Intell. Syst. Technol. 10, 1–34 (2019). https://doi.org/10.1145/3331449
54. Verenich, I., Nguyen, H., Rosa, M. La, Dumas, M.: White-box prediction of process performance indicators via flow analysis. ACM Int. Conf. Proceeding Ser. Part F1287, 85–94 (2017).

Co-innovation in a University-Industry Partnership: A Case Study in the Field of Process Mining

Fareed Zandkarimi[1], Janina Nakladal[2], Josua Vieten[3], Jerome Geyer-Klingeberg[4]

[1] University of Mannheim, Chair of Enterprise Systems, L15 1-6, 68161 Mannheim, Germany
zandkarimi@uni-mannheim.de

[2,3,4] Celonis SE, Academic Alliance Department, Theresienstraße 6, 80333 Munich, Germany
{j.nakladal, j.vieten, j.geyerklingeberg}@celonis.com

Abstract. Evolving technologies require continuous innovation and skill adaptation in the industry, often resulting in a gap between practical demand, academic research, and teaching curricula. Although both industry and academia share the common goal of being at the forefront of technological development, they are using different approaches and mindsets. Strategic alliances between the two parties have the potential to take advantage of both approaches and to unleash greater and innovative value. This paper presents a case study from a successful long-term collaboration between an academic chair and a software company in the field of process mining. We share insights from our case and present an agile cooperation framework for co-innovation between industry and academia. This incorporates the objective to ensure broad and sustainable adoption and further development of innovative technologies. In addition, possible conflicts of interest are tackled and discussed. The insights are equally valuable for representatives from higher education institutions and companies.

Keywords: University-industry collaboration, strategic alliances, co-innovation, data science, process mining, case study.

1 Introduction

In today's era of digitization and innovation, technology is evolving rapidly, offering countless opportunities for new business models, products and services. The sources of innovation are not limited to research laboratories and development teams, but can be established wherever people are creative, open-minded, and skilled.

The cooperation between universities and industry (U-I) plays an important role in the successful identification and implementation of innovation [1,2,3]. The symbiosis achieved by U-I partnerships can foster a powerful collaboration cycle starting with a problem occurring in practice, solutions generated by joining academic and applied research in projects, and, finally, the application of research results in industry and academia. The successful implementation of this cycle comes with opportunities and challenges. While there exist several studies on U-I cooperation in engineering, medicine, and computer science [4], little is known for partnering in the field of process mining.

In this paper, we present a case study about a successful U-I cooperation for research and teaching in the field of process mining. The partnership aims at leveraging the symbiosis of improving data science skills in the form of joint university courses and cooperative research projects, also involving students being trained in a collaborative teaching approach. Next to the lessons learned from this cooperation, we share challenges and transfer our best practices into a general framework that might serve as guidance for future U-I partnerships.

After a review of the related literature, we describe our case study, including different stages of the U-I cooperation, followed by the presentation of the cooperation framework, and finally, a discussion and conclusions.

2 Related Literature

Collaborative U-I research has been proven to be mutually beneficial. Although organizing the collaboration comes with challenges, benefits prevail. For instance, companies' R&D expenditures per person are much lower when they collaborate with universities, and this effect remains even after the collaboration ends [3]. For universities, collaborative research projects generate higher value than other commercialization activities such as patenting and licensing [5]. However, universities tend to focus on theoretical aspects and the aim of universities to generate knowledge with long-term value is sometimes conflicting with the short-term orientation of enterprises [1,3]. This raises barriers to collaboration. According to [2], these hurdles are differently perceived by different stakeholders. Key barriers include the lack of understanding of different norms and environments in academia and industry, bureaucracy, and intellectual property rights. Various success factors can be identified for a fruitful cooperation between industry and academia. These can be grouped into institutional factors, relationship factors, and output factors [6]. Several analytical frameworks have been developed to classify and evaluate collaboration levels as well as success factors, including frameworks for the effective management of collaborative R&D projects [7], or frameworks for complex dynamic knowledge networks [8].

There is also a strand of literature on co-creation and co-innovation. For instance, [9] propose five pillars of co-innovation: collaboration, coordination, convergence, complementarity, and co-creation. According to further literature, co-creation is based on methods and tools, processes, innovation, and culture [9]. Other authors highlight the benefits of distributed innovation processes in large organizations, as opposed to single 'inventors' in a traditional sense [10]. These structures, however, can only be applied in large institutions with significant R&D resources and programs. For U-I cooperation, one could think of a distributed innovation process across a consortium in which competencies of different partners complement each other. This may involve transferring the entrepreneurial paradigm to academic structures [11].

3 Case Study

3.1 Background about Partners

The academic partner in this case study is the Chair of Enterprise Systems at University of Mannheim[1]. Founded in 1967, the university offers undergraduate and graduate programs in business administration, economics, law, social sciences, humanities, mathematics, computer science and information systems. The focus of the research group is to study the design, implementation, and application of enterprise systems and to provide insights into their impact on organizations.

The industry partner is a software company based in Munich and founded in 2011. The firm has 1000 employees located in 10 offices worldwide. The products are various software applications in the field of process mining. Process mining is an analytical discipline that gathers and analyzes information available in common enterprise IT systems to solve problems related to business processes [12]. It is an interdisciplinary and rather young field that combines business and IT skills. In many innovation projects (in-house and in cooperation with partners and customers), the company shapes the progress of the technology [13-16]. The firm established an Academic Alliance, which is a department in charge of the global education and research program.[2] This program leverages more than 350 university partners and provides services like software access, teaching and learning material, applied research projects, and joint thesis supervision. The Academic Alliance co-creates innovation in education and research, also by exploring new forms of collaboration [17].

3.2 Goals of the Partners

Initially, there are two independent parties with their missions and goals. Before entering a cooperation, both parties should be aware of these conditions and find their answers to the question of why cooperation makes sense and what purpose it serves. The general goals in the cooperation presented here are the following:

Objectives of Academic Partner. Relevant and applied teaching, groundbreaking research, being an incubator for inspiration and innovation, publishing research.

Objectives of the Industry Partner. Train tomorrow's workforce, attractive employer marketing, technology leadership, product development, access to latest research, joint papers and conference contributions.

3.3 Evolution of the Partnership

Initiation. The two parties come together in the initiation phase. In the case presented here, both partners independently defined the mission and selected the appropriate partner. Previous experiences and cooperation projects have shown that long-term

1 https://www.bwl.uni-mannheim.de/en/hoehle/
2 https://www.celonis.com/academic-alliance

collaborations have a higher chance of being fruitful for both groups. Therefore, at the beginning of new collaborations, a decision-making process with several must-have requirements for each step is recommended.

Selecting the Industry Partner. The process of selecting a partner begins with the rationale of whether an industrial partner is a necessity. In this case, the decision was clear as the Chair has a strong focus on enterprise systems in teaching and research. The next challenge was the selection of the partner. Many companies are keen to work with universities, with the marketing aspect of their collaboration being the guaranteed minimum. Regardless of the quality of the potential results, industry partners appreciate the opportunity to present their company, products, and services at universities where their potential employees or customers are studying. Therefore, it is the responsibility of the academic team to assess the qualifications of the industrial partner and to justify their potential benefits for their research projects and students. However, there is no trade-off in making this decision, i.e., the high willingness of companies to collaborate does not necessarily imply lower potential value for universities.

In a next step, the research interest and content of teaching courses formed the base for a list of expectations to be met by the prospective industry partner. The following requirements were derived from previous studies and a market analysis:

- Technology- and innovation-driven, knowledge producer
- Working on new and trending technology in the context of enterprise systems
- Offering a software product not too specific to a market niche, nor too generic leaving no chance for challenging the students to learn to work with it
- Accessibility in the classroom and for research, having a dedicated academic support team as well as an academic licensing program
- Openness to collaborate in research projects

This list led to checking Germany's most successful start-ups and the 10 German Unicorns (as of Feb. 2018). [18] emphasizes the positive role of vicinity in U-I collaborations. Therefore, geographic distance was included as another requirement. Several companies were screened based on their publicly available information on their websites and further information was collected by contacting them directly (e.g., whether they offer academic licenses or not). The process mining market leader has been selected as the offered services matched with all requirements.

Selecting the Academic Partner. From an industry perspective, growing a business fast and building on innovative technology requires a strong and agile workforce, which is scarce, especially in the area of data analytics and other technical roles. The theoretical teaching of data science within the higher education system does not always include important practical skills and applied knowledge. It also faces a time lag between the first appearance of innovative technologies and methods and the corresponding adaptation of curricula. In addition, students of all disciplines require data-driven skills and thinking. Companies can take responsibility and supplement theoretical lectures with practical problems, tools, and applied knowledge. Additional research collaborations help to innovate and find faster solutions for customer-related problems. In the case presented here, the industry partner has a dedicated Academic Alliance department to address these topics and build symbiotic relationships between the academic community and industry.

When selecting an academic partner for a long-term, sustainable and mature partnership, the following criteria were used:

- Existing curriculum and research agenda based on business analytics and business processes
- Technology and innovation-driven mindset
- Modular and flexible curriculum with the possibility to incorporate new content and new formats for teaching
- Dedicated personnel eager to learn about process mining and to become an expert
- Openness for joint research projects, co-create content
- International relations, experienced with industry collaboration, embedded in the research community, thought leader

The University of Mannheim and the Chair of Enterprise Systems perfectly met all the criteria and also provided a collaborator who had previously come into contact with process mining and was able to dedicate a PhD position to it, while at the same time striving for innovative teaching. In particular, the openness and agility of the academic chair ensured an easy start to the partnership, rapid success stories, and a growing partner relationship.

Socialization and Growth. The collaboration began with teaching and integrating practical business software into courses via guest lectures and case studies. The quick 'harvest' of such 'low-hanging fruits' allows to quickly verify that the fundamental objectives of both parties are met. In addition, it enables both partners to get to know each other, understand each other's needs and find a common working mode. The core of the collaboration are the partner managers on each side. While the academic partner has a dedicated research assistant for the partnership, the industry partner provides an Academic Alliance Manager for the collaboration. Both have several points of contact and are preparing a roadmap for joint projects. Building on feedback and mutual trust, it has been possible to create larger formats, try out innovative teaching methods and involve more parties and roles like other researchers or product managers and recruiters. The official course evaluations have been securing these decisions as almost all students appreciated incorporating an industry partner directly in their courses. This attitude was reaffirmed later when a significant number of students applied for industry-related research projects. From an academic perspective, students were more active in the course forum, rated the courses very positively in the standard evaluation and performed remarkably better in learning and applying their knowledge in the case study task in comparison to previous years, which could be seen in their actual results but also in the simultaneously increased difficulty of the given tasks. Table 1 visualizes the evolution of the partnership and the increasing number of projects, as well as involved roles. This trend is complying with the recommendation to start with smaller projects and let the growth evolve naturally. This is especially true for partnerships in the area of a young discipline like process mining, which is characterized by innovation – jointly in practice and theory, non-existent role models, rapid technological progress, and small communities of research and application. It also applies specifically to U-I collaborations, in which both parties come from different cultural worlds and have to combine their methodologies, project management approaches, and objectives.

Maturity. The closeness of cooperation can be measured not only in numbers but also in the depth of the relationship, which becomes clear through mutual support. After many joint projects, one can speak not only of cooperation but also of partnership. This means that both sides have equal rights and understand each other, even taking over parts of the other world. This is expressed in mutual recommendations, invitations to events of the respective partner (academic conferences, industry fairs), and uncomplicated, quick help with inquiries. Thus, in this case study, the academic partner facilitates the international expansion of the industrial Academic Alliance by exchanging contacts and strong recommendations, and the industry partner provides easy access to its customers and partners for academic projects. In addition, mutual feedback for teaching materials and trainings, but also the software product is exchanged. Based on this kind of partnership, many promising innovation projects for both parties will be developed in the future.

3.4 Lessons Learned

Success Factors. The presented successful U-I collaboration rooted in five areas:

1. Envisioning the collaboration as a process: Using process mining as a rather new technology in teaching and research without comprehensive best practices, the cooperation had to be perceived as a step-by-step development over time, which did not have to be perfectly defined from the beginning and was still designed for the long-term. Within the partnership, each project is managed as an iterative process and with an agile partner management approach. This ensures openness, flexibility, rapid integration of feedback and continuous growth.
2. Defining a clear "Why": While each party must define its "Why" of U-I collaborations, the presented case had a common reasoning: to equip the future generation with specific data science skills and to accelerate innovation in the area of process mining. This not only created relevance for both partners, but each project was therefore very focused and enthusiastic. A clear mission also ensures the effective definition of long-term and short-term goals, which can be manifested in clear, mutually beneficial success parameters.
3. Establishing ownership: Agile management requires separation of work and clear ownership for topics. It is a matter of ensuring that the expert for a certain part of the common project works on this task. If additional stakeholders were required or could benefit from a project, they were called in while managing the increased complexity.

Table 1. Evolution of the U-I partnership.

Semester	Projects	Outcomes for university	Outcomes for industry	Involved roles	#Trained students	Research projects	Beneficial roles
Spring 2018	Guest lecture and case study in a master's degree course* for MBA and information system students focusing on software product management and design.	Covering the practical aspects of the course with the academic license. Gaining practical experience from design experts.	Receiving insights about the software from the students' perspective, creating enthusiasm.	Academic Alliance Manager (AAM), Research Assistant (RA)	60	0	2
Fall 2018	Guest lecture, online training, course-integrated training, experts open talk [20] in a master's degree course* for MBA and informatics students focusing on business analytics and process mining.	Managing a complex case study with the help of experts of the company, license provision, the software attracts students to apply for projects more than before.	Students analyze customer-related datasets and provide new insights; the company recruits two students.	AAM, RA, university Chair Holder (Prof), Solution Engineer team (SET), Talent Acquisition Team (TAT)	80	User journey analysis	4
Spring 2019	Guest lecture, experts talk, product design challenge, master thesis research project.	Students were given an opportunity to work with real product design challenges, receiving feedback from industry experts.	Being able to solve problems and receive product feedback via newly created research projects, student projects and thesis topics.	AAM, RA, Prof, CE, SET, TAT, Product Management (PM), Product Design (PD), Quality Engineering (QE), master thesis students (MTS)	45	User behavior clustering, recommendation agents for analytical cloud solutions, visualization techniques for recommendation agents	10-12
Fall 2019	Guest lecture, real-world case-study challenge, experts talk, research project on a real business problem.	Involving the students in the joint research project and providing them with real business analysis problems as a unique opportunity, receiving expert feedback on research project.	Defining a joint research project and introducing the technology to a new market, expanding the market by integrating the outcomes of the research project.	AAM, RA, Prof, CE, SET, TAT, Business Partner (BP), MTS	80	Complex process analysis in process mining, trace clustering, concept drift business value analysis	12-15
Spring 2020	Connecting students and product managers very closely for immediate innovative solutions, defining team projects, multiple company lectures by experts.	Direct involvement in product innovation and development of new features, very applied learning and internship-similar experience.	Offering academic support to clients who have complicated business problems. Involving 4 large companies into the collaboration.	AAM, RA, Prof, CE, SET, TAT, PM, PD, QE, User experience manager (UX), team of project students, MTS	90	5 ongoing research projects which mainly or partially rely on the provided technology/contacts of industry partner.	10-15

*this course is one of the fundamental courses of the chair and is offered regularly every year. Module catalogue is available in [19].

In addition, the dedicated project managers on each side were able to continuously drive and monitor the success, considering the cultural requirements and possibilities of each side. Above that, management attention, support and commitment, e.g. from professors and CEOs, are crucial, especially for larger projects requiring fast decision making, such as collaborative research projects.

4. Building a long-term partnership: A close cooperation is kept alive by trust and loyalty, which also results from satisfaction. The partners in this case study achieved this by celebrating success and sharing results, for example through social media or at conferences. In addition, joint networks were established independent of specific projects, which serve as a pool of potential employees and increase mutual understanding – leveraging the full potential of U-I collaboration.
5. Communicating effectively and efficiently: The presented partnership is driven by personal contact and cross-project commitment. One success factor is clear and proactive communication via several digital channels and at least one personal meeting every three months, which is possible due to the short geographical distance.

Challenges. In addition to the success factors mentioned above, several challenges had to be overcome. Process mining as a technology is constantly evolving, while at the same time there is little material and experience available, especially for teaching. Many resources must be created from scratch, software training and products are often in a transitional stage. This requires a high level of flexibility and creativity from both project partners and the exercise of different roles from product creator to tester. Moreover, the cooperation is strongly based on the personal commitment and work of the two project managers, especially during the initiation phase. There is a constant risk of fluctuation or a change of priorities, especially in a dynamic start-up. To ensure that this does not lead to a loss of purpose, institutionalization measures must be taken as the partnership matures. Moreover, due to the different focus of university and academia on long-term scientific and short-term commercial results, respectively, different results need to be produced in order to achieve the goals of both parties. More generally, however, this is above all a challenge in maintaining motivation and focus over time, since U-I cooperation projects are never the sole task of both sides. Finally, U-I partnerships form a bridge between two worlds with different rules and regulations imposed on the other side. A certain amount of time must be devoted to managing intellectual property rights and ensuring compliance.

As the case partnership matures and has a certain level of success, a cooperation framework is to be presented in the next step. This framework shall serve to share best practices, facilitate institutionalization, and alleviate some of the challenges mentioned above. As a first step, previously published frameworks in the area of cooperation and partner management were examined in order to adapt the input to the case study and complement it with specific experiences in the area of agility.

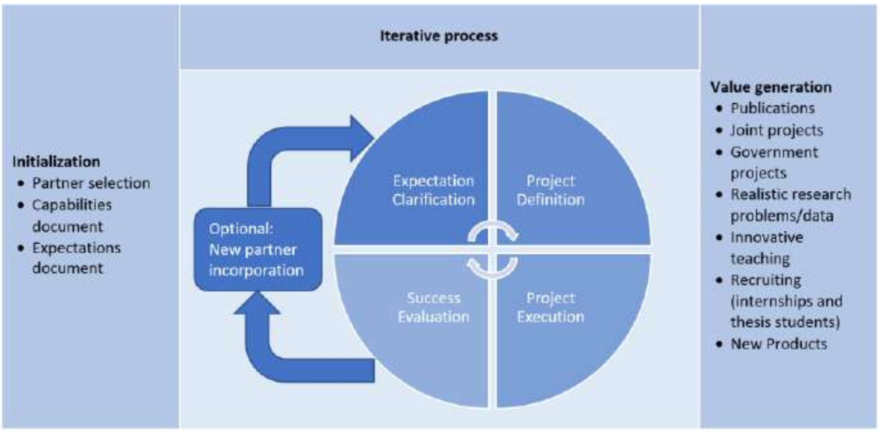
4 Framework

Based on the previous literature (Section 2), an agile project management approach with a continuous innovation process is used. Based on these aspects, an iterative innovation framework can be developed, which is presented in the following section.

Building a successful U-I collaboration requires planning, commitment, and constant alignment. These requirements are addressed in the previous studies, e.g. [1,2], however the emphasize on *an ongoing alignment component through iterative method* is presented in this framework. The certain key success factor in this case study is the mutual agreement in the beginning on having a long-term rather than a short-term collaboration. This also resulted in some exclusivity and clear focus on creating win-win-situations and success.

A generic model helps to explain the collaboration dynamics. As depicted in Figure 1, 4+1 phases summarize the strategic approach in our collaboration. Each quadrant is a mandatory part of the model, while incorporating new partners is an optional part that may happen after finishing one round of collaboration. Inspired by agile principles, this framework has an indispensable iterative nature that makes it distinct from previous U-I collaboration frameworks. After each iteration, an increased number of partners and projects can be expected.

Fig. 1. Agile U-I collaboration framework.



The introduced iteration model starts after the selection of a new partner. It is expected that each iteration will result in value for all parties involved. Value generation can take more than one round, which is not surprising if the projects are long-term. Detailed information on each phase is provided below:

Initialization. The most critical outcome of this phase is the partner selection decision, which was explained in Section 2.3. Decision makers need to compile expectations and capabilities lists. These lists should facilitate the selection process and help them to gain a full understanding of their position. This phase is not meant to be iterative, however, the aforementioned documents are subject to change in the future.

Expectation clarification. While general expectations have been defined in the initialization phase, the two partners can start the U-I collaboration with a deep-dive into specific project expectations whose definition is crucial for success. From the second iteration onwards, if new partners enter the partnership, expectations must be clarified again.

Project definition. Based on the available capabilities and expectations, one can expect various plans to pursue. Projects should be listed and scored in terms of the required resources, potential value, alignment with the organizations' goals, etc. Selecting the project(s) depends on the available resources of all parties and their interest in executing the project(s), which also depends on the generated value. For this, measurable goals and performance indicators must be defined. A best practice advice is to start small and grow big as the time investment in the first iterations can be greater than the value result. This ratio will be reversed as the partnership matures. A common understanding of each parties' members, resources, structure, agility, and commitment are the most important aspects for bigger projects.

Project execution. Selecting a project is about 'doing the right thing' while executing is about 'doing the thing right'. Academic projects generally have an explorative component expected to satisfy the scientific-contribution mission of academia. Consequently, the execution of U-I joint projects is likely to involve unexpected circumstances. Constant monitoring of the projects from all parties is the most crucial factor in this phase is clear communication to manage expectations and build trust.

Success evaluation. Performance indicators are necessary to assess the degree of success of the joint project(s). Several measurements could be nominated in connection with our U-I projects, e.g. publications, patents, data sets, analyses, and public or private projects of third parties. Usually the indicators were already defined, and the evaluation phase then includes reviews, feedback loops and the collection of best practices and improvement measures. While monitoring the number of publications and other results is important, it is crucial to also look at their quality.

New partner incorporation. This is an optional part of the model based on necessity or possible added value. When there is an opportunity or added value by involving new partners, existing parties should consider the potential benefits and costs before making this decision. It is likely that several types of partners will join the existing environment, e.g. an academic team from a different chair or university for a research project, a company interested in carrying out pilot projects or data collection, or a public sector or government organization.

This described model, together with the success factors mentioned in section 3.4, provides a framework that institutionalizes new initiatives of the presented partnership but also serves as a guidance for U-I collaborations in the area of data science.

5 Discussion

Several avenues for further research as well as limitations regarding the above-presented framework can be seen. First, the framework was derived from a single case study. It is therefore difficult to draw generalizations. The applicability for other cases must be proven by future U-I collaborations. Second, the results of this case study are strongly influenced by the short geographical distance between the two partners. Accordingly, the concepts might be different if applied to remote partnerships. Future research should also examine the role of platforms that are useful for remote research partnerships. Third, this case refers to a specific research/industry field, namely process mining. Future U-I partnerships might prove if the results are also valid for other disciplines and how the concepts must be adjusted for other areas. Fourth, a crucial success factor for the partnership is the background of the partner managers on both sides. On the academic side, the key person formerly worked in the same company and knows both industry and academia very well. On the industry side, the company established a dedicated department for university collaborations, which has experience from a large set of U-I partnerships. In a different setting with other partners, the timely reproduction of the results presented here might be difficult as it usually takes some time in the early stage of a partnership to develop a common understanding for each other.

6 Conclusion

This case study described the collaboration between an academic chair and a software company in the field of process mining. The development of the partnership together with the benefits and challenges that occurred along the journey were outlined. Based on this case study, an agile cooperation framework for collaboration between academia and industry is presented. This case study displays the status quo of the cooperation. Together, the two parties are eager to further expand their partnership. In the next step, they will apply for a public grant to gather funding for a joint research project on intelligent procurement processes. In the long-term, both parties aim at building a competence center for process mining. In the future, both parties will apply the best practices they gained from their collaboration and use the agile framework as a foundation for cooperation with other companies/universities.

References

1. Valentin, E.M.M.: University-Industry Cooperation: A Framework of Benefits and Obstacles. Industry and Higher Education 14(3), 165-172 (2000).
2. Siegel, D.S.; Waldman, D.A.; Atwater, L.E.; Link, A.N.: Toward a model of the effective transfer of scientific knowledge from academicians to practitioners: qualitative evidence from the commercialization of university technologies. Journal of Engineering and Technology Management 21, 115–142 (2004).
3. Scandura, A.: University- industry collaboration and firms' R&D effort. Research Policy 45(9), 1907-1922 (2016).

4. Ankrah, S.N.; Burgess, T.F.; Grimshaw, P.; Shaw, N.E.: Asking both university and industry actors about their engagement in knowledge transfer: What single-group studies of motives omit. Technovation 33, 50-65 (2013).
5. Perkmann, M.; King, Z.; Pavelin, S.: Engaging excellence? Effects of faculty quality on university engagement with industry. Research Policy 40(4), 539-552 (2011).
6. Rybnicek, R.; Königsgruber, R.: What makes industry–university collaboration succeed? A systematic review of the literature. Journal of Business Economics 89, 221-250 (2019).
7. Barnes, T.A.; Pashby, I.R.; Gibbons, A.M.: Collaborative R&D projects: a framework for effective management. In: Proceedings of the 2000 IEEE International Conference on Management of Innovation and Technology, Singapore (2000).
8. Carayannis, E.G.; Alexander, J.: Winning by Co-Opeting in Strategic Government-University-Industry R&D Partnerships: The Power of Complex, Dynamic Knowledge Networks Carayannis, E.G. & Alexander, J.: The Journal of Technology Transfer 24, 197-210 (1999).
9. Saragih, H.S.; Tan, J.D.: Co-innovation: a review and conceptual framework. International Journal of Business Innovation and Research 17(3), 361-377 (2018).
10. Eschenbächer, J.; Siefert, M.; Thoben, K.-D.: Improving distributed innovation processes in virtual organisations through the evaluation of collaboration intensities. Production Planning & Control 22(5-6), 473-487 (2011).
11. Lawrence, D.; David, K.: University-industry collaboration: Grafting the entrepreneurial paradigm onto academic structures. European Journal of Innovation Management 10(3), 316-332 (2007).
12. van der Aalst, Wil M.P.: Process Mining – Data Science in Action. 2nd edn, Springer, Heidelberg (2016).
13. Veit, F.; Geyer-Klingeberg, J.; Madrzak, J.; Haug, M.; Thomson, J.: The Proactive Insights Engine: Process Mining meets Machine Learning and Artificial Intelligence. In: Proceedings of the BPM Demo Track and BPM Dissertation Award co-located with 15th International Conference on Business Process Management, Barcelona, Spain (2017).
14. Geyer-Klingeberg, J.; Nakladal, J.; Baldauf, F.; Veit, F.: Process mining and robotic process automation: A perfect match. In: Proceedings of the Dissertation Award, Demonstration, and Industrial Track at BPM 2018 co-located with 16th International Conference on Business Process Management, Sydney, Australia (2018).
15. Badakhshan, P.; Bernhart, G.; Geyer-Klingeberg, J.; Nakladal, J.; Schenk, S.; Vogelgesang, T.: The Action Engine - Turning Process Insights into Action. Proceedings of the Business Process Management Demo Track and BPM Dissertation Award, Vienna, Austria (2019).
16. Vogelgesang, T.; Kaufmann, J.; Becher, D.; Seilbeck, R.; Klenk, M.: Celonis PQL: A Query Language for Process Mining. In: Polyvyanyy, A. (ed.): Process Querying Methods. Springer (2020, Forthcoming).
17. Nakladal, J.; Geyer-Klingeberg, J.; Stierle, M.; Matzner, M.: From Technical Product Training to Sustainable Education for Students-A Strategic Alliance Approach to Applied MOOCs. In: Proceedings of EMOOCs 2019. Work in Progress Papers of the Research, Experience and Business Tracks, Naples, Italy (2019).
18. Petruzzelli, A.M.: The impact of technological relatedness, prior ties, and geographical distance on university- industry collaborations: A joint-patent analysis. Technovation 31(7), 309-319 (2011).
19. University of Mannheim online module catalogue, https://www.bwl.uni-mannheim.de/media/Fakultaeten/bwl/Dokumente/Studium/MMM/Modulkatalog_MMM_Englisch.pdf, pp. 60-96, Last accessed 2020/05/01
20. LinkedIn, https://www.linkedin.com/posts/fareed-z-91746682_celonis-mannheimbusinessschool-celonis-activity-6539466325411012608-UQVn, Last accessed 2020/05/01

Sprecher der mit der EMISA assoziierten Arbeitskreise und Fachgruppen

Prof. Dr. Michael Fellmann *AK Semantische Technologien im Geschäftsprozessmanagement* Universität Rostock Institut für Informatik Albert-Einstein-Str. 22 18057 Rostock Tel: +49 - 0381 498-7440 E-Mail: michael.fellmann@uni-rostock.de	**Prof. Dr. Günther Specht** *Arbeitskreis Grundlagen von Informationssystemen* Universität Innsbruck Databases and Information Systems Technikerstrasse 21a A-6020 Innsbruck Tel: +43 512 507 53220 E-Mail: guenther.specht@uibk.ac.at
Horst Kremers *Comittee on Data for Science and Technology (CODATA-Germany)* Postfach 200548 D-13515 Berlin Tel: +49 - 172 3211738 E-Mail: office@horst-kremers.de	**Prof. Dr. Markus Nüttgens** *AK Geschäftsprozessmanagement mit Ereignisgesteuerten Prozessketten* Universität Hamburg WISO Fakultät, Wirtschaftsinformatik Von-Melle-Park 9 D-20146 Hamburg Tel: +49 – 40-428382792 E-Mail: markus.nuettgens@wiso.uni-hamburg.de

Fachexperten der EMISA

Prof. Dr. Jörg Desel FernUniversität in Hagen Lehrgebiet Softwaretechnik und Theorie der Programmierung D-58084 Hagen Tel: +49 - 2331-987 2609 E-Mail: joerg.desel@fernuni-hagen.de	**Prof. Dr. Dr.hc. Heinrich Mayr** Universität Klagenfurt Institut für Angewandte Informatik Universitätsstraße 65-67 A-9020 Klagenfurt am Wörthersee Tel: +43 - 463 2700 3732 E-Mail: Heinrich.Mayr@aau.at
Prof. Dr. Gottfried Vossen Universität Münster Lehrstuhl für Informatik Leonardo-Campus 3 D-48149 Münster Tel. +49- 251 83-38151 E-Mail: vossen@uni-muenster.de	**Prof. Dr. Andreas Oberweis** Karlsruher Institut für Technologie Institut für Angewandte Informatik und Formale Beschreibungsverfahren (AIFB) Kaiserstraße 89 D-76133 Karlsruhe Tel.: +49 - 721 608 44516 E-Mail: oberweis@kit.edu
Prof. Dr. Mathias Weske Universität Potsdam Hasso-Plattner-Institut für Softwaresystemtechnik Prof. Dr. Helmert-Str. 2-3 D-14482 Potsdam Tel: +49 (0)331-5509191 E-Mail: mathias.weske@hpi.uni-potsdam.de	**Dr.-Ing. Hansjürgen Paul** Institut Arbeit und Technik Forschungsschwerpunkt WISDOM Munscheidstraße 14 D-45886 Gelsenkirchen Tel. +49-209-1707229 E-Mail: paul@iat.eu
Prof. Dr. Stefan Strecker FernUniversität in Hagen Enterprise Modelling Research Group Universitätsstr. 41 D-58097 Hagen Tel: +49 2331 987 4430 E-Mail: stefan.strecker@fernuni-hagen.de	

EMISA-Leitungsgremium

Das im Mai 2018 gewählte EMISA-Leitungsgremium besteht aus den Sprechern der assoziierten Arbeitskreise und Fachgruppen (siehe vorige Seite) sowie folgenden gewählten Mitgliedern:

Prof. Dr. Jan Mendling *(Sprecher des Leitungsgremiums, Redaktion EMISA FORUM)* Wirtschaftsuniversität Wien Department of Information Systems & Operations Institute for Information Business Welthandelsplatz 1 A-1020 Wien Tel: +43-1-31336 5200 E-Mail: jan.mendling@wu.ac.at	**Prof. Dr. Agnes Koschmider** *(Stellv. Sprecherin)* Christian-Albrechts-Universität zu Kiel Institut für Informatik Arbeitsgruppe Wirtschaftsinformatik (Process Analytics) Hermann-Rodewald-Str. 3 24118 Kiel Tel: +49 431 880-6387 E-mail: ak@informatik.uni-kiel.de
Dr. Dirk Fahland TU Eindhoven Architecture of Information Systems Group NL-5600 MB Eindhoven Tel. +31 40 2474804 E-Mail: d.fahland@tue.nl	**Dr. Henrik Leopold** Kuehne Logistics University Großer Grasbrook 17 D- 20457 Hamburg Universität Potsdam, Hasso-Plattner-Institut für Softwaresystemtechnik Prof. Dr. Helmert-Str. 2-3 D-14482 Potsdam E-Mail: henrik.leopold@the-klu.org
Dr. Judith Michael RWTH Aachen University Department of Computer Science Ahornstraße 55 D-52074 Aachen Tel. +49 (241) 80-21323 E-Mail: michael@se-rwth.de	**Prof. Dr. Eric Proper** Luxembourg Institute of Science and Technology IT for Innovative Services department 5, avenue des Hauts-Fourneaux L-4362 Esch-sur-Alzette Tel: +352 275 888 1 e.proper@acm.org
Prof. Dr. Ulrich Reimer FHS St. Gallen Hochschule für Angewandte Wissenschaften Institut für Informations- und Prozess-Mgmt. Rosenbergstrasse 59 CH-9000 St. Gallen Tel. +41-71 226 17 46 E-Mail: ulrich.reimer@fhsg.ch	**Prof. Dr. Manfred Reichert** Universität Ulm Fakultät f. Ingenieurwissenschaften u. Informatik Institut für Datenbanken und Informationssysteme James-Franck-Ring, Geb. O27 D-89069 Ulm Tel: +49-731-5024135 E-Mail: manfred.reichert@uni-ulm.de
Prof. Dr. Matthias Weidlich Humboldt-Universität zu Berlin Institut für Informatik Rudower Chaussee 25 D-12489 Berlin Tel: +49 (0)30 2093 3143 E-Mail: matthias.weidlich@hu-berlin.de	