

SAP Academic Community Conference 2025 (D-A-CH)

Business Process Management drives Digital
Transformation

08. und 09. September 2025

Hasso Plattner Institut

Potsdam

**Konferenzband
der SAP Academic Community Conference 2025 (D-A-CH)**

**Uta Mathis, Hans-Jürgen Scheruhn, Stefan Weidner,
Holger Wittges, André Biener (Hrsg.)**

DOI: 10.14459/2025md1796257

Hasso-Plattner-Institut
Universität Potsdam
Prof.-Dr.-Helmert-Straße 2-3, 14482 Potsdam

<https://portal.ucc.ovgu.de/sap-acc-dach-2025/>



VORWORT

Die digitale Transformation verändert die Landschaft der globalen Wirtschaft. Dies zwingt Unternehmen dazu, ihre Geschäftsmodelle und Arbeitsprozesse nachhaltig anzupassen. In diesem sich schnell verändernden Umfeld spielen Schlüsseltechnologien wie Künstliche Intelligenz (KI), Big Data und Cloud-basierte ERP-Lösungen eine entscheidende Rolle. Diese Technologien ermöglichen es Unternehmen nicht nur, effizienter zu arbeiten, sondern auch, ihre Dienstleistungen und Produkte effektiver auf die Bedürfnisse ihrer Kunden abzustimmen. Ein wichtiger Hebel diese Veränderung zu gestalten, liegt in der Erschließung von neuen Technologien für das Business Process Management (BPM). Automatisierung, Process Mining und Process Simulation ermöglichen eine präzisere Analyse und Optimierung von Geschäftsprozessen. Dies wird zunehmend begleitet von einer Umstellung auf Public Cloud-Angebote und die zunehmende Autonomisierung des Datenaustauschs mit Kunden und Lieferanten.

Die SAP Academic Community Conference (ACC) 2025 der DACH-Region am Hasso Plattner Institut der Universität Potsdam steht daher unter dem Motto „Business Process Management drives Digital Transformation“ und bildet den Rahmen für einen Austausch über den Umgang mit Herausforderungen im Lehr- und Lernumfeld sowie dem Zusammenspiel mit der Anwendung von SAP-Lösungen. Die Konferenz bietet eine Plattform zur Diskussion wissenschaftlicher Erkenntnisse sowie anwendungsorientierter Berichte aus Lehre und Forschung.

Die SAP ACC 2025 (D-A-CH) ist in fünf Themenbereiche (Tracks) aufgeteilt. Track 1: Business Process Management behandelt aktuelle Entwicklungen bei der Dokumentation, Analyse, Optimierung und Management von Geschäftsprozessen, wie Process Mining und Simulation, RPA, Low-Code-Plattformen, KI-Unterstützung oder Dashboarding. Track 2: Technology fokussiert sich auf die Untersuchung der neuesten Technologien und deren Anwendung in der Geschäftswelt, einschließlich Künstlicher Intelligenz, Robotics, und der SAP Business Technology Plattform. Track 3: Digitale Transformation/Management begleitet die Diskussion über Strategien zur digitalen Transformation von Unternehmen, die Neuausrichtung von Geschäftsmodellen und die Integration digitaler Technologien in alle Unternehmensbereiche – einschließlich der Auswirkungen auf Positionierung, Unternehmenskultur (Change Management, Kompetenzen) und Organisation. Track 4: Didaktik widmet sich innovativen Lehrmethoden und didaktischen Konzepten zur Kompetenzvermittlung, mit besonderem Fokus auf interaktive, anwendungsorientierte Lernformate in der SAP-Ausbildung sowie der Integration moderner Technologien in die Lehre. Track 5: Internationalisierung beleuchtet, wie Unternehmen SAP-Systeme und -Prozesse so gestalten können, dass sie sowohl multinational konsistent als auch flexibel genug sind, um regionale Anforderungen zu erfüllen.

SAP Academic Community Conference 2025 (D-A-CH)

Business Process Management drives Digital Transformation

Für insgesamt fünf Tracks wurde zur Einreichung von folgenden Beitragsarten aufgerufen: wissenschaftliches Paper (8-10 Seiten), wissenschaftliches/praktisch orientiertes Short-Paper (3-5 Seiten), rein praktisch orientierte Präsentation (bis zu 15 Folien). Bereits zuvor publizierte Beiträge wurden nicht akzeptiert. Wir erhielten **67** Einreichungen, die alle im Double-Blind-Peer-Review begutachtet wurden und von denen wir nach Überarbeitung **59** Beiträge akzeptiert haben.

Ihre

Uta Mathis

Hans-Jürgen Scheruhn

Stefan Weidner

Holger Wittges

André Biener

zusammen mit dem Organisationsteam der ACC 2025 (D-A-CH)

Potsdam, im September 2025

SAP Academic Community Conference 2025 (D-A-CH)

Business Process Management drives Digital Transformation

Track 1: Business Process Management 1

Identifying Opportunities for Artificial Intelligence in Business Process Management	2
Community Sessions: An Overview of the SAP University Alliances Learning Portfolio with different focus: Business Process Management and Application Development and Coding	10
Global Bike mit SAP Signavio und SAP LeanIX	11
Brücken bauen zwischen Theorie und Praxis: Erfolgreiche Integration von SAP LeanIX in Lehre und Forschung am Beispiel eines S/4-Vorprojekts	12
Erfolgreiche Zusammenarbeit Hochschule Harz mit msg und Siemens in Forschung und Lehre im Bereich BPM und SAP S/4HANA	13
An Evaluation of Process Mining Techniques for People Analytics.....	15
Einsatz von Signavio Reports in der Lehre	20
Prozessorientierte Simulation als duale Studentin der DRV Bund.....	29
Process Comparisons: An Industry Perspective	30
Einordnung der SAP Best Practice Modelle in den Enterprise Online Guide	39
ESG meets EOG – Integration von Environment Social Governance (ESG) in die Technologie- und Netzwerk-Map des Enterprise Online Guide (EOG).....	52
Process Mining-basiertes Dashboard für Nachhaltigkeitsmanagement.....	62
Integration von Prozess- und Organisationsmodellen in SAP Signavio.....	72
Digital Twins in Supplier Risk Management: Building Dynamic Capabilities to drive Supply Chain Resilience & Process Excellence in Supplier Risk Management	82
Towards a Social Sustainability Insights Tool for Operations	90
Modellierung mit EOG @ SAP Signavio. Eine studentische Perspektive	91
Positive Veränderungen von Geschäftsprozessen in der Radverkehrsplanung durch die Bereitstellung unterstützender Dashboards	92

Track 2: Technology 93

Next Level Requirements Engineering NLP and ML Usage in the Software-Development-Life-Cycle	94
SAP ECC to S/4HANA Data Migration: Process Analysis and AI-Supported Optimization	102
Tiny Models, Big Impact: A Federated Learning Approach to Monitoring Depression Treatment via Smartphones and Wearables	111
Operating an automated SAP Helpdesk: Running a BERT-based Support Ticket Classifier on IBM Power10 Hardware	119
Visuelle Gestaltung eines GenAI-Assistenzsystems für die SAP Build-Umgebung: Ein Design-Science- Research-Ansatz.....	130

Track 3: Digitale Transformation/Management..... 139

Leveraging AI and Micro-Credentials in Education Insights from the EduAid Use Case Review	140
Wie betriebliche Weiterbildung skalierbar wird: Ein Zielbild digitaler SAP-Schulungen.....	151
Framing Labour Inclusion Through Micro-Credentials: A Theoretical Synthesis at the Intersection of Sustainability, Digital Transformation, and Human Capability	155
Social Architecture - The near future of distributed acting and learning	171
Störungen in Produktionssystemen Eine strukturierte Literaturanalyse zur Resilienzsteigerung	173
Disrupting the Chicken-and-Egg Barrier in Green Hydrogen: A Distributed-Ledger Subsidy Scheme with Purpose-Bound Tokens	181
Emerging Technologies and Artificial Intelligence in Business Software	186
Künstliche Intelligenz in der Internen Revision – Eine Studie gegenwärtiger Methoden.....	187
ERPfinder: Self-Service-Plattform zur Erstellung von ERP-Spezifikationen für KMU	198
Inclusive Pedagogy for Learners with Hearing Impairments: A Comparative Study of Uganda and Germany	206

SAP Academic Community Conference 2025 (D-A-CH)

Business Process Management drives Digital Transformation

Track 4: Didaktik 207

Hochschulübergreifende Unternehmenssimulation mit ERPsim: Konzeption, Herausforderungen und Lessons-learned	208
Individuelle SAP-Ausbildung mit LEGO: Innovative Lehrmethoden	222
Opening the Blackbox: A Socio-Technical Simulation Game for Teaching Resource-Efficient Information Logistics in End-to-End Value Streams	230
Competence-Based Assessment for SAP4school: Adaption and Pilot Testing	238
PDF oder Video? – Ein empirischer Vergleich zweier Lernformate für E-Learning-basierte SAP-Endanwenderschulungen	246
Nachhaltiger Lernerfolg durch Simulation? Ein Feldversuch zum Vergleich von ERPsim Manufacturing und PP-Fallstudie	247
Vom Lernhindernis zum Lernpfad – Microlearning als Schlüssel für effektive SAP-Endanwenderschulungen	248
SAP-Lehrkonzept „Personalwesen mit ERP-Systemen im Kontext der öffentlichen Verwaltung in Deutschland“	249
NextGen Modelers – Praxisorientierte Lehre von Prozessmodellierung	250
Improving Learning Outcomes: A Multi-Stimuli Approach on Teaching ERP Systems in Higher Education	251
Bridging Theory and Practice in ERP Education: A Didactic Toolbox for Sustainable SAP S/4HANA Teaching	260
Towards a Next-Generation Data Center Curriculum: Teaching AI Hosting Skills on IBM Power	265
Effective teaching with integrated analytics case studies & Tools	276
Teaching Enterprise Software Engineering in the Cloud – An Exemplary Case	277
Experiences with WatsonX Integration in Higher Education: Lessons from a TUM Course	278

Track 5: Internationalisierung 279

Evaluating Large Language Models for Causal Hypothesis Generation: Detecting Training Data Contamination	280
Strengthening International Collaboration through Erasmus+ KA171: Insights from East African Partnerships	291
Digital Marketing as a Driver of Business Process Innovation in Higher Education: A Case Study of Tanzanian Institutions	292
Smart Cities in Developing Countries: A Context-Adaptive Framework for Sustainable Urban Development	293
Toward Sustainable Digital Infrastructure: Decentralized Resource Utilization Using Existing Edge Systems	302
Leveraging German Expertise in Water Utility for Rwandan Education	304

TRACK 1: BUSINESS PROCESS MANAGEMENT

Track Chair: Prof. Dr. Hans-Jürgen Scheruhn

Identifying Opportunities for Artificial Intelligence in Business Process Management

B. Wurm¹, T. Zimmermann^{2,3}, P. Badakhshan^{3,4}, & S. N. Z. Naqvi⁵

¹LMU Munich School of Management, Munich, Germany

²University of Münster, Münster, Germany

³ERCIS-European Research Center for Information Systems, Münster, Germany

⁴SAP, Munich, Germany

⁵SAP, Walldorf, Germany

Abstract: Artificial intelligence (AI) has tremendous potential for business applications, yet its practical use in Business Process Management (BPM) remains underexplored. Current research on the link between AI and BPM is often either purely conceptual or narrowly focused on specific technologies, leaving a gap in empirical evidence regarding its holistic application. This study addresses this gap by investigating how AI can be used across the entire BPM lifecycle. We conducted nine semi-structured interviews with industry experts at the intersection of AI and BPM, analyzing the data against the established BPM lifecycle framework. Our findings identify significant opportunities in each phase, such as using AI to discover processes from unstructured data, generating redesign suggestions, and enabling autonomous process monitoring. A key insight is that AI can help overcome persistent BPM challenges like siloed knowledge and resource-intensive manual analysis. The study also reveals overarching strategic shifts, including the democratization of BPM by making tools more accessible to business users and a paradigm shift towards agent-based, outcome-focused process management.

Keywords: Artificial Intelligence, Business Process Management, Large Language Models, Agentic BPM, Strategic BPM.

1. Introduction

Companies are acknowledging the transformative potential of Artificial Intelligence (AI), but face difficulties in realizing it. A recent survey by the Boston Consulting Group has found that "companies face numerous challenges when implementing AI initiatives, with around 70% stemming from people- and process-related issues" (Boston Consulting Group 2024). This calls for a better understanding of how AI can be used in business processes.

Next to technical research, there is first research that explores how AI can contribute to the further development of Business Process Management (BPM). For example, Vidgof et al. (2023) analyze how Large Language Models (LLMs) can be used in BPM. However, their analysis is purely conceptual. Further, the paper by Vu et al. (2025) explores agentic BPM based on 22 interviews with BPM practitioners. Thus, research in this area is either conceptual or focuses on one AI technology, only. To address this limitation, we pose the following research question.

How can Artificial Intelligence be used for Business Process Management?

We answer this question through a qualitative-inductive study design. In particular, we conducted nine semi-structured interviews (Myers 2019) with experts in the area of AI and BPM. We analyzed the interviews against the BPM Lifecycle (Dumas et al. 2018) and identified for each of its phases multiple application scenarios. Our research offers implications for BPM in a world where AI drives more and more aspects of business.

The remainder of this paper is organized as follows. First, we summarize related work. Second, we explain our methodological approach. Third, we present our findings. Fourth, discuss them. Finally, we conclude this paper by summarizing its key contributions to the body of knowledge.

2. Related Work

2.1. The Business Process Management Lifecycle

Our analysis is structured along the traditional BPM Lifecycle, a key framework that outlines the stages of managing and improving business processes (Dumas et al. 2018). The lifecycle provides a systematic approach, beginning with *Process Identification*, where an organization's business processes are defined, categorized, and prioritized for improvement. This is followed by *Process Discovery*, where the "as-is" state of a selected process is documented, often through techniques like interviews, workshops, or process mining. In the *Process Analysis* phase, the as-is model is examined to identify weaknesses and their root causes. Based on these insights, the *Process Redesign* phase aims to create an improved "to-be" process model. Subsequently, *Process Implementation* involves managing the organizational change and technical development required to put the new process into practice. Finally, *Process Monitoring and Control* ensures that the redesigned process performs as expected by tracking relevant data, allowing for continuous improvement.

2.2. AI in Business Process Management

So far, BPM research has mainly integrated AI into technical tasks like cycle time forecasting (Gdowska et al. 2025), but without pointing to new application scenarios. Several conceptual frameworks offer high-level views. For instance, Dumas et al. (2023) characterize AI-augmented systems as adaptable and context-sensitive, while Rosemann et al. (2024) identify three essential drifts: a shift from automation to autonomization, from simplification to sophistication, and from transformational to conversational interaction. Similarly, Feuerriegel et al. (2024) list potential uses for generative AI, such as automating routine tasks and supporting idea generation, but these frameworks often lack empirical validation.

Vidgof et al. (2023) provide a detailed blueprint for integrating LLMs into BPM. They conceptualize how conversational agents can elicit knowledge and use semantic queries to generate process models from unstructured inputs. Central to their proposal is a prompt engineering approach that weaves together various data sources. Their framework, however, remains theoretical, and they note challenges like factual inaccuracies.

An empirical study on agentic BPM by Vu et al. (2025) found that practitioners expect efficiency gains but also have concerns about bias, opaque decision logic, and the need for human-in-the-loop controls. Their work, however, focuses narrowly on Robotic Process Automation (RPA) style task automation. To this point, research has remained either highly conceptual or focused on specific tools, detached from broader practical application. Our study

addresses this gap by empirically investigating AI opportunities across all phases of the BPM lifecycle.

3. Method

To investigate how AI can be used for the management of business processes, we employ a qualitative-inductive study design (Myers 2019). In particular, we conducted interviews with experts at the intersection of BPM and AI to understand how this technology changes how we approach the management of business processes. This approach is suitable, as there is little knowledge on how AI will impact BPM and is in line with previous research, e.g. in the area of process mining (Badakhshan et al. 2022). We next describe our data collection, followed by our data analysis.

3.1. Data Collection

We conducted nine semi-structured in-depth interviews (Myers 2019) with domain experts. We recruited the interview participants through the professional network of the authors. In particular, we were looking for people who have expertise at the intersection of Artificial Intelligence and BPM. Among others, the interviewees included product managers, product developers, and researchers in residence.

All interviews were guided by a semi-structured interview questionnaire (Myers 2019) that we devised in the author team. The questionnaire was organized around three key topics: 1) the personal background and experience of interviewees, 2) a short discussion of the BPM lifecycle, 3) current limitations of and the potential of AI for the phases of the lifecycle, and 4) the overall impact of AI on BPM.

The first interview was conducted by the core author team. All other interviews were conducted individually. All interviews lasted about an hour and were subsequently transcribed.

3.2. Data Analysis

We analyzed the interview data using the MAXQDA Software for qualitative analysis. Based on the first interviews, the first two authors derived a coding scheme that was then applied to the remaining data. The coding scheme was flexible, such that it allowed for the emergence of new codes and the modification of existing codes as far as necessary.

The coding scheme was structured along the BPM lifecycle. For each of the lifecycle phases, we identified 1) challenges and 2) opportunities of AI. Furthermore, the coding scheme included codes regarding how AI changes overarching aspects of BPM, such as how strategic aspects change through the introduction of AI. The second author then applied the coding scheme and regularly discussed codes and insights with the first author. Due to confidentiality agreements, the full interview transcripts and detailed coding cannot be made publicly available. We present our findings next.

4. Findings

In this section, we present our findings. First, we describe the challenges that we identified in contemporary BPM. Second, we present opportunities for AI in BPM. Third, we discuss the strategic implications of AI for the management of business processes.

Tab. 2: Summary of Findings Source: Authors

Phase	Limitations	Opportunities
Process Identification	- Dispersed information hinders finding novel processes. "Cold start" challenge with ill-defined processes. - Overreliance on poor data and vague inputs.	- Analyze unstructured data to find process patterns. - Generate starting points from vague inputs. - Identify rare or high-impact processes.
Process Discovery	- Inconsistent interview details - Focus on idealized, not actual, processes. - Undocumented critical knowledge. - Technical hurdles	- Automate knowledge capture (e.g. chatbots). - Convert human explanations/actions to models. - Act as a reliable domain knowledge repository.
Process Analysis	- Time-consuming manual data interpretation. - Scarcity of specialized resources. - Gap between analysis & actionable improvements. - Reliance on unvalidated expertise.	- Provide deeper, actionable insights. - Automatically highlight key patterns. - Offer proactive analysis with explanations. - AI-enhanced simulation.
Process Redesign	- Difficulty assessing organizational feasibility. - High complexity of change management. - Siloed knowledge prevents idea reuse. - Hindrance from legacy systems.	- Act as a collaborative partner. - Proactively suggest improvements. - Use what-if modeling to show change impact. - Translate natural language to formal models.
Process Implementation	- Difficulty translating design to operations. - Poor management of human factors. - Lack of organizational buy-in. - Challenges reconfiguring IT systems.	- Auto-generate training materials. - AI agents execute IT configurations or create apps. - Cognitive RPA and autonomous execution.
Monitoring & Controlling	- Inadequate training. - Integrating heterogeneous data. - Ineffective conformance checking. - Reliance on symptom-based KPIs. - User information overload. - Real-time monitoring is often unfeasible.	- Provide continuous, intelligent oversight. - Holistic conformance checking. - Real-time & predictive monitoring. - Proactive alerts & autonomous root-cause analysis.

4.1. Identification of current Limitations

The interviews highlighted the pervasive limitations of BPM, mainly the compartmentalization of process knowledge and data within organizational silos, which hinders end-to-end visibility and leads to incomplete process views. Current practices often rely on manual, resource-

intensive methods that are costly and cannot be scaled. A narrow focus on isolated processes, neglecting interdependencies, further fragments understanding. Persistent data quality issues and difficulties in interpreting complex data also impede effectiveness, alongside misalignments from varied stakeholder perspectives and documentation purposes. These general challenges compound phase-specific issues.

In *Process Identification*, discovering novel or cross-functional processes is hampered by dispersed information and the "cold start" challenge with ill-defined processes. Overreliance on inadequate system data and vague user input further complicate initial efforts.

During *Process Discovery*, inconsistent interview details and a workshop focus on idealized rather than actual processes impede coherent understanding. Much critical knowledge remains undocumented, while technical hurdles, such as complex Extract, Transform, and Load pipeline creation for process mining, persist.

The *Process Analysis* phase is marked by time-consuming manual data interpretation and a scarcity of specialized resources for techniques such as process mining. A significant gap remains in translating analytical findings into actionable improvements, compounded by the reliance on deep, yet often unvalidated domain expertise.

For *Process Redesign*, key challenges include assessing the organizational feasibility of changes and managing the complexity of the change. Siloed knowledge that inhibits idea reuse, coupled with legacy systems and difficulties in validating redesign efficacy, further impede progress.

Process Implementation faces hurdles in translating designs to operations and managing human factors. Securing organizational buy-in, reconfiguring IT systems, and mitigating unforeseen ripple effects are critical, often compounded by inadequate training.

Process Monitoring and Controlling is challenged by integrating heterogeneous data and effective conformance checking. Reliance on symptom-based Key-Performance-Indicators, user information overload, and the difficulty of isolating change limit its utility, with real-time monitoring often being unfeasible.

4.2. Identification of Opportunities for AI in the BPM lifecycle

Our interviews with practitioners indicate a strong consensus on the transformative impact of AI. Key anticipated benefits include managing greater complexity, automating routine and cognitive tasks, democratizing BPM through enhanced knowledge accessibility, and accelerating time-to-value.

In *Process Identification*, AI is expected to enhance the recognition and prioritization of processes by analyzing unstructured data, such as emails and documents, to automatically surface potential process patterns. Practitioners highlighted AI's ability to generate viable starting points from vague user inputs and expand focus beyond high-frequency processes to rare or high-impact ones. Such advancements are supported by techniques like natural language processing (NLP), LLMs, and generative AI.

For *Process Discovery*, practitioners foresee AI substantially automating knowledge capture. This includes using chatbots for initial stakeholder interviews and converting human explanations or on-screen behaviors directly into structured process models. AI is also seen as a reliable domain knowledge repository capable of inferring logic from poorly documented

solutions. These capabilities are expected to be based on LLMs, task mining, and information extraction.

During *Process Analysis*, AI is poised to provide deeper, more actionable insights. It can automatically highlight salient patterns in complex process maps and offer proactive analysis with clear explanations, empowering generalist users. AI-enhanced simulation, incorporating intelligent agents, is anticipated to overcome challenges in deriving meaningful analytical outcomes. These functionalities are supported by techniques such as AI-enhanced process mining, LLMs, and AI agents.

In *Process Redesign*, AI is envisioned as a collaborative partner. It can proactively suggest improvements based on organizational experience and best practices, and use what-if modeling to communicate the potential effects of changes. Practitioners also foresee AI presenting various change options with pros and cons, and translating natural language specifications into formalized process models. This phase is facilitated by simulation, LLMs, and generative AI.

For *Process Implementation*, AI is considered crucial for operationalizing redesigned processes. It can generate training materials and documentation, and AI agents could execute configuration changes in IT systems or automatically create ad-hoc applications from process models. Cognitive RPA and autonomous process execution were also prominent themes. LLMs, AI agents, and intelligent automation platforms are central to these advancements.

Within *Process Monitoring and Controlling*, AI is expected to provide continuous, intelligent oversight. It can offer a holistic view for conformance checking by analyzing both structured logs and unstructured content. This is complemented by real-time and predictive monitoring capabilities that provide proactive insights, intelligent alerts, and autonomous root-cause analysis. Core techniques include generative AI, anomaly detection, and causal inference.

4.3. Impacts of AI on Overall BPM

AI is fundamentally reshaping BPM by converting raw data into actionable intelligence, thereby breaking down organizational silos. This shift democratizes BPM by lowering entry barriers for business users and evolving experts' roles toward strategic oversight, with a continued need for human-in-the-loop governance. A key paradigm shift is toward agent-based interaction, where users engage intelligent agents rather than complex tools, shifting focus from lifecycle steps to business outcomes. BPM systems themselves may evolve into intelligent coaching platforms or even systems that generate processes based on specified goals. This evolution requires robust data and knowledge integration to fuel AI insights, with governance over these knowledge bases remaining paramount for trustworthy performance. New operational paradigms are also emerging, with AI acting as a "sparring partner" for strategic validation and enabling a move towards higher abstraction focused on outcomes over explicit steps. However, significant challenges persist, including managing AI's developmental trajectory, the risk of introducing new complexities, and critical ethical considerations regarding security, transparency, and accountability.

5. Implications

Our research offers three key implications for research. First, we identify current limitations in the phases of the BPM lifecycle (Dumas et al. 2018). In bringing these issues to light, we represent practitioners' perspectives and identify challenges for future research to explore.

Second, we identified opportunities for AI in each phase of the BPM lifecycle (Dumas et al. 2018). For example, in process redesign, AI can be used to generate process improvement ideas. Academics researching technical aspects of BPM and software developers can build on these opportunities and develop respective software artifacts and solutions.

Third, we examined how AI changes BPM on a general level. Among other things, we found that AI can contribute to the democratization of BPM, as fewer skills will be required to conduct BPM-related tasks. AI is further expected to reduce barriers between departments as knowledge becomes more easily accessible. This makes the execution of BPM programs more feasible.

6. Conclusion

In this paper, we reported on an interview-based study that examines the potential of AI for the management of business processes throughout the phases of the BPM lifecycle. Most notably, AI can be used to generate process improvement ideas and leads to a democratization of BPM. Future research can build on our study to further identify and assess the impact of AI on the management of business processes.

Disclosure of Interests. Peyman Badakhshan and Syeda Noor Zehra Naqvi are employed by SAP Signavio.

7. References

- Badakhshan, P., Wurm, B., Grisold, T., Geyer-Klingenberg, J., Mendling, J. & vom Brocke, J. (2022): Creating business value with process mining. In: *The Journal of Strategic Information Systems* 31(4), 101745.
- Boston Consulting Group (BCG). (2024): *Ai adoption report 2024: Struggles in achieving and scaling value*. Retrieved from <https://www.bcg.com/press/24october2024-ai-adoption-in-2024-74-of-companies-struggle-to-achieve-and-scale-value>.
- Dumas, M., Rosa, M.L., Mendling, J. & Reijers, H.A. (2018): *Fundamentals of Business Process Management, Second Edition*. Springer.
- Dumas, M., Fournier, F., Limonad, L., Marrella, A., Montali, M., Rehse, J.R., Accorsi, R., Calvanese, D., De Giacomo, G., Fahland, D., Gal, A., La Rosa, M., Völzer, H. & Weber, I. (2023): Ai-augmented business process management systems: A research manifesto. In: *ACM Transactions on Management Information Systems* 14(1).
- Feuerriegel, S., Hartmann, J., Janiesch, C. & Zschech, P. (2024): Generative ai. In: *Business & Information Systems Engineering* 66(1), 111-126.
- Gdowska, K., Gómez-López, M.T. & Rehse, J. (eds.) (2025): *Business Process Management Workshops BPM 2024 International Workshops, Krakow, Poland, September 1-6, 2024, Revised Selected Papers, Lecture Notes in Business Information Processing, vol. 534*. Krakow, Poland: Springer.
- Myers, M.D. (2019): *Qualitative Research in Business and Management*. London, UK: SAGE Publications Ltd.

- Rosemann, M., Brocke, J.v., Van Looy, A. & Santoro, F. (2024): Business process management in the age of ai three essential drifts. In: *Information Systems and e-Business Management*, 1-15.
- Vidgof, M., Bachhofner, S. & Mendling, J. (2023): Large language models for business process management: Opportunities and challenges. In: *Proceedings of the International Conference on Business Process Management*. Springer, 107-123.
- Vu, H., Klievtsova, N., Leopold, H., Rinderle-Ma, S. & Kampik, T. (2025): Agentic business process management: The past 30 years and practitioners' future perspectives. In: *arXiv preprint arXiv:2504.03693*.

Community Sessions: An Overview of the SAP University Alliances Learning Portfolio with different focus: Business Process Management and Application Development and Coding

Philip Träutlein, Yuri Kupriianov
SAP SE

Abstract: These two community sessions aim to provide a comprehensive overview of the SAP University Alliances Learning Portfolio, with a particular emphasis on Business Process Management (BPM) and Application Development and Coding. Educators will gain insights into the various learning offerings designed to meet diverse academic use cases.

During both sessions, we will cover free learning offers, including exploration materials that introduce different cloud solutions, providing a solid foundation for educators and students alike. Additionally, we will highlight hands-on learning opportunities offered through self-paced content formats, designed for practical learning. The sessions will also cover academic experiences that involve utilizing live systems for hands-on practice and experiential learning.

Furthermore, we will explore low-cost academic packages supported by the University Competence Center (UCC) and Academic Competence Centers (ACC) partners. These packages include system and content support, offering a more integrated and supported learning experience at a low cost.

A key focus of the first session in the BPM Track will be the teaching and learning portfolio dedicated to Business Process Management (BPM). We will discuss the resources available to enhance your teaching capabilities and student outcomes in the area of BPM. This portion of the session will provide detailed information on the available resources, their implementation, and how they can benefit your teaching practices.

The key focus of the session in the Technology Track will be on the curriculum and resources available for application development and coding. We will highlight the latest tools and techniques to ensure you are equipped to stay at the forefront of technological advancements. This segment will cover the implementation of these resources and how they can enrich your teaching practices and prepare your students for the future.

These sessions offer detailed information on the available resources and their implementation, emphasizing practical insights and the benefits for educators. Join us to gain a deeper understanding of the diverse offerings within the SAP University Alliances Learning Portfolio.

We look forward to your participation

Global Bike mit SAP Signavio und SAP LeanIX

Hans Scheruhn¹, Milan Schulze², Jan Langhof¹
Hochschule Harz, Deutschland; 2: SAP SE

Abstract: „BPM drives Digital Transformation“: Um das Motto der SAP ACC2025 mit Leben zu erfüllen, müssen Business Process Management (BPM), Application Life Cycle Management (ALM) und Enterprise Architecture Management (EAM) bei einer Umsetzung mit einem ERP-System eng einander greifen. Dieser praktische Beitrag soll am Beispiel der in der SAP University Alliance (UA) weit verbreiteten Demo-Firma „Global Bike Group“ im Kontext des Enterprise Online Guide (EOG) detailliert aufzeigen, wie das BPM-Tool SAP Signavio gemeinsam mit dem EAM-Tool SAP LeanIX in der Lehre und Forschung strukturiert einsetzbar sind. Als Ausgangspunkt dient dabei die seit vielen Jahren bewährte Lead-to-Cash (L2C)-Fallstudie aus dem Bereich Sales and Distribution (SD) des SAP University Competence Center (UCC) Magdeburg mit SAP S/4HANA. Zusätzlich unterstützt der erstellte Prototyp auch „The Bike Company“ der SAP Learning Journey aus dem aktuellen Angebot der SAP University Alliance (UA) sowie die Modelfirma „EOBikes“ des SAP Value Accelerators.

Zunächst erweitert der Ansatz im Kontext von EOG eine Visualisierung der horizontalen Navigation im SAP LeanIX Relationsexplorer um eine vertikale Ausprägung. In beiden Fällen folgt der Ansatz dem Meta-Model V4.0 von LeanIX für alle Komponenten des „Inventories“ und setzt die nunmehr zweidimensionale Navigation ähnlich einem „Google Maps for Enterprise“ anschaulich um. Dieses erleichtert die Nutzung von LeanIX noch einmal deutlich und lässt sich auch von den beteiligten Stakeholdern aus den Businessbereichen durch eine gemeinsame Sprache zielführend in gemeinsamen SAP LeanIX „Reports“ und „Diagrams“ einsetzen.

Zusätzlich zeigt der Prototyp für alle 8 EOG-Maps auf jeweils 4 Zoomstufen eine Integration bzw. Aufteilung mit SAP Signavio. Aus einer EOG-Übersichtskarte heraus kann der Nutzer geführt (Guide) zwischen SAP LeanIX und SAP Signavio wechseln um die Beispielunternehmen aus allen Sichten und aus unterschiedlichen Detaillierungsgraden zu erkunden. Im nächsten Schritt können sie dann ein neues Unternehmen anteilig in den beiden Tools abbilden.

Eine erfolgreiche Umsetzung konnte mit ca. 10 Studierenden des dualen Studienganges IT-Management im Fachbereich Verwaltungswissenschaften an der Hochschule Harz im SoSe2025 durchgeführt werden. Dabei kamen die für die SAP UA freigegeben Academic Versionen von SAP Signavio und SAP LeanIX gemeinsam zum Einsatz.

Brücken bauen zwischen Theorie und Praxis: Erfolgreiche Integration von SAP LeanIX in Lehre und Forschung am Beispiel eines S/4-Vorprojekts

Max Graf von Hardenberg¹, Tobias Bretschneider², Tim Müller²

¹: msg systems AG, ²: Hochschule Harz

Abstract: Um das Motto der SAP Academic Community Conference 2025 – „BPM Drives Digital Transformation“ – mit Leben zu füllen, ist eine enge Verzahnung von Business Process Management (BPM), Application Lifecycle Management (ALM) und Enterprise Architecture Management (EAM) im Kontext von ERP-Systemen essenziell. Was in der Praxis erfolgreicher SAP-Beratungsunternehmen bereits gelebt wird, ist in vielen Hochschulen noch nicht vollständig etabliert.

Dieser Beitrag präsentiert ein gelungenes Kooperationsprojekt zwischen der Hochschule Harz und der msg Group, das aufzeigt, wie sich Forschung und Lehre im SAP-Umfeld praxisnah verbinden lassen. Ziel ist es, Studierende frühzeitig in reale Transformationsprojekte einzubinden. Im Rahmen von Lehrveranstaltungen und anschließenden Werkstudententätigkeiten wirken Studierende aktiv an der Gestaltung von sogenannten „Vorprojekten“ mit, bei denen SAP LeanIX und reale Kundenanforderungen im Mittelpunkt stehen.

Ausgangspunkt bildet die Nutzung von Business Capabilities, die mithilfe des Enterprise Online Guide (EOG) und des neuen Ansatzes der msg effizient mit SAP Best Practices verknüpft werden. In einem konkreten Projekt gelang es Studierenden des Fachbereiches Automatisierung und Informatik, eine strukturierte Verbindung zwischen SAP LeanIX, Signavio und den SAP Best Practices herzustellen. Dabei wurden Business Capability Reference Contents genutzt, um eine inhaltliche Gegenüberstellung und Konsolidierung zu ermöglichen.

Das Projekt zeigt beispielhaft, wie Hochschulen durch enge Zusammenarbeit mit der Wirtschaft einen substanziellen Beitrag zur digitalen Transformation leisten können.

Erfolgreiche Zusammenarbeit Hochschule Harz mit msg und Siemens in Forschung und Lehre im Bereich BPM und SAP S/4HANA

Christian Reiter¹, Ulrich Förster², Hans-Jürgen Scheruhn³

¹: msg; ²: Siemens; ³: Hochschule Harz

Abstract: „BPM drives Digital Transformation“: Um das Motto der SAP ACC2025 mit Leben zu erfüllen, müssen Business Process Management (BPM), Application Life Cycle Management (ALM) und Enterprise Architecture Management (EAM) bei einer Umsetzung mit einem ERP System eng einander greifen. Was von erfolgreichen Beratungs-Unternehmen im Bereich SAP heute täglich gelebt werden muss, steht vielen Hochschulen möglicherweise noch bevor. Dieser praktische Beitrag soll am Beispiel der Hochschule Harz im Zusammenschluss mit den Unternehmen msg Group und Siemens eine fast 30-jährige erfolgreiche Zusammenarbeit im Bereich Forschung und Lehre aufzeigen. Zahlreiche Empfehlungen mit „Dos“ und „Dont's“ sollen dabei helfen, auch an anderen Hochschulen bestehende oder neue Partnerschaften mit Unternehmen nachhaltig in Richtung einer stärkeren Integration von BPM und SAP S/4HANA auszubauen.

Als Ergebnis haben die Autoren ein Business Modell für die Hochschule Harz aus Sicht eines Lehrenden und Forschenden entwickelt. Mit diesem - jedem Start-up Unternehmen bekannten - Ansatz können sich Hochschulen als Ordnungsrahmen auf 9 typische Bestandteile konzentrieren.

Hochschulen vermitteln als einer der sogenannten „Schlüsselaktivitäten“ wie der „Lehre“ bewährte Konzepte, z.B. im Kontext der Konferenz, und können mit SAP -Lösungen deren Relevanz praktisch nachweisen. Sie können im Zuge einer möglichen digitalen Transformation von Forschung und Lehre aber zukünftig auch noch stärker Vordenker sein und leichter eigene (digitale Produkte) als „Wertbeitrag“ für Ihre Kunden entwickeln.

Unser Business-Modell hebt deutlich unsere Studierenden als „Kunden“ hervor. Das gilt besonders für die an Hochschulen angesiedelten Promotionsstudiengängen. Unsere „Kunden“ bringen sich mit ihrem Commitment und einem wertvollen Teil ihrer Lebenszeit in unsere Hochschulen ein.

Weiterhin wird deutlich, wie Erfolg nur mit „Schlüsselpartnern“ in einem erweiterten Eco System gelingen kann. In unserem Business Modell wird klar, wie diese (z.B. msg und Siemens) sich in einem geregelten Ordnungsrahmen letztlich zum Wohle der Studierenden in allen 9 Bestandteilen einbringen.

Das Business Modell beantwortet auch die Frage, wie Hochschulen ihre „Kundenbeziehungen“ pflegen können. Der schnelle Wandel des Wissens und der KI-Wissensträger erfordert immer stärker ein gemeinsames Lernen mit den Studierenden, z.B. im Rahmen von Abschlussarbeiten. Das sollte immer auf „Augenhöhe“ erfolgen, z.B. auch durch den gemeinsamen Besuch von eigenen Konferenzen am Hochschulstandort oder der SAP ACC2025.

Auch die „Marktkanäle“ zum Erreichen unserer Studierenden kann man dem Business Model entnehmen. Hier spielt im Kontext einer digitalen Transformation die Nutzung eines erweiterten Eco-System mit Studierenden und Partnern eine zunehmende Rolle.

Bei der „Kostenstruktur“ verdeutlicht das Business Model die enorme Wichtigkeit einer Einhaltung der Compliance, unterstützt durch professionell aufgestellte Dezernate z.B. im Bereich „Personal“ und „Haushalt“ einer Hochschule. Diese machen auch einen wichtigen Teil unserer „Schlüsselressourcen“ aus. Die Zuordnung der Kosten einer Ausgründung oder einer Konferenz im öffentlichen Bereich sind nicht zu unterschätzen. Das Ganze wird getragen von dem Commitment jedes Mitwirkenden.

Einnahmen sind zunächst wiederum das Commitment der (Promotions-)Studierenden, die sich mit einem wertvollen Teil ihrer Lebenszeit in unsere Hochschulen einbringen als eigentliche Währung der Studiengebühren. Daraus können Hochschulen vorzeigbare Auftragsforschung gemeinsam mit ihnen entwickeln, gute Drittmittelprojekte mit geeignetem Personal besetzen, leichter die Hürden von DFG-Großgeräteanträgen nehmen und letztlich ein gesuchter Partner für Drittmittel-, Auftragsforschung und Sponsoring-Geber werden.

Eine besondere Anerkennung erfahren die Autoren aktuell durch die Bereitstellung ihres „Wertebeitrages“ in den SAP Value Accelerator, welcher ab sofort auch außerhalb der Hochschule sichtbar und nutzbar wird.

An Evaluation of Process Mining Techniques for People Analytics

L. Buck & H. Sperling
University of Mannheim, Mannheim, Germany
SAP SE, Walldorf, Germany

Abstract: Process mining (PM) techniques have grown increasingly popular in both academia and industry practice. While researchers have extended their application from business to human-centric processes, the domain of people analytics has so far received scarce attention in the PM literature. Building on existing work in educational process mining, this ongoing case study interprets professional careers as sequences to evaluate the potential of PM to extract insights for people analytics. We systematically analyze how PM can help fulfill the analytics needs of the hiring, career development, and retention phases in the employee lifecycle. This involves the creation of event logs from HR master data, mapping career data attributes to predefined formats, as well as exploring the application of process discovery and conformance algorithms. Preliminary results indicate significant potential of PM for people analytics.

Keywords: Process mining, careers, people analytics.

1. Introduction

People analytics is essential for effective business operations. Amidst the changing nature of work, analyzing employee career data remains instrumental to meet shifting workforce needs and evolving business contexts (Mirchandani, 2025). Nonetheless, approaches to people analytics remain limited (Angrave et al., 2016). Traditionally, analysis of employee career data leverages techniques related to machine learning and optimal matching analysis (Biemann et al., 2020). While effective for prediction and comparison purposes, these methods often overlook the temporal and procedural nature of career development (Akasheh et al., 2024). With an explicit focus on procedural elements, process mining (PM) offers a powerful tool to extract insights from a wide array of processes. Traditionally involving product-oriented or service-based data from business processes, researchers are increasingly evaluating the application of PM for domains outside operational workflows (Garcia et al., 2019). For example, PM has proven valuable for the analysis of patient flows in the healthcare domain (Mans et al., 2015) or educational pathways to analyze student and learning success (Potena et al., 2024). Since people analytics involves quantitatively analyzing employees' digital traces for data-driven management (Gal et al., 2020), PM offers potential for application in this domain.

2. Related Work

PM aims to compare event data, which refers to actual observed behavior, to process models, which are either manually created or automatically discovered (van der Aalst, 2016). Research increasingly explores the application of PM within organizational contexts, highlighting its cross-industry relevance (Eggers et al., 2021). This includes the management of Human Resources (HR) processes, with recent literature on PM revealing a growing focus on human-

centric approaches. Central themes include resource allocation, work satisfaction, employee behavior, and educational PM (Goel et al., 2023; Hotel et al., 2020; Kim et al., 2022). Educational PM addresses learner-centric trajectories, focusing on monitoring, evaluating, and recommending individualized learning strategies in educational contexts (Potena et al., 2024; Wambsganss & Schmitt, 2024).

Social scientists have studied careers using techniques such as optimal matching, predictive machine learning, and deep learning. These approaches are used to track transitions across roles and industries and to forecast outcomes such as turnover and career success (Akasheh et al., 2024; Biemann et al., 2020). These techniques allow a thorough understanding of the dynamics and behaviors underlying careers across industries and backgrounds, but often lack interpretability and fail to focus on the process- and time-based nature of career trajectories. Interpreted as career sequences, career paths are individual successions of occupational jobs (Joseph et al., 2012). Given this procedural nature of career data, the application of PM presents a promising avenue for analysis, addressing limitations within current approaches to career analytics. However, its application in that context has so far remained limited to student careers (e.g., Cameranesi et al., 2017).

3. Case Study Design and Preliminary Results

This ongoing case study follows a three-step process to explore the application of PM techniques on career data in people analytics. In the first step, we collaborated with a multinational company based in Germany to understand real-world HR master data structures. We then manually constructed a small dummy data set of sample career sequences based on real-world HR master data structures. Focusing on master data as a basis for people analytics ensures scalability and generalizability of our approach, as HR master data structures tend to be similar across organizations. To enable more robust and convincing experimental analysis using a realistic, high-volume synthetic dataset, we then constructed a larger career dataset using TimeGAN data synthetization, a technique that allows the synthetization of time-series data while preserving temporal dynamics as well as feature relationships (Yoon et al., 2019). Our final dataset comprises 2015 employee career records spanning five years, including a unique employee identifier and yearly job information, as summarized in Tab. 1. Alignment of career sequences to a common starting point enables cohort-based comparisons of careers.

Tab. 1: Variable Overview

Variable	Type	Description/format
Employee_ID, Gender, Country, Job_Function	Categorical - Nominal	Female/male (int); country (string); business function (string)
Timestamp	Temporal	YYYY-MM-DD (date)
Year_of_Birth, Tenure	Numerical - Discrete	YYYY (int), Tenure in years (int, 1-23)
Job_Level	Categorical - Ordinal	Hierarchical with career type information (string, M – Managerial, E – Expert)

In the second step, we transformed our dataset into event logs in XES format, suitable for the application of PM techniques. Tab. 2 details the mapping schema used to translate the synthetic HR dataset into event log formats suitable for different HR analyses. In the last step, we apply PM techniques to the career sequences contained in the event logs and evaluate the outcome.

Specifically, we employ variant analyses, transition analyses, conformance checking, and bottleneck analyses.

Tab. 2: XES key to HR data mapping.

XES key	HR variable
trace:concept:name	Employee_ID
event:concept:name	Job and/or Job Level
time:timestamp	Timestamp

The highlighted path in Fig. 1 depicts a software engineering career with job level promotions after year 2 and 4. In the first year of observation (Year 1), 17 individuals joined as job level 1 software engineers. During Year 2, nine employees remain in the same Job_Level and Job_Function, while others transition to other jobs or turn over. This approach enables the tracking of behaviors following the entry into an organization and facilitates the assessment of the appropriateness of initial hiring decisions, as well as subsequent internal movements across job functions.

4. Discussion and Future Work

Depending on the assignment of unique event identifiers to the HR variables, PM allowed us to track employee careers across both job and job levels. Assigning the unique event identifiers to the variable *Job_Function* and subsequently applying process exploration allowed us to derive insights into the distribution of employees across job functions as well as the complexity of career paths. Moreover, variant analyses resulted in the identification of common career path sequences within the organization to provide insights into frequent progression patterns and typical career trajectories. The mere visualization of hierarchical job levels using PM provided limited value, since hierarchical movements tend to follow a discrete and continuous route.

However, analyzing job levels in conjunction with career types provided insights into potential promotion differences between different careers. In the course of this ongoing case study, we plan to use variant analysis as a potentially well-suited PM technique to reveal temporal patterns in promotion behavior by providing insights into the time to promotion between Job_Levels. Similarly, conformance checking might help in assessing the alignment of actual career paths with predefined promotion models, uncovering deviations or adherence. Transition analyses might be useful to quantify the timing and likelihood of promotions for each Job_Level transition.

To provide a generalizable approach for PM in people analytics, our ongoing research leverages real-world HR master structures that tend to be similar across organizations. However, not all organizations might track information on career progression and job tenure in a similar fashion. Future research should explore the application of PM on real-world corporate data in different organizational settings and incorporate the possibility of intra-year job or job level switches. Moreover, the incorporation of a skill-based framework may provide deeper insights into career progression dynamics and can address limitations associated with the analysis of hierarchical features such as Job_Levels. This opens up new opportunities to understand career mobility on

an organizational as well as an individual level. Last, the integration of career data with process analytics tools such as SAP Signavio provides a promising avenue for future study.

5. References

- Angrave, D., Charlwood, A., Kirkpatrick, I., Lawrence, M., & Stuart, M. (2016). HR and analytics: Why HR is set to fail the big data challenge. *Human Resource Management Journal*, 26(1), 1–11. <https://doi.org/10.1111/1748-8583.12090>
- Aysolmaz, B., Nemeth, M., & Iren, D. (2021). A method for objective performance benchmarking of teams with process mining and DEA. *ECIS 2021 Proceedings*.
- Biemann, T., Mühlenbock, M., & Dlouhy, K. (2020). Going the distance in vocational behavior research: Introducing three extensions for optimal matching analysis based on distances between career sequences. *Journal of Vocational Behavior*, 119, 103399. <https://doi.org/10.1016/j.jvb.2020.103399>
- Eggers, J., Hein, A., Böhm, M., & Krcmar, H. (2021). No Longer Out of Sight, No Longer Out of Mind? How Organizations Engage with Process Mining-Induced Transparency to Achieve Increased Process Awareness. *Business & Information Systems Engineering*, 63(5), 491–510. <https://doi.org/10.1007/s12599-021-00715-x>
- Cameranesi, M., Diamantini, C., Genga, L., & Potena, D. (2017). Students' careers analysis: a process mining approach. *Proceedings of the 7th International Conference on Web Intelligence, Mining and Semantics (WIMS '17)*. <https://doi.org/10.1145/3102254.3102270>
- Gal, U., Jensen, T. B., & Stein, M.-K. (2020). Breaking the vicious cycle of algorithmic management: A virtue ethics approach to people analytics. *Information and Organization*, 30(2), 100301. <https://doi.org/10.1016/j.infoandorg.2020.100301>
- Garcia, C. D. S., Meinheim, A., Faria Junior, E. R., Dallagassa, M. R., Sato, D. M. V., Carvalho, D. R., Santos, E. A. P., & Scalabrin, E. E. (2019). Process mining techniques and applications – A systematic

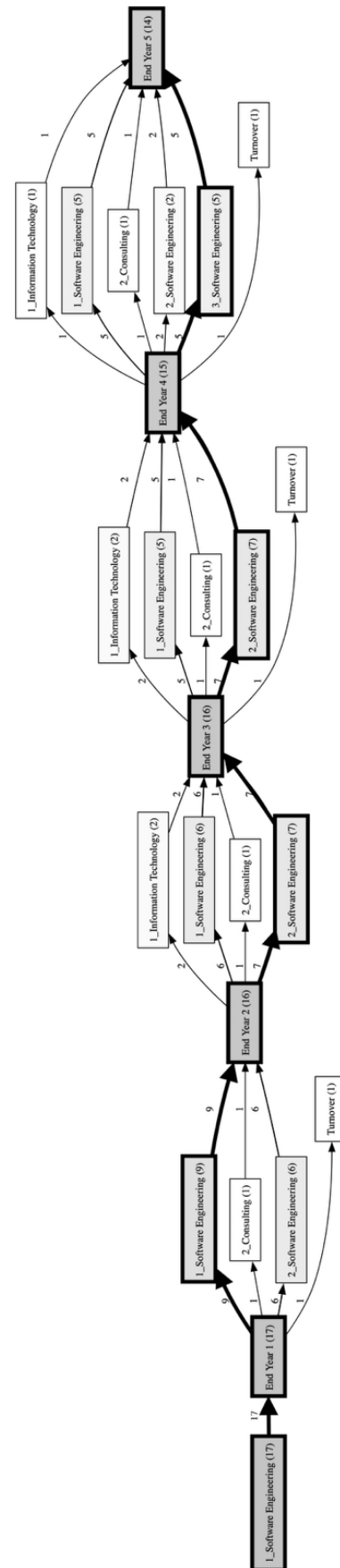


Fig. 1: PM visualization of the careers of newly hired junior software engineers.

- mapping study. *Expert Systems with Applications*, 133, 260–295.
<https://doi.org/10.1016/j.eswa.2019.05.003>
- Goel, K., Fehrer, T., Röglinger, M., & Wynn, M. T. (2023). Not Here, But There: Human Resource Allocation Patterns. In C. Di Francescomarino, A. Burattin, C. Janiesch, & S. Sadiq (Eds.), *Business Process Management* (Vol. 14159, pp. 377–394). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-41620-0_22
- Hofman, M., Grela, G., Orzelska, P., & Banaś, J. (2024). Adapting to Change: Employees Ambidexterity as a Driver for Operational Adaptability and Organizational Development. In C. Di Ciccio, W. Fdhila, S. Agostinelli, D. Amyot, H. Leopold, M. Krčál, M. Malinova Mandelburger, G. Polančič, K. Tomičić-Pupek, K. Gdowska, T. Grisold, P. Sliž, I. Beerepoot, R. Gabryelczyk, & R. Plattfaut (Eds.), *Business Process Management: Blockchain, Robotic Process Automation, Central and Eastern European, Educators and Industry Forum* (Vol. 527, pp. 250–266). Springer Nature Switzerland.
https://doi.org/10.1007/978-3-031-70445-1_16
- Hotel, O., Gzara, L., Verjus, H., & Triaa, W. (2020). Competency Cataloging and Localization to Support Organizational Agility in BPM. In A. Del Río Ortega, H. Leopold, & F. M. Santoro (Eds.), *Business Process Management Workshops* (Vol. 397, pp. 60–69). Springer International Publishing. https://doi.org/10.1007/978-3-030-66498-5_5
- Joseph, D.; Boh, W. F.; Ang, S.; & Slaughter, S. A. (2012). The Career Paths Less (or More) Traveled: A Sequence Analysis of IT Career Histories, Mobility Patterns, and Career Success. *MIS Quarterly*, 36(2), 427–452. <https://doi.org/10.2307/41703462>
- Kim, J., Comuzzi, M., Dumas, M., Maggi, F. M., & Teinemaa, I. (2022). Encoding resource experience for predictive process monitoring. *Decision Support Systems*, 153, 113669.
<https://doi.org/10.1016/j.dss.2021.113669>
- Mans, R. S., Van Der Aalst, W. M. P., & Vanwersch, R. J. B. (2015). *Process Mining in Healthcare: Evaluating and Exploiting Operational Healthcare Processes*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-16071-9>
- Mirchandani, N. (2025, April 25). *Data-Driven HR: Redefining Business Success In A Volatile Market*. Forbes.
<https://www.forbes.com/councils/forbeshumanresourcescouncil/2023/04/25/data-driven-hr-redefining-business-success-in-a-volatile-market/>
- Potena, D., Genga, L., Basta, A., Mercati, C., & Diamantini, C. (2024). Evidence-Based Student Career and Performance Analysis with Process Mining: A Case Study. In J. De Smedt & P. Soffer (Eds.), *Process Mining Workshops* (Vol. 503, pp. 349–360). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-56107-8_27
- Tang, W., & Matzner, M. (2020). Creating Humanistic Value with Process Mining for Improving Work Conditions—A Sociotechnical Perspective. *ICIS 2020 Proceedings*.
- van der Aalst, W. (2016). *Process Mining*. Springer Berlin Heidelberg.
<https://doi.org/10.1007/978-3-662-49851-4>
- Wambsganss, T., & Schmitt, A. (2024). Enhancing Personalized Learning Through Process Mining: A Taxonomy and Design Patterns Approach. *Business & Information Systems Engineering*. <https://doi.org/10.1007/s12599-024-00901-7>

Einsatz von Signavio Reports in der Lehre

Uta Mathis

Hochschule Esslingen, Esslingen, Deutschland

Abstract: Dieser Beitrag analysiert den didaktischen Einsatz von Reports zur Prozessanalyse und -optimierung in einer Lehrveranstaltung für technische Betriebswirte mit Fokus auf SAP Signavio.

Grundlage bilden Methoden des Business Process Management (BPM) und betriebswirtschaftliche Kennzahlen zur Prozessanalyse und -verbesserung. Darauf aufbauend wird ein praxisorientierter Lehransatz vorgestellt, mit dem die Studierenden die Methoden nachvollziehen und anwenden können. Anhand einer Fallstudie wird vorgestellt, wie Signavio-Reports in der Lehre eingesetzt werden können. Dieser Beitrag ist eine Ergänzung des Beitrags der Hochschule Harz, in dem der Einsatz der Prozesssimulation in Signavio vorgestellt wird.

Gestaltung von Prozessen erfordert sowohl Methodenwissen als auch praktische Anwendung. Prozessberichte, die auf Prozessmodellen (BPMN oder EPK) beruhen, spielen hier eine zentrale Rolle. In diesem Beitrag wird an einem Beispiel gezeigt, wie ein Prozessmodell mit Signavio-Reports aus betriebswirtschaftlicher Sicht analysiert werden kann und die Studierenden angeregt werden können, Lösungsbeispiele zu entwickeln

Keywords: Business Process Management, Kennzahlen, Prozessanalyse, Signavio-Reports.

1. Einleitung

Die Vermittlung von BPM in der Lehre erfordert sowohl Methodenwissen als auch die Möglichkeit für die Studierenden, dieses abstrakte Wissen praktisch in einem Fallbeispiel anzuwenden.

Zur praktischen Durchführung von Prozessanalysen stehen verschiedene digital unterstützte Methoden wie z.B. Process Mining, Simulation und Reports zur Verfügung.

Der Fokus in diesem Beitrag liegt auf der Vermittlung von BPM in der Lehre aus Sicht der Betriebswirtschaft mit einem Fokus auf Kosten- und Ressourcenkennzahlen sowie der Betrachtung von Verantwortlichkeiten im Prozessablauf.

2. Theoretischer Hintergrund

Prozessmanagement ist ein Ansatz zur nachhaltigen Optimierung der Wettbewerbsfähigkeit von Unternehmen. Gestaltungsgegenstand sind die Geschäftsprozesse, die als End-to-End-Prozesse vom Kundenbedürfnis bis zur Befriedigung des Kundenwunsches definiert werden. Klassische Ziele, die dabei verfolgt werden, sind z.B. Schaffung von Transparenz, Standardisierung und Verschlankeung von Prozessen sowie Kostensenkung. Typische Maßnahmen, die dabei angewendet werden, sind Verkürzung von Durchlaufzeiten und Verringerung des Ressourcenverbrauchs durch Vermeidung von nicht wertschöpfenden

Tätigkeiten, Fehlervermeidung und Reduzierung von IT-technischen und organisatorischen Schnittstellen.(Gadatsch 2023; Dumas u.a. 2018)

Eine strukturierte Basis zur Prozessanalysen bilden Prozessmodelle, wie z.B. BPMN oder EPKs. Prozessmodellierung ist ein Lehrziel im Kontext der Lehreinheit von BPM, aber erst die Analyse von Attributen, wie z.B. Zeiten und Kosten, schafft eine Entscheidungsgrundlage für eine betriebswirtschaftliche Analyse und Bewertung von Prozessen. (Mendling 2014; Recker & Mendling 2016; Röglinger u.a. 2017; Dumas u.a. 2018, Ley u.a. 2012)

SAP Signavio bietet umfangreiche Prozessanalysemethoden Process Mining, Prozesssimulation (vergleiche Beitrag von Hans Scheruhn, Hochschule Harz) und Reports. In diesem Beitrag liegt der Fokus auf dem Einsatz von Signavio Reports. Diese gliedern sich in (SAP Signavio 2025):

- Analyseberichte (Kosten- und Ressourcenverbrauch)
- Matrizen (RACI, Handover, IT-Und Dokumentnutzung)
- Konventions- und Metrikprüfungen (Qualitätssicherung, Governance)

Die pädagogisch-didaktisch Grundlage bilden erfahrungsbasiertes und projekt-basiertes Lernen. (Costa et al 2020; Hommel 2019; Gerogiannis & Fitsilis 2006)

3. Methodik

Für die Integration von Reports in die Vorlesung Prozessmanagement wurde ein Lehrkonzept entwickelt, das aktivierende Lernmethoden unterstützt. Basis bildet ein Lernzyklus, der aus einer zweifachen Dialektik Aktion - Reflektion und Erfahrung – Abstraktion besteht und mehrfach durchlaufen werden soll (Kauffeld & Othmer 2019; Schmidpeter & Kolb 2018):

- Präsenzvorlesung mit BPM-Grundlagen (Mathis 2025): Nutzen und Definition von Prozessmanagement und Prozessen, historische Entwicklung, Prozessmodellierung, Ziele und Kennzahlen von Prozessmanagement
- Online-Module zu Signavio Reports: Anwendungsvideos zu einzelnen Übungen
- Praktische Übungen in Kleingruppen von 2-3 Personen, begleitet und ergänzt durch Video-Tutorials, (SAP Signavio, SAP Learning Journey, YouTube Drumm)
- Reflexion und Peer-Feedback: Gruppenergebnisse stellen sich Gruppen gegenseitig vor, diskutieren und bewerten die Ergebnisse anhand vorgegebener Kriterien
- Präsentation und Diskussion der Ergebnisse: Nach Überarbeitung werden Endergebnisse vor der gesamten Gruppe präsentiert
- Diskussion der Lernerfahrung und Vergleich der Ergebnisse der Lernzielerreichung mit und ohne Lerneinheit Signavio-Reports

Die methodische Gestaltung stützt sich auf Best Practice aus der BPM-Lehre und Erfahrungen mit virtuellen Prozesslaboren.(Kauffeld & Othmer 2019; Becker u.a. 2012)

Die Unterrichtseinheit wurde für Studierende der Studiengänge der Technischen Betriebswirtschaft im 4. Semester entwickelt. Lernziele sind:

- Anwendung von Modellierungsmethoden (EPK oder BPMN)
- Verbindung von betriebswirtschaftlichen Grundlagen mit Prozessanalysemethoden

- Erkennen von Interpretationsmöglichkeiten von Prozesskennzahlen
- Selbstständige Entwicklung von Lösungsstrategien.

Die Studierenden verfügen durch ihr Studium über betriebswirtschaftliche Kenntnisse, wie z.B. Kosten- und Budgetrechnung, Controlling, Kapazitätsberechnungen. Im Rahmen der Vorlesung wurden die Modellierungsmethoden EPK und BPMN vermittelt und in Signavio bereits geübt.

4. Fallbeispiel: Prozessanalyse „Werkstück bearbeiten“ mit Signavio Reports

In einer Lerneinheit modellieren die Studierenden einen textlich beschriebenen Prozess „Werkstück bearbeiten“ in Signavio. Die Studierenden können wählen, ob sie den Prozess als EPK oder BPMN abbilden wollen.

Zu dem Prozess sind verschiedene Zeit-, Kosten- und Ressourcenattribute sowie Mengen und Verantwortlichkeiten gegeben. Die gepflegten Attribute können gleichermaßen für eine Prozesssimulation genutzt werden und vice versa (vgl. Beitrag Scheruhn, Hochschule Harz). Die Studierenden sollen mit der Übung zum einen eine Operationalisierung der Begriffe Prozesseffizienz und -effektivität kennen lernen (vgl. Ley u.a. 2012; Dumas u.a. 2021) und zum anderen auf dieser Basis beispielhaft eine quantitative Bewertung der Prozesse durchführen.

4.1. Berechnung der Durchlaufzeiten

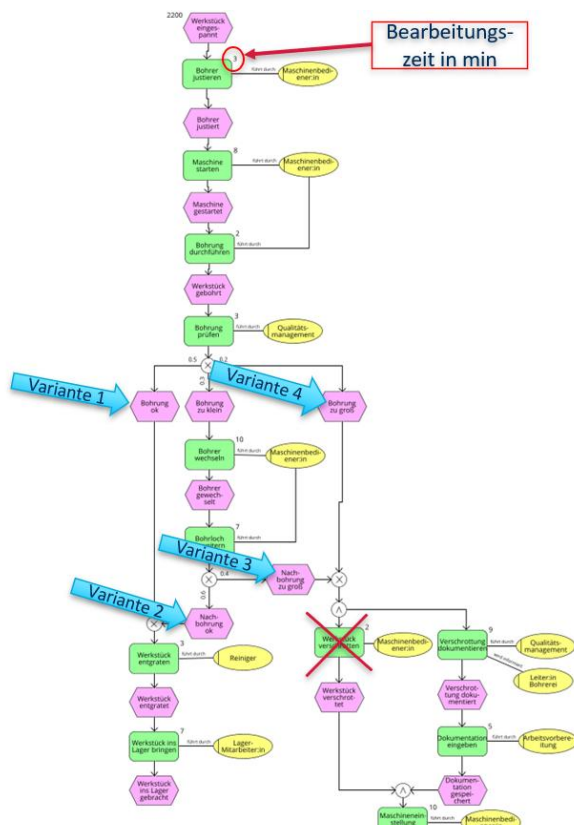


Abbildung 1: Beispiel EPK zur Berechnung der Durchlaufzeiten

Der erste Aufgabenschritt besteht darin, die theoretischen Prozessdurchlaufzeiten für ein Werkstück zu berechnen. Zur Berechnung der theoretischen Durchlaufzeit (DLZ) sind für jede Aktivitäten durchschnittliche Bearbeitungszeiten gegeben (oben rechts angezeigt).

Bei paralleler Bearbeitung (UND-Verknüpfung) kann der nächste Prozessschritt erst begonnen werden, wenn beide Schritte abgeschlossen sind. Bei Prozessvarianten, die durch Oder (XOR oder inkl. Oder) entstehen, gibt es verschiedene Möglichkeiten, wie der Prozess bis zum Ende verlaufen kann mit entsprechend unterschiedlichen DLZ. Bei Oder-Verzweigungen muss die DLZ für jede Variante berechnet werden. Im ersten Schritt werden die Wahrscheinlichkeiten nicht berücksichtigt. Später werden die Wahrscheinlichkeiten berücksichtigt. (Dumas u.a. 2021)

Aufgabe	Ausführungszeit in [min]	Prozessvarianten			
		1 Bohrung ok	2 Bohrung zu klein; Nachbohrung ok	3 Bohrung zu klein; Nachbohrung zu groß	4 Bohrung zu groß
Bohrer justieren	3	3	3	3	3
Maschine starten	8	8	8	8	8
Bohrung durchführen	2	2	2	2	2
Bohrung prüfen	3	3	3	3	3
Bohrer wechseln	10		10	10	
Bohrloch erweitern	7		7	7	
Werkstück entgraten	3	3	3		
Werkstück ins Lager bringen	7	7	7		
Werkstück verschrotten	2	Schritt entfällt, parallele Bearbeitung, kürzerer Pfad			
Verschrottung dokumentieren	9			9	9
Dokumentation eingeben	5			5	5
Maschineneinstellung anpassen	10			10	10
Summe in [min]		26	43	57	40

Abbildung 2: Berechnung der Durchlaufzeiten für alle Varianten in dem Beispiel

Über die Anzahl der vier Varianten im Beispielprozess, die durch die zwei XOR-Verknüpfungen entstehen, sehen die Studierenden, wie schnell die Variantenanzahl bei geplanten und ungeplanten Prozessumwegen steigt.

Durch die Varianten soll der Zusammenhang zu einer „Standardisierung von Prozessen“, d.h. einer Reduzierung von Varianten, verdeutlicht werden.

Zusätzlich soll die unterschiedliche Länge der Aufgaben und deren Einfluss auf die gesamte DLZ diskutiert werden.

Für Schritte, die besonders lange dauern, wie z.B. „Bohrer wechseln“ und „Maschineneinstellung anpassen“, sollen Ideen entwickelt werden, wie diese verkürzt werden können.

4.2. Bewertung der Prozessqualität und Prozesseffizienz

Im Anschluss an die Durchlaufzeitberechnung soll die Prozessqualität daten- und faktenbasiert mit gegebenen Eintrittswahrscheinlichkeiten analysiert und bewertet werden. Die Prozessqualität wird wie folgt berechnet (Gadatsch 2025):

$$\text{Prozessqualität in \%} = \frac{\text{Summe (Anzahl abgeschlossene Aufträge ohne Nacharbeit)}}{\text{Summe Anzahl aller abgeschlossener Aufträge}} * 100$$

Hier werden die Mengen pro Jahr (2200 Stück) und die Eintrittswahrscheinlichkeiten mitberücksichtigt. Prozessqualität = 50% (ergibt sich im Beispiel direkt aus der Eintrittswahrscheinlichkeit). Die Prozesseffizienz soll auf Basis einer Kategorisierung der Aufgaben in Nutz-, Support-, Blind- und Fehlleistungen oder vereinfacht wertschöpfende und nicht wertschöpfende Tätigkeiten berechnet werden.(Arndt 2021)

$$\text{Prozesseffizienz in \%} = \frac{\text{Nutzleistung}}{(\text{Nutz-} + \text{Support-} + \text{Blind-} + \text{Fehlleistung})} * 100$$

Aufgabe	Ausführungszeit in [min]	Leistungsarten			
		Nutz	Stütz	Blind	Fehl
Bohrer justieren	3		3		
Maschine starten	8		5		
Bohrung durchführen	2	2			
Bohrung prüfen	3		3		
Bohrer wechseln	10			10	
Bohrloch erweitern	7			7	
Werkstück entgraten	3	3			
Werkstück ins Lager bringen	7		7		
Werkstück verschrotten	2				2
Verschrottung dokumentieren	9				9
Dokumentation eingeben	5				5
Maschineneinstellung anpassen	10			10	
Summe in [min]		69	18	27	16

Abbildung 3: Mögliche Kategorisierung der Aufgaben nach Leistungsarten

Mit Nutzleistungen wird der Wert des Produkts erhöht. Stützleistungen, sind notwendig, erhöhen den Wert des Produktes allerdings nicht, z.B. Prüf-, Rüst- oder Transportvorgänge. Blindleistungen sind Korrekturvorgänge und Nacharbeiten oder Transportvorgänge zwischen (Puffer-)Lagern. Fehlleistungen sind alle Vorgänge, die Werte vernichten und zur Verschrottung des Produkts führen bzw. im Rahmen der „Wertvernichtung“ durchgeführt werden müssen.

Nutz- und Stützleistungen sollen beschleunigt, Blind- und Fehlleistungen sollten eliminiert werden. Beispielhaft kann eine Einstufung der Aufgaben folgendermaßen aussehen:

Die Prozesseffizienz ergibt hier $\text{Prozesseffizienz} = 5/69 \cdot 100 = 7,24\%$

Bei dieser Teilübung sollen der Beitrag der einzelnen Aufgaben zum Gesamtergebnis reflektiert, Problembereiche identifiziert und mit Zahlen untermauert werden. Zur weiteren Bewertung können die Mengen und Eintrittswahrscheinlichkeiten herangezogen werden.

4.3. Berechnung des Kosten- und Ressourcenbedarfs

Auf Basis der Kosten- und Ressourcenbedarfsrechnung von Signavio soll ergänzend dazu die Prozessqualität kosten- und ressourcenseitig analysiert und bewertet werden.

Als Ergebnis können hier die Kosten und der Ressourcenbedarf für die Aufträge ohne Fehler und Fehlerkosten für fehlerhafte Aufträge berechnet und diskutiert werden.

Was kostet der Prozess im Jahr, wenn 2200 Stück hergestellt werden?

Report:	Kostenrechnung (Prozess)								
Datum:	14.08.2025								
Zeit:	15:50:07								
Name:	Werkstück bohren und weiterbearbeiten (kennzahlenbsp. Maschbed.)								
Häufigkeit (pro Jahr):	2200								
Task	Einsatzfaktor	Kosten pro Ausführung	Arbeitsvorbereitung Kosten	Bohrerei Kosten	Lager Bohrteile Kosten	Qualitätsmanagement Kosten	Reinigung Kosten	Verschrottung Kosten	Gesamt Kosten
Bohrer justieren	1,00	4,00 €		8.800,00 €					8.800,00 €
Maschine starten	1,00	10,60 €		23.320,00 €					23.320,00 €
Bohrung durchführen	1,00	1,50 €		3.300,00 €					3.300,00 €
Bohrung prüfen	1,00	6,00 €				13.200,00 €			13.200,00 €
Bohrer wechseln	0,30	13,30 €		8.778,00 €					8.778,00 €
Bohrloch erweitern	0,30	9,30 €		6.138,00 €					6.138,00 €
Werkstück entgraten	0,68	3,00 €					4.488,00 €		4.488,00 €
Werkstück ins Lager bringen	0,68	3,50 €			5.236,00 €				5.236,00 €
Werkstück verschrotten	0,32	2,00 €						1.408,00 €	1.408,00 €
Verschrottung dokumentieren	0,32	9,00 €						6.336,00 €	6.336,00 €
Dokumentation eingeben	0,32	12,50 €	8.800,00 €						8.800,00 €
Maschineneinstellung anpassen	0,32	15,00 €		10.560,00 €					10.560,00 €
Summen		45,62 €	8.800,00 €	60.896,00 €	5.236,00 €	13.200,00 €	4.488,00 €	7.744,00 €	100.364,00 €
Leistungsarten									
Nutz	Stütz	Blind	Fehl						

Abbildung 4: Kostenreport aus SAP ® Signavio kombiniert mit der Kategorisierung in Leistungsarten

Über die Kombination der Leistungsarten mit dem Kostenreport in Signavio kann die Prozessqualität und können die Prozesskosten quantitativ bewertet werden. Die Studierenden sollen auf Basis von Zahlenmaterial und einem Wechsel von absoluten und relativen Zahlen ihre Prozessbewertungen argumentativ untermauern. Z.B. beträgt das Verhältnis der Kosten der Nutzleistungen zu den Gesamtkosten $= 7788 \text{ €} / 100.364 \text{ €} \cdot 100 = 7,8\%$. Während 42.000 € pro Jahr für Nacharbeit (Blindleistung) und Verschrottung (Fehlleistung) aufgewendet werden.

Weitere Ansätze können die Ermittlung von Kostentreibern und Ressourcenengpässen sein. Als Kostentreiber werden die Aufgaben verstanden, die im Prozess die höchsten Kosten verursachen. Bottlenecks können durch die Erhöhung der Menge besser sichtbar gemacht werden.

4.4. Analyse der Verantwortlichkeiten

Nachdem der Prozess aus Kosten- und Ressourcensicht analysiert wurde, sollen die Studierenden eine RACI-Matrix zur Verantwortlichkeitsklärung und -bewertung mit Signavio erstellen.

Zunächst sollen die Studierenden reflektieren, was es bedeutet, einen Verantwortlichkeitsübergang zu haben. Gelernt und vorher diskutiert wurden, z.B. Liege- und Wartezeiten zwischen den Schritten. Fehlende Verantwortung für das Gesamtergebnis, fehlende Transparenz für die Entstehung des Gesamtergebnisses, mögliche Kommunikationsprobleme an den Schnittstellen. Anhand des Ergebnisses sollen die Studierenden die organisatorische Gestaltung des Prozessablaufs diskutieren und bewerten sowie Verbesserungsvorschläge erarbeiten.

Report:	Verantwortlichkeitsübergaben			
Datum:	28.07.2025			
Zeit:	12:22:00			
Benutzer:	Uta Mathis			
Arbeitsbereich:	Hochschule Esslingen (Uta Mathis)			
Diagramm	vom	zu	Quellelement	Zielelement
Werkstück bohren und weiterbearbeiten (kennzahlenbsp. Maschbed.)	Arbeitsvorbereitung	Maschinenbediener:in	Dokumentation eingeben	Maschineneinstellung anpassen
	Maschinenbediener:in	Reiniger	Bohrloch erweitern	Werkstück entgraten
	Maschinenbediener:in	Qualitäts-management	Bohrloch erweitern	Verschrottung dokumentieren
	Maschinenbediener:in	Qualitäts-management	Bohrung durchführen	Bohrung prüfen
	Qualitäts-management	Reiniger	Bohrung prüfen	Werkstück entgraten
	Qualitäts-management	Arbeitsvorbereitung	Verschrottung dokumentieren	Dokumentation eingeben
	Qualitäts-management	Maschinenbediener:in	Bohrung prüfen	Bohrer wechseln
	Qualitäts-management	Maschinenbediener:in	Bohrung prüfen	Werkstück verschrotten
	Reiniger	Lager-Mitarbeiter:in	Werkstück entgraten	Werkstück ins Lager bringen

Abbildung 5: Darstellung der internen Verantwortungsübergabe in einem Prozess mit SAP® Signavio

Anhand der Verantwortungsübergaben kann gezeigt werden, zwischen welchen Personen ggf. Kommunikationsprobleme auftreten können und welche Aufgaben logisch gesehen besser ein und derselben Person zugeordnet werden sollten. In der Regel sollten Mitarbeiter ihre eigenen Arbeitsergebnisse selbst prüfen (Geiger/ Kotte 2008), im Beispiel prüft das Qualitätsmanagement das Arbeitsergebnis des Maschinenbedieners.

5. Ergebnisse

Die Anwendung der Reports wurde mit den Studierenden zum einen diskutiert und zum anderen wurden Klausurergebnisse der Aufgaben zu Kennzahlen im Prozessmanagement miteinander verglichen.

Das Feedback der Studierenden war positiv bzgl. des Erkenntnisgewinns. Aussagen waren:

„Wir haben besser verstanden, was ein Prozess ist und wie Prozesse strukturiert abgebildet werden können.“

„Wir konnten besser verstehen, wie Prozesskennzahlen zustande kommen.“

„Wir haben die Auswirkung von Prozessänderungen, z.B. Reduzierung von Varianten, besser verstanden.“

„Wir haben (besser) nachvollziehen können, wie Kostentreiber berechnet werden.“

Einige Studierende hatten mit der korrekten Modellierung Probleme. Diese fanden zwar die Syntax- und Metrikprüfung hilfreich, waren aber mit dem Ergebnis und der Darstellung überfordert. Manche Studierende fanden die Vorbereitung der Prozesse und das Befüllen mit den Attributen als (sehr) aufwändig.

Positiv zu erwähnen ist noch, dass bei den Präsentationen der Übungsergebnisse Querverbindungen zu anderen Vorlesungen hergestellt wurden und Kennzahlen, wie z.B. Overall Equipment Effectiveness (OEE) einbauten.

Für zwei Gruppen konnten die Klausurergebnisse miteinander verglichen werden. In der ersten Gruppe Wintersemester (WS) 2023/24 wurde die Möglichkeit der Reports zwar angesprochen und gezeigt, aber keine verpflichtende Übung dazu durchgeführt. Im darauffolgenden Sommersemester (SS) 2024/24 wurde die Übung nach der vorgestellten Methodik durchgeführt.

	nur Hinweis	mit Übung
	WS 2023/24	SS 2024/24
Anzahl Teilnehmer (Klausur)	27	33
Durchschnittl. erreichte Punktzahl	6,4 von 15	8 von 11
Erreichte Punkte in %	43	72
Übung gar nicht bearbeitet in %	7,4	6

Der Vergleich der Klausurergebnisse bei der gleichen Aufgabe zur Prozessanalyse zeigt bei den Studierenden, die die Fallstudie mit den Reports durchgeführt hatten, eine höhere Antwortquote der Aufgabe überhaupt und bessere Ergebnisse (höhere erreichte Punktzahlen). Da die Aufgaben nicht vollkommen identisch sind und auch die maximale Punktzahl unterschiedlich ist, können die Ergebnisse nur vorsichtig als positive Tendenz gewertet werden.

6. Fazit und Ausblick

Der Einsatz von Fallstudien mit den Signavio Reports bietet eine wirkungsvolle praxisorientierte Ergänzung zur Vermittlung von Kennzahlen und deren Zusammenhängen und Auswirkungen im Prozessmanagement.

Die Vorbereitung der Prozessmodelle und das Befüllen mit Attributen ist einerseits relativ zeitaufwändig, andererseits können so die Objekteigenschaften und deren Attribute noch einmal reflektiert werden (Funktionen/ Aufgaben sind Träger von Zeiten und Kosten). Anhand der Reports können Studierende Kennzahlen im Prozessmanagement besser nachvollziehen und verstehen. Durch die Ausarbeitung von Präsentationen, in denen eigene Lösungsansätze vorgestellt werden, können Querverbindungen zu anderen Vorlesungsinhalten hergestellt werden und die Querschnittsfunktion von BPM kann herausgestellt werden.

E-Learning-Einheiten und Video-Tutorials können technische Hürden reduzieren. Feedback-Runden und Diskussionen innerhalb und zwischen den Gruppen können aktivierend wirken.

Zukünftige angewandte Forschung im Bereich BPM in der Lehre sollte sich auf eine stärkere Verknüpfung betriebswirtschaftlicher Methoden und Methoden und Tools für Prozessmanagement fokussieren. Dazu zählt auch der Einsatz der Simulationsfunktion (vgl. Scheruhn, HS Harz).

7. Literatur

- Arndt, H.: Supply Chain Management, 8. Aufl., Springer Fachmedien Wiesbaden GmbH, 2021 [https://doi.org/10.1007/978-3-658-34406-1_5]
- Bake, B.; Sieboth, P.; Scheruhn, H.: Erfahrungsbericht aus Studentischer Sicht, Die Arbeit mit EOG, Signavio und Simulation, Beitrag SAP ACC 2025
- Becker, J., Kugeler, M., Rosemann, M.: Process Management. 4. Aufl., Springer Wiesbaden 2013
- Becker, J.; Kugeler, M.; Rosemann, M.: Prozessmanagement, 7. Aufl., Springer Wiesbaden 2012
- Costa, C. J.; Aparicio, M.; Raposo, J.: Determinants of management learning performance in ERP context, Helion 6 (2020), e03689 [<https://doi.org/10.1016/j.heliyon.2020.e03689>]
- Drumm, C.: Informationssysteme, Videos zum Modul Informationssysteme an der FH Aachen, [<https://www.youtube.com/playlist?list=PLI09U8aTDcv1eIkxyPKNAKKmqPJR3RC0o>] [abgerufen 28.07.2025]
- Dumas, M., La Rosa, M., Mendling, J., Reijers, H. A.: Fundamentals of Business Process Management. Springer, 2018 [<https://doi.org/10.1007/978-3-662-56509-4>]
- Dumas, M., La Rosa, M., Mendling, J., Reijers, H. A.: Grundlagen des Geschäftsprozessmanagements. Springer, 2021 [<https://doi.org/10.1007/978-3-662-58736-2>]
- Gadatsch, A.: Grundkurs Geschäftsprozessmanagement, 10. Aufl., Springer Wiesbaden, 2023 [<https://doi.org/10.1007/978-3-658-40298-3>]
- Gadatsch, A.: Geschäftsprozesse analysieren und optimieren, 2025, Springer Wiesbaden [https://doi.org/10.1007/978-3-658-47782-0_1]
- Geiger, W.; Kotte, W.: Handbuch Qualität, 5. Aufl., Vieweg Wiesbaden 2008
- Gerogiannis, V. C.; Fitsilis, P.: Project-based Learning as a Teaching Method for ERP Systems, 3rd International Conference on Enterprise Systems and Accounting (ICESAcc'06), 26-27 June 2006, Santorini Island, Greece 2006
- Hommel, M.: Lernhandeln mit integrierter Unternehmenssoftware, 2019, TU Dresden, Habilitationsschrift 2019
- Kauffeld, S.; Othmer, J. (Hrsg.): Handbuch Innovative Lehre, Springer Wiesbaden 2019
- Ley, T.; Jurisch, M.; Wolf, P.; Krcmar, H.: Kriterien zur Leistungsbeurteilung von Prozessen – Ein State-of-the-Art. Tagungsband der Multikonferenz Wirtschaftsinformatik (MKWI), Braunschweig 2012 [<http://www.digibib.tu-bs.de/?docid=00047607>]
- Mathis, Uta: Prozessmanagement/ Methoden des Prozessmanagements, Teil des Moduls Wirtschaftsinformatik 3/ Digitalisierung 1: Prozessmanagement und ERP-Systeme, Skript Hochschule Esslingen 2025
- Mendling, J.: Metrics for Process Models. Springer, 2014

- Recker, J., Mendling, J.: The state of the art of business process management research as published in the bpm conference. Business & Information Systems Engineering 2016 [DOI: 10.1007/s12599-015-0411-3]
- Röglinger, M.; Seyfried, J.; Stelzl, S.; zur Mühlen, M.: Cognitive Computing: What's in for Business Process Management? BPM Conference 2017 Barcelona, Spain, September 2017
- SAP (Hrsg.): SAP Signavio Process Manager - Benutzerhandbuch, 2025
[<https://help.sap.com/doc/10ae12665b494798a8a332bcc195689d/SHIP/de-DE/sap-signavio-process-manager-user-guide-de.pdf>] [abgerufen 28.07.2025]
- SAP (Hrsg.): SAP ® Signavio: Value Cookbook, Digital Library public
[<https://www.signavio.com/downloads/white-papers/unlock-the-hidden-value-in-your-business-processes-with-our-value-cookbook/>] [abgerufen 28.07.2025]
- SAP (Hrsg.): Learning Journey – Understanding Reports. SAP Learning, 2025
[https://learning.sap.com/learning-journeys/design-business-processes-with-sap-signavio-solutions/using-standard-reports_b585f735-bc69-4aa7-9ed0-30a3b8c2ea05] [abgerufen 28.07.2025]
- Strasser, J.: Schnelleinstieg in SAP® Signavio. Espresso Tutorials GmbH, 2023
- Strasser, J., Sokollek, M., Sängler, M., Spierling, M., Schönwälder, M.: Business Process Transformation mit SAP Signavio, SAP PRESS, 2023
- Schmidpeter, R.; Kolb, M.: Wirtschaft im Wandel – Neue Anforderungen an die Managementausbildung, Springer Nature 2018, in: Raueiser, M.; Kolb, M. (Hrsg.): CSR und Hochschulmanagement, Management-Reihe Corporate Social Responsibility, Springer Gabler Berlin 2018 https://doi.org/10.1007/978-3-662-56314-4_2
- van der Aalst, W. M. P.: Process Mining: Data Science in Action. 2. Aufl., Springer 2021
- vom Brocke, J., Rosemann, M. (Hrsg.): Handbook on Business Process Management. Band 1: Introduction, Methods, and Information Systems. Springer 2015
- Weske, M.: Business Process Management: Concepts, Languages, Architectures. 4. Aufl., Springer 2019

Prozessorientierte Simulation als duale Studentin der DRV Bund

Lisa Iljin¹, Patricia Sieboth², Benjamin Bake²

¹: ITZ Bund, ²: Hochschule Harz

Abstract: Dieser praktische Beitrag soll aufzeigen, wie das BPM-Tool SAP Signavio im Kontext des Enterprise Online Guides (EOG) in der Lehre einsetzbar ist. Als Ausgangspunkt dient dabei die L2C-Fallstudie des SAP University Competence Center (UCC) Magdeburg mit SAP S/4HANA. Eine erfolgreiche Umsetzung erfolgte im Wintersemester 2024 mit dreizehn Studierenden des dualen Studiengangs IT-Management Verwaltungsinformatik am Fachbereich Verwaltungswissenschaften der Hochschule Harz. Dabei kam die Academic Edition von SAP Signavio zum Einsatz.

Als Prüfungsleistung hat die Autorin ihr Praxisunternehmen DRV-Bund im Kontext von EOG abgebildet. Nach dem anschließenden Wechsel in das EOG-Team hat sie eine bestehende Simulationsfallstudie übernommen, methodisch weiterentwickelt, eigenständig durchgeführt und dokumentiert.

Aufbauend auf dem Order-to-Cash-Prozess der Global-Bike-Fallstudie konnte sie mithilfe von SAP Signavio Simulation eigene Szenarien erstellen und die Auswirkungen unterschiedlicher Prozessvarianten analysieren. Besonders eindrücklich war für die Autorin die Analyse von Bottlenecks und deren gezielte Auflösung durch Anpassung von Eingabeparametern (z. B. Ressourcen, Bearbeitungsdauer, Systemoptimierung mit FIORI) im Kontext EOG.

Die strukturierte Arbeit mit dem EOG-Ansatz hat der Autorin die Prozesszusammenhänge nicht nur technisch, sondern auch betriebswirtschaftlich erschlossen. Insgesamt ermöglichte die Kombination aus Modellierung, Simulation und EOG-Methodik einen fundierten Einblick in die Potenziale digitaler Transformation, sowohl aus technischer als auch betriebswirtschaftlicher Perspektive.

Process Comparisons: An Industry Perspective

G. Berg, T. Kampik, A. Rebmann, M. Rafiei, S. Hadian, R. Blümel & L. Kirchdorfer
SAP SE, Walldorf, Germany

Abstract: We present an experience summary highlighting the importance of process comparisons for effective data-driven business process management. Specifically, five central comparison tasks are crucial: (A) event log to event log, (B) event log to process model, (C) event log to best practice model, (D) process model to process model, and (E) process model to best practice model. For each of the tasks, we discuss application examples and finally highlight challenges that require future research.

Keywords: Business Process Management, Process Comparison, Conformance Checking

1. Introduction

The usage of process mining, often combined with process modeling, is broadly recognized as a critical part in Business Process Management (BPM) (Srivastava et al. 2025). Both process mining and modeling require various forms of process representations, such as process models and event logs, enabling the analysis of real-life processes to uncover areas for improvement. These representations are not only subject to individual analysis but are also compared to each other. Traditional examples include conformance checking, i.e., comparing a process model to execution data from an event log to assess their conformance (van der Aalst 2016), and model comparison to identify differences between “to-be” and “as-is” models (Dijkman et al. 2011).

Despite research on diverse process comparison methods, there is no clear overview of the concrete comparisons required by practitioners. This overview is essential because the ambiguity that practitioners encounter when their stakeholders want to *compare processes* is set to increase as more flavors of *process comparisons* are established in practitioners' toolboxes. For example, comparing customer models with best practice models offered by vendors or consultants requires some standardization. Existing comparisons are based on data type distinctions: model-to-log, log-to-log, and model-to-model. However, in practice, these manifest themselves with distinct flavors and practical implications that remain underexplored.

To shed light on this problem, this paper presents an experience summary highlighting five central comparison tasks that are crucial for effective data-driven BPM (cf. Figure 1): (A) event log to event log, (B) event log to process model, (C) event log to best practice model, (D) process model to process model, and (E) process model to best practice model. For each use case, we explain their underlying rationale, provide an overview of what practitioners already use, and highlight challenges that require research and innovation efforts.

5 Process Comparisons in the Signavio Suite

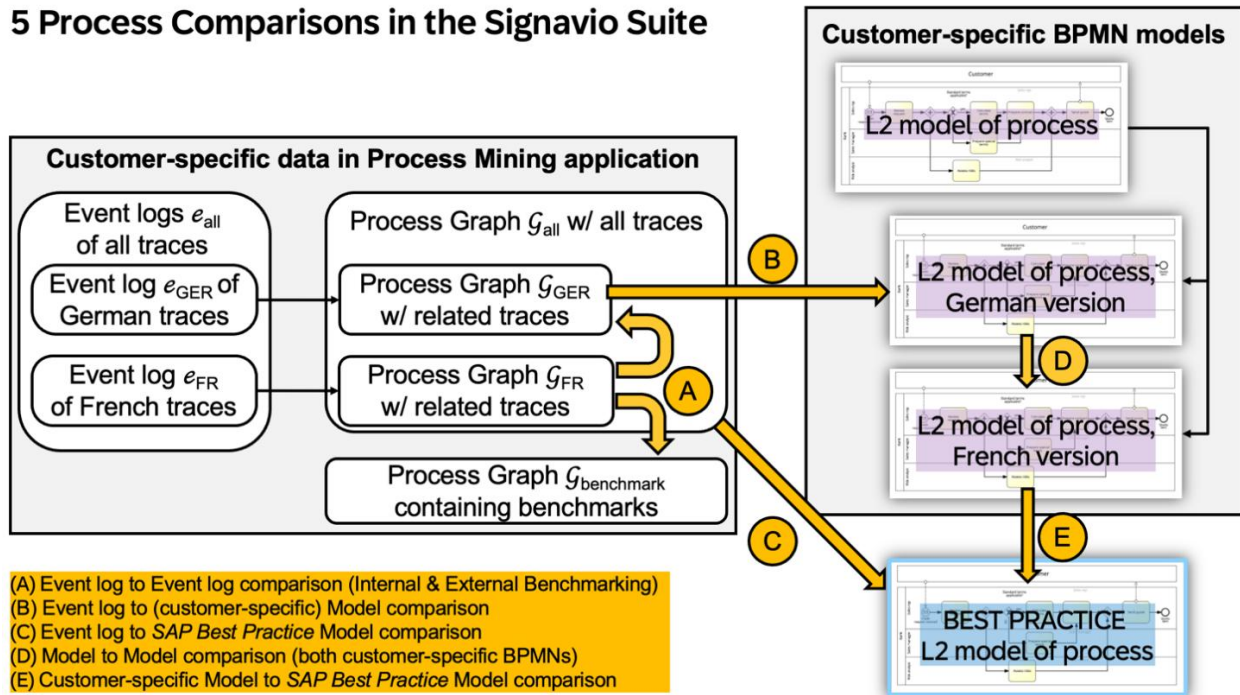


Figure 1: Depending on which datasets are available to a practitioner, up to five distinct process comparisons are possible.

2. Comparisons

Depending on which capabilities of a modern BPM suite are used, customers may have (1) up to thousands of BPMN process models describing their processes, or (2) Directly-Follows Graphs (DFGs) (van der Aalst 2016), so-called process graphs, reflecting their actual process execution, or both datasets. Additionally, process mining vendors may provide external benchmarking capabilities for process graphs along with benchmarking datasets (Berg et al. 2025, Abb et al. 2025). Finally, while each customer can have customer-specific models, vendors and consulting companies also provide standardized process models (often BPMNs), sometimes referred to as *best practice models*. Thus, between modeling and mining capabilities, the datasets of a customer and standardized benchmarks/best practices must be considered for potential comparisons.

In this section, we explain the different process comparisons. Specifically, for each comparison type, we describe what it is about, why a customer would need it, which tools and related methods are already available for the respective use case, and which challenges should be addressed to increase future impact.

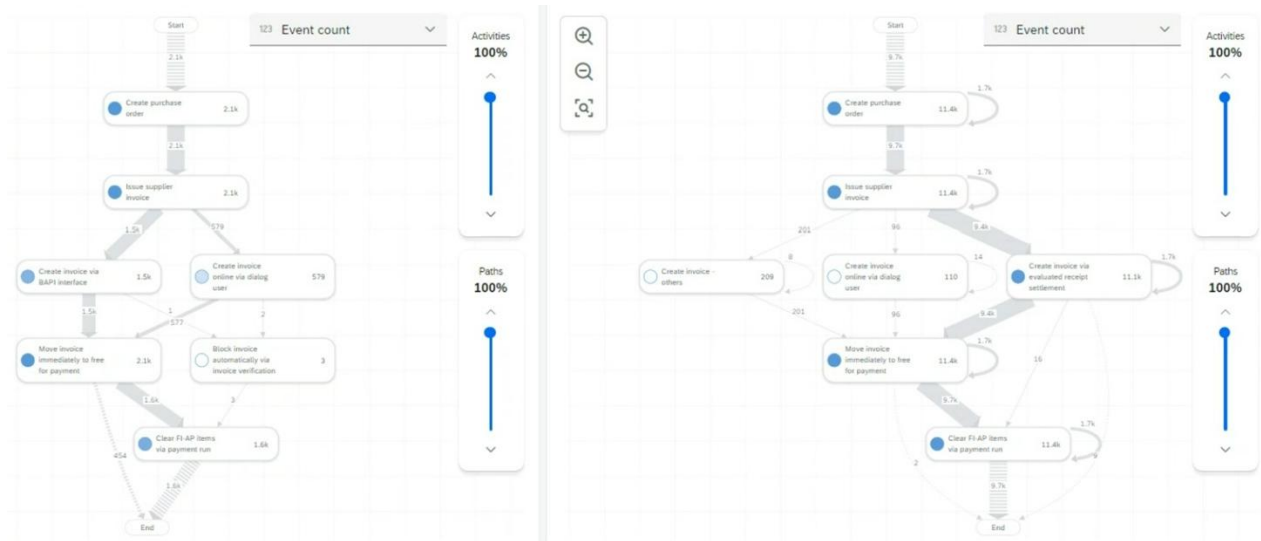


Figure 2: As-is approach for (A), inspecting two graphs next to each other.

2.1. Event Log to Event Log Comparison (A)

What it is: Compare process executions via process discovery graphs either internally (e.g., for company codes) or externally against industry peers.

Why customers need this: *Internally*, customers are interested in identifying the most effective process variants across, e.g., different countries or product groups (Taymouri et al. 2021). *Externally*, customers aim to achieve the same by understanding which best-in-class execution paths are quicker or desirable based on related process metrics.

What is available for this use case:

A1 The most commonly used version is a side-by-side comparison in two browser tabs to inspect the same underlying data using dedicated filters to, e.g., compare the execution in different company codes (Fig. 2).

A2 Abb et al. (2025) presented a comparison, which allows to compare either internal company codes or even selected internal variants with best-in-class execution versions of industry peers. By combining process graphs and coloring them accordingly, users can easily identify the better-performing process variant (Berg et al. 2025), cf. Fig. 3.

What challenges need to be addressed:

- Behavioral and performance differences between process (executions) in different organizations is often purposeful, even when it pertains to the same process (such as order-to-cash) and industry (such as retail). For example, a high-end grocery store would expect different behaviors and performance than a discounter. Controlling for purposeful difference requires domain expertise and is crucial to avoid mismatched comparisons.
- Process graph benchmarking is a topic that is academically underexplored. A key reason for this is the absence of benchmark-able open data, that features event log collections of the execution of one process in many organizations. Creating such collections, for example based on openly available data of software engineering processes on platforms like GitHub could provide the data foundations for process graph benchmarking

research.

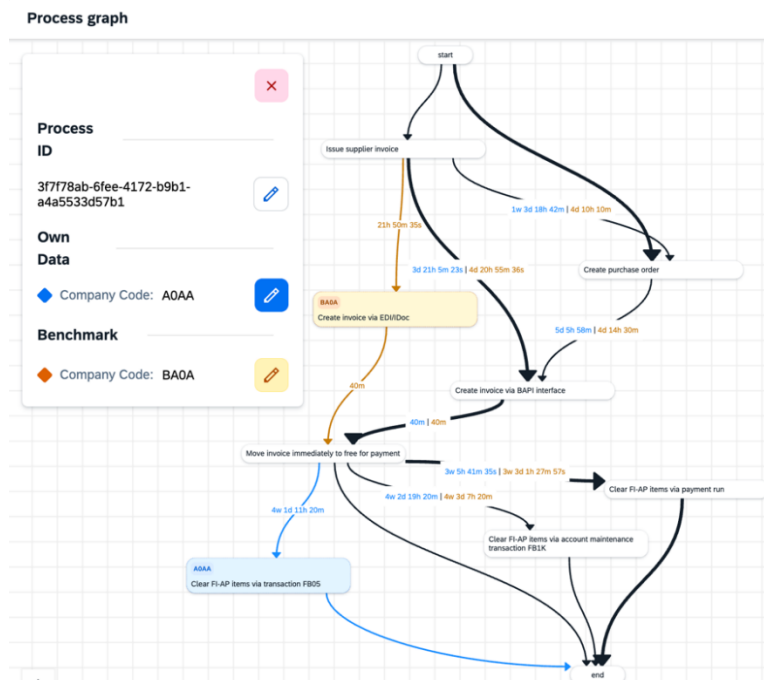


Figure 3: Process graph overlay comparison for (A) as described in Berg et al. (2025).

2.2. Event Log to Customer-Specific Model Comparison (B)

What it is: Comparing the process executions captured in an event log to the normative, modeled version of the process that resides in a modeling workspace. The BPMN model is customer-specific and cannot be assumed to be standardized in any way, e.g., by using standard activity labels or document names.

Why customers need this: Analyzing to what extent the actual execution of a process complies with its normative or descriptive model enables customers to find commonalities and discrepancies (van der Aalst 2016). Both results can be used to resolve compliance issues, as well as to kickoff process improvements.

What is available for this use case:

B1 Nowadays, BPM suites offer *conformance widgets* which overlay event log traces on BPMN models. However, this requires a mapping of customer-specific models and activities to the events captured in the process mining application.

B2 A user can also generate a BPMN model capturing common variants from the event log to use the model-to-model comparison (cf. Section 2.4).

What challenges need to be addressed:

- Traditional conformance checking is based on imperative process models, such as BPMN models and Petri nets (Carmona et al. 2018). However, imperative models, while often readily available in organizations, tend to be too rigid, overwhelming users with false positives and hence require systematic translation into declarative constraints as a prerequisite for *relaxation*.
- In practice, it is crucial to use *queries* as the fundamental abstractions for conformance checking: both special-purpose (such as SAP Signavio's SIGNAL) and generic query

languages (such as SQL) are crucial and prevalent tools for querying process data. Accordingly, further exploring the relationship between imperative and declarative process models and query languages is a highly relevant research direction.

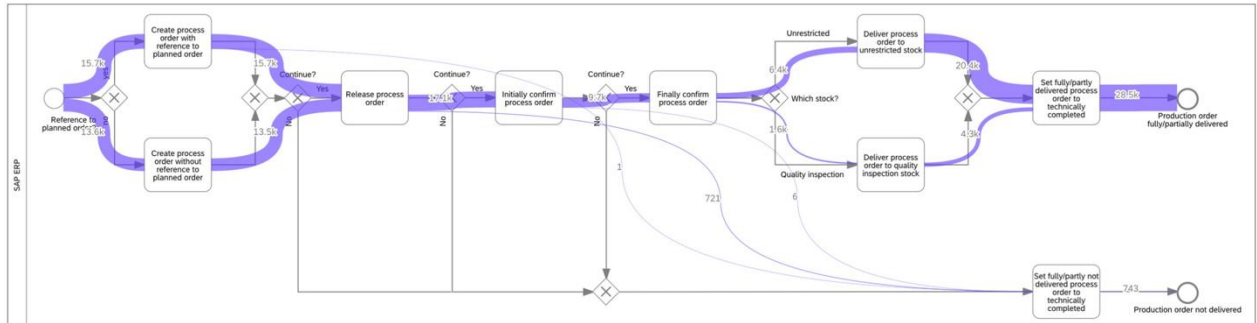


Figure 1: For (B) and (C), mapping events to BPMN activities enables visual conformance checks based on normative process models.

2.3. Event Log to Best Practice Model Comparison (C)

What it is: The same as (B), except that a well-known standard set of BPMN models provided by a vendor or partners is used to evaluate the *distance* to this standard. *Thus, the models are standardized and known beforehand.*

Why customers need this: Customers require guidance when it comes to business transformations, e.g., when migrating from the on-premises SAP ERP to S/4HANA Public Cloud—which processes are already *close enough*, which ones need to be *changed fundamentally*.

What is available for this use case:

- C1 As (B1) in Section 2.2, conformance widgets can be used (Fig. 4). However, the need for *one-off mappings* disappears if the mapping between events and standard activities is provided by the process mining templates.
- C2 The same *simplification option* (B2) exists as in (B)—by generating a BPMN model, a model-to-model comparison is possible (cf. Section 2.4).
- C3 Also, standardized process analyses via SAP Signavio Process Insights can be used to analyze *how* standardized specific activities are executed, specifically in an SAP ERP system. This analysis leverages the so-called *best-run score* (SAP SE 2024) to show the distance to the *best practice processes*.
- C4 Recent academic work demonstrates promising results regarding the automatic extraction of conformance checking queries from best practice model collections (Rebmann et al. 2024), as well as the use LLMs that have been fine-tuned using best practice models for conformance checking (Busch et al. 2024).

What challenges need to be addressed:

- In its current form, best practice models cannot be meaningfully applied for conformance checking without sophisticated pre-processing, as mentioned above. Accordingly, both the conformance checking toolbox and the *content* used for conformance checking need to evolve to better match customer needs.
- Tracing back conformance violations to best practice content, and providing automatic support for interpretation, root-cause analysis, and problem resolution are long-standing

challenges.

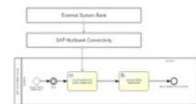
Process Comparator

First Process

Search by keyword

Select First Process

Import Bank Statements (MBC)

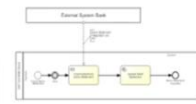


Second Process

Search by keyword

Select Second Process

Import Bank Statements (API)



Differences

Compute

Output

High-level comparison: The processes are similar with common parts.

List of differences:

1. The First Process includes a CollapsedPool component with the name "SAP Multibank Connectivity", while the Second Process does not have this component.
2. The First Process includes a MessageFlow component with a source index of 0 and a target index of 2, while the Second Process includes a MessageFlow component with a source index of 4 and a target index of 2.
3. The First Process includes a Task component with the name "Upload Bank Statement" and a source index of 7, while the Second Process includes a Task component with the same name but a different source index of 3.
4. The First Process includes an EndNoneEvent component with the name "Bank Statement Imported" and a source index of 1, while the Second Process includes the same component with a different source index of 6.
5. The First Process includes a Lane component with the name "System" and a source index of 6.

Process 1



Process 2



LLM output

Highlevel comparison:

The processes are completely different in terms of steps and structure.

List of differences:

1. The first process starts with "Contract signed" while the second begins with "Individual objectives".
2. The first process includes "Prepare desk & equipment", absent in the second process.
3. The second process contains "Generate document" which is not in the first process.
4. The first process has a "Plan field assignment" step not found in the second.
5. The second process has a conditional "Accept evaluation" step which the first lacks.
6. The first process includes "Create email and other accounts", missing in the second.
7. The second process includes "Evaluate Objectives" step, not present in the first process.
8. The first process involves "Conduct Training", missing in the second process.
9. The second process contains "Review for alignment", a step not in the first process.
10. The final step in the first process is "Review progress," which differs from the second's "Accepted" step.

Figure 2: For (D) & (E), we explored two LLM-based comparisons—an approach to compare BPMN models to each other, providing textual output (left) and an approach which compares images of models to describe their differences (right).

2.4. Customer-Specific Model-to-Model-Comparison (D)

What it is: Comparison that takes two BPMN models to highlight their differences either visually or as a textual description of the text. Both approaches hardly scale beyond one-to-one comparisons, struggling when, e.g., looking for the best match in a whole set of models. For this, a distance metric (e.g. *79% similarity*) is required to ensure a holistic overview. For (D), both models are customer-specific and cannot be assumed to be standardized in any way.

Why customers need this: Customers need to compare and manage different versions of the same process (e.g., by country or product to sell) to identify the desired commonalities and purposeful differences.

What is available for this use case:

- D1 Most BPM suites offer simple comparison views providing visual feedback on where two models differ (cf. Fig. 6).
- D2 Recently, the exploration of LLM-based process comparisons became viable, extending simpler model matching techniques (Cayoglu et al. 2013) and similarity measures (Schoknecht 2017) previously used. These automatically take models as input, match their activities and provide (usually textual) output on where these models differ. For the examples shown in Fig. 5, the comparisons can be done by converting either BPMN models or images to text and extracting text embeddings to calculate the cosine similarity between them.
- D3 Voelter et al. (2024) extracted multi-sets of elements such as roles, activities, and events from pairs of BPMN models to calculate an aggregated score across the similarity of these sets of elements.

What challenges need to be addressed:

- Traditional semantic process model comparisons focus primarily on control flow aspects that are not necessarily of relevance in real-world scenarios, where one may, e.g., be interested in aspects like underlying system configuration or automation rates. Thus, comparison approaches should be devised that focus on what matters in practice as opposed to what matters given academic representational biases (e.g., on control-flow).
- Also, the integration and trade-off between formal and natural language semantics for different comparison use cases has not been solved, yet.

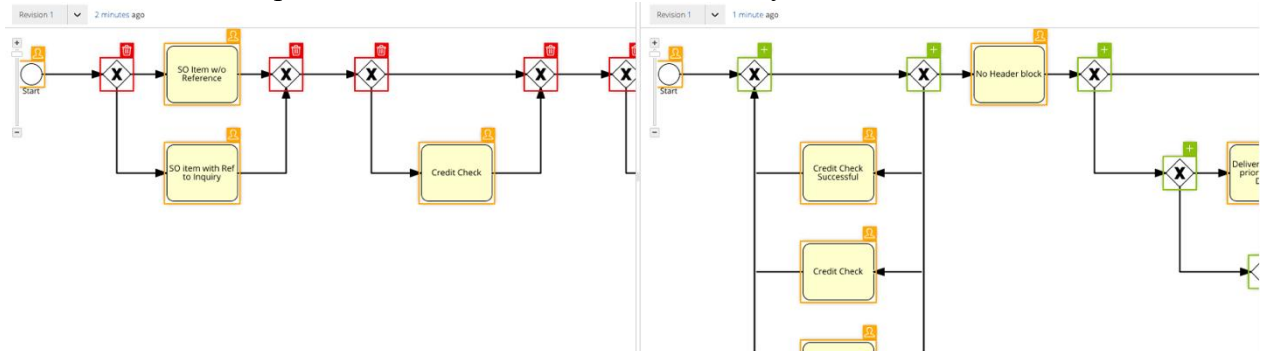


Figure 3: For (D) & (E), different versions of a model can be compared, the results are provided visually to the user.

2.5. Customer Model to Best Practice Model Comparison (E)

What it is: The same comparison as in (D) (Section 2.4) applies, except for the fact that one model is customer-specific while the other is a *standard model* by a vendor or a partner.

Why customers need this: Customers can evaluate the differences between their current way of executing a process and the proposed *best practice* way of doing the same. This includes identifying the best-matching standard process for any given customer process. Additionally, this comparison also visualizes the impact of adopting these standards (i.e., how many changes would be required, and which roles would be affected).

What is available for this use case: *All the options provided for (D) in Section 2.4 apply here as well.* Additionally, extracting declarative constraints from collections of (imperative) best practice process models (Rebmann et al. 2024) helps identifying which of these constraints are already fulfilled by models of a customer.

What challenges need to be addressed:

- Often, best practice models primarily reflect assumptions about “good” process behavior that domain experts, such as specialized consultants, have. However, the expectations are not necessarily grounded in data—designing methods for doing so can be considered a crucial research challenge. Also, best practice models that are grounded in data may require novel representation approaches.
- Similarly, the life cycle of managing best practice models has not been systematically studied. Certainly, well-maintained, continuously updated, and nuanced best practice models provide more value than stale and poorly designed sketches. Best practice model collections tend to contain both the former and the latter class of models, typically on a scale. Providing transparency with respect to domain-agnostic as well as

generic model quality is, thus, a key challenge.

3. Conclusion and Future Work

In this paper, we have laid out the practical importance of model-based, data-driven, and hybrid process comparisons, both within and across organizations for modern BPM. We have described five process comparison types that have been successfully deployed in practice. Still, several challenges exist that inhibit actionability and we have outlined research challenges for each comparison type. Notably, we advocate for research on data-driven BPM that (i) transcends organizational boundaries—most importantly to further facilitate process benchmarking—and (ii) focuses on representations that are underexplored in BPM research but common in practice, such as database queries and configuration files.

4. References

- Abb, L., Rafiei, M., Kampik, T., Rehse, J. (2025): Identifying process improvement opportunities through process execution benchmarking. In: BPMDS 2025, Vienna, Austria, Proceedings, volume 558 of LNBIP, pages 19-35. Springer, 2025.
- Berg, G., Kampik, T., Baier, S. (2025): External benchmarks for process mining graphs, March 27 2025. US Patent App. 18/473,505.
- Busch, K., Kampik, T., Leopold, H. (2024): xsemad: Explainable semantic anomaly detection in event logs using sequence-to-sequence models. In Business Process Management - 22nd International Conference, BPM 2024, Proceedings, volume 14940 of LNCS, pages 309-327. Springer, 2024.
- Carmona, J., van Dongen, B., Solti, A., Weidlich, M. (2018): Conformance checking. Switzerland: Springer, 56:12, 2018. Springer.
- Cayoglu, U., Dijkman, R., Dumas, M., Fettke, P., García-Bañuelos, L., Hake, P., Klinkmüller, C., Leopold, H., Ludwig, A., Loos, P. and others. (2013): Report: The process model matching contest 2013. In International Conference on Business Process Management, pages 442-463. Springer, 2013.
- Dijkman, R., Dumas, M., van Dongen, B., Käärik, R., Mendling, J. (2011): Similarity of business process models: Metrics and evaluation. Information Systems, 36(2):498-516, 2011. Semantic Integration of Data, Multimedia, and Services.
- Rebmann, A., Kampik, T., Corea, C., van der Aa, H. (2024): Mining constraints from reference process models for detecting best-practice violations in event log. CoRR, 2407.02336, 2024.
- SAP SE. (2024): Plug and gain 101: #1 what is a best-run score? <https://community.sap.com/t5/technology-blog-posts-by-sap/plug-and-gain-101-1-what-is-a-best-run-score/ba-p/13621174>, 2024. Accessed 2025-07-29.
- Schoknecht, A., Thaler, T., Fettke, P., Oberweis, A., Laue, R. (2017): Similarity of business process models—a state-of-the-art analysis. ACM Computing Surveys (CSUR), 50(4):1-33, 2017. ACM New York, NY, USA.

- Srivastava, T., Kerremans, M., Sugden, D. (2025): Gartner magic quadrant for process mining platforms. Technical report, Gartner Research, 2025.
- Taymouri, F., La Rosa, M., Dumas, M., Maggi, F. (2021): Business process variant analysis: Survey and classification. *Knowledge-Based Systems*, 211:106557, 2021. Elsevier.
- Van der Aalst, W. (2016): *Process Mining: Data Science in Action*. Springer Publishing Company, Incorporated, 2nd edition, 2016.
- Voelter, M., Hadian, R., Kampik, T., Breitmayer, M., Reichert, M. (2024): Leveraging generative vision models for extracting process models from documents. In *Business Process Management Workshops*, volume 534 of LNBIP, pages 271-282. Springer, 2024.

Einordnung der SAP Best Practice Modelle in den Enterprise Online Guide

Christiane Endter¹, Luca Gußmag², Hans-Jürgen Scheruhn²

¹msg Systems ag, Deutschland

²Hochschule Harz, Wernigerode, Deutschland

Abstract – „BPM drives Digital Transformation“: Um das Motto der SAP ACC2025 mit Leben zu erfüllen, müssen Business Process Management (BPM) und Enterprise Architecture Frameworks (EAF) bei einer Umsetzung mit einem ERP System eng einander greifen. Dieses Paper untersucht die Anwendung und Weiterentwicklung von SAP Best Practice Modellen im Kontext moderner Geschäftsprozessmodellierung. Im Fokus steht die Frage, inwieweit eine Einordnung dieser Modelle in das Enterprise Online Guide (EOG)-Framework möglich und sinnvoll ist. Dazu werden Herausforderungen von SAP Best Practices identifiziert und analysiert. Aufbauend darauf wird ein strukturierter Lösungsansatz entwickelt, der eine Integration in das Enterprise Online Guide-Framework unterstützt und die Verständliche der Best Practice Modelle steigert.

1. Einleitung

1.1. Motivation

Ein zentraler Bestandteil der SAP-Systemlandschaft ist das Geschäftsprozessmanagement. Es dient der systematischen Modellierung, Analyse und Optimierung von Geschäftsprozessen. Das Ziel des Geschäftsprozessmanagements ist es, die Effizienz, Transparenz und Konsistenz unternehmensinterner Abläufe zu steigern. [vgl. Dumas et al., 2021]

Unternehmen greifen dabei häufig auf die von SAP bereitgestellten Best Practice Modelle zurück. Diese stellen standardisierte Referenzen dar und dienen als Leitfaden für die Einführung und Anpassung von Geschäftsprozessen. Die Modelle ermöglichen es, Prozesse entlang bewährter Muster zu gestalten, was insbesondere bei der Einführung von SAP-Lösungen wie S/4HANA von strategischer Relevanz ist.

1.2. Problem

Während die Verwendung von SAP Best Practices stark nachgefragt wird, stellt sich die Frage, wie diese am effektivsten in den jeweiligen Unternehmenskontext integriert werden können. [vgl. Spontoni et. al., 2023] Besonders unter Verwendung einer bestehenden Unternehmensarchitektur mit individuellen Modellierungskonventionen ist eine Einordnung der SAP Best Practices zu prüfen. Dadurch ergeben sich zwei Forschungsfragen (Tabelle 1), die anhand des Design-Science-Research Ansatzes nach Peffers [vgl. Peffers et. al., 2007] untersucht werden sollen.

ID	Forschungsfragen
1	Welche Herausforderungen entstehen bei der Verwendung von SAP Best Practices?
2	Bietet die Einordnung von SAP-Best Practice Modellen in das Enterprise Online Guide Framework einen Mehrwert in der Verständlichkeit und Lesbarkeit?

Tabelle 1: Forschungsfragen

Um die Forschungsfragen beantworten zu können, werden im folgenden Kapitel zunächst die begrifflichen Grundlagen beschrieben.

2. Grundlagen

2.1. BPM-Lebenszyklus

Der Business Prozess Management-Lebenszyklus (BPM-Lebenszyklus) hat das Ziel Prozesse zu verstehen, zu analysieren und damit einhergehend stetig zu verbessern. Abbildung 1 zeigt den Lebenszyklus unterteilt in sechs Phasen.

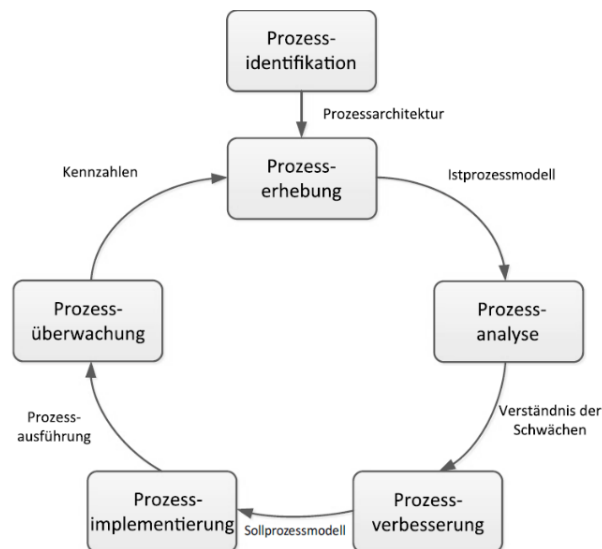


Abbildung 1: BPM-Lebenszyklus[vgl. Dumas et. al., 2021]

Phase eins ist die Prozessidentifikationsphase. Wird ein Problem in einem Unternehmen festgestellt, so müssen zunächst die relevanten Prozesse identifiziert werden. Ziel ist eine angepasste Prozess-Architektur, auf deren Grundlage die relevanten Prozesse nach Verbesserungspotenzialen untersucht werden können. Folgend werden die Prozesse in der Prozesserhebung dokumentiert. Dabei kommen typischerweise die Ist-Prozessmodelle zum Einsatz. Wurden in einem Unternehmen noch keine Prozesse dokumentiert, können die SAP Best Practices als Basis zur Ist-Dokumentation verwendet werden. Nach vollständiger Dokumentation der Ist-Situation folgt die Phase der Prozessanalyse. Dabei werden die zu Beginn erkannten Probleme genauer untersucht und spezifische Prozesskennzahlen ausgewertet. Der Output dieser Phase ist eine strukturierte Sammlung von Problemen, die in

der Phase der Prozessverbesserung gelöst werden sollen. Um die ermittelten Probleme zu lösen, können bei der Dokumentation konkrete Änderungsoptionen festgelegt werden. Das Ziel ist die Erstellung von Soll-Prozessmodellen, die den optimierten Prozess dokumentieren. Auf Basis der neu definierten Prozesse beginnt in der Phase der Prozessverbesserung die Umsetzung im System. Zuletzt folgt ein Monitoring der Prozesse in der Phase der Prozessüberwachung. Werden innerhalb des Monitorings Engpässe, wiederkehrende Fehler oder ähnliche Abweichungen identifiziert, die einen größeren Aufwand bedeuten oder zunächst genauer untersucht werden müssen, beginnt der Lebenszyklus erneut. [vgl. Dumas et. al., 2021]

2.2. SAP-Best Practices

SAP Best Practices sind vorkonfigurierte und sofort verfügbare Lösungen mit Templates für Geschäftsprozesse, Test-Skripte und Dokumentationen. [vgl. SAP, 2025]

Diese Best Practices unterstützen vor allem dabei, die Geschäftsprozesse optimal abzubilden sowie die Standardfunktionen schneller zu verstehen und das Umsetzungsrisiko von beispielsweise S/4HANA- Transformationen so gering wie möglich zu halten. [vgl. Munkacsi et. al., 2024] Die Best Practices setzen sich strukturell aus 4 Ebenen zusammen, die in Abbildung 2 zu visualisiert sind.

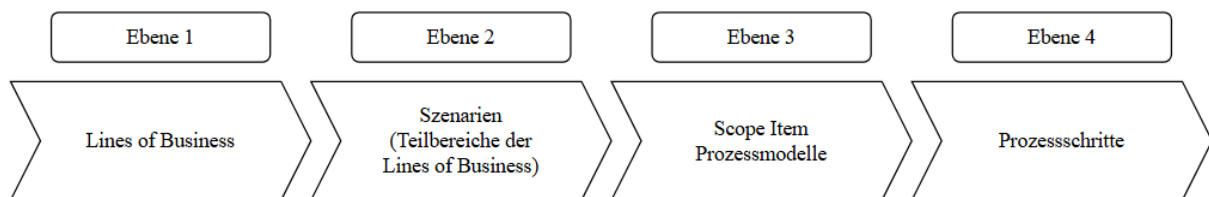


Abbildung 2: Struktur der SAP-Best Practice

Auf oberster Ebene werden die Best Practices in ihren „Lines of Business“ unterteilt. Diese stehen für klar abgegrenzte, fachliche Geschäftsbereiche innerhalb der Organisation. Jeder dieser Geschäftsbereiche umfasst auf Ebene zwei mehrere Szenarien, die typische betriebliche Abläufe innerhalb des Geschäftsbereichs abbilden. Die Szenarien werden auf Ebene drei in Scope Items unterteilt. Ein SAP Best Practice Scope Item ist ein vordefinierter Prozessbaustein, der einen konkreten Teilprozess beschreibt. Er enthält vorkonfigurierte Inhalte, wie BPMN-Prozessmodelle, welche den jeweiligen Lösungsprozess abbilden sowie dazugehörige Testskripte, die den Prozess validieren.

2.3. Enterprise Online Guide

Der Enterprise Online Guide (EOG) ist ein akademisch fundiertes Enterprise Architecture Framework, entwickelt von Hans-Jürgen Scheruhn. [vgl. Scheruhn, 2013] Das Framework verbindet betriebswirtschaftliche und technische Aspekte, die innerhalb eines Unternehmens relevant sind und stellt diese in Modellen dar. Dabei kann unter anderem zwischen Geschäftsmodellen und Aufbauorganisationen sowie Prozess- und Datenmodellen unterschieden werden. Das dabei entstehende Ergebnis ist eine Informationsarchitektur, die alle relevanten Informationen eines Unternehmens visualisiert und miteinander verknüpft. [vgl. Scheruhn, 2022]

Konkret wird der EOG, wie in Abbildung 3 dargestellt, in Form einer tabellarischen Übersicht visualisiert. Der Detaillierungsgrad der enthaltenen Modelle folgt einer vierstufigen Hierarchiestruktur. Aus horizontaler Perspektive lassen sich drei übergeordnete Mappen

unterscheiden, die jeweils unterschiedliche Unternehmensbereiche abbilden: die *Sustainability Map*, die *Intelligence Map* und die *Ecosystem Map*. Diese übergeordneten Maps beinhalten weitere Modelltypen. So dokumentiert die *Intelligence Map* unter anderem prozessuale und datenbezogene Sichten, während die *Sustainability Map* beispielsweise Unternehmensziele und strategische Ausrichtungen konkretisiert. [vgl. Scheruhn, 2022]

	SUSTAINABILITY MAP				INTELLIGENCE MAP		ECOSYSTEM MAP	
	Motivation	Organization	Function	Process	Application	Data	Technology	Network
1_Corporate Management	M1	O1	F1	P1	A1	D1	T1	N1
2_Area Management	M2	O2	F2	P2	A2	D2	T2	N2
3_Workplace	M3	O3	F3	P3	A3	D3	T3	N3
4_Experience	M4	O4	F4	P4	A4	D4	T4	N4

©Prof. Dr.-Ing. Hans-Jürgen Scheruhn

Abbildung 3: EOG-Matrix

3. Forschungsmethode

Dieses Paper folgt den Phasen des Design Science Research Vorgehens nach Peffers. [vgl. Peffers et. al., 2007] Im ersten Kapitel wurde das Problem identifiziert und beschrieben. Darauf aufbauend wurden die in Tabelle 1 dargestellten Forschungsfragen formuliert. Diese dienen als Grundlage für die Entwicklung, Demonstration und anschließende Evaluierung des Lösungsansatzes. Die Designphase umfasst die Analyse der Herausforderungen von SAP Best Practice Modellen und die anschließende Entwicklung eines Lösungsansatzes zur strukturierten Einbindung der Modelle in den Enterprise Online Guide. In der Demonstrationsphase wird das Vorgehen anhand eines konkreten Beispiels veranschaulicht. Die abschließende Evaluation analysiert die Wirksamkeit des Ansatzes im Hinblick auf die definierten Anforderungen aus Tabelle 2.

Anforderungsnummer	Anforderung
1	Erfassung der Herausforderungen von SAP-Best Practices
2	Einordnung der SAP-Best Practices in die Enterprise Online Guide Konventionen und Strukturen

Tabelle 2: Anforderungen

4. Analyse der Herausforderungen von SAP-Best Practices

Um Herausforderungen im Umgang mit SAP Best Practices möglichst praxisnah zu analysieren, wurde im Rahmen der Bachelorarbeit „*Einordnung der SAP-Best Practice Modelle in das Enterprise Online Guide Rahmenwerk*“ [vgl. Gußmag, 2025] ein Experteninterview durchgeführt. Die Ergebnisse wurden mit den Erkenntnissen aus einer projektbezogenen Befragung eines Beratungsunternehmens verglichen. Darüber hinaus wurden

Umfrageergebnisse einer Studie des IT-Onlinemagazins in Kooperation mit der Scheer Group einbezogen und wiedergegeben. Die Erkenntnisse werden in Tabelle 3 dargestellt.

	Experten-Interview	Weitere Forschungsergebnisse	
Kategorien	Befragung eines Senior SAP Consultants	Convista-Befragung [vgl. Convista, 2024]	IT-Onlinemagazin Umfrage [vgl. it-onlinemagazin, 2023]
Relevanz	“Best Practices bieten eine gute Grundlage, für die Anpassung bestehender Prozesse an SAP-Standards”	“SAP Best Practices dienen als Orientierungshilfe, können jedoch nicht alle individuellen Anforderungen abdecken”	“Nutzung von SAP Best Practice Referenzprozessen ist für 59% der befragten Unternehmen wichtig”
Standardisierungsgrad	“Unternehmen zeigen zunehmendes Interesse an der Standardisierung ihrer Prozesse”	“SAP Best Practices sind vielmehr eine vorgefertigte Konfiguration[...]um bei der Implementation von S/4Hana zu unterstützen”	“Kunden setzen verstärkt auf die Best Practices von SAP und nutzen den Umzug nach S/4HANA als Gelegenheit zur Prozessoptimierung“
Individualisierung	“Die hohe Anzahl an Best Practices erfordert eine Selektierung und Individualisierung relevanter Best Practice Modelle”	“Jedes Unternehmen hat individuelle Prozesse, die von den generischen Best Practices abweichen können. Es ist daher wichtig, diese Unterschiede zu erkennen und anzupassen.”	„Bei der Wahl der Migrationsstrategie sollte man berücksichtigen, ob die SAP Best Practice Prozesse später tatsächlich genutzt werden können“

Tabelle 3: Herausforderungen von SAP Best Practices

Es zeigt sich, dass die SAP Best Practices eine gute Grundlage für S/4HANA Transformationen und Neueinführungen bieten, wobei die Modelle stark an die individuellen Prozessabläufe der Unternehmen angepasst werden müssen. Aufgrund der hohen Standardisierung der Prozesse können diese zudem zur Prozessverbesserung verwendet werden, vor allem wenn ein möglichst naher Standard das Ziel ist. Die SAP Best Practices sind daher auch eine ideale Grundlage für die Phase vier des BPM-Lebenszyklus.

5. Einordnung der SAP-Best Practices in den EOG

Nachdem die Herausforderungen von SAP Best Practice Modellen dargestellt wurden, stellt sich die Frage, wie die Modelle optimal in die Unternehmensprozessarchitektur eingebettet werden können. Dazu ist es notwendig, die Strukturen und Konventionen der Best Practice Modelle mit denen des Enterprise Online Guide Frameworks abzugleichen. Ziel ist die Entwicklung eines methodischen Vorgehens, das eine Integration der SAP Best Practices in einen einheitlichen, End-to-End-orientierten Kontext ermöglicht. Durch die Einbindung der Scope Items in die Prozesslandkarte des EOGs soll eine Vernetzung mit dem gesamten Framework gewährleistet werden.

Eine wesentliche Erkenntnis, die sich bei der Analyse der SAP Best Practice Modelle ergeben hat, liegt in ihrer starken Modularisierung. Die Modelle sind weitgehend isoliert und oftmals ohne klaren Bezug zu vor- oder nachgelagerten Prozessen modelliert. Ein übergreifender Zusammenhang ist nur schwer zu erkennen. Die Identifikation dieser Zusammenhänge muss in jedem Unternehmen individuell und mit Eigenaufwand vorgenommen werden.

Vor diesem Hintergrund stellt die Überführung der Best Practice Modelle in das EOG-Framework einen strukturierten Lösungsansatz dar. Dabei werden die Scope Items in eine End-to-End-Prozessstruktur eingebettet, die es ermöglicht, SAP-spezifische und nicht-SAP-spezifische Aktivitäten gemeinsam abzubilden. Wie auch das Experteninterview bestätigt,

bestehen reale Geschäftsprozesse nicht ausschließlich aus SAP-Prozessschritten. Vielmehr enthalten sie zusätzliche Aktivitäten, die unabhängig von SAP-Systemen durchgeführt werden, aber dennoch Teil des Gesamtprozesses sind.

Die Zielsetzung, für die Integration der Best Practice Modelle in den End-to-End Kontext ist es, die SAP-spezifischen Teilprozesse einem konkreten Best Practice Scope Item zuzuordnen. Die Scope Item Prozesse können aus den Szenarien abgeleitet und anschließend der zweiten Ebene der EOG-Prozess Map zugeordnet werden.

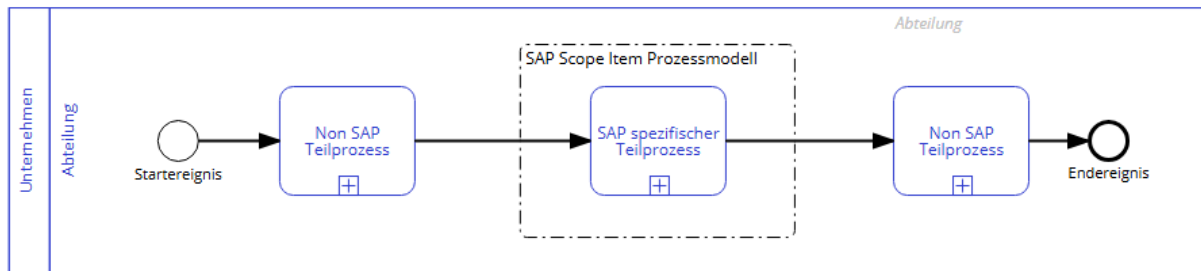


Abbildung 4: Exemplarischer End-to-End-Prozess

Diese Sichtweise spiegelt sich auch in Abbildung 4 wieder. Innerhalb eines End-to-End-Prozesses werden verschiedene Teilprozesse modelliert, die aufeinander aufbauen. Jeder Prozessschritt besitzt, mit Ausnahme des ersten und letzten, einen definierten Vorgänger und Nachfolger. Die Strukturierung erfolgt im EOG auf der zweiten Hierarchiestufe; der Abteilungsebene. In dieser werden die Teilprozesse einem organisatorischen Kontext zugeordnet. Der zentrale Mehrwert liegt in der klaren Zuordnung von SAP Scope Items zu konkreten Teilprozessen innerhalb des End-to-End-Prozesses. Anders als in den Best Practices, in denen Scope Items fragmentiert dargestellt werden, entsteht durch die End-to-End-Struktur eine kohärente Prozesskette, die alle beteiligten Organisationseinheiten berücksichtigt.

Durch die Einbettung wird nicht nur das Verständnis einzelner Scope Items erleichtert, sondern auch die Transparenz und Nachvollziehbarkeit der einzelnen Modelle im Gesamtprozess erhöht. Voraussetzung hierfür ist die Anwendung klar definierter Modellierungskonventionen, wie sie im EOG vorgegeben sind. [vgl. Scheruhn, 2023] Diese stellen sicher, dass sowohl Rollen, als auch Prozessschritt-Referenzen konsistent modelliert werden.

Element	SAP-Best Practice Modell	EOG-Konvention	EOG-Glossareintrag
Pool	Bezeichnung des Scope Items	Unternehmensname	1_Org_Unit
Lane	Bezeichnung der Rolle	Abteilung / Bereich	2_Org_Unit
Rolle		Rollenbezeichnung	3_Role
Ereignis	Inkonsistente Konvention	Objekt-„zu“-Infinitiv Verb-Stil	3_Event
Aktivität	Inkonsistente Konvention	Objekt-Verb-Stil	3_Activity
Dokument	Nicht dokumentiert	Dokumentname	3_Document

Tabelle 4: Mapping der Elementtypen zwischen SAP-Best Practices und den Enterprise Online Guide

Das Ergebnis ist ein integriertes Prozessmodell, in dem die Best Practice Scope Items in einen logischen und verständlichen Zusammenhang gebracht werden. Im Folgenden zeigt Tabelle 4 ein Mapping der Elementtypen aus den Best Practice BPMN-Modellen mit den EOG-Modellierungskonventionen sowie den zugehörigen Glossareinträgen.

Eine Erkenntnis aus der Gegenüberstellung ist, dass SAP Best Practice Modelle ausschließlich aus einer rollenbasierten Perspektive modelliert sind. Hierin liegt jedoch ein wesentlicher Kritikpunkt. Die Modelle sind nicht in eine abteilungsübergreifende Prozesssicht eingebettet, sondern verfolgen eine Sichtweise, die den organisatorischen Kontext weitgehend ausblendet. Tabelle 4 verdeutlicht zusätzlich, dass die aktuelle Modellierung der Best Practice Modelle inkonsistent ist. Sowohl bei der Darstellung von Ereignissen als auch bei der Benennung von Aktivitäten fehlen verbindliche Konventionen. Formulierungen variieren, wodurch die Verständlichkeit erschwert wird. Auch der Dokumentenfluss wird innerhalb der Scope Item Modelle nicht abgebildet, was die Nachvollziehbarkeit und Vollständigkeit der Prozesse einschränkt.

Um die konsistente Modellierung der BPMN-Modelle zu gewährleisten, kann auf die Konventionen der dritten Hierarchiestufe des EOG-Frameworks zurückgegriffen werden. Neben der Strukturierung des BPMN-Pools in drei Ebenen, welche die Unternehmensführungs-, Abteilungs- und Arbeitsplatzebene untergliedern, setzt das Framework ebenfalls auf eine konsistente Benennung der Ereignisse und Aktivitäten. Dabei werden die Aktivitäten in einem „Objekt-Verb-Stil“ modelliert und die Dokumente den spezifischen Prozessschritten zugeordnet, um sowohl den Prozess- als auch den Datenfluss abzubilden. Darüber hinaus werden den Objekten EOG-Glossareinträge zugeordnet, sodass die Vernetzung in andere Modelle gewährleistet wird.

6. Demonstration

Um die Anwendung der Konventionen anschaulich und praxisnah darzustellen wird das Vorgehen anhand eines exemplarischen End-to-End-Prozesses durchgeführt. Dabei bietet es sich an den Source-to-Contract-Prozess als Beispiel zu wählen, da dieser sehr strukturiert visualisierbar ist.

Der Prozess beginnt üblicherweise mit der Identifizierung eines Bedarfs und endet mit der Unterzeichnung eines Lieferantenvertrages. Der Fokus liegt dabei vor allem in der Optimierung der Lieferantenauswahl sowie dem Aufbau und Erhalt langfristiger Partnerschaften mit wichtigen Lieferanten. Das Verfahren setzt sich sowohl aus transaktionalen sowie manuellen Aktivitäten zusammen, die nicht zwingend an ein ERP System gebunden sind. [vgl. Hofbauer, 2023] Der gesamte End-to-End-Prozess lässt sich in fünf Prozessbausteine unterteilen, die sich in Abbildung 5 wiederfinden.

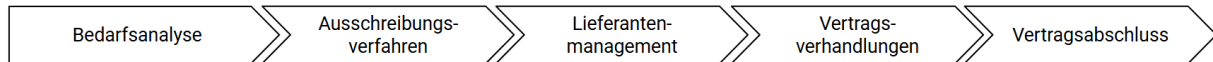


Abbildung 5: Source-to-Contract-Prozess

Die Integration der Best Practice Modelle in den End-to-End-Prozess beginnt auf der ersten Ebene der EOG-Prozess Map. Diese Ebene bildet die End-to-End-Prozesse des Unternehmens strukturiert ab. (vgl. Abbildung 6)

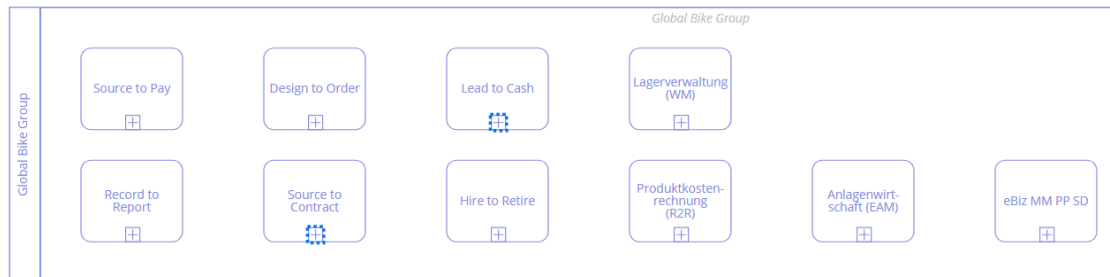


Abbildung 6: EOG-Prozessebene 1

Möchte man den Detaillierungsgrad nun erhöhen und den End-to-End-Prozess aus einer abteilungsbezogenen Perspektive betrachten, kann in die zweite Ebene navigiert werden. Im Beispiel des Source-to-Contract-Prozesses zeigt sich, dass der End-to-End Prozess zwei Abteilungen involviert. Dies wurde im Rahmen von Workshop-Analysen ermittelt, indem jede Prozessaktivität analysiert und einer verantwortlichen Abteilung zugeordnet wurde.

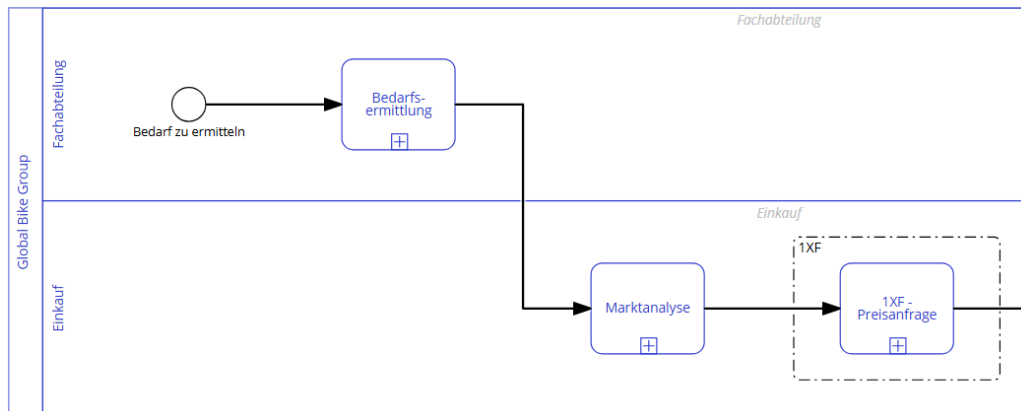


Abbildung 7: Teilprozesse des Source-to-Contract-Prozesses in EOG-Prozessebene 2 (Auszug)

Abbildung 7 zeigt die Teilprozesse zwischen der Fachabteilung, die den Bedarf identifiziert und der Einkaufsabteilung, die sich um die Abwicklung des Bedarfes kümmert. Den Hauptteil des End-to-End-Prozesses übernimmt die Einkaufsabteilung, nachdem die Informationen aus der Fachabteilung übermittelt wurden.

Wichtig ist an dieser Stelle hervorzuheben, dass die Bedarfsermittlung und die Marktanalyse Teilprozesse darstellen, die nicht zwingend an ein SAP-System gebunden sind. Dennoch können die Prozesse auf einer feingranularen Ebene modelliert werden. Eine Einordnung des SAP Best Practice Contents erfolgt zu diesem Zeitpunkt noch nicht. Der zentrale Aspekt, welcher in dem vorherigen Abschnitt erarbeitet wurde, stellt fest, dass die Best Practice Scope Items nun in den End-to-End-Prozess integriert werden können, wodurch sie einen konkreten Bezug zum Gesamtprozess erhalten. Anhand des Source-to-Contract-Beispiels folgt auf die

Marktanalyse ein Teilprozess, welcher im SAP-System durchgeführt wird. Dieser beschreibt die Preisanfrage an einen Lieferanten und lässt sich aus dem Scope Item „1XF“ ableiten. Durch die Integration in das P2-Modell erhält das Scope Item einen konkreten Kontext, der eine eindeutige Zuordnung ermöglicht. Um das Modell aus arbeitsplatzorientierter Perspektive zu betrachten, kann auf die dritte Ebene des EOGs navigiert werden.

Abbildung 8 zeigt einen Auszug des Scope Item Modells *1XF - Preisanfrage*. In dem Pool wird die Bezeichnung des Scope Items beschrieben. Innerhalb der Lane werden die Prozessschritte der zuständigen Rolle zugeordnet.

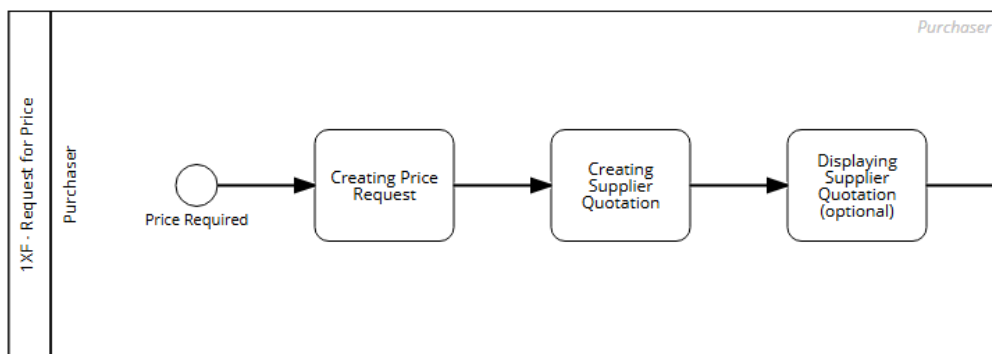


Abbildung 8: SAP Best Practice Scope Item Modell "1XF-Preisanfrage" (Auszug)

Ordnet man das Modell nun in die EOG-Struktur ein, wird das Modell in 3 Hierarchiestufen abgebildet. (vgl. Abbildung 9) Die erste beschreibt die Bezeichnung des Unternehmens. Anschließend werden auf der zweiten Hierarchiestufe die relevanten Abteilungen für den Prozess abgebildet. Zuletzt findet sich auf der dritten Hierarchiestufe die Rolle, welche die Prozessschritte ausführt. Die Bezeichnung der Rolle kann hierbei aus dem Best Practice Modell entnommen werden oder ergibt sich, wie die Information der Abteilung aus Prozesserhebungen. Innerhalb der Lane kann anschließend der Prozessfluss visualisiert werden. Der Ausschnitt des Modells in Abbildung 9 beinhaltet die Erstellung einer Preisanfrage sowie das Anlegen und Anzeigen eines Lieferantenangebots.

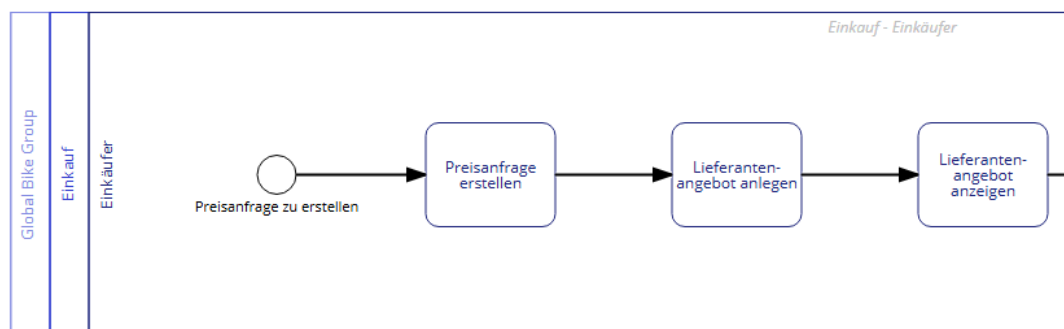


Abbildung 9: SAP Best Practice Scope Item Modell "1XF-Preisanfrage" in EOG-Prozessebene 3 eingeordnet (Auszug)

7. Evaluierung

Die Evaluierung orientiert sich am Vorgehensmodell von Sonnenberg und vom Brocke [vgl. Sonnenberg, 2012] Zur Bewertung der in Kapitel 2 definierten Anforderungen wurden

geeignete Evaluierungskriterien ausgewählt. Sonnenberg und vom Brocke stellen hierfür eine Vielzahl möglicher Kriterien zur Verfügung. Für den Kontext dieses Papers bietet es sich an, zwischen den Kriterien *Durchführbarkeit* der Anforderungen, *Einfachheit* der Umsetzung sowie *Nützlichkeit* des Vorgehens zu unterscheiden. Tabelle 4 stellt die Anforderungen dar und ordnet sie den Kriterien zu. Die Bewertung erfolgt nach den Faktoren: + *erfüllt*, o *teilweise erfüllt*, - *nicht erfüllt*

ID	Beschreibung	Durchführbarkeit	Einfachheit	Nützlichkeit
1	Erfassung der Probleme von SAP Best Practices	+	+	+
2	Einordnung der SAP Best Practices in die EOG-Konventionen und Strukturen	+	o	+

Tabelle 5: Evaluierung der Anforderungen

Die Auswertung erfolgt auf Basis der Analyse aus Kapitel 4 sowie den Erkenntnissen aus Kapitel 5 und 6. Um die Probleme der SAP Best Practices zu erarbeiten, wurden zentrale Herausforderungen in einem Experteninterview erfasst. Die Erkenntnisse konnten durch Erfahrungen aus weiteren Studien und Befragungen ergänzt werden. Die Durchführung des Interviews ist aufgrund der gewonnenen Expertise eines fachkundigen Experten sehr gut durchführbar.

Die Identifikation der Herausforderungen stellt einen problemorientierten Ausgangspunkt für die Forschung dar und bildet einen essenziellen Einstieg in die Thematik. Aus den identifizierten Herausforderungen konnten im weiteren Verlauf Kategorien abgeleitet werden, auf deren Basis ein Lösungsansatz entwickelt wurde. Dieser soll zur Verbesserung der Verständlichkeit und Lesbarkeit von SAP Best Practice Modellen beitragen. Die Nützlichkeit des Vorgehens ist somit als erfüllt einzustufen.

Die Einordnung der SAP Best Practice Modelle erfolgt auf der zweiten und dritten Hierarchieebene des EOGs und betrifft somit Prozesse auf Abteilungs- und Arbeitsplatzebene. Dabei werden Prozessaktivitäten angepasst und konsistent modelliert. Der Lösungsansatz wurde in Kapitel 5 konzeptioniert und in Kapitel 6 dargestellt. Die Durchführbarkeit gilt damit als gegeben. Eine wesentliche Herausforderung der Einordnung liegt im hohen Implementierungsaufwand. Um die Konventionen des EOGs gezielt anzuwenden, müssen die Beziehungen zwischen verschiedenen Scope Items sowie ergänzenden Non-SAP-Prozessen identifiziert werden. Dies erfordert eine detaillierte Analyse, häufig in Form von Prozesserhebungs-Workshops, um die Teilprozesse zu einem vollständigen End-to-End-Prozess zusammenzuführen. Dies erfordert einen hohen Aufwand, wodurch die Einfachheit nur teilweise erfüllt ist.

Die strukturierte Einbettung der Scope Items in den Gesamtprozess verbessert das Verständnis für die praktische Nutzung der SAP Best Practices. Es wird klarer, wie die Modelle im operativen Geschäft angewendet werden können und welchen Beitrag sie zu konkreten Geschäftsprozessen leisten. Kapitel 5 belegt zudem den hohen Nutzen der EOG-Konventionen im Zusammenhang mit den Best Practice Modellen. Neben einer konsistenteren Modellierung ermöglicht die Integration eine zielgerichtete Anpassung an die unternehmensspezifischen Anforderungen. Ein zentraler Vorteil liegt in der Vernetzung mit anderen Modellen innerhalb

des EOG-Frameworks. Dadurch entsteht eine durchgängige Ontologie, die das Verständnis bereichsübergreifend fördert und die Kommunikation zwischen Abteilungen verbessert. Während individuell entwickelte Unternehmensprozesse stark variieren können, bieten die EOG-Konventionen in Verbindung mit den SAP Best Practices eine wiederverwendbare, strukturierte Grundlage. Dies fördert die Abstimmung zwischen Fachbereichen und schafft Einheitlichkeit bei der Dokumentation von Prozessen. Die Kategorie Nützlichkeit ist daher als vollständig erfüllt zu bewerten

Nachdem die Evaluierung der Anforderungen abgeschlossen ist, können nun die Forschungsfragen beantwortet werden. Die erste Forschungsfrage lautete: „*Welche Herausforderungen entstehen bei der Verwendung von SAP Best Practices?*“ Zur Beantwortung dieser Frage wurde ein Experteninterview durchgeführt und mit weiteren Forschungsergebnisse verglichen. Dabei zeigte sich, dass der Zusammenhang zwischen SAP Best Practice Modellen aufgrund ihrer Modularisierung nur schwer nachvollziehbar ist. Zudem erfordert die praktische Anwendung der Modelle häufig einen erheblichen Anpassungsaufwand, da die bereitgestellten Prozesse oft nicht den tatsächlichen Abläufen im Unternehmen entsprechen. Daraus ergibt sich die Notwendigkeit der Individualisierung.

Die zweite Forschungsfrage beinhaltete die Fragestellung: „*Ist eine Einordnung der SAP Best Practice Modelle in die EOG-Konventionen möglich?*“ Wie sich durch die Konzeption in Kapitel 5 sowie die Demonstration in Kapitel 6 herausgestellt hat, können die SAP Best Practice Scope Items in eine End-to-End-Prozessstruktur integriert werden, und die Anwendung der EOG-Konventionen ist möglich. Allerdings handelt es sich dabei um einen komplexen Prozess, der zu Beginn einen hohen Planungsaufwand erfordert und eine Vielzahl an Informationen benötigt. Ein positiver Aspekt in diesem Zusammenhang ist, dass die Komplexität durch zunehmende Erfahrung deutlich sinkt, während die Umsetzbarkeit gleichzeitig steigt. Verwendete Best Practices können auf Grundlage etablierter Strukturen und Referenzen aus früheren Projekten wiederverwendet werden, was die Herausstellung des Zusammenhangs zwischen End-to-End-Prozessen und Scope Items vereinfacht. Durch die EOG-Konventionen und die Ergebnisse des Papers lässt sich der Schluss ziehen, dass die SAP Best Practices anpassbar sind und eine Einordnung der EOG-Konventionen in die Best Practice Modelle einen Mehrwert bietet.

8. Fazit

Die Forschungsergebnisse zeigen, dass SAP Best Practices trotz ihrer Standardisierung mit einer Reihe praxisrelevanter Herausforderungen verbunden sind. Insbesondere der hohe Anpassungsaufwand erfordert eine gezielte Individualisierung. Die im Rahmen des Papers durchgeführte Analyse belegt, dass eine Einordnung der Best Practices in den Enterprise Online Guide nicht nur möglich, sondern auch sinnvoll ist. Der entwickelte Lösungsansatz zur Einbettung der Scope Items verbessert die strukturelle Konsistenz und fördert das unternehmensweite Verständnis für Geschäftsprozesse. Trotz des initialen Implementierungsaufwands lassen sich durch Wiederverwendung und zunehmende Projekterfahrung langfristig Effizienzgewinne erzielen. Insgesamt zeigt sich, dass die Verknüpfung von SAP Best Practices mit dem Enterprise Online Guide eine vielversprechende Grundlage für eine verbesserte Prozessmodellierung bietet und einen Mehrwert für die Praxis darstellt.

9. Literatur

- Convista Consulting AG (2024) Warum SAP Best Practices im S/4HANA-Projekt niemals die vollständige Lösung sein können. Abgerufen am 28.7.2025 von <https://www.convista.com/impulse/thema/s4hana/warum-sap-best-practices-im-s4hana-projekt-niemals-die-vollstaendige-loesung-sein-koennen/>
- Dumas, M., La Rosa, M., Mendling, J., Reijers, Hajo A. (2021) Grundlagen des Geschäftsprozessmanagements. S. 7, S. 19ff., S. 26 Springer Berlin, Heidelberg. DOI: <https://doi.org/10.1007/978-3-662-58736-2>
- Gußmag, L. (2025) Einordnung der SAP-Best Practice Modelle in das Enterprise Online Guide Rahmenwerk, S. 29ff.
- Hofbauer, G., Kececioğlu, S. (2023) Digitalisation of the Source-to-Contract Process: Evaluation and Proof of Concept. S. 8ff. Abgerufen am 28.7.2025 von https://www.thi.de/fileadmin/daten/Working_Papers/thi_workingpaper_68_Hofbauer_Gueanter.pdf
- It-onlinemagazin (2023) SAP S/4HANA Studie: Überraschende Zahlen zu Best Practices, Prozessanalysen, Cloud-Nutzung und Ressourcenengpässen, Abgerufen am 28.7.2025 von <https://it-onlinemagazin.de/sap-s-4hana-studie-ueberraschende-zahlen-zu-best-practices-toolgestuetzter-prozessanalyse-cloud-nutzung-und-ressourcenengpaessen/>
- Munkacsı, I., Orosz, T. G., Alexy, M. (2024) Implementation Challenges of Industry-Specific ERP System Solutions by Utilizing Best Practices. S. 4, Springer, Singapore. DOI: https://doi.org/10.1007/978-981-97-3442-9_32
- Peffer, K., Tuunanen, T., Rothenberger, M. A., Chatterjee, S. (2007) A design science research methodology for information systems research. S. 282-305. Abgerufen am 28.7.2025 von https://www.researchgate.net/publication/284503626_A_design_science_research_methodology_for_information_systems_research
- SAP (2025) Abgerufen am 28.7.2025 von https://help.sap.com/docs/SUPPORT_CONTENT/sm/3518046512.html?locale=de-DE
- Scheruhn, H.-J., Ackermann, D., Braun, R., Förster, U. (2013) Repository-Based Implementation of Information Pyramid: A Study Based on an ERP Case Study. S. 446 ff, Springer, Berlin, Heidelberg. DOI: https://doi.org/10.1007/978-3-642-39265-8_50
- Scheruhn, H.J., Schulze, M., Fotteler, M., Mendling, J., Bake, B. (2023) Verbesserung der ERP-Lehre im Kontext des BPM-Lebenszyklus. S. 174ff. Abgerufen am 28.7.2025 von <https://mediatum.ub.tum.de/doc/1719876/document.pdf>
- Scheruhn, H.-J., Reiter, C., Bayramli, E. (2022) Von Process Mining hin zu Enterprise Mining. S. 228ff. DOI: <https://doi.org/10.14459/2022md1685828>
- Sonnenberg, C., Brocke, J. v. (2012) Evaluations in the Science of the Artificial – Reconsidering the Build-Evaluate Pattern in Design Science Research. S. 11ff. Springer, Berlin, Heidelberg. DOI: https://doi.org/10.1007/978-3-642-29863-9_28
- Spontoni, E., Krawczyk, M., Buenen, M., Kumar, VJ. (2023) State of Worldwide Business Assurance for SAP solutions – 2023. S. 28 ff. Abgerufen am 28.7.2025 von

SAP Academic Community Conference 2025 (D-A-CH)

Business Process Management drives Digital Transformation

<https://www.sogeti.com/wp-content/uploads/sites/3/2024/10/state-of-worldwide-business-assurance-for-sap-solutions-2023.pdf>

ESG meets EOG – Integration von Environment Social Governance (ESG) in die Technologie- und Netzwerk-Map des Enterprise Online Guide (EOG)

E. Bayramli²), V. Bekiri¹), S. Stöckler¹), K. Oettmeier¹), R. Hänggi¹), & H. Scheruhn³)

¹) OST – Ostschweizer Fachhochschule, St. Gallen, Schweiz

²) msg systems ag, München, Deutschland

³) Hochschule Harz, Wernigerode, Deutschland

Abstract: In den letzten Jahren sind sowohl das Thema Nachhaltigkeit als auch die Komplexität von IT-Landschaften von Randthemen zu strategischen Hauptthemen geworden, getrieben durch ESG-Regulierungen, Energiekrisen und digitale Transformationsinitiativen. Die nachhaltige Gestaltung von Produkten, Dienstleistungen und IT-Infrastrukturen hat für Unternehmen eine hohe Priorität erlangt. Weiterführend zur Veröffentlichung der ESG meets EOG (Bekiri et al. 2024), werden in diesem Beitrag die übrigen zwei Maps des Enterprise Online Guide (EOG) in Verbindung mit Environment Social Governance (ESG)-Aspekten beleuchtet. Um Organisationen ganzheitlich bei der Einführung eines integrierten Nachhaltigkeits- und Technologiemanagements zu unterstützen, wird die Systematik des Enterprise Architecture Management (EAM) durch die Kombination von SAP LeanIX, einem führenden Werkzeug für EAM, mit dem EOG genutzt.

Die *Technology*- und *Network*-Maps des EOG ermöglichen es, nachhaltige IT-Architekturen zu planen, während LeanIX dessen architektonische Planung unter Berücksichtigung von Green-IT-Metriken wie Energieverbrauch, CO₂-Emissionen der IT-Infrastruktur unterstützt. Durch die Ergänzung der Sustainability Balanced Score Card mit den letzten beiden EOG-Perspektiven können Unternehmen ihre Technologieportfolios nicht nur nach Effizienz, sondern auch nach Nachhaltigkeitskriterien optimieren. Somit dient die Verbindung von LeanIX und EOG als umfassendes Navigationswerkzeug für eine holistische Nachhaltigkeits- und Technologieplanung, die es ermöglicht, ESG-Ziele systematisch in der gesamten IT-Landschaft zu verankern und messbar zu machen. Die Vorgehensweise wird anhand eines kurzen Fallbeispiels eines Produktionsprozesses an der SmartFactory@OST dargestellt.

Keywords: Enterprise Architecture Management, Nachhaltigkeitsmanagement, Enterprise Online Guide, Corporate Social Responsibility, Environmental Social Governance

1. Einleitung

Laut einer Umfrage von SAP aus dem Jahr 2021 nannten 32% der befragten Führungskräfte in Fertigungsunternehmen die nachhaltigere Gestaltung von Produkten und Dienstleistungen als ihre wichtigste Priorität für Umsatzsteigerungen (SAP Signavio Team 2023). Um diese Umgestaltung voranzutreiben, müssen vor allem die Prozesse nicht nur ökonomisch, sondern auch ökologisch und sozial optimiert werden. Das Prozessmanagement spielt eine entscheidende Rolle bei der Umsetzung von Nachhaltigkeitsinitiativen in Unternehmen, da die Prozesse die zentralen Ausführungsträger betrieblicher Leistungen sind und somit eine systematische Berücksichtigung von Nachhaltigkeitsaspekten gewährleisten können. (FACTORY net und Karl Wagner 2020)

Eine Visualisierung verschiedener aktueller Herausforderungen für Firmen, um am Markt bestehen zu können, zeigen u.a. (Kirchhoff et al. 2024, S. 5) auf. Sie geben das Thema «Environment, Social, Governance (ESG)» als einen von vier Haupt-Herausforderungen an, aber auch in den Bereichen «Geopolitik/Geoökonomie» und «Gesellschaftlicher Wandel» finden sich Corporate Social Responsibility (CSR)-Themen, wie Klimawandel, Umwelt- und Sozialthemen oder Arbeit 4.0. All diesen Herausforderungen müssen sich Unternehmen täglich stellen und auf neue Anforderungen durch drastischen Wandel in den Umwelteinflüssen gerüstet sein.

Der EOG ist ein EAM-Framework und dient als Grundlage für die Umsetzung dieser Arbeit. Die Fokussierung auf die letzten beiden Maps des EOG – *Technology* und *Network* – eröffnet besonders interessante Perspektiven für die Implementierung von ESG-Strategien mit LeanIX als Enterprise Architecture Management Tool. Diese beiden finalen Dimensionen des EOG vervollständigen das ganzheitliche Framework für die Nachhaltigkeitsplanung und bieten konkrete Ansatzpunkte für die technologische Operationalisierung von ESG-Zielen. (Bekiri et al. 2024) Dadurch entsteht die folgende Forschungsfrage „Wie können die Technology- und Network-Maps des Enterprise Online Guide (EOG) systematisch mit ESG-Kriterien verknüpft werden, um eine ganzheitliche Nachhaltigkeitsplanung in der IT-Architektur zu ermöglichen?“ und dementsprechend ist das Forschungsziel auch die Vervollständigung des EOG-Frameworks zur Integration von ESG in IT-Architekturen unter Verwendung von SAP LeanIX.

1.1. ESG-Dimensionen und deren Zusammenwirken

Die Implementierung von Nachhaltigkeitsmanagement in Unternehmen basiert auf den drei miteinander verflochtenen Dimensionen der ESG-Kriterien. Die *ökologische Dimension* umfasst weit mehr als nur CO₂-Reduktion und zielt auf die umfassende Minimierung von Umweltbelastungen durch Schadstoffe, Ressourcenverbrauch, Kreislaufwirtschaft und Abfallproduktion ab. Gleichzeitig erfordert die *soziale Dimension* hohe Standards in Arbeitssicherheit, fairen Arbeitsbedingungen und Gleichstellung, die sich sowohl auf interne als auch externe Stakeholder erstrecken. Die *ökonomische Dimension* integriert effizienten Ressourceneinsatz und langfristige Geschäftsmodelle, wobei bereits bestehende Prozessparameter wie Kosten und Zeit teilweise abgedeckt werden. (Hinrichs 2021, S. 34–35) (Schneider 2015, S. 872) (Podleisek et al. 2025)

Während im Bereich des Green Business Process Management bereits Ansätze zur Messung ökologischer Aspekte wie Treibhausgasemissionen existieren, bleibt die soziale Nachhaltigkeit und deren Wechselwirkungen mit ökologischen Faktoren im Prozessmanagement weitgehend unberücksichtigt. Diese Integration aller drei Dimensionen bildet die Grundlage für ein ganzheitliches Nachhaltigkeitsmanagement. (Recker et al. 2012) (Schoormann et al. 2017, S. 2)

1.2. Systematische Entwicklung des Nachhaltigkeitsmanagements

Der Aufbau eines effektiven Nachhaltigkeitsmanagements orientiert sich an einem vierstufigen Reifegradmodell von CSR 0.0 bis CSR 3.0, das nach oben offen konzipiert ist. Entscheidend für den langfristigen Erfolg ist die Erreichung mindestens der Stufe CSR 2.0, auf der Nachhaltigkeitsthemen nicht als zusätzliche parallele Aufgaben, sondern als integrale Bestandteile der Unternehmensführung etabliert werden. Erst diese systematische Integration in bestehende Managementsysteme gewährleistet nachhaltige Auswirkungen und vermeidet die

Behandlung von Nachhaltigkeit als isoliertes Addendum. (Schneider 2015)(Schneider und Schmidpeter 2015, S. 25–37);(Schaltegger et al. 2006, S. 10).

Eine einfache und kostengünstige IT-Integration ist von zentraler Bedeutung für ein erfolgreiches Nachhaltigkeitsmanagement. Dadurch wird die Digitalisierung ein integraler Baustein der Implementierung. Die Operationalisierung erfolgt durch eine erweiterte Sustainability Balanced Scorecard, die verschiedene Perspektiven des Unternehmens strukturiert abbildet und messbare Ziele definiert. Diese methodische Herangehensweise schafft die notwendige Verbindung zwischen strategischen Nachhaltigkeitszielen und operativer Umsetzung. (Figge et al. 2002; Hansen 2010)

Um eine Brücke zwischen dieser und weiteren notwendigen Methoden und Modellen zur Operationalisierung von ESG zu schlagen, wird nachfolgend ein kleiner Überblick der EOG von (Scheruhn et al. 2023) als übergeordnetes Management-Tool gegeben und der Fokus dann auf die letzten beiden Maps Technology und Network gesetzt.

2. Der Enterprise Online Guide – EOG

Der EOG entwickelte sich aus der Vision von (Scheruhn et al. 2024, S. 3), ein systematisches Navigationsinstrument für die Unternehmensführung zu entwickeln. Das Framework basiert auf bewährten Enterprise Architecture-Ansätzen, insbesondere dem The Open Group Architecture Framework (TOGAF) (vgl. z.B. The Open Group 2022) und dem Zachman Framework (Zachman 1987).

2.1. EOG Maps und Abstraktionsebenen

Die vertikale Gliederung des EOG umfasst vier hierarchische Abstraktionsebenen, beginnend mit der Unternehmensführung (Corporate Management) an der Spitze, gefolgt von der Bereichsleitung (Area Management), der Arbeitsplatzebene (Workplace) und schließlich der Erfahrungsebene (Experience). Horizontal betrachtet bietet das Framework acht verschiedene Sichtweisen, auch Maps genannt, auf die Organisation, wodurch eine durchgängige und strukturierte Betrachtungsweise ermöglicht wird.

	SUSTAINABILITY MAP				INTELLIGENCE MAP		ECOSYSTEM MAP	
	Motivation	Organization	Function	Process	Application	Data	Technology	Network
1. Corporate Management	Sustainability Balanced Scorecards, Digital Business Model Canvas.	Company Group	Business Capability Tree (Process oriented).	End-to End-Processes like Source-to-Pay (S2P), Lead-to-Cash (L2C), R2R.	IT Systems	Big Data	System landscape of extended Ecosystem.	Cloud Service-Models, extended Ecosystem.
2. Area Management	External / internal customer-/ supplier-relationships.	Area Departments with Persons and Positions.	Assignment of Business Capabilities to SAP Solution Capabilities.	Area Process across departments.	IT System Landscape	Integration OLTP with OLAP	Topfloor / Shopfloor / Data Center.	Social and technical network-infrastructure.
3. Workplace	Customer Journey	Process oriented organizational structure with roles	Business Capability characteristics (like "Create Order").	Process across workplaces	Single Step (Use Case)	Data Model	Software, Hardware, Robots (RPA), Actuators, Sensors.	Licences, permissions, Router, Security
4. Experience	Key Performance Indicator (KPI) and Customer Experience.	Resources, working contract (work schedule) and Employee Experience.	Service Level Agreement (SLA) like Duration, Costs.	Process Performance Indicator (PPI)	User Experience	Tables	Platform-certificate (e.g. RAMI 4.0 conform).	Supplier Experience, BSI 200-2 / ISO 27001.

Abb. 1: Enterprise Online Guide (EOG) (Quelle: (Scheruhn et al. 2023))

Diese systematische Herangehensweise schafft unterschiedliche Perspektiven auf das Unternehmen, wobei sich das vorliegende Dokument auf die beiden abschließenden Sichtweisen *Technology* und *Network* konzentriert (vgl. (Scheruhn et al. 2024, S. 3)). Die spezifische Fokussierung auf die technologischen und netzwerkbezogenen Aspekte ergänzt das umfassende Verständnis der Unternehmensarchitektur um wesentliche infrastrukturelle Komponenten:

- **Technology:** Das Ziel ist die Abbildung technologischer Dienste bzw. der IT-Landschaft in Form von Hardware oder Software, die für Applikationen und Daten benötigt werden. Dadurch werden die Planung und Umsetzung von IoT-Anwendungen sowie modernen Produktions- und Automatisierungstechnologien im Unternehmenskontext unterstützt.
- **Network:** Sie beschreibt die Struktur, Interaktion und Kommunikation zwischen physischen und logischen Knoten innerhalb einer Enterprise-Architektur. Sie umfasst die Netzwerkinfrastruktur, Kommunikationsbeziehungen, sowie das orchestrierte Zusammenwirken von Hardware, Software, Maschinen, KI-Systemen und menschlichen Akteuren.

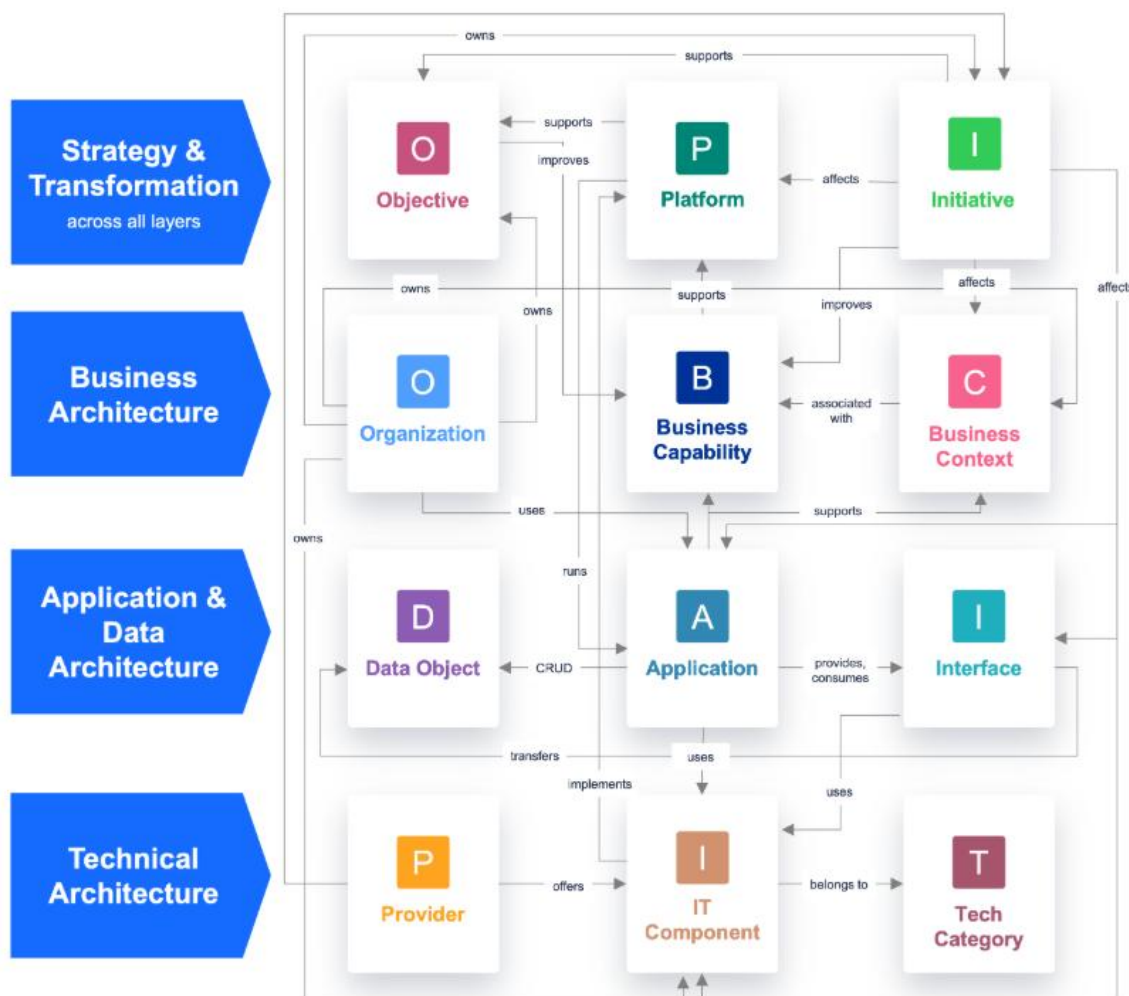


Abb 2: Meta Model v4 von LeanIX (Quelle: (SAP 2025))

(über mehrere Abstraktionsebenen hinweg) als auch horizontal (über verschiedene Maps hinweg) navigieren.

Dadurch lassen sich beispielsweise auf eine Kennzahl bezogene technologische Auswirkungen in der *Technology Map* (etwa eine konkrete IT-Komponente) oder relevante Lieferantenbeziehungen in der *Network Map* (z. B. ein betroffener Softwareanbieter) identifizieren und nachvollziehen.

	SUSTAINABILITY MAP				INTELLIGENCE MAP		ECOSYSTEM MAP	
	Motivation	Organization	Function	Process	Application	Data	Technology	Network
1_Corporate Management	Double Materiality Assessment, ESG Goals, CSR Strategy	Corporate ESG Responsibility, Chief Sustainability Officer	Assignment of ESG activities to Business Capabilities	Assignment of CSR KPIs to End-to-End Processes (e.g. P2P, O2C,...)	ESG Goals mapped to IT Applications	ESG Data Lake	Tech Platform	ESG Governance Network
2_Area Management	ESG-Goals of External / Internal Stakeholders	ESG Responsibility, Organizational Unit	Assignment of ESG activities to Business Capabilities	Assignment of CSR KPIs to Area Process (Across Departments)	ESG Goals mapped to Application-Modules	ESG Data Reports	IoT Monitoring	Social and technical network-infrastructure.
3_Workplace	Workplace ESG Objectives	ESG Responsibility, Organizational Unit	Assignment of CSR KPIs to Business Objects (e.g. Sales Order or Customer)	Assignment of CSR KPIs to Process Across Workplaces	ESG Goals mapped to SW-Functions	ESG Data Model	Tech Components	Licences, permissions, Router, Security, Media breaks..
4_Experience	CSR Key Performance Indicators (KPI)	Ecological, Economical, Social Effects	Measurement Rules for CSR KPIs	Process Performance Indicator (PPI) including CSR KPIs	CSR Measurement	ESG relevant Tables with Key-/ foreign key relationships.	Platform-certificate (e.g. RAMI 4.0 conform).	Supplier Experience, BSI 200.2 / ISO 27001, Blockchain.

Abb. 4: EOG mit ESG/CSR-relevanten Inhalten (Quelle: (Eigene Darstellung))

Erweiterung der *Technology Map*

Die *Technology Map* des EOG fokussiert sich auf die Abbildung aller technologischen und sozialen Plattformen, die in einem Unternehmen prozessorientiert eingesetzt werden. Dabei geht es nicht nur um die reine Technologie, sondern auch um deren Konfiguration und Customizing. Im Nachhaltigkeitskontext umfasst diese Map auf der Corporate Management-Ebene strategische Cloud-ESG-Plattformen wie SAP Analytics Cloud und Sustainability Footprint Management, die für unternehmensweites ESG-Reporting und Compliance eingesetzt werden können. Auf der Area Management-Ebene werden IoT-Sensoren und Monitoring-Systeme abgebildet, die den Energie- und Ressourcenverbrauch in verschiedenen Geschäftsbereichen überwachen. Die Workplace-Ebene fokussiert auf digitale Zwillinge und Simulationstechnologien, die nachhaltige Produktentwicklung unterstützen, während die Experience-Ebene KI-gestützte Dashboards und Predictive Analytics für Real-time Sustainability Monitoring bereitstellt.

Erweiterung der *Network Map*

Die Network-Map des Enterprise Online Guide beschreibt, wie Menschen und Systeme im Unternehmen mittels Technologien prozessorientiert sowohl technisch als auch sozial miteinander vernetzt sind. Sie bildet die Kommunikations- und Kollaborationsstrukturen ab, die für die effektive Zusammenarbeit in nachhaltigen Transformationsprozessen erforderlich sind. Auf der Corporate Management-Ebene werden ESG-Governance-Netzwerke dargestellt, die Stakeholder-Management und Compliance-Reporting mit externen Partnern und Regulierungsbehörden ermöglichen. Die Area Management-Ebene fokussiert auf Lieferanten-Nachhaltigkeitsnetzwerke, die durch digitale Plattformen Sustainability Scorecards und Supplier Collaboration unterstützen. Auf der Workplace-Ebene werden interne Mitarbeiter-

Communities wie Green Teams und Nachhaltigkeits-Trainings-Netzwerke abgebildet, während die Experience-Ebene Kunden-Feedback-Netzwerke für Sustainability Ratings und Green Product Reviews umfasst. Diese Vernetzung schafft die Basis für eine ganzheitliche Nachhaltigkeitsstrategie, die alle Stakeholder-Gruppen einbezieht.

2.4. Fallbeispiel eines Produktionsprozesses an der SmartFactory@OST

Als konkretes Fallbeispiel wird hier die SmartFactory@OST hergenommen, welche schon bei der ersten Veröffentlichung verwendet wurde (Bekiri et al. 2024). Eine kurze Beschreibung des Settings ist in (Stöckler et al. 2021) und (Hänggi et al. 2021) zu finden. Bei der SmartFactory@OST handelt es sich um eine kunststoffverarbeitende virtuelle Fabrik mit real existierenden Maschinen, Produkten und Prozessen. Daher eignet sich diese Umgebung auch als Fallbeispiel für ESG-Themen. Grundidee in der SmartFactory@OST ist, im Sinne der Nachhaltigkeit möglichst auf Kreislaufwirtschaft zu setzen. Durch die Systematik des EOGs lassen sich u.a. die vorhandenen Datenquellen identifizieren und mit den definierten Zielen verknüpfen, fehlende KPI-Grundlagen erkennen und auch die technische Basis dieser abbilden.

Abb. 5 zeigt den strukturierten Ablauf der ESG-Integration innerhalb des EOG-Frameworks anhand eines Produktionsprozesses an der SmartFactory@OST. Der Prozess beginnt mit der strategischen Definition nachhaltiger Unternehmensziele, die sich durch konkrete ESG-Kennzahlen und Nachhaltigkeitsobjektive manifestieren. Diese strategischen Vorgaben werden in bereichsspezifische Nachhaltigkeitsinitiativen übersetzt, wobei ESG-Abteilungen und Kreislaufwirtschaftsansätze implementiert werden. Auf der Workplace-Ebene erfolgt die operative Umsetzung durch die Rolle des Chief Sustainability Officers (CSO), der Produktionsaufträge mit Sekundärrohstoffen koordiniert und nachhaltige Geschäftsprozesse etabliert. Die Experience-Ebene bildet schließlich den Endpunkt der Operationalisierung, indem konkrete ESG-Datenmodelle und Compliance-Metriken in SAP Tabellen wie "ZZ_ESG_REC_Y_PER" erfasst werden. Die Technology-Spalte zeigt eine technologische Hierarchie von strategischen Cloud-Plattformen und Enterprise-Systemen auf Corporate-Ebene über Hosting-Lösungen auf Area-Ebene bis hin zu spezifischen Server- und Middleware-Komponenten auf Workplace-Ebene, die schließlich in konkreten Informationssicherheits-Managementsystemen auf Experience-Ebene münden. Die Network-Spalte bildet parallel dazu die Vernetzungsarchitektur ab, beginnend mit externen Rechenzentren und internen Rechnern auf strategischer Ebene, über Produktionsnetzwerk-Segmente und Datacenter-Firewalls auf operativer Ebene bis hin zur strukturierten Gebäudeverkabelung auf Erfahrungsebene. Beide Spalten verdeutlichen die technische Infrastruktur, die zur Umsetzung der ESG-Ziele von der strategischen Planung bis zur operativen Implementierung erforderlich ist. Diese Integration vernetzt technologische und soziale Komponenten der ES-Governance durchgängig von strategischen Zielen bis zu operativen Kennzahlen. Dieser mehrstufige Ansatz gewährleistet eine durchgängige Verfolgbarkeit von strategischen Nachhaltigkeitszielen bis hin zu messbaren operativen Kennzahlen und schafft damit die Basis für eine systematische ES-Governance im gesamten Unternehmen.

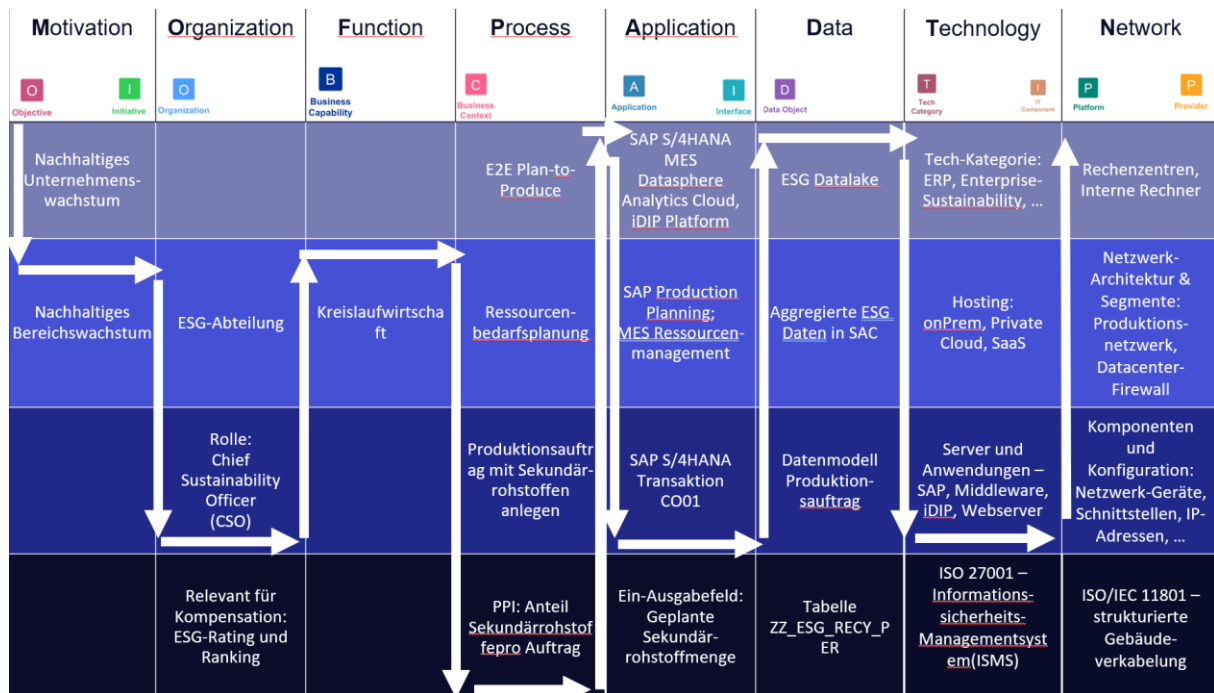


Abb. 5: Fallbeispiel eines Produktionsprozesses anhand der SmartFactory@OST

3. Zusammenfassung und Ausblick

In diesem Beitrag wurde gezeigt, wie die *Technology* und *Network* Map im EOG den ganzheitlichen Nachhaltigkeitsansatz aus (Bekiri et al. 2024) über LeanIX vervollständigen. Das EOG kann nun als Rahmen für die Entwicklung, Durchführung und Steuerung von ESG-Maßnahmen genutzt werden. Die grundlegende Systematik des EOGs erlaubt eine integrative Sicht auf alle drei ESG-Dimensionen – Ökonomie, Ökologie und Soziales. Die Forschungsfrage kann mit dem vorgeschlagenen Beitrag beantwortet werden. Es sind keine Ergänzungen oder Veränderungen an der ursprünglichen Definition des EOG notwendig. Der EOG bildet damit ein solides Framework für eine durchgängige Orientierung über ein gesamtes Unternehmen hinweg, wobei besonders die Verlinkung der einzelnen Elemente von der Strategie bis hinunter zu den Datenfeldern hervorgehoben werden muss. Dies ermöglicht eine langfristige Ausrichtung der gesamten Tätigkeiten mit einer klaren Führung über die Zieldefinitionen, die ohne Beschränkung finanzieller oder ökologischer bzw. sozialer Natur sein können. Das Meta Model v4 von LeanIX bietet darüber hinaus weitere umfangreiche Querverbindungen zwischen den verschiedenen EOG-Maps, die über die hier betrachteten Technology und Network Maps hinausgehen und in zukünftigen Veröffentlichungen näher untersucht werden sollten. Damit können Unternehmen den Nachhaltigkeitsaspekt tiefgreifend nach dem CSR-Reifegradmodell im Unternehmen verankern. In weiterer Hinsicht ist eine Umsetzung anhand des Fallbeispiels bzw. Fallstudie zur Validierung des Frameworks erstrebenswert. Auch die Integration des Smart Factory Frameworks von Budde et al. 2023 in das EOG-Setup bietet weitere konkrete Ansatzpunkte, insbesondere Use Case 10 "Sustainability and Carbon Footprint Tracking" für die systematische Erfassung von CO₂-Emissionen und Ressourcenverbrauch in der SmartFactory@OST. Durch die Kombination des Smart Factory Navigator mit der EOG-Systematik lässt sich eine durchgängige Nachhaltigkeitsstrategie von der strategischen Planung bis zur operativen Umsetzung in der Kreislaufwirtschaft realisieren. (Budde et al. 2023)

4. Literaturverzeichnis

- Bekiri, Valmir; Bayramli, Elnur; Stöckler, Stefan; Scheruhn, Hans-Jürgen; Bretschneider, Tobias (2024): ESG meets EOG – Environment Social Governance strukturieren mit dem Enterprise Online Guide (EOG). In: Uta Mathis, Nicole Ondrusch, Dietmar Kilian, Alexander Redlein, Klaus Turowski und Holger Wittges (Hg.): Proceedings of the SAP Academic Community Conference 2024 (D-A-CH). SAP Academic Community Conference 2024 (D-A-CH). TU Wien: Technical University of Munich, S. 185–195. Online verfügbar unter <https://www.tuwien.at/index.php?eID=dumpFile&t=f&f=224163&token=6138ba853fdb9fc1a22c635a501410b2e6390e6b>, zuletzt geprüft am 10.10.24.
- Budde, Lukas; Hänggi, Roman; Friedli, Thomas; Rüedy, Adrian (2023): Smart Factory Navigator. Cham: Springer International Publishing.
- FACTORY net; Karl Wagner (2020): Warum Nachhaltigkeit und Prozessmanagement untrennbar verbunden sind. Hg. v. WEKA Industrie Medien GmbH. Online verfügbar unter <https://factorynet.at/artikel/warum-nachhaltigkeit-und-prozessmanagement-untrennbar-verbunden-sind/>, zuletzt aktualisiert am 23.12.2021, zuletzt geprüft am 01.05.2024.
- Hänggi, Roman; Nyffenegger, Felix; Ehrig, Frank; Jaeschke, Peter; Bernhardsgrütter, Raphael (2021): Smart Learning Factory – Network Approach for Learning and Transfer in a Digital & Physical Set up. In: Felix Nyffenegger, José Ríos, Louis Rivest und Abdelaziz Bouras (Hg.): PRODUCT LIFECYCLE MANAGEMENT ENABLING SMART. 17th ifip wg 5.1, Bd. 594. [S.l.]: Springer (IFIP Advances in Information and Communication Technology), S. 15–25.
- Hinrichs, Bernd (2021): Nachhaltigkeit als Unternehmensstrategie. Roadmap für nachhaltiges Wirtschaften und Innovation. 1. Auflage. Freiburg, München, Stuttgart: Haufe Group. Online verfügbar unter <https://www.haufe.de/>.
- Kirchhoff, Klaus Rainer; Niefünd, Sönke; Pressentin, Julian A. von (2024): ESG: Nachhaltigkeit als strategischer Erfolgsfaktor. Wiesbaden: Springer Fachmedien Wiesbaden.
- Podleisek, André; Luban, Katharina; Hänggi, Roman (2025): Circular Economy: The Next Level of Company Success. Cham: Springer Nature Switzerland.
- Recker, Jan; Rosemann, Michael; Hjalmarsson, Anders; Lind, Mikael (2012): Green business process management. Towards the sustainable enterprise. Berlin, Heidelberg: Springer (Springer Management/Business for professionals).
- SAP (2025): Meta Model v4 - LeanIX. Hg. v. SAP. Online verfügbar unter <https://help.sap.com/docs/leanix/ea/meta-model?locale=en-US>.
- SAP Signavio Team (2023): Nachhaltigkeit ist für die diskrete Fertigung zu einer wichtigen Priorität geworden. Können Sie Ihre Prozesse an diese Herausforderung anpassen? Hg. v. SAP Signavio.
- Schaltegger, Stefan; Herzig, Christian; Kleiber, Oliver; Klinke, Torsten; Müller, Jan (2006): Nachhaltigkeitsmanagement in Unternehmen. Von der Idee zur Praxis: Managementansätze zur Umsetzung von Corporate Social Responsibility und Corporate

Sustainability. Unter Mitarbeit von Peter Franz. Deutsches Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit (BMU). Lüneburg.

Scheruhn, Hans-Jürgen; Schulze, Milan; Fotteler, Christine; Mendling, Jan; Bake, Benjamin (2023): Verbesserung der ERP-Lehre im Kontext des BPM-Lebenszyklus. In: Proceedings of the SAP Academic Community Conference 2023 (D-A-CH), S. 174–185.

Scheruhn, Hans-Jürgen; Weidner, Stefan; Müller, Simone (2024): Fallstudie: Enterprise Online Guide - Einführung. Enterprise Online Guide (EOG). Hg. v. Hochschule Harz. SAP ACC Magdeburg.

Schneider, Andreas (2015): Reifegradmodell CSR. Eine Begriffserklärung und -abgrenzung. In: Andreas Schneider und René Schmidpeter (Hg.): Corporate Social Responsibility. Verantwortungsvolle Unternehmensführung in Theorie und Praxis. Berlin: Springer Gabler, 2015.

Schneider, Andreas; Schmidpeter, René (Hg.) (2015): Corporate Social Responsibility. Verantwortungsvolle Unternehmensführung in Theorie und Praxis. Berlin: Springer Gabler, 2015.

Schoormann, Thorsten; Behrens, Dennis; Knackstedt, Ralf (2017): Sustainability in Business Process Models: A Taxonomy-Driven Approach to Synthesize Knowledge and Structure the Field.

Stöckler, Stefan; Hänggi, Roman; Bernhardsgrütter, Raphael; Baumgarten, Christoph (2021): Industrie 4.0 begreifbar machen – Die SmartFactory@OST. In: Lennart Brand, Karin Gräslund, Dietmar Kilian, Helmut Krcmar, Klaus Turowski und Holger Wittges (Hg.): Proceedings of the SAP Academic Community Conference 2021 DACH. bridging sustainability & digital innovation. Zeppelin Universität Friedrichshafen, 13. und 14. September 2021. Technische Universität München, Fakultät für Informatik. München, 208-224.

The Open Group (2022): The TOGAF® Standard, 10th Edition -- Introduction and Core Concepts. 1st ed. Hertogenbosch: Van Haren Publishing. Online verfügbar unter <https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=29222965>.

Process Mining-basiertes Dashboard für Nachhaltigkeitsmanagement

A. McGee, S. Stöckler & V. Bekiri

Institut für Informations- und Prozessmanagement (IPM), St. Gallen, Schweiz

Abstract: Diese Arbeit adressiert die Herausforderung, ökologische Nachhaltigkeitskennzahlen systematisch in die Analyse betrieblicher Prozesse zu integrieren, indem der eXtensible Event Stream (XES)-Standard systematisch um umweltbezogene Attribute erweitert wird. Ziel ist es, Nachhaltigkeitsdaten wie Energieverbrauch, Materialeinsatz, Emissionen und Abfall direkt aus Event Logs abzuleiten und innerhalb eines Process-Mining-Tools automatisiert auszuwerten. Auf Basis eines synthetischen Modellprozesses der industriellen Fertigung wird ein vollständiges Event Log erzeugt, das zusätzlich Nachhaltigkeitsattribute auf Event-Ebene enthält. Die Berechnung relevanter, GRI-konformer Kennzahlen erfolgt über formelbasierte Ausdrücke innerhalb der Plattform SAP Signavio Process Intelligence, wobei sowohl absolute als auch relative Indikatoren berücksichtigt werden.

Zur benutzerfreundlichen Anzeige wird ein Dashboard entwickelt, das eine thematisch strukturierte und prozess(schritt)bezogene Visualisierung der Kennzahlen ermöglicht. Die Ergebnisse zeigen, dass durch die gezielte Erweiterung des XES-Standards und die Nutzung bestehender Process-Mining-Technologien eine effektive, datenbasierte Analyse ökologischer Auswirkungen auf Prozessebene möglich ist. Damit leistet die Arbeit einen Beitrag zur methodischen und technischen Verbindung von Nachhaltigkeitsberichterstattung, digitalem Prozessmanagement und dem Aufbau eines Digital Twin of an Organisation (DTO).

Keywords: Nachhaltigkeitsmanagement, Process Mining, eXtensible Event Stream, XES, Dashboard, GRI, ESRS, Digital Twin of an Organisation, DTO

1. Einleitung

Im Zuge zunehmender globaler Herausforderungen durch den Klimawandel entwickelt sich nachhaltiges Wirtschaften von einem freiwilligen Engagement zu einer unternehmerischen Verpflichtung, die regulatorisch verankert und gesellschaftlich eingefordert wird. (Furchheim et al., 2023; Hizirolu & Dogan, 2025; Schaltegger, Hansen & Lüdeke-Freund, 2016) Insbesondere die Einführung der Corporate Sustainability Reporting Directive (CSRD) (Directive (EU) 2022/2464, 2022) und die damit verbundenen European Sustainability Reporting Standards (ESRS) (EFRAG, 2022) erhöhen den Druck auf Unternehmen, ihre Umweltwirkungen nachvollziehbar und standardisiert zu erfassen und zu berichten.

Das Wissen um den ökonomischen Nutzen von Prozessoptimierung ist inzwischen schon sehr gut bekannt und wird immer weiterverbreitet. Entsprechende Methoden werden bereits in vielen Unternehmen angewendet (vgl. z.B. (Opitz et al., 2024)). Ebenfalls immer intensiver genutzt werden die Möglichkeiten des Process Mining, die sich als wirksames Instrument etablieren, um Abläufe zu rekonstruieren oder Ineffizienzen aufzudecken (Garcia et al., 2019). Für Nachhaltigkeitsbetrachtungen fehlt es jedoch oft an Know-how, technischen Ressourcen oder methodischer Unterstützung, um diese entlang der operativen Wertschöpfung systematisch zu messen. (Setyaningsih, Widjojo & Kelle, 2024)

Hier setzt die vorliegende Arbeit an: Sie untersucht, wie Prozessdaten – sogenannte Event Logs – zur Erhebung und Analyse von Nachhaltigkeitskennzahlen verwendet werden können, damit diese in einem bedarfsgerechten Dashboard visualisiert werden können. Die Grundidee besteht darin, aus den verwendeten Informationssystemen (über alle Ebenen der Automatisierungspyramide) strukturierte Zeitreihen zu extrahieren, die Prozessinstanzen abbilden und dadurch Transparenz über reale Abläufe schaffen. Diese digitale Spur lässt sich nicht nur zur Effizienzsteigerung und somit zur ökonomischen Verbesserung, sondern auch zur Bewertung ökologischer Wirkungen nutzen – vorausgesetzt, relevante Nachhaltigkeitsdaten sind integraler Bestandteil der Event Logs. (van der Aalst & Carmona, 2022)

Und in genau dieser Kombination zwischen klassischem Process Mining und Nachhaltigkeitsmanagement liegt die bisherige Forschungslücke: Zwar existieren zahlreiche Konzepte zur Nachhaltigkeitsbewertung und ebenso eine breite Forschungsbasis zu Process Mining, doch die systematische Verknüpfung beider Bereiche steht noch am Anfang. Ziel dieser Arbeit ist es daher, eine fundierte methodische, technisch realisierbare und durchgängige Lösung zur Erhebung, Transformation und Darstellung ökologischer Nachhaltigkeitskennzahlen auf Basis von Event Logs aufzuzeigen. Daraus ergeben sich die Forschungsfragen, die in Abbildung 1 Zielsetzungen zugeordnet sind.

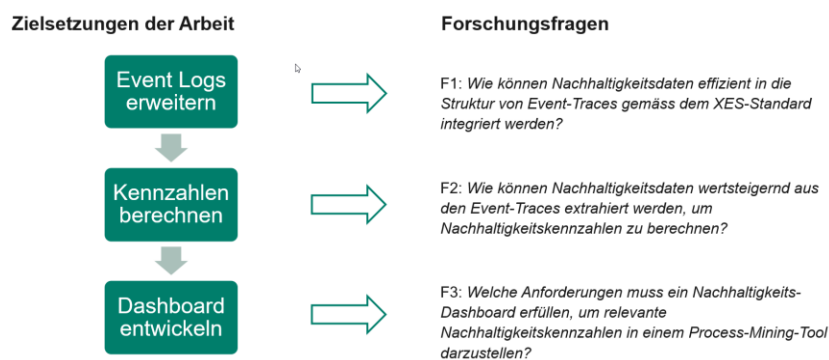


Abbildung 1. Zielsetzungen und Forschungsfragen (Eigene Darstellung)

2. Nachhaltigkeitsmanagement – Anforderungen und Praxis

Das Konzept der Nachhaltigkeit (NH) umfasst drei Dimensionen: ökonomisch, ökologisch und sozial (Purvis, Mao & Robinson, 2019). Die Berichterstattung zu diesem Thema zielt folglich darauf ab, die wesentlichsten Auswirkungen einer Organisation auf Wirtschaft, Umwelt und Gesellschaft systematisch offenzulegen ("GRI 1: Grundlagen 2021," 2023). Das NH-Reporting hat sich in den letzten Jahrzehnten zu einem wichtigen Instrument für Unternehmen entwickelt, um transparent die Auswirkungen in diesen Dimensionen darzustellen und Verantwortung im Sinne einer nachhaltigen Entwicklung zu übernehmen (Setyaningsih et al., 2024). Der Fokus liegt dabei nicht nur auf Transparenz, sondern zunehmend auch auf strategischer Steuerung, Vergleichbarkeit und der Kontrolle der Wirksamkeit von Massnahmen (Hahn & Kühnen, 2013). Dadurch können Unternehmen das Vertrauen der Kunden, allfälligen Investoren sowie der Öffentlichkeit stärken und ihre Nachhaltigkeitsleistungen strukturiert darstellen (Schaltegger et al., 2006).

Eine zentrale Rolle in der Praxis spielen hierbei die international am häufigsten verwendeten GRI-Standards der Global Reporting Initiative (Frey & Ashauer, 2025; Hahn & Kühnen, 2013). Diese Arbeit, im Speziellen der dabei entwickelte Prototyp, nutzt – ohne Beschränkung der

Allgemeingültigkeit –dieses Frameworks. Dabei kann auch auf die Interoperabilität der Standards, z.B. mit den in den EU verpflichtend anzuwendenden ESRS, verwiesen werden. Üblicherweise gliedern sich die Frameworks in branchenspezifischen und themenspezifischen Standards (siehe Abbildung 2. GRI-Standards: Gliederung). Diese sind jeweils in einzelne Angaben unterteilt, welche wiederum aus einzelnen Anforderungen bestehen, wobei die Kennzahlen in der Abstraktionsebene der Anforderungen enthalten sind. Solche Strukturen bieten eine flexible, international anwendbare Grundlage für freiwillige Berichterstattung ("GRI 1: Grundlagen 2021," 2023).

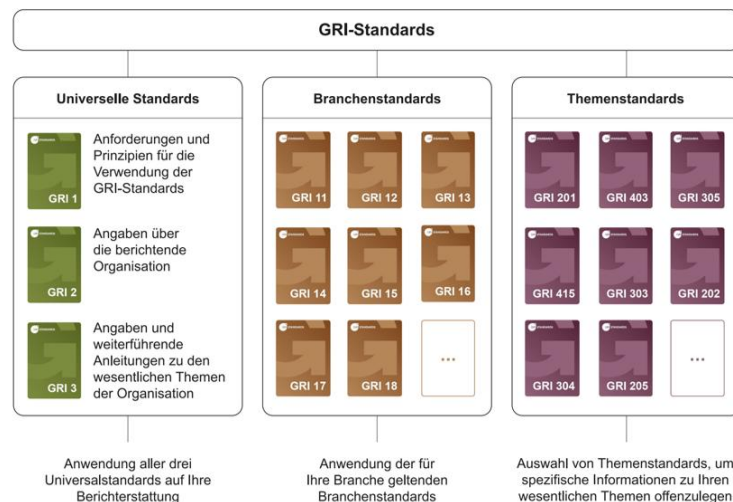


Abbildung 2. GRI-Standards: Gliederung (Quelle: ("GRI 1: Grundlagen 2021," 2023))

Die Hierarchie, wie aus den Standards einzelne auf Kennzahlen basierte Anforderungen top-down ermittelt werden, ist in Abbildung 3 dargestellt.

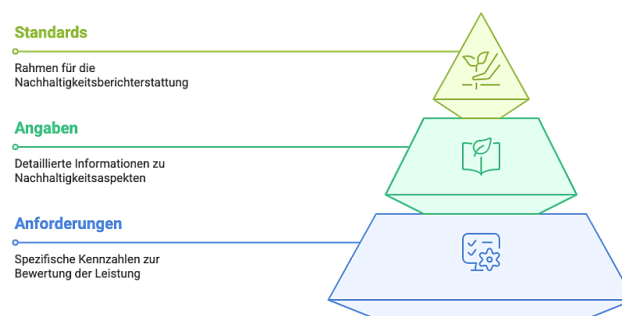


Abbildung 3. GRI-Standards: Hierarchie (Eigene Darstellung)

Die Kennzahlen zur Bewertung der Leistung im Nachhaltigkeitsbereich müssen aus den Daten der Prozesse gewonnen werden. Um das exemplarisch zu demonstrieren, wird ein Modellprozess, der die wesentlichen Schritte der Produktion von Glasflaschen abbildet, herangezogen. Abbildung 4 zeigt diesen Prozess als BPMN-Modell.

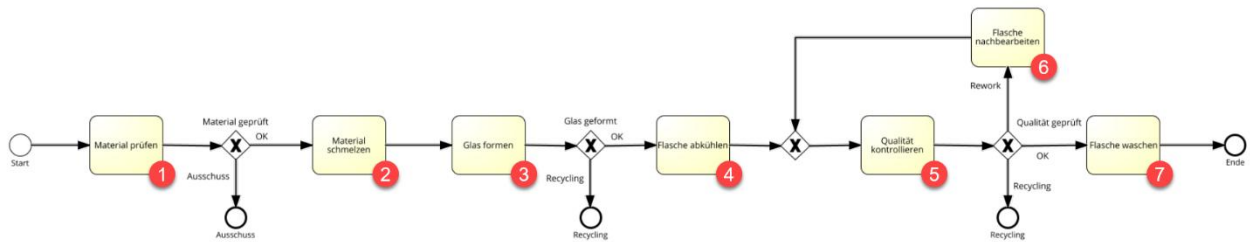


Abbildung 4. Musterprozess "Produktion Glasflaschen" (Eigene Darstellung)

Für diesen Prozess können GRI-Standards angewendet werden, die sich mit den Materialien, der Energie, dem Wasser/Abwasser, den Emissionen und dem Abfall beschäftigen (vgl. GRI-Standards 301, 302, 303, 305, 306). Daraus ergeben sich wiederum die folgenden 13 Angaben und damit die zugehörigen Anforderungen/Kennzahlen (vgl. Abbildung 3. GRI-Standards: Hierarchie (Eigene Darstellung), die gemessen und berichtet werden müssen:

Tabelle 1. Nachhaltigkeitskennzahlen zum Musterprozess „Glasflasche“

Angabe	Anforderung
301-1 Eingesetzte Materialien nach Gewicht/Volumen	Gesamtgewicht der zur Herstellung benötigten Materialien
301-2 Eingesetzte recycelte Ausgangsstoffe	Prozentsatz der recycelten Ausgangsstoffe
302-1 Energieverbrauch innerhalb der Organisation	Stromverbrauch
302-1 Energieverbrauch innerhalb der Organisation	Wärmeenergieverbrauch
302-1 Energieverbrauch innerhalb der Organisation	Kühlenergieverbrauch
302-1 Energieverbrauch innerhalb der Organisation	Gesamter Energieverbrauch innerhalb der Organisation
302-3 Energieintensität	Energieintensitätsquotienten der Organisation
303-5 Wasserverbrauch	Gesamter Verbrauch von Wasser
305-2 Indirekte THG-Emissionen (Scope 2)	Bruttovolumen der indirekten energiebedingten THG-Emissionen (Scope 2)
305-4 Intensität der Treibhausgasemissionen	THG-Intensitätsquotienten der Organisation
306-3 Angefallener Abfall	Gesamtgewicht des anfallenden Abfalls
306-4 Von Entsorgung umgeleiteter Abfall	Gesamtgewicht des von der Entsorgung umgeleiteten Abfalls
306-5 Zur Entsorgung weitergeleiteter Abfall	Gesamtgewicht des zur Entsorgung weitergeleiteten Abfalls

3. Erweiterung der Event-Trace-Struktur

Für die Erfassung der Prozessdaten hat sich der XES-Standard (IEEE Standard IEEE Std 1849™-2023) durchgesetzt, der auch von den marktgängigen Process Mining-Tools als Basis verwendet wird. Diese standardisierte Form der Datenhaltung erlaubt es auch, dass Teile der Event-Streams aus unterschiedlichen Anwendungen ermittelt und anschliessend zu einem zusammengehörenden Trace verbunden werden. Da der XES-Standard auf einem offenen, erweiterbaren Format basiert, kann die Struktur anwendungsspezifisch gestaltet werden (van der Aalst & Carmona, 2022).

Die vorliegende Erweiterung operationalisiert dabei den konzeptionellen Ansatz von Hizioglu & Dogan (2025) und überführt diesen in eine strukturierte Event-Datenbasis im XES-Format. Dabei werden – wie in (Graves, Koren & van der Aalst, 2023) bereits vorgeschlagen – die Nachhaltigkeitsdaten als Attribute den einzelnen Events im Log zugeordnet.

Die Attributstruktur gliedert sich in zwei Ebenen: Zum einen werden prozesslogische Attribute (z.B. id, name, timestamp, ...) definiert, die eine strukturierte Prozessdarstellung und das Zusammenführen von Daten aus verschiedenen Quellsystemen ermöglichen. Zum anderen

werden Nachhaltigkeitsattribute eingeführt, die ökologische Wirkdimensionen abbilden und eine spätere Kennzahlenberechnung ermöglichen. Die NH-Attribute werden auf Event-Ebene (vgl. Abbildung 5) verortet, damit sie einzelnen Prozessschritten direkt zugeordnet werden können.

Wie im XES-Standard vorgesehen, werden die nachhaltigkeitsbezogenen Attribute in einer eigenen, neuen Extension (vgl. (IEEE Standard IEEE Std 1849™-2023, S. 16) oder Abbildung 5) *sustainability* zusammengefasst, sind alle im Float-Format und verwenden international gebräuchliche Einheiten (Kilowattstunden, Kilogramm, Liter), wie sie auch im zugehörigen GRI-Standard definiert sind. Die jeweiligen Nachhaltigkeitsdaten werden immer in derselben Einheit, die im Attributnamen codiert ist, in den Traces geführt. Dies ermöglicht die Verarbeitung in beliebigen Process Mining-Tools, obwohl auch geschachtelte Attribute (also Wert und Masseinheit als Kombination) laut IEEE-Standard möglich wären (siehe Abbildung 5).

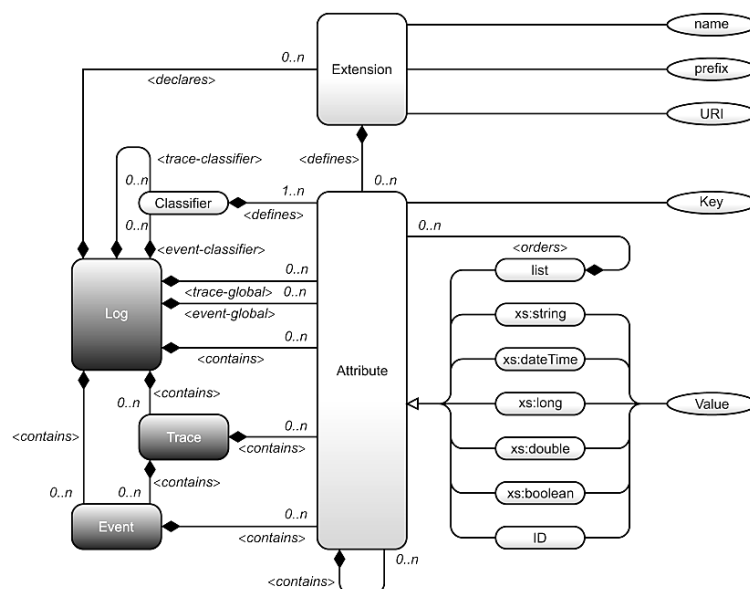


Abbildung 5. IEEE-Standard Datenstruktur für XES, Quelle: (IEEE Standard IEEE Std 1849™-2023)

Die folgende Tabelle 2 listet alle Attribute, sowie deren Beschreibung und einer Zuordnung zu den Schritten im Musterprozess (rote Referenznummern im BPMN-Diagramm).

Tabelle 2. Nachhaltigkeitsbezogene Attribute (Extension: sustainability) – Beispiele aus dem Musterprozess „Glasflasche“

Attribut	Beschreibung	Prozess-schritte	Format	Mögliche Werte
Material_input_kg	Verbrauchte Glasmenge pro Flasche im Schmelzprozess.	1	Float (kg)	0.495 – 0.505 kg
Energy_kwh	Energieverbrauch des Prozessschritts bezogen auf eine Flasche.	1-7	Float (kWh)	0.117 – 0.8 kWh
Water_l	Wasserverbrauch pro Flasche im jeweiligen Prozessschritt.	7	Float (l)	0.9 – 1.1 L
Waste_kg	Menge an Ausschussmaterial in kg.	1	Float (kg)	0.495 – 0.505 kg
Recycling_kg	Materialmenge, welche für Recycling aus dem Prozess entnommen wird.	3 & 5	Float (kg)	0.495 – 0.505 kg

Mit diesen Definitionen ist die Forschungsfrage 1 beantwortet: XES erlaubt domänenspezifische Erweiterungen, wobei die gewählte Codierung der Mengeneinheiten im Attributnamen die Verwendung in klassischen, aktuell verfügbaren Process Mining Tools gewährleistet.

4. Nachhaltigkeits-KPIs aus Event-Traces berechnen

Über die Aggregation der Attributwerte in den Traces lassen sich unterschiedliche Perspektiven auf verschiedene Verursacher einnehmen, wie dies auch am Beispiel des Musterprozesses exemplarisch ausgeführt wird:

- *Event-Ebene* = KPIs für einen einzelnen *Prozessschritt*. Z.B. der Wärmeenergieverbrauch beim Schmelzen im Modellprozess (GRI 302-1).
- *Trace-Ebene* = KPIs für einen vollständigem *Prozessdurchlauf*. Relevant für Kennzahlen mit Referenzeinheit (z.B. pro Produkt, im Musterprozess: Flasche) als Bezugsgröße (GRI 302-3, GRI 305-4).
- *Log-Ebene* = KPIs für eine *Menge von Prozessinstanzen*, z.B. Log eines Arbeitstages. Beispiele sind Gesamtenergieverbrauch, Gesamtmenge von Abfall oder recycelter Materialien (GRI 301-2, GRI 306-3).

Während einige Kennzahlen direkt in ihrer finalen Einheit im Event Log hinterlegt sind, erfordern andere Umrechnungen in das vorgegebene Zielformat. Hierbei handelt es sich um die folgenden Typen von Operationen (siehe auch Abbildung 6):

- Einheitenumrechnungen, etwa bei der Berechnung von Treibhausgasemissionen aus Energieverbrauch, z.B. kWh zu CO₂-equivalent
- Normalisierung auf die definierte Einheit, z.B. pro Produkt
- Berechnungen mit mehreren Variablen, um Vergleiche oder Verteilungen zu ermitteln, z.B. Prozentsatz der recycelten Ausgangsstoffe

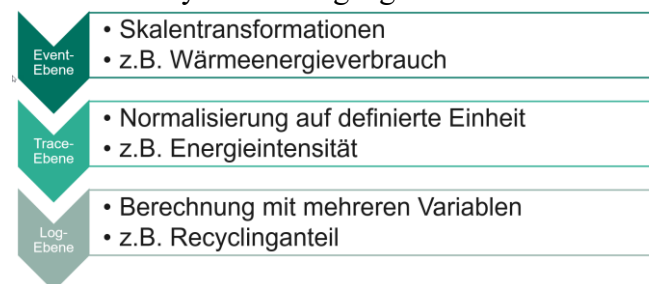


Abbildung 6. Transformationsvarianten (Eigene Darstellung)

Tabelle 3 zeigt Beispiele des Musterprozesses für die Verarbeitung mehrerer Attribute:

Tabelle 3. Berechnung aggregierter Kennzahlen – Beispiele aus dem Musterprozess „Glasflasche“

Kennzahl	Aggregationsebene	Berechnung
Recycling-Anteil (GRI 301-2)	Log-Ebene	$\frac{\sum \text{recycelte Menge}}{\sum \text{verwendete Menge}} * 100\%$
Energieintensität (GRI 302-3)	Trace-Ebene	$\frac{\sum \text{Energieverbrauch}}{\text{Anzahl produzierte Flaschen}}$
Indirekte THG-Emissionen ¹ (GRI 305-2)	Log-Ebene	$\text{Stromverbrauch} * \text{Emissionsfaktor}^2$
Emissionsintensität (GRI 305-4)	Trace-Ebene	$\frac{\sum \text{Indirekte Treibhausgasemissionen}}{\text{Anzahl produzierte Flaschen}}$

Damit ist Forschungsfrage 2 beantwortet und der konzeptionelle Aufbau der Datenbasis für das Dashboard geschaffen.

5. Konzeption eines Nachhaltigkeits-Dashboards

Für eine erste prototypische Umsetzung des Dashboards wird das vorhandene SAP-Signavio System verwendet, mit dem aufbauend auf die ermittelten Traces die Visualisierung umgesetzt wird. Diese Wahl soll aufzeigen, wie mit einem marktgängigen, weitverbreiteten Tool ein solches Dashboard aufgebaut werden kann und welche Limitationen dabei zu beachten sind.

Mithilfe der ermittelten Transformationsformeln (wie z.B. oben in Tabelle 3) werden im Mining-Tool die KPIs errechnet und dargestellt. Dazu werden – soweit es in den grafischen Limitationen des Werkzeuges möglich ist – die Hinweise zur effektiven Informationsvermittlung nach (Nussbaumer Knaflitz, 2015) beachtet. Hier zeigt sich, dass die Dashboard-Konzeption stark vom verwendeten Tool abhängt und nicht alle Varianten optimal umgesetzt werden können. Weitere Studien sollen künftig aufzeigen, wie dies verbessert werden kann.

Um zumindest zwei Darstellungsformen zu zeigen, sind in Abbildung 7 Screenshots des implementierten Prototyps aufgezeigt: Einerseits die klassische Kennzahlendarstellung von SAP-Signavio und andererseits ein Balkendiagramm für den Energieverbrauch gegliedert nach Prozessschritten.

¹ THG ... Treibhausgas

² Emissionsfaktor

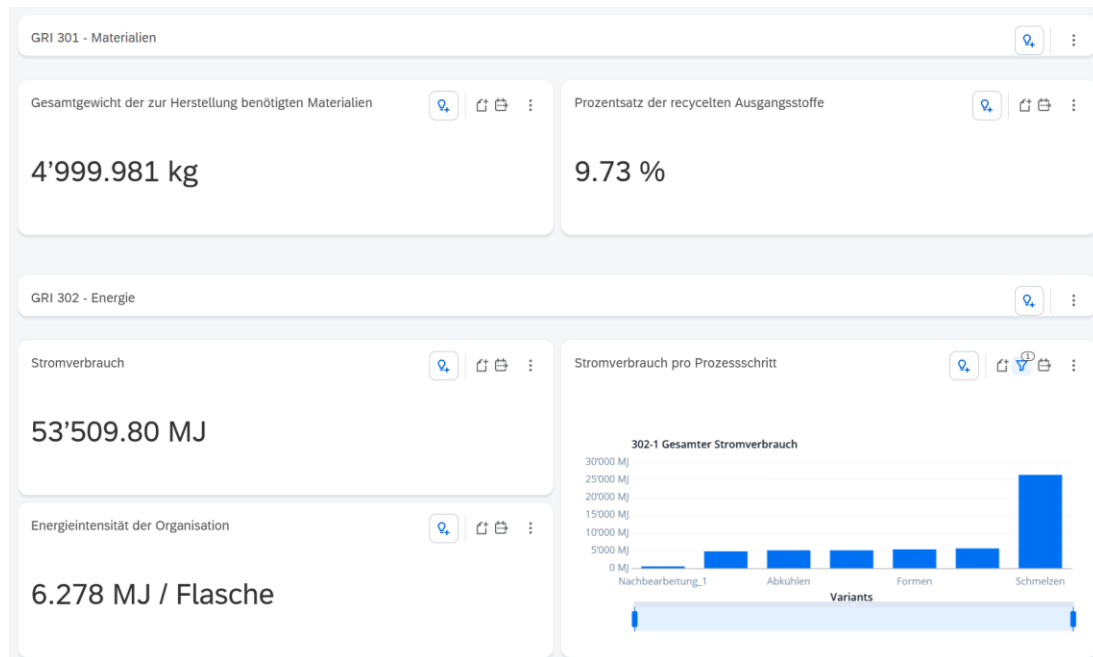


Abbildung 7. Dashboard – Beispiel (Eigene Darstellung)

Abbildung 7 zeigt als Antwort für Forschungsfrage 3 auch, dass die KPIs nach den GRI-Gruppen zusammengefasst werden können und durch die strukturierte Datenerfassung und -verarbeitung Drill-Down-Funktionalitäten zur Verfügung stehen. Damit sind die Anforderungen sowohl für eine laufende Steuerung als auch ein Reporting nach GRI- oder alternativ ESRS-Standard erfüllt.

6. Fazit und Ausblick

Die vorliegende Arbeit leistet einen innovativen Beitrag zur methodischen und technischen Verknüpfung der Themenbereiche Nachhaltigkeit, Prozessmanagement und datenbasierter Analyse. Durch die Erweiterung des XES-Standards um ökologische Attribute, die Transformation dieser Daten in standardkonforme Kennzahlen sowie deren visuelle Aufbereitung in einem Process-Mining-Tool wurde eine durchgängige, praxisnahe Lösung für die Nachhaltigkeitsanalyse auf Prozessebene konzipiert, prototypisch umgesetzt und evaluiert.

Mit der Verknüpfung von Process Mining und Nachhaltigkeit entsteht eine neue Perspektive für das betriebliche Nachhaltigkeits-Controlling. Der datenbasierte, prozessorientierte Ansatz erlaubt eine stärkere Integration ökologischer Zielgrößen in das operative Management durch die Abbildung des Unternehmens im Sinne eines Digital Twins of the Organisation (DTO).

Im Rahmen künftiger Arbeit sollen weitere Aspekte, wie z.B. eine Erweiterung der Mining-Logik auf Berechnungen von KPIs, von denen nur Start- und Endzeitpunkte in den Traces zu finden sind, oder die Anwendung von KI-Methoden zur Erkennung von Mustern in den Logs sowie zur Simulation von Prozessanpassungen. Auch die Darstellungsformen im Dashboard sollen weiter ausgearbeitet werden.

7. Literaturverzeichnis

- Directive (EU) 2022/2464 Amtsblatt der Europäischen Union (European Parliament 14.12.2022).
- EFRAG. (2022). European Sustainability Reporting Standards ESRS: General Requirements.
- Frey, B. & Ashauer, P. (2025). Pflichtwissen zu freiwilligen Nachhaltigkeitsberichtsstandards (Fin.Connect.Basics No. 5). Mühlheim. Abgerufen von <https://www.fin-connect-nrw.de/studien/pflichtwissen-zu-freiwilligen-nachhaltigkeitsberichtsstandards>
- Furchheim, P., Heierli, R., Stallone, V., Rettenmund, M. & Greschner, I. M. (2023). Swiss Sustainable Consumption Study 2022 : Befragung zum Stand von nachhaltigem Konsum in der Schweiz.
- Garcia, C. d. S., Meinheim, A., Faria Junior, E. R. et al. (2019). Process mining techniques and applications – A systematic mapping study. *Expert Systems with Applications*, 133, 260–295. doi:10.1016/j.eswa.2019.05.003
- Graves, N., Koren, I. & van der Aalst, W. M. (2023). ReThink Your Processes! A Review of Process Mining for Sustainability. In 2023 International Conference on ICT for Sustainability (ICT4S) (S. 164–175). IEEE.
- (2023). GRI 1: Grundlagen 2021. In Global Sustainability Standards Board (GSSB) (Hrsg.), Konsolidierte GRI-Standards. Global Reporting Initiative, GRI. GRI-Standards : Global Sustainability Standards Board (GSSB).
- Hahn, R. & Kühnen, M. (2013). Determinants of sustainability reporting: a review of results, trends, theory, and opportunities in an expanding field of research. *Journal of Cleaner Production*, 59, 5–21. doi:10.1016/j.jclepro.2013.07.005
- Hiziroglu, O. A. & Dogan, O. (2025). Sustainability Performance Through a Business Process Mining Based Conceptual Framework Integrating GRI Metrics. In S. Tiwari, F. Ortiz-Rodriguez, M.-A. Sicilia & T. R. Chhetri (Hrsg.), *Communications in Computer and Information Science. Artificial Intelligence: Towards Sustainable Intelligence* (S. 1–15). Cham: Springer Nature Switzerland.
- IEEE Standard IEEE Std 1849™-2023 (Version 2023). IEEE Standard for eXtensible Event Stream (XES) for Achieving Interoperability in Event Logs and Event Streams. (2023). Piscataway, NJ, USA: IEEE.
- Nussbaumer Knafllic, C. (2015). *Storytelling with Data: A Data Visualization Guide for Business Professionals* (1. Aufl.). New York, NY: John Wiley & Sons.
- Opitz, C., Simon, A., Savu, S., Bernstorff, J., Knedelstorfer, C., Schnägelberger, S. & Schwarz, F. (2024). *Prozessexzellenz durch Kompetenzaufbau und digitale Werkzeuge: Prozessmanagement & Analytik Studie 2024*. Frankfurt/Main.
- Purvis, B., Mao, Y. & Robinson, D. (2019). Three pillars of sustainability: in search of conceptual origins. *Sustainability Science*, 14(3), 681–695. doi:10.1007/s11625-018-0627-5
- Schaltegger, S., Hansen, E. G. & Lüdeke-Freund, F. (2016). Business Models for Sustainability. *Organization & Environment*, 29(1), 3–10. doi:10.1177/1086026615599806

- Schaltegger, S., Herzig, C., Kleiber, O., Klinke, T. & Müller, J. (2006).
Nachhaltigkeitsmanagement in Unternehmen: Von der Idee zur Praxis:
Managementansätze zur Umsetzung von Corporate Social Responsibility und Corporate
Sustainability. Lüneburg.
- Setyaningsih, S., Widjojo, R. & Kelle, P. (2024). Challenges and opportunities in
sustainability reporting: a focus on small and medium enterprises (SMEs). *Cogent
Business & Management*, 11(1). doi:10.1080/23311975.2023.2298215
- van der Aalst, W. & Carmona, J. (Hrsg.). (2022). *Lecture Notes in Business Information
Processing*: Bde. 448. *Process mining handbook*. Cham: Springer.

Integration von Prozess- und Organisationsmodellen in SAP Signavio

Benjamin Bake¹, Christian Biermann² und Hans-Jürgen Scheruhn³
SNP Deutschland GmbH Heidelberg¹/ msg systems ag München²/ Hochschule Harz Wernigerode³

Abstract: In Transformationsprojekten mit SAP S/4HANA weist die Qualität der erstellten Prozessmodelle häufig Defizite auf, die auf eine unzureichende Integration mit der prozessorientierten Aufbauorganisation des jeweiligen Unternehmens zurückzuführen sind. Unter anderem zeigt die Untersuchung an einem Large Process Model Dataset im akademischen Bereich, dass neben BPMN-Modellen nur ein geringer Teil aller erstellten Modelltypen ein Organigramm repräsentieren, weit abgeschlagen hinter anderen Modelltypen wie Wertschöpfungsketten, EPKs, UML und ArchiMate.

Den Zusammenhang zwischen den zu untersuchenden Modelltypen konnten die Autoren im Rahmen einer Masterarbeit aus dem Metamodell des Enterprise Architecture Frameworks TOGAF ableiten. Im methodischen Teil der Arbeit definieren die Autoren zunächst die in TOGAF fehlenden Abstraktionsebenen durch Mapping mit dem Metamodell des Enterprise Online Guide (EOG). Anschließend entwickeln sie ergänzende Funktionalitäten entlang des Metamodells zur anschaulicheren Darstellung der Beziehungen zwischen den beiden Modelltypen „BPMN“ und „Organigramm“ und setzen diese in SAP Signavio am Beispiel der Demo Firma Global Bike um. Dabei ist die Academic Edition von SAP Signavio sowie die vom UCC Magdeburg gehosteten Lösungen für SAP S/4 HANA zum Einsatz gekommen.

Keywords: Business Process Management (BPM); Enterprise Architecture (EA); SAP Signavio; BPMN 2.0; TOGAF; Enterprise Online Guide (EOG); Prozessmodellierung, Organisationsmodellierung.

1. Grundlagen der Modellierung in Signavio

SAP Signavio ist eine Business Process Management (BPM)-Suite, die verschiedene Modellierungssprachen zur Verfügung stellt (Strasser et al. 2024). Im Zentrum steht die Modellierungssprache Business Process Model and Notation (BPMN 2.0) zur Abbildung von Prozessen (Dumas et al. 2021). Die SAP Signavio Business Transformation Suite unterstützt darüber hinaus die Verwaltung von Glossaren und Attributen zur Verbindung von Objekttypen verschiedener Modelltypen, wie z. B. des Organigramms oder des Entity-Relationship-Modells (Scheruhn et al. 2024).

1.1. EAF-Struktur und ihre Abbildung in Signavio

Die Enterprise-Architecture-Framework (EAF)-Struktur dient als konzeptionelles Gerüst zur Modellierung der Beziehungen zwischen strategischen, organisatorischen, prozessualen und technischen Elementen (Lankhorst 2017). In SAP Signavio kann die EAF-Struktur partiell abgebildet werden. Die Abbildung 1 zeigt auf der linken Seite die Enterprise Online Guide (EOG) Matrix (Scheruhn et al. 2023) und auf der rechten Seite die Umsetzung in SAP Signavio. Dabei ist die Glossarstruktur als farblicher Rahmen für die jeweiligen Objekttypen abgebildet.

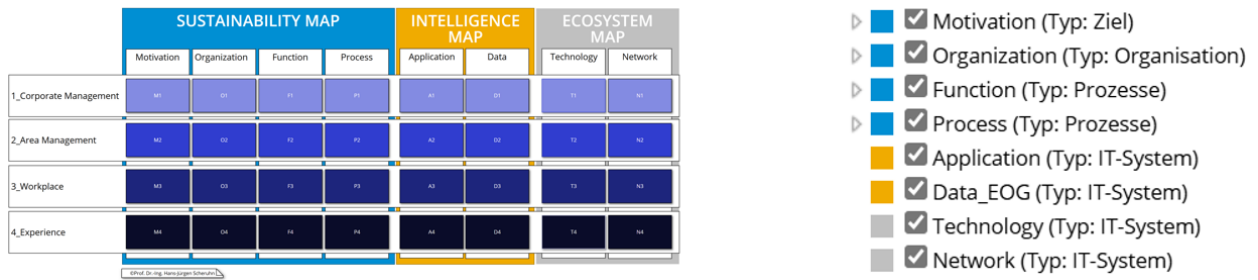


Abbildung 1 Enterprise Online Guide Matrix und Glossarstruktur Signavio (Scheruhn et al. 2023)

1.2. Objekttypen und deren Repräsentation

SAP Signavio differenziert nicht zwischen den Objekttypen "Positionen" und "Rollen" in Organigrammen (Abbildung 2). Diese Unterscheidung ist jedoch im Kontext von SAP S/4HANA erforderlich. Das HCM-Modul im S/4/HANA unterscheidet entsprechend zwischen „Planstelle“ („Position“ in Signavio) und „Stelle“ („Rolle“ in Signavio) (*SAP Help o. D.*). Diese Einschränkung wollen die Autoren im weiteren Verlauf dieser Arbeit über die Glossarstruktur auflösen.

Stelle

Definition

Hilfsmittel zum Anlegen von Planstellen.

Während **Planstellen** die konkreten, von Inhabern zu besetzenden Positionen in einem Unternehmen darstellen (z.B. Sachbearbeiter im Einkauf) sind Stellen allgemeine Klassifikationen von Funktionen in einem Unternehmen (z.B. Sachbearbeiter), die durch die Zuordnung von Eigenschaften näher zu bestimmen sind. Sie liefern damit Stellenbeschreibungen, die für mehrere Planstellen mit vergleichbaren Aufgaben und Eigenschaften gelten.

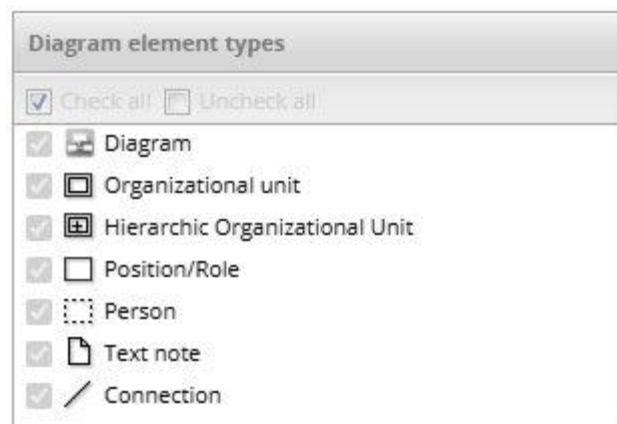


Abbildung 2 Objekttypen Organigramm in SAP Signavio (SAP Help, SAP Signavio)

Die Abbildung 3 zeigt die entsprechende Struktur am Beispiel des Enterprise Online Guide. Der Abbildung 3 ist zu entnehmen, dass die beiden Objekttypen nicht eindeutig getrennt werden. Organigramme repräsentieren prozessorganisatorische Einheiten (Organization Map:

Corporate Management, Area Management und Workplace), während BPMN-Diagramme die eigentliche Prozesssicht (Process Map: Corporate Management, Area Management und Workplace) darstellen. Die jeweiligen Objekttypen für diese Diagramme sind in der Abbildung 3 in 4 unterschiedlichen Blautönen farblich als Hierarchie ausgeprägt und entsprechen im Enterprise Online Guide der Organization- bzw. Process Map.

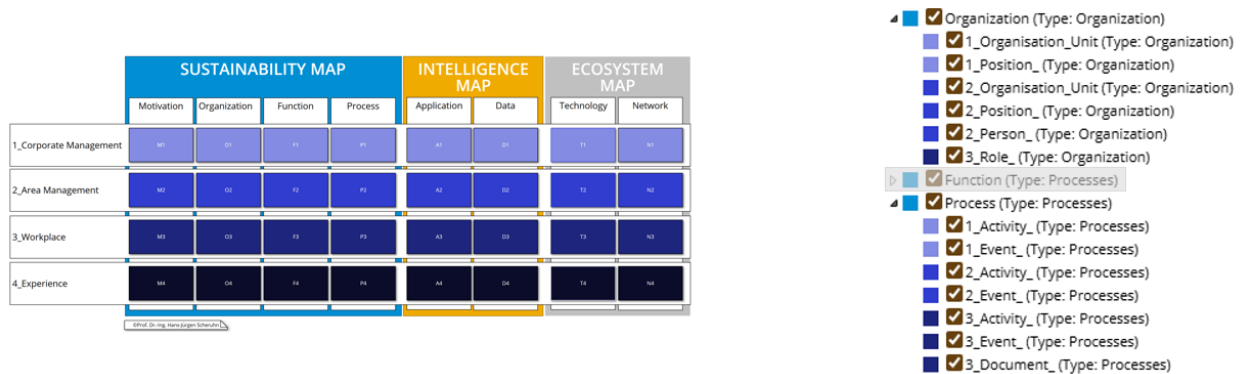


Abbildung 3 Enterprise Online Guide Matrix und Glossarstruktur für die Organisation- und Prozess-Map (Scheruhn et al. 2023)

2. Methodik

Dieses Kapitel beschreibt das methodische Vorgehen auf Basis des Design-Science-Research-Ansatzes, der zur Entwicklung und Evaluation des Artefakts herangezogen wurde.

2.1. Zielsetzung der Analyse

Das Forschungsziel der vorliegenden Arbeit ist es, die Verbindung zwischen BPMN-Modellen und Organigrammen im Kontext von Unternehmensarchitekturen zu strukturieren. Grundlage hierfür bilden sowohl das Metamodell von TOGAF als auch der Enterprise Online Guide, die zentralen Konzepte der Organisations- und Prozessmodellierung integrieren und deren Beziehungen zueinander explizit darstellen (Scheruhn et al. 2023). Die prototypische Umsetzung dient der Validierung der Abbildbarkeit dieser Konzepte in SAP Signavio.

2.2. Methodisches Vorgehen

Die Methodik dieser Arbeit folgt dem Design-Science-Research-Ansatz nach (Peppers et al. 2007). Ziel dieser Methode ist die Entwicklung und Evaluation innovativer Artefakte zur Lösung praxisrelevanter Probleme. Dieser Prozess besteht aus den folgenden sechs Phasen:

- **Problemidentifikation und Motivation:** Die mangelnde Integration von Organisations- und Prozessmodellen in SAP Signavio wird als zentrales Problem identifiziert.
- **Zieldefinition für die Lösung:** Es soll die Verbindung zwischen Organisation und Prozess gezeigt werden.
- **Entwicklung:** Auf Basis des bestehenden Frameworks EOG wird die Verbindung zwischen Prozessen und Organigrammen gezeigt. Grundlage ist die Enterprise Architecture mit Fokus auf die Maps „Prozess“ und „Organisation“, wie sie in der Masterarbeit des Erstautors konkretisiert wurde (Bake et al. 2024).

- **Demonstration:** Die Anwendbarkeit des Metamodells wird am Beispiel der fiktiven Organisation Global Bike gezeigt.
- **Evaluation:** Beschreibt die Verwendung der Objekttypen in Prozess, Organisation und deren Beziehungen.
- **Kommunikation:** Die Ergebnisse werden im Rahmen dieser wissenschaftlichen Publikation sowie im Kontext der SAP Academic Community diskutiert.

3. Herleitung des Metamodells

Das entwickelte Metamodell vom Enterprise Online Guide basiert auf TOGAF (The Open Group 2024). Die Abbildung 4 zeigt die direkte Gegenüberstellung der TOGAF- und EOG-Metamodells für die Prozesssicht. Während das EOG ein Ebenen-Konzept abbildet, verzichtet TOGAF auf Ebenen in seinem Metamodell (Lankhorst 2017).

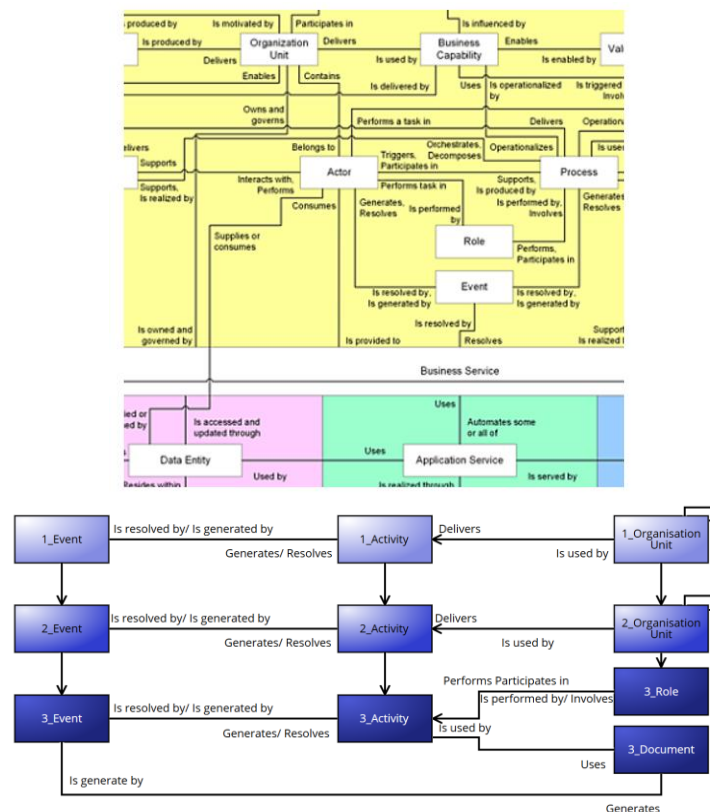


Abbildung 4 TOGAF-Metamodell (Ausschnitt Business Architektur) und Prozess-Map EOG (The Open Group 2024, Eigene Darstellung)

In der Organization Map (Abbildung 5) bildet das EOG-Metamodell zusätzlich den Objekttyp „Position“ ab. Das EOG-Framework unterscheidet im Vergleich explizit zwischen Stellen (Position) und Rollen (Role). Wie von SAP S/4HANA gefordert.

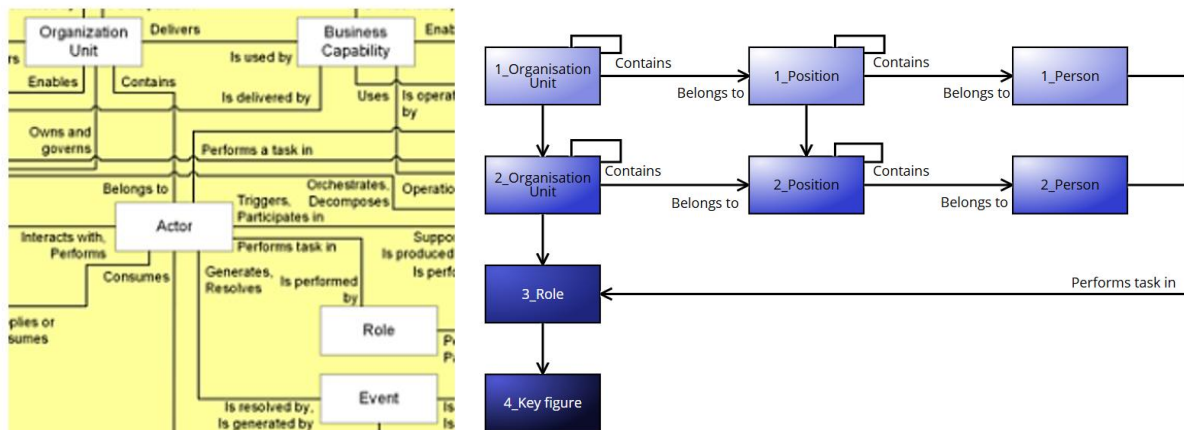


Abbildung 5 TOGAF Metamodell (Ausschnitt Business Architektur) und Organisation Map EOG (The Open Group 2024, Eigene Darstellung)

3.1. Modelltypen in Signavio

Organigramme in SAP Signavio dienen der Abbildung der prozessorientierten Aufbauorganisation eines Unternehmens. Dabei werden Organisationseinheiten, Abteilungen, Stellen und Rollen modelliert. Es wird jedoch keine strikte Trennung zwischen Position und Rolle über den Objekttyp vorgenommen, was aus Sicht der Autoren die Analyse von Verantwortlichkeiten erschwert. Im Kontext von EOG konnten die Autoren daher über das Glossar je eine Kategorie für die Stelle und eine für die Rolle erfolgreich einführen.

3.2. Beziehungen zwischen Organisation und Prozess-Map

In einer Untersuchung an einem Large Process Model Data Set beteiligter Hochschule mit SAP Signavio (Abbildung 8) sind im Vergleich zu den verwendeten BPMN-Modellen (60 %) weniger als 0,5 % aller Diagramme vom Typ Organigramm (Sola et al.).

In SAP Signavio besteht jedoch die Möglichkeit, zwischen verschiedenen Modelltypen zu navigieren. Beispielsweise kann mit dem Framework von EOG aus einem Prozessmodell zu einem zugehörigen Organigramm gewechselt werden. Dies erleichtert die Analyse von Verantwortlichkeiten und Abläufen. In Abbildung 6 sind die Verbindungen entsprechend ihrer Hierarchiestufe in EOG farblich markiert.

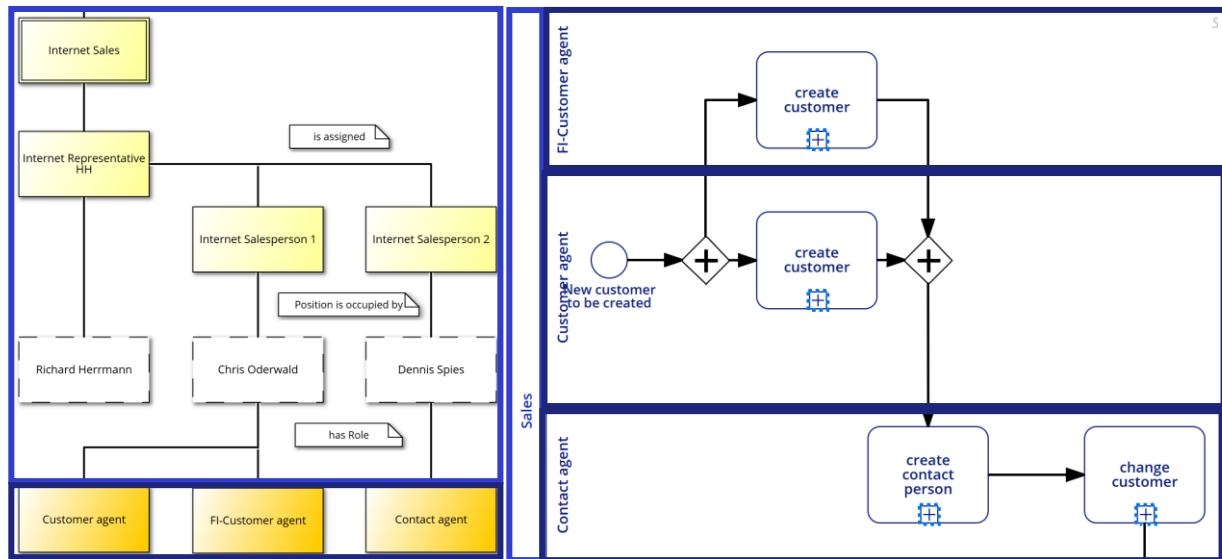


Abbildung 6 Beispiel aus dem EOG-Content zum Organigramm und Prozessmodell (Eigene Darstellung)

3.3. Grenzen der Modellintegration in Signavio

SAP Signavio liefert außerhalb der BPMN-Modellierung eingeschränkte Funktionalitäten zur direkten Abfrage von Beziehungen zwischen Objekttypen unterschiedlicher oder gleicher Modelltypen. Zwar können Objekt-Attribute genutzt werden, um Beziehungen zu dokumentieren, jedoch könnte eine systematische Auswertung innerhalb des Tools oder eine automatische Anzeige der Beziehungen über Kanten eine sinnvolle Erweiterung darstellen. Dadurch wird die Integration in eine prozessorientierte Aufbauorganisation unmittelbar sichtbar. Eine mögliche Lösung zeigt die Abbildung 7. In dem Beispiel wurde die Hauptattribute im Glossareintrag „Global Bike Group“ um dafür vorgesehene Kundenattribute erweitert. Damit lassen sich innerhalb des gesamten Workspace in allen Modellen, in denen die

Global Bike Group verwendet wird, die direkten Beziehungen aus dem Organigramm darstellen. So wird der Informationsfluss aus den Diagrammen verbessert.

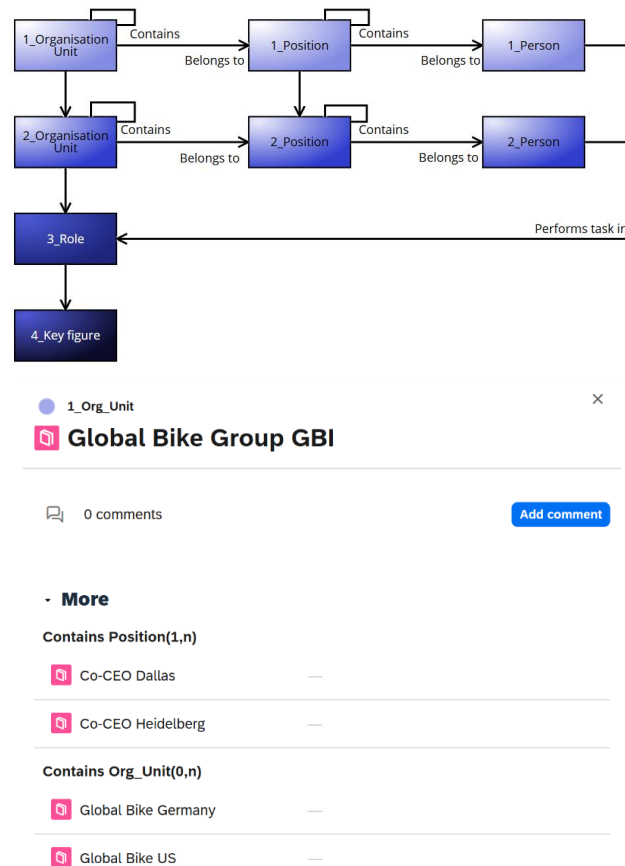


Abbildung 7 Beispiel TOGAF-Metamodell aus Abbildung 5 in Signavio (Eigene Darstellung)

4. Limitierung

Die zentrale Limitierung dieser Arbeit besteht darin, dass das entwickelte Metamodell ein theoretisches Konstrukt darstellt, dessen praktische Anwendbarkeit bislang nicht empirisch überprüft wurde. Eine Validierung in realen Anwendungsszenarien ist zwar vorgesehen, konnte im Rahmen dieser Untersuchung jedoch noch nicht umgesetzt werden. Das Modell basiert auf etablierten Frameworks wie TOGAF sowie auf den praktischen Erfahrungen der Autoren. Aufgrund der begrenzten Verfügbarkeit einschlägiger Literatur zu vergleichbaren Metamodellen ist interne Validität eingeschränkt. Darüber hinaus lässt sich ein potenzieller Bias infolge subjektiver Einschätzungen und der begrenzten Anzahl an beteiligten Autoren nicht vollständig ausschließen.

Die Generalisierbarkeit der Ergebnisse ist eingeschränkt, da das Metamodell primär im Kontext von ERP- und insbesondere SAP-Systemen entwickelt wurde und somit nur bedingt auf andere Domänen übertragbar ist. Zudem ist zu berücksichtigen, dass sich das Forschungsfeld der Prozess- und Sicherheitsmodellierung dynamisch weiterentwickelt, wodurch einzelne Konzepte oder Frameworks zukünftig angepasst oder obsolet werden könnten.

5. Ausblick

Die Untersuchung zeigt, dass SAP Signavio eine flexible Plattform zur Modellierung von Organisations- und Prozessstrukturen bietet. So können zusätzlich explizite Verbindung zwischen Objekttypen, der Organisations- und der Prozess-Sicht durch systematische Nutzung von Objekt-Attributen im SAP Signavio Glossar definiert werden. Zukünftige Entwicklungen könnten eine stärkere Integration dieser Sichten ermöglichen, etwa durch erweiterte Funktionalitäten, semantische Modellierungen innerhalb oder zwischen verschiedenen Modelltypen. Für die weitere Forschung ist eine Evaluierung außerhalb des akademischen Umfelds vorgesehen, die in Kooperation mit den Unternehmen des Erst- und Zweitautors erfolgen soll. Diese soll anhand von Fallstudien in realen Projektszenarien die Praxistauglichkeit des Metamodells überprüfen.

6. Fazit

Die Arbeit hat gezeigt, dass SAP Signavio durch den Einsatz zusätzlicher Attribute und Glossareinträge eine strukturierte Integration von Organisations- und Prozessmodellen im Kontext des EOG-Frameworks ermöglicht. Auf dieser Basis konnten die im Forschungsziel definierten Anforderungen – insbesondere die Abbildung der Verbindungen zwischen BPMN-Modellen und Organigrammen im Umfeld von SAP S/4HANA erfolgreich umgesetzt werden. Damit trägt das entwickelte Metamodell dazu bei, die strukturellen Lücken bestehender Frameworks wie TOGAF zu schließen.

Zu den wesentlichen Ergebnissen dieser Arbeit zählen:

- Erfolgreiche prototypische Umsetzung,
- Vollständige Kompatibilität zu SAP S/4HANA HCM,
- Strukturierte Integration zwischen Organisationen und Prozessen,
- Visualisierung semantischer Beziehungen zwischen abhängigen Objekttypen,
- Grundlage für eine ontologiegestützte KI

Für eine vollständig integrierte Unternehmensarchitektur ist vorgesehen, die Forschung im Kontext von EAM-Tools wie SAP LeanIX weiterzuführen. Zudem zeigt sich, dass das EOG-Framework die bestehenden Defizite in der Hierarchisierung von TOGAF erfolgreich ergänzt.

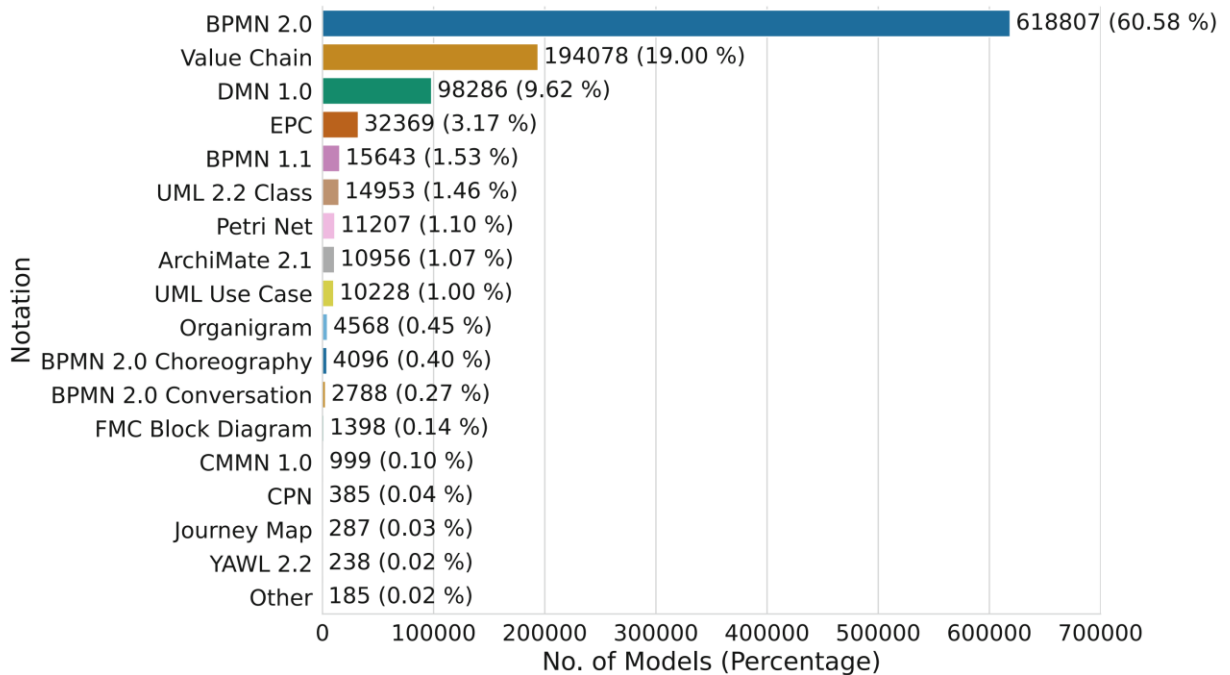


Abbildung 8 Vergleich Verwendung BPMN2.0 und Organigramm (Sola et. al 2023)

7. Literatur

Bake, B., Scheruhn, H.-J., & Bläß, E., (2024). Bewertung eines Enterprise-Architecture-Frameworks: Eine vergleichende Betrachtung von TOGAF und EOG (Masterarbeit). Hochschule Harz, Wernigerode.

Dumas, M.; La Rosa, M.; Mendling, J.; Reijers, H. A. (2021): Grundlagen des Geschäftsprozessmanagements. Berlin, Germany: Springer Vieweg.

Lankhorst, M. (2017): Enterprise Architecture at Work: Modelling, Communication and Analysis (4th ed.). Springer.

Peffer, K.; Tuunanen, T.; Gengler, C. E.; Rossi, M.; Hui, W.; Virtanen, V.; Bragge, J. (2007): Design Science Research Process: A Model for Producing and Presenting Information Systems Research. Online verfügbar unter 10.48550/arXiv.2006.02763, zuletzt aktualisiert am 02.2006, zuletzt geprüft am 25.08.2025.

SAP Help o. D. https://help.sap.com/doc/saphelp_autoid70/7.0/de-DE/4b/6a43cfcafe11d2b49d006094b9c9b4/content.htm?no_cache=true, zuletzt geprüft am 25.08.2025

Scheruhn, HJ., Mendling, J. (2024). Teaching Enterprise Resource Planning Systems and Business Process Management in Tandem Using the Enterprise Online Guide. In: Di Ciccio, C., et al. Business Process Management: Blockchain, Robotic Process Automation, Central and Eastern European, Educators and Industry Forum. BPM 2024. Lecture Notes in Business Information Processing, vol 527. Springer, Cham. https://doi.org/10.1007/978-3-031-70445-1_27

Scheruhn, H.-J.; Schulze, M.; Fotteler, C.; Mendling, J.; Bake, B. (2023): Verbesserung der ERP-Lehre im Kontext des BPM-Lebenszyklus. Konferenzband der SAP Academic

Community Conference 2023 (D-A-CH) Hg. v. Mathis, U.; Ondrusch, N.; Kilian, D.; Krcmar, H.; Turowski, K.; Weidner, S. und Wittges, H. Online verfügbar unter <https://doi.org/10.14459/2023md1719876>, zuletzt aktualisiert am 2023.

Sola, D., Warmuth, C., Schäfer, B., Badakhshan, P., Rehse, JR., Kampik, T. (2023). SAP Signavio Academic Models: A Large Process Model Dataset. In: Montali, M., Senderovich, A., Weidlich, M. (Hrsg.) Process Mining Workshops. ICPM 2022. Lecture Notes in Business Information Processing, Bd. 468. Springer, Cham.

Strasser, J.; Sokollek, M.; Sanger, M.; Spierling, M.; Schoenwalder, M.; Murra, M.; (2024): SAP Signavio Business Process Transformation. Springer

The Open Group (2024): The TOGAF® Standard, 10th Edition. TOGAF® Standard—TOGAF® Standard—Architecture Content. Hg. v. The Open Group. Online verfügbar unter <https://publications.opengroup.org/c220>, zu-letzt aktualisiert am 22.12.2024, zuletzt geprüft am 23.12.2024.

Digital Twins in Supplier Risk Management: Building Dynamic Capabilities to drive Supply Chain Resilience & Process Excellence in Supplier Risk Management

Katrin Oettmeier, Wolfgang Groher, Valmir Bekiri & Stefan Stöckler
Eastern Switzerland University of Applied Sciences, St. Gallen, Switzerland

Abstract: In this paper, we develop a concept how digital twin (DT) technology can be applied in supplier risk management (SRM) to drive the identification and analysis of supplier-related risks by systematically leveraging both, internal and external data sources. We thereby also show a way how large language models (LLMs) can be integrated in a DT architecture for SRM to enable dynamic business process modelling. The insights generated from the DT can not only be used for a proactive management of supplier risks, but also as a source for continuously reviewing SRM processes and providing impulses for process adaptations. With that, business process excellence in SRM can be improved. We argue that DTs in SRM can help to build a dynamic capability for fostering supply chain resilience.

Keywords: digital twin, supplier risk management, supply chain resilience, business process management.

1. Introduction

Digital twins (DT), as virtual replicas of physical assets and processes, offer transformative potential for enhancing risk management in supply chains. By leveraging Internet of Things (IoT) and Artificial Intelligence (AI) technologies, DTs can continuously collect and analyze data from different sources, offering insights into the current state of the supply chain and predicting future events (Adeyinka et al. 2022; Espinosa-Jaramillo et al. 2024). The integration of DTs with predictive analytics tools allows for the identification of potential risks and disruptions before they occur. This proactive approach enables organizations to implement mitigation strategies, reducing the impact of unforeseen events (Enyejo et al. 2024; Zhang et al. 2023). Machine learning algorithms within DTs can analyze historical data to predict delays and optimize decision-making processes, thus enhancing the overall resilience of the supply chain (Greis et al. 2021).

Despite their potential for bolstering supply chain resilience e.g. against disruptions, little knowledge exists on the application of DTs in supply processes (Bhandal et al. 2022; Guo & Mantravadi 2025). Currently, supply chain resilience and risk management and the usage of DTs within this context is still considered an emerging research stream (Bhandal et al. 2022).

Supplier Risk Management (SRM) is an intricate process that encompasses both dynamic and static elements: The dynamic elements comprise i.a. risk flow, changes in risk attributes, and risk conduction laws (Zhang 2011). They reflect the evolving nature of risks. Static SRM elements include risk sources, risk senders, risk takers and risk triggers – the fixed aspects of risk management. Consequently, effective SRM needs to acknowledge both, dynamic and static aspects (Zhang 2011). Due to their flexible nature, it can be assumed that DTs have the potential to accommodate the dynamic as well as the static SRM elements and hence enable a holistic approach for managing supplier risks.

While a few concepts of supply chain DTs exist (e.g. Busse et al. 2021; Ivanov & Dolgui 2021; Wang et al. 2020), only one study focusses specifically on DTs in supplier management, namely: the supplier selection process (Cavalcante et al. 2019). Currently, there are no conceptualizations of DTs in SRM, although the supply disruptions witnessed during the COVID 19 pandemic have demonstrated the fragility of contemporary supply chains. Moreover, as DT technology has the potential to significantly automate processes in supply chain operations (Espinosa-Jaramillo et al. 2024), investigating its value for enhancing SRM processes seems to be an interesting field for both, research and practice alike.

In the context of business process management, concepts for high-fidelity DT models are scarce (Fornari et al. 2025). Particularly the usage of Large Language Models (LLMs) to retrieve process relevant information from extensive organizational documents, thereby establishing a connection between static documentation and dynamic process modelling, is considered as a promising new research area (Fornari et al. 2025).

The present paper seeks to address the before-mentioned gaps in the literature by answering the following research questions:

- RQ1: How can a high-level digital twin architecture for SRM look like?
- RQ2: How can LLMs be integrated in a digital twin architecture for SRM to enable dynamic business process modelling?

2. Methodology

In this paper, we use Design Science Research (DSR) as a methodology to answer the before-mentioned research questions. In DSR, an artifact is created to address a relevant, unresolved problem (Hevner & Chattjee 2010). The artifact should thereby provide a solution for the defined issue and be derived from existing knowledge (Hevner & Chattjee 2010). We employed a problem-centered approach as supply issues caused by disruptive events like the COVID 19 pandemic depicted the starting point for our investigations.

As theoretical underpinning for our research, we build on the concept of dynamic capabilities. It refers to the way how firms integrate, develop and adapt their internal and external competencies in response to fast and unpredictably changing surroundings (Teece et al. 1997). We assume that the deployment of DTs in SRM can help firms to detect potential supplier-related issues faster and adapt accordingly. As such, it may help create a dynamic capability to foster resilience and risk management in supply (Bhandal et al. 2022).

3. Digital Twin for a Holistic Supplier Risk Management

In this chapter, requirements for a DT in SRM will be pointed out. Subsequently, a high-level architecture for such a respective DT solution will be presented.

3.1. Requirements for a digital twin for SRM

In the sense of a holistic SRM, a wide variety of supply risk types needs to be considered to gain comprehensive insights in a supplier's situation. Based on an extensive review of the literature, Yoon et al. (2018) identified 15 different supply risk factors, which are relevant for supplier selection and risk mitigation. Tab. 1 classifies these supply chain risk types according

to potential data sources to assess this risk. Our classification suggests that only 2 supply risk types can be analyzed relatively easily through structured data from business information systems like ERP systems. Most supply risk factors tend to rely on semi- or unstructured data sources. Thus, a DT solution for a holistic SRM would need to fulfill the following requirement:

- REQ1: The DT for SRM should enable the processing of structured, semi-structured and unstructured data.

Tab. 1: Supply risk types for a holistic SRM. (own illustration based on Yoon et al. 2018)

Type of data	Structured	Semi-structured	Unstructured
Potential data sources	e.g. ERP, QMS	e.g. supplier audit ratings, 4D / 8D / A3 reports, e-mails with suppliers	e.g. company reports, press releases, ESG reports, geodata, legal amendmends, feedbacks on employer rating platforms
Supply risk types	• Poor quality • Late delivery	• Supplier failure • Supply disruption • Poor supplier service • Supplier's technology risk	• Dispersed geographical location • Uncertain capacity • Supplier's financial distress / supplier's economic risk • Lack of supplier involvement • Supplier's information risks • Supplier's management risk • Supplier's societal risk • Supplier's environmental risk • Supplier's ethical risk

SRM is typically a discrete process, whereby supplier risks are assessed sporadically through audits or performance indicators (e.g. on-time delivery). To step up, a DT for SRM should allow for a continuous monitoring and prediction of supplier risks. This implies that besides real-time data processing capabilities, the solution also requires analytical capabilities:

- REQ2: The DT for SRM should enable real-time descriptive and predictive data analysis of potential supplier risks.

To move from a mere monitoring of the supplier risk status towards a system, which proactively supports SRM process excellence, the DT solution should also include prescriptive elements. For example, if a supplier's risk suddenly changes from low to high, the system should not only signal this new supplier risk assessment but also propose an adequate SRM process to match the new situation. This could mean to change the to-be SRM process for the supplier from a low risk SRM process path towards a high-risk process path, which involves a closed-loop monitoring of and frequent check-ins with the supplier.

Moreover, it would be desirable that the DT solution also has the capability not only to match supplier assessments with adequate SRM processes, but also to propose SRM process changes when needed – e.g. when new legal requirements such as amendments to the Supply Chain Due Diligence Act call for a general adaptation of all SRM processes. These requirements can be summarized as follows:

- REQ3: The DT for SRM should enable data-driven propositions for adaptations in SRM processes.

3.2. Concept of a digital twin for SRM

Based on the requirements outlined in chapter 3.1., we have created a concept for a DT for SRM. As illustrated in Fig. 1, the proposed solution closes the loop in the SRM process management cycle by combining process analysis (a reactive element) with AI-powered supplier risk identification and propositions for SRM process adaptations (a proactive element). It combines structured, semi-structured and unstructured data (REQ1).

The SRM process analysis functionality will provide insights in the current as-is SRM processes and leverage process-mining technologies for that. Process deviations can be detected by comparing the pre-defined SRM process, which is visualized e.g. via a BPMN diagram, with the (typically rather structured) data from the real-life SRM process. The foundation of the proactive element in the DT solution is formed by an AI agent and a *Retrieval-Augmented Generation (RAG)* system. The RAG system taps a vector database, which consolidates data from different unstructured, external data sources (e.g. press releases from suppliers, changes in commercial register entries) and enables precise, context-sensitive responses in real-time (REQ2). Meanwhile, the *AI agent* serves as the key interface for the interaction with the user. It utilizes LLMs and proposes a switchover to an alternative SRM process path or adaptations to the overall SRM process when a specific supplier's risk score or the overall risk factor landscape calls for it (REQ3).

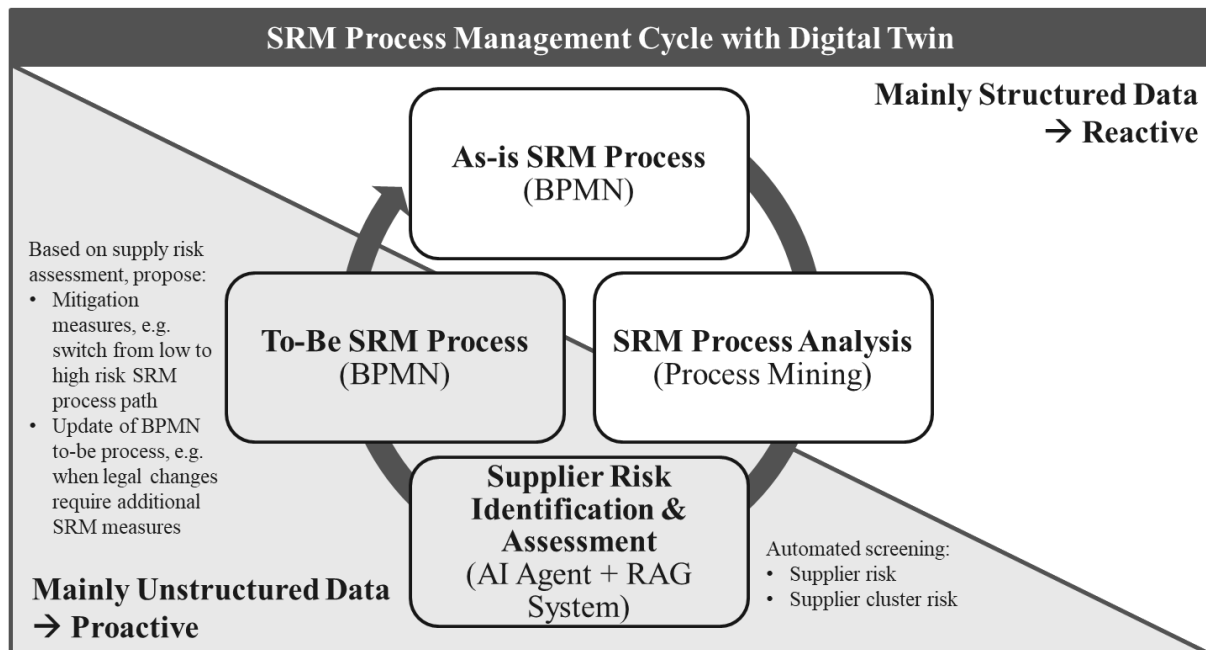


Fig. 1: SRM Process Management Cycle with a Digital Twin. (own illustration)

Fig. 2 shows our proposed high-level architecture of a DT for SRM. It builds on the DT concept from Busse et al. (2021) for multi-modal supply chains and transfers it to the context of SRM. Supplier risk data from various internal and external sources is consolidated through an interface module. In a prior step, unstructured data, which typically stems from external data sources, is accessed and connected via a RAG system. The interface module feeds into a supplier risk module, where supplier risk analyses are done. A reporting module shows the outcomes of the risk assessment, such as supplier risk scores.

Business stakeholders, e.g. strategic supply managers, receive recommendations for changes or improvements in the SRM process through an AI-powered SRM process recommendation engine. The recommendation engine utilizes LLMs and builds on information generated from the supplier risk module, a simulation and a SRM optimization module.

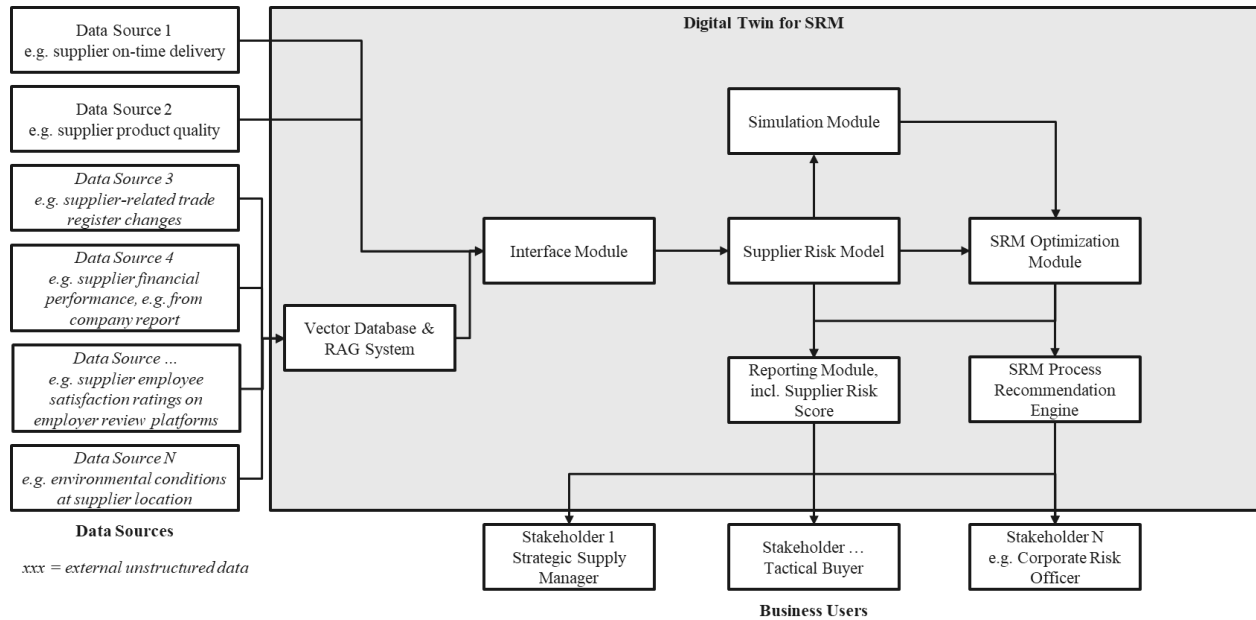


Fig. 2: High-level architecture of a Digital Twin for SRM. (own illustration based on Busse et al. 2021)

To operationalize the 15 supply risk types (based on Yoon et al. 2018), we propose specific metrics, extraction methods, and feature schemas that can be integrated into the DT architecture. For *structured risks* such as ‘Poor quality’ or ‘Late delivery’, quantitative KPIs are extracted from ERP systems, e.g., defect rates (metric: percentage of defective deliveries, threshold: >5%) or delivery delays (metric: average delay in days, validated through time series data). For *semi- and unstructured risks* like ‘Supplier’s financial distress’, the RAG component uses Natural Language Processing (NLP) techniques for extraction: Features include sentiment analysis of financial reports (e.g., via VADER sentiment score, negative < -0.05) and keyword matching (e.g., ‘insolvency’, ‘debts’). Validation is achieved through ground-truth labeling, using historical delivery failures as benchmarks (e.g., manually annotated datasets from past audits). For ‘Supplier’s environmental risk’, geodata-based features such as CO₂ footprint indicators (extracted from ESG reports via regex-based parsing) and metrics like environmental penalties (number per year) are used. Aggregation occurs in a weighted risk score model (e.g., Bayesian network) that accounts for uncertainties and is periodically calibrated through expert feedback.

We suggest a *step-by-step retrieval pipeline* based on a combination of external and internal data sources. The RAG system begins with indexing relevant unstructured data sources, such as press releases, ESG reports, and employer rating platforms (e.g., via APIs from sources like Indeed API or news aggregators like NewsAPI). The update frequency is dynamically controlled: For highly volatile risk factors like financial distress (e.g., based on real-time changes in commercial registers), hourly queries are performed, while less volatile factors (e.g., environmental risks) are updated daily. The retrieval phase uses vector-based search (e.g., with embeddings from models like Sentence-BERT) to extract context-relevant documents. Prompts for the LLM (e.g., based on GPT-4 or similar models) are designed to explicitly minimize

hallucinations. An example prompt could read: "Analyze the following text excerpt [inserted retrieval output] regarding the risk type 'Supplier's financial distress'. Provide only fact-based assessments, cite sources, and rate on a scale of 1-10 without making assumptions."

For *quality assurance*, we propose to integrate Reinforcement Learning from Human Feedback (RLHF) mechanisms, where supply managers' manually validated risk assessments serve as ground truth to iteratively fine-tune the model. Explainability is ensured through attention mechanisms in the LLM, which visualize the weighting of relevant input fragments.

Governance aspects include a policy engine that ensures data protection (e.g., GDPR-compliant anonymization) and bias mitigation (e.g., through diverse training data). Metrics for evaluating the components include precision / recall for risk detection (target: >85% based on simulated test data) and latency time for real-time analyses (<5 seconds per query for well-defined risk categories, <10 seconds for comprehensive analyses). This operationalization enables dynamic process modeling, as the LLM suggests alternative SRM process paths based on risk scores (e.g., threshold >7 for 'high risk'), such as switching to more intensive monitoring via automated audit triggers.

4. Discussion

This paper presents a first concept for a DT solution in SRM and thus paves the way for a more holistic management of supplier risk. It can enable firms to move from a predominantly reactive and static SRM, where supplier audits, supplier delivery performance and quality parameters depict the key sources of information, towards a proactive SRM, which combines different data sources for real-time monitoring as well as predictive and prescriptive analytics.

A benefit of a DT for SRM would be risk mitigation through dynamic, up-to-date insights in the risk profile of suppliers and a faster identification of needs for action. By utilizing the data to detect likely adaptation needs to SRM processes, the DT can also help to continuously improve SRM processes and drive operational excellence in supply. Thus, we believe that DTs in SRM enable firms to build a dynamic capability for mitigating supplier risks. We therefore join the canon of researchers, who argue that DTs can bolster supply chain resilience (e.g. Bhandal et al. 2022).

The DT concept developed in this paper has the potential to accommodate various supply risk types and thus to yield a multifaceted assessment of a supplier's risk situation in a largely automated manner. However, to obtain a robust DT solution for SRM, research still needs to address a large variety topics. Areas for further research include a compilation of reliable data sources and methods for measuring the 15 supply risk types and a configuration concept for the AI agent to ensure high-quality supplier risk data. To strengthen practical significance, we suggest a proof-of-concept implementation in a simulated environment, e.g., using tools like LangChain and PyTorch for RAG, LLM fine-tuning. Simulated results could e.g. demonstrate a 40% reduction in risk detection time (based on synthetic datasets with historical supply disruptions). A robust evaluation strategy may include KPIs such as early warning time (metric: hours until detection), precision / recall for risk events (>90% target), and cost reduction (e.g., through automated process adaptations, baseline: manual SRM). Future work may include a case study pilot in collaboration with industry partners to validate transferability and compare baselines against the status quo.

5. Conclusion

Based on a design science research approach, this paper crafts a theoretical concept for the application of DT technology in SRM. By combining real-time data from external and internal sources on a large variety of supply risk types, prescriptive analytics, and LLMs, the proposed DT solution can enable a holistic and proactive SRM. Firms may benefit from the solution by achieving higher levels of supply chain resilience and operational excellence in SRM. They can automate supplier risk monitoring and analysis activities, increase responsiveness to mitigate potential risks and obtain a source for continuously improving SRM processes. Through the specification of a high-level DT architecture for SRM, including a first operationalization of LLM/RAG components with governance and metrics, we create scientific added value and demonstrate practical applicability. The proposed concept paves the way for a robust implementation that may help firms to build dynamic capabilities and continuously optimize their SRM processes. Consequently, the idea of DTs in SRM offers promising avenues for practice as well as for research.

6. References

- Adeyinka, M.A., Solanke, B., Onita, F.B., & Agho, M.O. (2022): Integrating digital technologies for supply chain resilience in the energy industry. In: *International Journal of Scientific Research Updates* 3:1, 48-67.
- Bhandal, R., Meriton, R., Kavanagh, R.E. & Brown, A. (2022): The application of digital twin technology in operations and supply chain management: a bibliometric review. In: *Supply Chain Management* 27:2, 182-206.
- Busse, A., Gerlach, B., Lengeling, J.C., Poschmann, P., Werner, J. & Zarnitz, S. (2021): Towards Digital Twins of Multimodal Supply Chains. In: *Logistics* 5:25, 1-12.
- Cavalcante, I.M., Frazzon, E.M., Forcellini, F.A. & Ivanov, D. (2019): A supervised machine learning approach to data-driven simulation of resilient supplier selection in digital manufacturing, In: *International Journal of Information Management* 49, 86-97.
- Espinosa-Jaramillo, M., Zuta, M., Koneti, C., Jayasundar, S., Zegarra, S. & Carvajal-Ordoñez, V. (2024): Digital Twins in Supply Chain Operations Bridging the Physical and Digital Worlds using AI. In: *Journal of Electrical Systems* 20:10, 1764-1774.
- Enyejo, J.O., Fajana, O.P., Jok, I.S., Ihejirika, C.J., Awotiwon, B.O., & Olola, T.M. (2024): Digital Twin Technology, Predictive Analytics, and Sustainable Project Management in Global Supply Chains for Risk Mitigation, Optimization, and Carbon Footprint Reduction through Green Initiatives. In: *International Journal of Innovative Science and Research Technology* 9:11, 609-630.
- Fornari, F., Compagnucci, I., Callisto De Donato, M., Bertrand, Y., Beyel, H., Carrión, E., Franceschetti, M., Groher, W., Grüger, J., Kilic, E., Koschmider, A., Leotta, F., Li, C.-Y., Lugaresi, G., Malburg, L., Mangler, J., Mecella, M., Pastor, O., Riss, U. & Valderas, Pedro. (2025): Digital Twins of Business Processes: A Research Manifesto. In: *Internet of Things* 30:101477, 1-14.

- Guo, D. & Mantravadi, S. (2025): The role of digital twins in lean supply chain management: review and research directions, In: *International Journal of Production Research* 63:5, 1851-1872.
- Greis, N. P., Nogueira, M. L. & Rohde, W. (2021): Digital Twin Framework for Machine Learning-Enabled Integrated Production and Logistics Processes. In: Dolgui, A., Bernard, A., Lemoine, D., von Cieminski, G. & Romero, D. (eds): *Advances in Production Management Systems. Artificial Intelligence for Sustainable and Resilient Production Systems*. APMS 2021. IFIP Advances in Information and Communication Technology, Vol. 630. Springer, Cham.
- Hevner, A. & Chatterjee, S. (2010): Design Science Research in Information Systems. In: *Design Research in Information Systems: Theory and Practice*, Springer: Boston, MA, USA.
- Ivanov, D. & Dolgui, A. (2021): A digital supply chain twin for managing the disruption risks and resilience in the era of Industry 4.0, In: *Production Planning & Control* 32:9, 775-788.
- Yoon, J., Talluri, S., Yildiz, H. & Ho, W. (2018): Models for supplier selection and risk mitigation: a holistic approach, *International Journal of Production Research* 56:10, 3636-3661.
- Teece, D.J., Pisano, G. & Shuen, A. (1997): Dynamic capabilities and strategic management, In: *Strategic Management Journal* 18:7, 509-533.
- Wang, K., Xie, W., Wang, B., Pei, J., Wu, W. & Baker, M. (2020): Simulation-Based Digital Twin Development for Blockchain Enabled End-to-End Industrial Hemp Supply Chain Risk Management, *2020 Winter Simulation Conference (WSC)*. Orlando, FL, USA, 3200-3211.
- Zhang, J.G. (2011): Research on the elements and process of supply chain risk conduction. In: *Logistics Engineering and Management* 33:11, 139-151.
- Zhang, W., Yu, C., Xiao, J., Kin, D.C.H. & Zhong, R.Y. (2023): Digital-Twin Enabled Construction System For Supply Chain Risk Management. *2023 IEEE 19th International Conference on Automation Science and Engineering (CASE)*. Auckland, New Zealand, 1-6.

Towards a Social Sustainability Insights Tool for Operations

K. Oettmeier, V. Bekiri & S. Stöckler

Eastern Switzerland University of Applied Sciences, St. Gallen, Switzerland

In this presentation, we show a concept for a technical solution, which analyzes employee satisfaction ratings in combination with data from transactional business information systems to provide a novel source of impulses for improving company-internal social sustainability in operational jobs. While the concept of leveraging transactional business data for social sustainability insights is novel in operations management, it builds on advancements in HR analytics, which involves the evaluation of employee-related data to support solution discovery and evidence-based decision-making regarding HR topics.

Empirical evidence supports correlations between workload indicators from ERP systems and employee job satisfaction. Yet, no integrated tool systematically combines frequent employee feedback with transactional data for proactively managing social sustainability in operations. Our work addresses this by designing a high-level architecture grounded in Design Science Research, extending people analytics to operational contexts. The proposed social sustainability insights tool utilizes artificial intelligence technologies, especially pattern recognition. By correlating employee feedback with operational process data e.g. from ERP or CRM systems, the solution will determine potential influencing factors on employee satisfaction (e.g. relation between logistics team member satisfaction and delivery volumes). Managers can use these insights in discussions with their team members as a source for improvement ideas.

Modellierung mit EOG @ SAP Signavio. Eine studentische Perspektive

Yin Tsan^{1,2}, Simone Rochlitzer^{1,3}

¹: Hochschule Harz, ²: ITZ Bund, ³: Agentur für Arbeit

Abstract: Wie führt man Studierende aus dem ersten Semestern an SAP heran? Aus erster Hand und ohne Vorerfahrung mit SAP zeigt dieser Beitrag die erfolgreiche Herangehensweise und Lerneffekte an der Hochschule Harz.

Dieser praktische Beitrag soll aufzeigen, wie das BPM Tool SAP Signavio im Kontext des Enterprise Online Guides (EOG) in der Lehre einsetzbar ist. Als Ausgangspunkt diente dabei die HCM-Fallstudie des SAP University Competence Center (UCC) Magdeburg mit SAP S/4HANA. Eine erfolgreiche Umsetzung konnte mit ca. 13 Studierenden des dualen Studienganges IT-Management im Fachbereich Verwaltungswissenschaften an der Hochschule Harz im WiSe2024/25 durchgeführt werden. Dabei kam die Academic Edition von SAP Signavio zum Einsatz.

Als Prüfungsleistung hat die Autorin ihr Praxisunternehmen ITZ-Bund im Kontext von EOG abgebildet. Nach dem anschließenden Wechsel in das EOG-Team hat sie die vorliegende Modellierungsfallstudie weiterentwickelt und dokumentiert.

Aufbauend auf dem Hire-to-Retire (HCM) Prozess der Global-Bike-Fallstudie konnte sie in SAP Signavio eigene Glossareinträge erstellen. Diese ermöglichen eine horizontale Navigation über 8 verschiedene Maps sowie vertikal über 4 Zoomfaktoren. Besonders eindrücklich war für die Autorin die nahtlose Verbindung von Sustainability Balanced Scorecards mit BPMN-Diagrammen und Datenmodellen sowie SAP S/4HANA Anwendungen.

Die strukturierte Arbeit mit dem EOG-Ansatz hat der Autorin die Prozesszusammenhänge ähnlich einem Google Maps für Unternehmen nicht nur IT-technisch, sondern auch betriebswirtschaftlich erschlossen. Insgesamt konnte sie dadurch einen tiefen Einblick in die Potenziale digitaler Transformationen gewinnen – sowohl in der Theorie als auch in der praktischen Umsetzung.

Dieser Beitrag „Modellierung mit EOG @ SAP Signavio - Eine studentische Perspektive“ beleuchtet aus studentischer Sicht den Erfolgsweg mit allen Herausforderungen und Lerneffekten in einem Semester. Anhand des HCM Moduls wurden Modelle des Praxispartners ITZBund gestaltet und präsentiert. Dabei kam die Academic Edition von SAP Signavio zum Einsatz.

Positive Veränderungen von Geschäftsprozessen in der Radverkehrsplanung durch die Bereitstellung unterstützender Dashboards

J. Schering, J. Marx Gómez & F. Köster

Universität Oldenburg, Abteilung für Wirtschaftsinformatik VLBA, Oldenburg, Deutschland
Deutsches Zentrum für Luft- und Raumfahrt DLR, Institut für KI Sicherheit, St. Augustin,
Deutschland

Abstract: Der Radverkehrsplanung fällt eine zentrale Rolle bei der Transformation von Städten hin zu mehr Nachhaltigkeit zu. Der bislang etablierte Geschäftsprozess weist allerdings einige Hindernisse bei der Beschaffung wichtiger Datengrundlagen auf, da diese bislang dezentral bei unterschiedlichen Stakeholdern vorliegen. Obwohl Daten zu Verkehrsinfrastruktur, zur Unfalllage oder Verkehrsströmen im Entscheidungsprozess weiter an Bedeutung gewinnen, müssen diese bei der Erstellung von Maßnahmenplänen bei verschiedenen Fachämtern, Behörden oder der Polizei angefragt werden. In den häufigsten Fällen können nicht alle verfügbaren Informationen unter einem vertretbaren Aufwand verfügbar gemacht werden. Diese Gegebenheiten im heutigen Geschäftsprozess stehen einem effektiveren Ausbau der Fahrradinfrastruktur, der unter den wachsenden Herausforderungen der Klimaproblematik immer dringlicher wird, im Wege. In diesem Zusammenhang untersucht dieser Beitrag Änderungspotentiale im Geschäftsprozess, die sich durch die zentrale Bereitstellung zentraler Datengrundlagen für Stakeholder durch neue Dashboard-Lösungen ergeben. In insgesamt 17 Experteninterviews wurden die Herausforderungen des heutigen Prozesses umfassend untersucht. Anschließend versetzen wir uns in die Lage des Radverkehrsplaners, dem nun ein dynamisches Dashboard mit allen relevanten Datenquellen zentral zur Verfügung steht: Ein typischer Planungsfall wird am Beispiel der Stadt Osnabrück unter Einsatz eines neuen digitalen Planungstools beispielhaft durchlaufen. Ziel ist es, die verbesserten Aspekte im veränderten Arbeitsprozesses zu identifizieren, die sich aus der zentralen Datenbereitstellung ergeben können. Einen großen Vorteil bietet die Möglichkeit, Daten zur Infrastruktur mit Verkehrsstärken oder der Unfalllage überlagern zu können. Mögliche Limitationen des vorgestellten Dashboard-Ansatzes werden im Vortrag zum Abschluss diskutiert.

Keywords: Radverkehrsplanung, Geschäftsprozess, Dashboarding, KPIs

TRACK 2: TECHNOLOGY

Track Chair: Prof. Dr. Stefan Stöckler

Next Level Requirements Engineering NLP and ML Usage in the Software-Development-Life-Cycle

A. Rottensteiner, B. Seebacher, D. Kilian, C. Ploder
MCI – The Entrepreneurial School, Innsbruck, Austria

Abstract: The increasing complexity of modern software projects has amplified the challenges in Requirements Engineering (RE), a critical phase in the Software Development Life Cycle (SDLC). This study explores the transformative potential of Natural Language Processing (NLP) and Machine Learning (ML) in automating RE tasks, addressing issues such as ambiguity, inconsistency, and inefficiency in natural language requirements. Through a Systematic Literature Review (SLR) of 44 research papers published between 2015 and 2024, the study examines the role of NLP and ML in RE, evaluates the effectiveness of existing tools and techniques, and identifies the challenges hindering their widespread adoption. Key findings highlight the effectiveness of NLP techniques like Part-of-Speech (POS) Tagging and Named Entity Recognition (NER), as well as ML models such as Recurrent Neural Networks (RNNs) in automating tasks like requirement classification and semantic analysis. However, challenges such as data quality, model generalizability, and transparency limit their practical application. The study concludes with recommendations for future research, emphasizing the need for robust datasets, improved model explainability, and industry-friendly tools to bridge the gap between research and practice. These advancements have the potential to revolutionize RE processes, making them more efficient and reliable, and inspire optimism for the future of software development.

Keywords: Requirements Engineering, Natural Language Processing, Machine Learning, Automation in Software Development,

1. Introduction

In an era dominated by digital transformation, the development of high-quality software systems has become a cornerstone for both private and business sectors (Ribeiro et al., 2014). Despite advancements in software development methodologies, many projects still fail, experience significant delays, or exceed budgets. A primary reason for these failures is the increasing complexity of modern software systems and the challenges associated with Requirements Engineering (RE) (Sinkie et al., 2023; Unterkalmsteiner et al., 2015). As a critical phase in the Software Development Life Cycle (SDLC), RE involves the elicitation, documentation, analysis, validation, and management of requirements. However, the reliance on natural language for requirement documentation often introduces ambiguities, inaccuracies, and inconsistencies, further complicating the process (Sonbol et al., 2022).

To address these challenges, researchers have turned to Natural Language Processing (NLP) and Machine Learning (ML) as potential solutions for automating RE tasks. NLP, a subfield of artificial intelligence, focuses on enabling machines to understand and process human language, while ML leverages algorithms to learn patterns from data and make predictions. Together, these technologies offer promising avenues for improving the efficiency and accuracy of RE

processes, demonstrating their potential as viable solutions (Shah & Jinwala, 2015; Zhao, 2021).

This paper investigates the role of NLP and ML in automating RE based on a Systematic Literature Review (SLR) of 44 research papers published between 2015 and 2024. The study aims to address the following research questions (RQs), which are crucial for understanding the current state and future potential of NLP and ML in RE:

RQ1: What role do NLP and ML play in the automation of Requirements Engineering?

RQ2: What tools, techniques, and methods are currently used, and how effective are they?

RQ3: What challenges exist in integrating NLP and ML into the RE process?

By systematically analyzing the selected studies, this paper provides insights into the current state of the field, identifies practical tools and techniques, and highlights the barriers to widespread adoption. The findings aim to bridge existing research gaps and serve as a foundation for future advancements in NLP and ML for RE.

2. Methodology

To systematically explore the role of Natural Language Processing (NLP) and Machine Learning (ML) in automating Requirements Engineering (RE), this study employed a Systematic Literature Review (SLR) following the guidelines proposed by Kitchenham (2004). The SLR methodology ensures a structured and unbiased approach to identifying, selecting, and analyzing relevant research studies. The process was divided into key steps, which are explained in the following subsections.

2.1. Search Strategy

The search was conducted across four major academic databases: ACM Digital Library, IEEE Xplore, ScienceDirect, and Scopus. Keywords such as "Requirements Engineering," "NLP4RE," "Natural Language Processing," and "Machine Learning" were used to identify relevant studies. Boolean operators and advanced search filters were applied to refine the results. The search was limited to peer-reviewed articles published between 2015 and 2024 to capture the most recent advancements in the field (Kitchenham, 2004). This strategy was chosen to ensure that the study is based on the most current and relevant research in the field.

2.2. Inclusion and Exclusion Criteria

To ensure the relevance and quality of the selected studies, specific inclusion and exclusion criteria were applied. Studies were included if they focused on the application of natural language processing (NLP) and machine learning (ML) in requirements engineering (RE), were written in English, and had been published in peer-reviewed journals or conference proceedings. In contrast, studies were excluded if they were unrelated to RE, lacked empirical evidence, or focused solely on theoretical frameworks without any practical implementation.

2.3. Screening Process

The initial search yielded a total of 658 papers, which were then screened through a multi-stage process to ensure relevance and quality. In the first stage, papers were filtered based on their titles to determine their significance to the research topic. This was followed by an abstract screening phase, during which the abstracts of the remaining papers were reviewed to refine the selection further. Finally, a full-text review was conducted on the shortlisted papers to assess their alignment with the research questions and the previously defined inclusion criteria.

2.4. Quality Assessment

To minimize researcher bias and ensure methodological rigor, the selected papers were evaluated using established quality assessment tools, including the CASP Qualitative Studies Checklist and the JBI Critical Appraisal Tools. Each paper was scored based on criteria such as clarity of objectives, methodological soundness, and relevance to the research questions (Zhao et al., 2022).

2.5. Data Extraction and Synthesis

Key data points, including the tools, techniques, and methods discussed in each study, were extracted and synthesized for analysis. Metrics such as F1-Score, Precision, and Recall were used to evaluate the effectiveness of the proposed approaches. The findings were categorized to address the three research questions and identify trends, challenges, and gaps in the field.

By following this rigorous methodology, the study ensures a comprehensive and unbiased analysis of the role of NLP and ML in automating RE processes.

3. Results

This systematic literature review (SLR) synthesizes insights from 44 peer-reviewed studies (see Table 1) on the role of Natural Language Processing (NLP) and Machine Learning (ML) in automating Requirements Engineering (RE). The findings highlight three core aspects: their role in RE tasks, the tools and techniques employed, and challenges of integration, alongside observable trends in the research landscape.

NLP and ML have gained increasing relevance in RE, particularly since 2021/22, driven by the emergence of large language models (LLMs) such as ChatGPT and the growing complexity of software projects. These technologies address long-standing challenges of ambiguity, incompleteness, and inconsistency in natural language requirements (Sonbol et al., 2022; Shah & Jinwala, 2015). Specific contributions include Part-of-Speech (POS) tagging and Named Entity Recognition (NER) for reducing ambiguity (Shah & Jinwala, 2015; Zhao, 2021), semantic analysis to improve elicitation and validation processes (Ferrari et al., 2017), and ML models such as Recurrent Neural Networks (RNNs) and Support Vector Machines (SVMs) for classifying requirements into functional and non-functional categories (Binkhonain & Zhao, 2019; Gnanasekaran et al., 2021; Imam et al., 2021). Word2Vec has also been applied to support semantic similarity analysis and traceability (AlDhafer et al., 2022). Reported performance varies with dataset and context; for example, Ferrari et al. (2017) achieved an F1-score of 0.85 in defect detection through NLP-based pattern recognition.

Despite these advances, several limitations persist. Progress is hindered by the scarcity of high-quality, annotated datasets, with many studies relying on domain-specific corpora that limit model generalizability across projects (Sonbol et al., 2022; Zhao et al., 2022; Binkhonain & Zhao, 2019). Furthermore, explainability remains a critical issue, as deep learning approaches often function as “black boxes,” reducing stakeholder trust (Ferrari et al., 2017; Zhao, 2021). Practical adoption in industry is further constrained by implementation costs, system complexity, and the need for specialized expertise (Dalpiaz et al., 2018).

Publication trends confirm accelerating research activity, with the number of relevant studies doubling in 2022 (Zhao et al., 2022). Future research should prioritize the creation of robust, domain-independent datasets, improving model generalizability and scalability, and enhancing transparency to facilitate broader industrial adoption.

SAP Academic Community Conference 2025 (D-A-CH)

Business Process Management drives Digital Transformation

Entry No.	Author	Titel	DOI	Literature Review	Case Study	Survey	Experiment	Framework	Data Set	Mixed Methods (Case Study, Interviews, Survey, Experiment)
1	Iqbal et al.	A Bird's Eye View on Requirements Engineering and Machine Learning	10.1109/APSEC.2018.00015	X						
2	Alzayed and Al-Hunaidy	A Bird's Eye View of Natural Language Processing and Requirements Engineering	10.14569/UACSA.2021.0120512	X						
3	Abualhaija et al.	A machine learning-based approach for demarcating requirements in textual specifications	10.1109/RE.2019.00017							X
4	Parra et al.	A methodology for the classification of quality of requirements using machine learning techniques	https://doi.org/10.1016/j.infsof.2015.07.006							X
5	Vineetha and Samuel	A Multinomial Naïve Bayes Classifier for identifying Actors and Use Cases from Software Requirement Specification documents	10.1109/CONIT55038.2022.9848290				X			
6	Binkhomain and Zhao	A review of machine learning algorithms for identification and classification of non-functional requirements	https://doi.org/10.1016/j.eswax.2019.100001	X						
7	Perez-Verdejo et al.	A systematic literature review on machine learning for automated requirements classification	https://doi.org/10.1109/CONISOFT50191.2020.00014	X						
8	Mehraj et al.	A Tertiary Study on AI for Requirements Engineering	https://doi.org/10.1007/978-3-031-57327-9_10	X						
9	Umar and Lano	Advances in automated support for requirements engineering: a systematic literature review	https://doi.org/10.1007/978-66-023-00411-0	X						
10	Tikayat Ray et al.	Agile Methodology for the Standardization of Engineering Requirements Using Large Language Models	https://doi.org/10.3390/systems11070352				X			
11	Yeow et al.	An Automated Model of Software Requirement Engineering Using GPT-3.5	10.1109/ICETSS61505.2024.10459458				X			
12	AlDhafer et al.	An end-to-end deep learning system for requirements classification using recurrent neural networks	https://doi.org/10.1016/j.infsof.2022.106877				X			
13	Liu et al.	Artificial Intelligence in Software Requirements Engineering: State-of-the-Art	10.1109/IR54793.2022.00034	X						
14	Arora et al.	Automated Checking of Conformance to Requirements Templates Using Natural Language Processing	10.1109/TSE.2015.2428709		X					
15	Winkler and Vogelsang	Automatic classification of requirements based on convolutional neural networks	10.1109/REW.2016.021				X			
16	Li and Nong	Automatically classifying non-functional requirements using deep neural network	https://doi.org/10.1016/j.patrec.2021.108948				X			
17	Kaur and Kaur	BERT-CNN: Improving BERT for Requirements Classification using CNN	https://doi.org/10.1016/j.procs.2023.01.234				X			
18	Mullis et al.	Deep Neural Networks in Natural Language Processing for Classifying Requirements by Origin and Functionality: An Application of BERT in System Requirements	https://doi.org/10.1115/1.4063764				X			
19	Ferrari et al.	Detecting requirements defects with NLP patterns: an industrial experience in the railway domain	https://doi.org/10.1007/s10664-018-9596-7		X					
20	Mahmoud and Williams	Detecting, classifying, and tracing non-functional software requirements	https://doi.org/10.1007/s00766-016-0225-8				X			
21	Wang et al.	Enhancing automated requirements traceability by resolving polysemy	https://doi.org/10.1109/RE.2018.00053				X			
22	Sainani et al.	Extracting and Classifying Requirements from Software Engineering Contracts	10.1109/RE48521.2020.00026				X			
23	Sholih et al.	Generating BPMN diagram from textual requirements	https://doi.org/10.1016/j.jssci.2022.10.007				X			
24	Luitel et al.	Improving requirements completeness: automated assistance through large language models	https://doi.org/10.1007/978-66-024-00416-3				X			
25	Rajender Kumar Surani et al.	Intelligent Chatbot for Requirements Elicitation and Classification	https://doi.org/10.1109/RETECT46194.2019.0010607				X			
26	Romani et al.	Investigating ChatGPT's Potential to Assist in Requirements Elicitation Processes	https://doi.org/10.1109/SEAA60479.2023.00064				X			
27	Zamani et al.	Machine Learning in Requirements Engineering: A Mapping Study	https://doi.org/10.1109/REWS3955.2021.00023	X						
28	Zhao et al.	Natural Language Processing for Requirements Engineering: A Systematic Mapping Study	https://doi.org/10.1145/3444689	X						
29	Ferrari et al.	Natural Language Requirements Processing: A 4D Vision	https://doi.org/10.1109/MS.2017.4121207					X		
30	Hey et al.	NoBERT: Transfer Learning for Requirements Classification	https://doi.org/10.1109/RE48521.2020.00028				X			
31	Winkler et al.	Optimizing for recall in automatic requirements classification: An empirical study	https://doi.org/10.1109/RE.2019.00016				X			
32	Ferrari et al.	PURE: A Dataset of Public Requirements Documents	https://doi.org/10.1109/RE.2017.29						X	
33	Tamai and Anzai	Quality requirements analysis with machine learning	https://doi.org/10.5220/9906694502410248				X			
34	Calle Gallego and Zapata Jaramillo	QUARE: towards a question-answering model for requirements elicitation	https://doi.org/10.1007/s10515-023-00386-w							X
35	Dekhtyar and Fong	RE Data Challenge: Requirements Identification with Word2Vec and TensorFlow	https://doi.org/10.1109/RE.2017.26				X			
36	Nayak et al.	Req2Spec: Transforming Software Requirements into Formal Specifications Using Natural Language Processing	https://doi.org/10.1007/978-3-030-98464-9_8				X			
37	Shah and Jimwala	Resolving Ambiguities in Natural Language Software Requirements: A Comprehensive Survey	https://doi.org/10.1145/2815021.2815032			X				
38	Wang	Semantic information extraction for software requirements using semantic role labeling	https://doi.org/10.1109/PIC.2015.7489864				X			
39	Merten et al.	Software Feature Request Detection in Issue Tracking Systems	https://doi.org/10.1109/RE.2016.8				X			
40	Inam et al.	SVM Machine Learning Classifier to Automate the Extraction of SRS Elements	https://doi.org/10.14569/UACSA.2021.0120522				X			
41	Sorbol et al.	The Use of NLP-Based Text Representation Techniques to Support Requirement Engineering Tasks: A Systematic Mapping Review	https://doi.org/10.1109/ACCESS.2022.3182372	X						
42	Gnanasekaran et al.	Using recurrent neural networks for classification of natural language-based non-functional requirements	https://www.semanticscholar.org/paper/Using-Recurrent-Neural-Networks-for-Classification-Gnanasekaran-Chakraborty/144083c20488bd5910224b67681a07b34340				X			
43	El-Hajjimi et al.	Which AI Technique Is Better to Classify Requirements? An Experiment with SVM, LSTM, and ChatGPT	https://doi.org/10.48550/arXiv.2311.11547				X			
44	Alhoshan et al.	Zero-shot learning for requirements classification: An exploratory study	https://doi.org/10.1016/j.infsof.2023.107202				X			
				10	2	1	26	1	1	3

Table 1: Overview of the selected papers for the SLR

4. Discussion

The findings of this study highlight the transformative potential of Natural Language Processing (NLP) and Machine Learning (ML) in automating Requirements Engineering (RE). By addressing the three research questions, the discussion provides a critical analysis of the current state of the field, the effectiveness of existing tools and techniques, and the challenges that hinder their widespread adoption.

4.1. The Growing Role of NLP and ML in RE

The increasing complexity of software projects and the limitations of traditional RE methods have driven the adoption of NLP and ML. These technologies have demonstrated their ability to automate tasks such as ambiguity resolution, requirement classification, and semantic analysis, which were traditionally manual and error-prone. For instance, NLP techniques like Part-of-Speech (POS) Tagging and Named Entity Recognition (NER) have proven effective in reducing ambiguities in natural language requirements (Shah & Jinwala, 2015; Zhao, 2021). Similarly, ML models such as Recurrent Neural Networks (RNNs) and Word2Vec have shown promise in classifying and analyzing requirements (Binkhonain & Zhao, 2019; Ferrari et al., 2017). However, while these advancements are significant, their practical application remains limited to specific contexts, often constrained by the quality of input data and the complexity of the requirements.

4.2. Effectiveness of Tools and Techniques

The tools and techniques identified in this study have been evaluated using metrics such as F1-Score, Precision, and Recall, which provide insight into their performance. For example, Ferrari et al. (2017) reported an F1-score of 0.85 for detecting requirement defects using NLP patterns. Despite these promising results, the effectiveness of these tools is highly dependent on the availability of high-quality, annotated datasets. Moreover, while NLP techniques excel in preprocessing and analyzing requirement texts, their integration with ML models often requires domain-specific customization, limiting their scalability and generalizability (Sonbol et al., 2022).

4.3. Challenges in Practical Adoption

The integration of NLP and ML into requirements engineering (RE) processes continues to encounter several significant challenges. A primary concern is the limited availability of robust, annotated datasets necessary for effectively training ML models. Many existing studies rely on domain-specific datasets, which constrain the generalizability and broader applicability of these models across different contexts (Zhao et al., 2022). Another critical issue is the lack of transparency and explainability, particularly in deep learning approaches. These models often operate as black boxes, making it difficult for stakeholders to understand how decisions are derived—an essential requirement for practical adoption in real-world RE scenarios (Ferrari et al., 2017). Furthermore, although numerous tools have shown strong performance in academic research, their uptake in industrial settings remains limited. This gap is primarily attributed to factors such as high implementation costs, system complexity, and the necessity for specialized technical expertise (Dalpiaz et al., 2018).

4.4. Future Directions

The discussion underscores the need for future research to address these challenges. Future work should prioritize (1) developing robust, domain-independent datasets to enhance the generalizability of ML models, (2) improving the transparency and explainability of ML tools to build trust among stakeholders, and (3) bridging the gap between research and practice by creating cost-effective, user-friendly tools that can be easily adopted in industry settings.

In conclusion, while NLP and ML have shown significant potential in automating RE, their practical implementation requires addressing critical challenges related to data quality, model transparency, and industry adoption. By overcoming these barriers, these technologies can revolutionize RE processes, making them more efficient and reliable.

5. Bibliography

- AlDhafer, O., Ahmad, I. & Mahmood, S. (2022). An end-to-end deep learning system for requirements classification using recurrent neural networks. *Information and Software Technology*, 147. <https://doi.org/10.1016/j.infsof.2022.106877>
- Binkhonain, M. & Zhao, L. (2019). A review of machine learning algorithms for the identification and classification of non-functional requirements. *Expert Systems with Applications: X*, 1, 100001. <https://doi.org/10.1016/j.eswx.2019.100001>
- Dalpiaz, F., Ferrari, A., Franch, X. & Palomares, C. (2018). Natural Language Processing for Requirements Engineering: The Best Is Yet to Come. *IEEE Software*, 35(5), 115–119. <https://doi.org/10.1109/MS.2018.3571242>
- Ferrari, A., Dell’Orletta, F., Esuli, A., Gervasi, V. & Gnesi, S. (2017). Natural Language Requirements Processing: A 4D Vision. *IEEE Software*, 34(6), 28–35. <https://doi.org/10.1109/MS.2017.4121207>
- Gnanasekaran, R. K., Chakraborty, S., Dehlinger, J., & Deng, L. (2021, October). Using recurrent neural networks for classification of natural language-based non-functional requirements. In *REFSQ workshops* (Vol. 2857).
- Imam, A. T., Alhroob, A. & Jumah, W. (2021). SVM Machine Learning Classifier to Automate the Extraction of SRS Elements. *International Journal of Advanced Computer Science and Applications*, 12(3). <https://doi.org/10.14569/IJACSA.2021.0120322>
- Kitchenham, B. (2004). Procedures for performing systematic reviews. Keele, UK, Keele University, 33(2004), 1-26.
- Ribeiro, C., Farinha, C., Pereira, J. & Mira Da Silva, M. (2014). Gamifying requirement elicitation: Practical implications and outcomes in improving stakeholders collaboration. *Entertainment Computing*, 5(4), 335–345. <https://doi.org/10.1016/j.entcom.2014.04.002>
- Shah, U. S. & Jinwala, D. C. (2015). Resolving Ambiguities in Natural Language Software Requirements: A Comprehensive Survey. *ACM SIGSOFT Software Engineering Notes*, 40(5), 1–7. <https://doi.org/10.1145/2815021.2815032>
- Sinkie, M., Gronli, T. M., Midekso, D. & Lakhan, A. (2023). Joint Impact of Agents and Services in Enhancing Software Requirements Engineering. *Electronics*, 12(18), 3955. <https://doi.org/10.3390/electronics12183955>

- Sonbol, R., Rebdawi, G. & Ghneim, N. (2022). The Use of NLP-Based Text Representation Techniques to Support Requirement Engineering Tasks: A Systematic Mapping Review. *IEEE Access*, 10, 62811–62830. <https://doi.org/10.1109/ACCESS.2022.3182372>
- Unterkalmsteiner, M., Gorschek, T., Feldt, R. & Klotins, E. (2015). Assessing requirements engineering and software test alignment—Five case studies. *Journal of Systems and Software*, 109, 62–77. <https://doi.org/10.1016/j.jss.2015.07.018>
- Zhao, L., Alhoshan, W., Ferrari, A., Letsholo, K. J., Ajagbe, M. A., Chioasca, E.-V. et al. (2022). Natural Language Processing for Requirements Engineering: A Systematic Mapping Study. *ACM Computing Surveys*, 54(3), 1–41. <https://doi.org/10.1145/3444689>

SAP ECC to S/4HANA Data Migration: Process Analysis and AI-Supported Optimization

S. Goneswaran, H. Gruber, N. Nawratil & H. Beckmann
Heilbronn University, Heilbronn, Germany
U. K. Gangashetty, SAP SSC, Irland
J. Reinhardt, J. Bubeck, SAP Deutschland SE & Co. KG

Abstract: With SAP ECC approaching the end of its mainstream support, enterprises face complex, resource-intensive migrations to SAP S/4HANA. This paper proposes a RAG-based AI Assistant that streamlines data profiling and field mapping via conversational guidance. Utilizing SAP BTP, vector databases, and large language models, the assistant enables interactive access to system-specific knowledge, improving transparency and reducing manual effort across migration phases. A conceptual system framework and a proof-of-concept illustrate how the assistant can complement established SAP-centric workflows while highlighting current limitations, including data privacy considerations and the need for staged data pipelines when using external LLM services. The paper concludes with a discussion on future directions toward AI-enhanced, autonomous migration co-pilots, contributing to the vision of intelligent, agile ERP transformations through scalable AI integration.

Keywords: SAP S/4HANA, RAG, LLM, SAP BTP, AI-Assisted Data Migration

1. Introduction

With the approaching end of mainstream support for SAP ECC and the increasing adoption of SAP S/4HANA, many enterprises face complex and resource-intensive migration projects. Expert interviews and practical experiences reveal persistent challenges across the migration lifecycle, often resulting in delays, failed test cycles, and postponed go-lives, leading to significant operational and financial risks. Recent advancements in Artificial Intelligence (AI), particularly in Large Language Models (LLMs) and Retrieval-Augmented Generation (RAG), offer new opportunities to support and streamline SAP data migration processes. This paper explores how a RAG-based AI Assistant can improve transparency, reduce manual effort, and enhance data profiling and field mapping during migration. The central research question is: *How can AI improve the SAP ECC to S/4HANA data migration process?*

The following sub-questions are addressed:

- What are the key phases of the SAP ECC to S/4HANA data migration process?
- What are the main bottlenecks that reduce efficiency?
- How can an AI solution enhance or automate specific tasks in this process?

To answer these questions, the study follows a structured methodology combining expert interviews, a systematic literature review, and conceptual design. The literature analysis applies the model by Fettke (2006) and includes forward and backward search strategies, as recommended by Webster and Watson (2002). In alignment with the Design Science Research (DSR) methodology by Peffers et al. (2007), the AI solution is developed along the first four core DSR phases. The paper begins with a theoretical foundation and a current state analysis.

Based on the findings, it presents the conceptual AI solution, including an enterprise architecture model, a system framework within SAP BTP, and PoC mock-ups. The paper concludes with a reflection on key insights and outlines directions for future research.

2. Theoretical Background

This chapter introduces key concepts relevant to the context of migration data from the SAP ECC to S/4HANA, providing the foundation for the proposed solution and its integration within enterprise environments.

2.1. Large Language Models

LLMs have become a core technology in natural language processing (NLP), using large pre-trained transformer architectures to understand semantic patterns and language structure (Zhao et al. 2024). Unlike earlier models, LLMs operate with billions of parameters, enabling zero-shot and few-shot learning with minimal fine-tuning. Their scalability supports diverse use cases such as content generation, customer support, translation, and information retrieval, significantly improving productivity and user experience. (Gupta et al. 2025) LLMs are increasingly used in enterprise environments to support complex, knowledge-intensive processes through advanced NLP capabilities (Pasupuleti et al. 2024). Their strength lies in automating decisions and enabling context-aware data evaluation. MacMaster and Sinistore (2024) highlight their effectiveness in data quality assessments, tasks requiring subjective judgment and contextual reasoning. LLM-based chatbots are also improving operational efficiency by enabling human-like, context-aware interactions (Rani et al. 2024). Despite these benefits, challenges around transparency, interpretability, and data privacy remain and must be addressed to ensure responsible use (Pasupuleti et al. 2024). Overall, LLMs can transform enterprise workflows, but success depends on strategic deployment and ethical oversight.

2.2. Retrieval-Augmented Generation

Retrieval-augmented generation (RAG) enhances the accuracy of generative language models by grounding their outputs in external, factual sources (Lewis et al. 2020, Yu et al. 2025). It combines a retrieval module which identifies contextually relevant content via vector search and a generation module, typically an LLM, to produce informed responses and reduce the risk of hallucinations (Yu et al. 2025). In enterprise settings, RAG supports precise, data-intensive tasks by integrating structured and unstructured sources, improving real-time business-critical applications (Shi et al. 2025). Overall, RAG improves the reliability and contextual depth of generative AI outputs for enterprise use.

2.3. SAP Business Technology Platform

The SAP Business Technology Platform (BTP) offers a modular foundation for integrating AI, data, and process automation within enterprise workflows. It includes key services such as *SAP AI Core*, which manages AI model lifecycles including training, deployment, and monitoring, and the *Generative AI Hub*, which enables enterprises to experiment with and integrate LLMs via prompt-based interfaces. *SAP HANA Cloud* further extends BTP's capabilities through vector search and knowledge graphs, supporting semantic retrieval and data-driven AI scenarios like RAG. These components together enable the scalable and secure implementation of AI-driven solutions in SAP landscapes. (SAP SE 2025)

3. Current Analysis of the SAP ECC to S/4HANA Data Migration Process

This chapter analyzes the current SAP ECC to S/4HANA data migration process.

3.1. Key-Roles in SAP Data Migration

Successful SAP data migrations require collaboration among different roles. Each role has been assigned specific responsibilities to ensure that migration is carried out to the highest professional standards and is accurate and fully compliant. *Business Process Owner (BPO)* represents the customer and has in-depth knowledge of business processes, legal requirements, and domain-specific logic (e.g. payment modalities under German regulations). The BPO is responsible for selecting and prioritizing migration objects, providing source data, validating and approving migrated data, and supporting field mapping with their domain expertise. *Functional Team* is part of the project team. They have a strong understanding of S/4HANA processes and are able to translate BPO requirements into SAP-aligned processes. The team will then configure the target system, including company codes, purchasing organizations, and workflows. They also collaborate closely with the BPO team to create accurate field mappings. *Technical Team* is responsible for overseeing the technical aspects of the migration process. This includes data extraction, transformation, and loading using ETL tools, such as SAP Data Services. The team uses the SAP Migration Cockpit to perform migration and carry out technical validation. They also support testing cycles and execute the production migration during the cutover phase.

3.2. Process Description: SAP ECC to S/4HANA Data Migration

The following section analyzes the current SAP ECC to S/4HANA data migration process, referencing the BPMN model (see Fig. 1) developed during weekly SAP expert exchanges and process review meetings. The migration process starts with project initialization, where the BPO and Functional Team define scope and stakeholder coordination. The BPO assigns responsibilities for each migration object, ensuring accountability for data provision and validation. Relevant ECC tables and fields are identified and documented in a source inventory. The BPO then conducts data profiling to assess data quality, completeness, duplicates, and field usage. The outcome of this task is a profiling report, which provides transparency and supports data cleansing and mapping decisions. Based on this, the Functional Team selects migration objects and, together with the BPO, defines field mappings and transformation rules. The migration approach is chosen, either file-based with manual templates or ETL-based using SAP Data Services. In the ETL approach, the Technical Team builds, extracts, transforms, and loads data into Staging Tables before transferring it to S/4HANA. Subsequently, the Functional Team and BPO validate the data, resolve issues, and conduct User Acceptance Tests. After approval and formal sign-off by the BPO, the Technical Team performs the production data migration, ensuring operational readiness in S/4HANA.

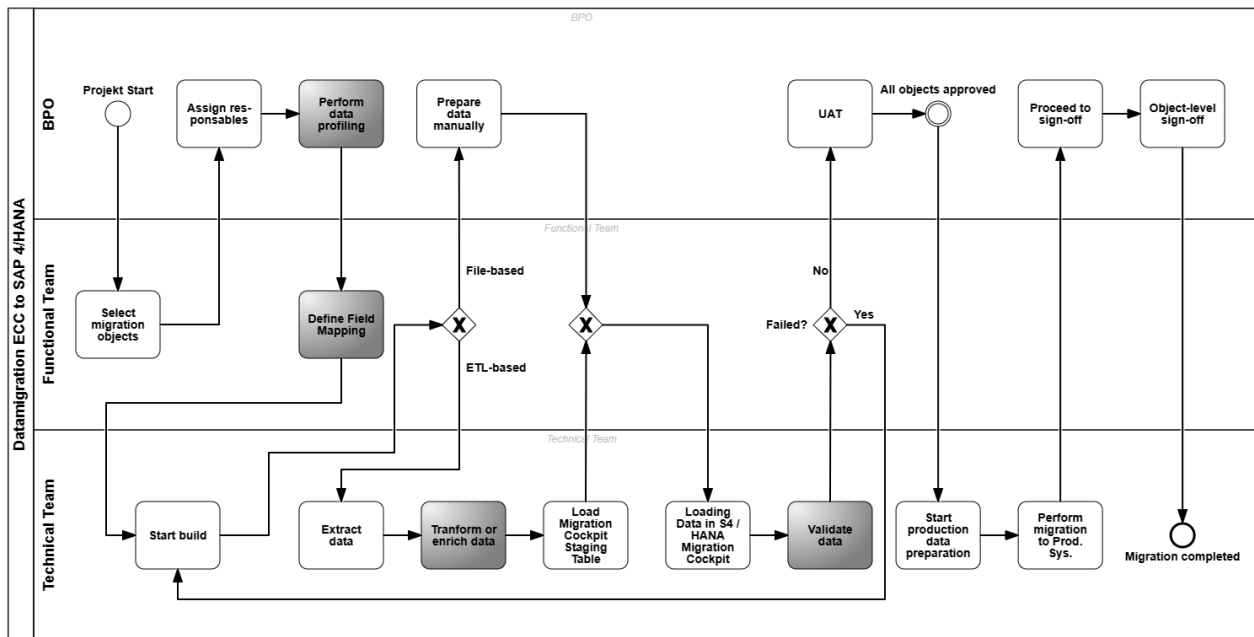


Fig. 1: BPMN Data Migration ECC to S4/HANA

3.3. Identified Bottlenecks in the SAP Data Migration Process

Several bottlenecks were identified in key steps of the SAP ECC to S/4HANA migration process, based on expert Interviews and process review meetings (see Fig. 1). These lead to inefficiencies, higher manual workload, and increased risk of delays. *Data profiling* is often manual or semi-manual, using exports and staging extracts instead of automated system-wide scans. It is frequently delayed until after field mapping, allowing incomplete or inconsistent data into later stages. Fragmented tools (e.g., MS Excel) and missing monitoring for duplicates or outliers cause quality gaps and delay mapping and cleansing. *Field mapping* is time-consuming and error-prone, requiring manual checks of hundreds of fields with varying rules. Free-text entries and unclear responsibilities add complexity. Without reusable knowledge bases, mappings remain inconsistent and project-specific. *Transforming and enriching data* is typically manual, with business rules implemented in ETL tools like SAP Data Services. This rigid setup requires manual updates for any mapping changes and lacks dynamic validation. Enrichment is hardcoded rather than automated, increasing maintenance. *Validation* often occurs too late, requiring rework and multiple load cycles. Errors are corrected manually, and feedback is milestone-based without real-time dashboards, leading to testing delays and resource loss. Experts further stressed that field mapping requires understanding the actual content and use of legacy data, not just technical matching. Manual data lookups during workshops were cited as a major bottleneck, underlining the need for earlier, systematic profiling to support accurate mappings.

4. Target State – Proposed Solution: AI Chat Assistant

In this chapter, the proposed solution is structured according to the Design Science Research (DSR) methodology by Peffers et al. (2007).

4.1. Phase 1: Identified Problem and Motivation

The need for AI in the SAP ECC to S/4HANA migration process stems from practical challenges identified in Chapter 3. Expert interviews and process reviews revealed major bottlenecks, especially in manual data discovery, time-consuming profiling, and error-prone field mapping. These tasks demand a deep understanding of legacy data, currently achieved through manual lookups in workshops. Experts stressed that field mapping is not just technical label matching but requires insight into data content and usage to meet business and compliance needs. Without automated support, early migration phases are prone to delays, errors, and high coordination effort across BPO, Functional, and Technical Teams.

4.2. Phase 2: Objectives of the Solution

Based on the identified challenges in the SAP ECC to S/4HANA migration process, this solution aims to address key bottlenecks through clearly defined objectives. It seeks to *enhance transparency* by giving BPOs and Functional Teams immediate, context-aware access to legacy data without relying on manual lookups. It further aims to *reduce manual effort* by providing an AI assistant that supports data profiling and mapping preparation, saving time and resources. To *improve data quality awareness*, the solution enables early detection of inconsistencies to support accurate and business-aligned migration planning. A key priority is to ensure seamless integration into existing SAP workflows, supporting all involved teams without disrupting established processes. Finally, the approach *leverages SAP-native technologies*, including SAP BTP, vector databases, RAG, and LLMs, to provide scalable, cloud-based support for enterprise migration projects.

4.3. Phase 3: Design and Development

The proposed RAG-based AI Assistant is designed for seamless integration into SAP-centric migration workflows, providing interactive, context-aware support throughout the SAP ECC to S/4HANA process. As shown in the Conceptual System Framework (see Fig. 2), the solution runs entirely on SAP BTP, ensuring compatibility with SAP environments.

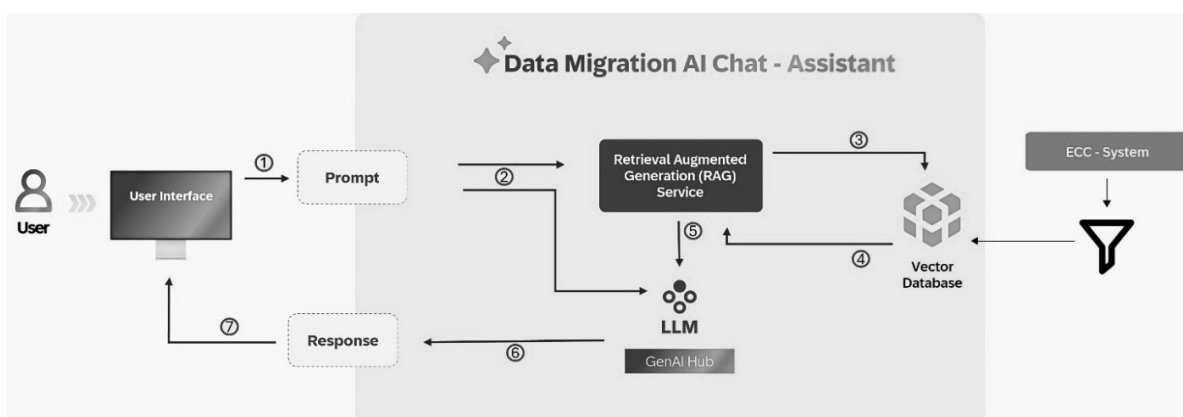


Fig. 2: Data Migration Chat-Assistant

Users submit natural language prompts via a chat UI. These are processed by the RAG service on SAP BTP, which generates embeddings and queries a vector database containing chunked ECC data. Relevant text segments are retrieved, combined with the original prompt, and sent to the LLM via SAP GenAI Hub. The LLM responds with clear, data-backed answers. This

approach embeds the AI Assistant into existing migration workflows, reducing manual effort and improving data transparency and understanding.

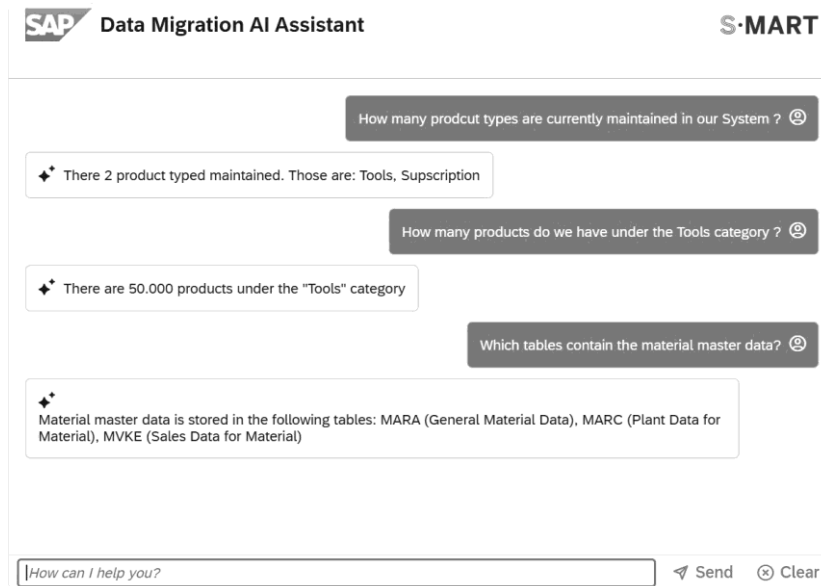


Fig. 3: PoC - Data Migration AI Assistant

4.4. Phase 4: Demonstration

To illustrate the proposed AI Assistant, a conceptual Proof of Concept (PoC) mockup (see Fig. 3) was developed. It visualizes how the assistant could support data profiling and field mapping in SAP ECC to S/4HANA migration. The assistant, shown within an SAP-style chat UI, provides direct, context-rich responses based on ECC data. This allows users to clarify questions without manual system navigation. The mockup demonstrates how the assistant accelerates mapping workshops, supports BPOs and Functional Teams with data insights, reduces manual lookups, and increases transparency. As system access was granted late by SAP, implementation of the PoC has only recently begun. No live integration or deployment has taken place yet.

4.5. Enterprise Architecture Model for the Target State

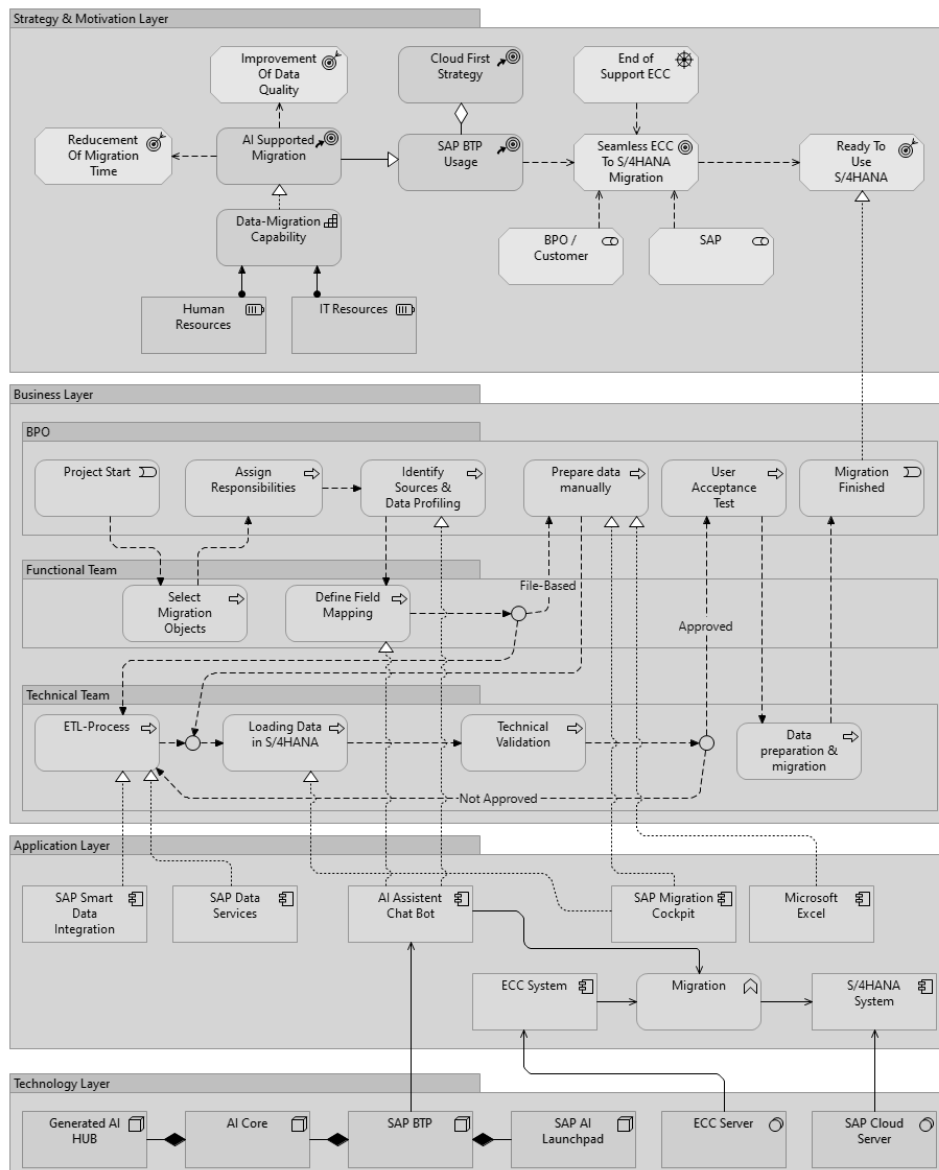


Fig. 4: Enterprise Architecture Model (ArchiMate)

An Enterprise Architecture Model (see Fig. 4), developed using ArchiMate, visualizes the AI-supported target state for SAP ECC to S/4HANA migration. It ensures alignment with business goals such as reducing migration time, improving data quality, and supporting a cloud-first strategy. At the *Strategy Layer*, the model positions the AI Assistant as a strategic enabler responding to drivers like ECC phase-out and SAP BTP adoption. In the *Business Layer*, the assistant complements existing roles by supporting BPOs and Functional Teams with interactive profiling and mapping insights. The *Application Layer* shows integration with tools like SAP Data Services and Migration Cockpit, while the *Technology Layer* includes SAP BTP, SAP AI Core, Launchpad, GenAI Hub, and vector databases. This layered architecture ensures scalable, SAP-native integration without disrupting current workflows, guiding a smooth, intelligent enterprise transition.

5. Discussion & Critical Reflection

The developed RAG-based AI Assistant shows strong potential for SAP ECC to S/4HANA migration, but key limitations and open questions remain regarding practical implementation. Currently, LLMs are not hosted locally or within SAP infrastructure; instead, queries are sent to external providers (e.g. OpenAI, Anthropic). As a result, prompt-based queries containing potentially sensitive ECC data may leave the enterprise environment, raising data privacy and compliance concern, especially with personal or business-critical information. One mitigation strategy could involve staging and aggregating data before submission to reduce exposure risks during retrieval. Furthermore, this work remains a conceptual design and PoC, with open questions about its performance in real-world scenarios, across complex landscapes, and within enterprise governance frameworks. Aspects such as user acceptance, operational efficiency, and process integration still require validation through practical pilot projects to ensure secure and scalable deployment. This paper examined the integration of a RAG-based AI Assistant into the SAP ECC to S/4HANA migration process, showing how it can improve transparency, reduce manual effort, and support data profiling and field mapping. Leveraging SAP BTP, the solution offers a more efficient and data-driven migration process. The assistant provides interactive access to system knowledge, improving data quality and reducing redundant coordination. However, the work remains at a conceptual and PoC stage. Practical implementation requires further research, particularly in data privacy, integration, and scalability. Future work could extend the assistant into an autonomous agent for rule generation and anomaly detection. Research could also explore custom model hosting (e.g. via vLLM) to improve data sovereignty and control. Further studies should focus on implementing and evaluating the PoC in SAP BTP environments, analyzing performance, scalability, and user acceptance to offer practical guidance for enterprise adoption.

6. Literature

- Fettke, P. (2006). State-of-the-Art des State-of-the-Art. In: *WIRTSCHAFTSINFORMATIK* 48(4).
- Gupta, R., Tiwari, S. & Chaudhary, P. (2025). Large language models. In: *Lecture notes on data engineering and communications technologies*, 81-102.
- Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S. & Kiela, D. (2020). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. In: *arXiv (Cornell University)*.
- MacMaster, S. & Sinistore, J. (2024). Testing the use of a large language model (LLM) for performing data quality assessment. In: *The International Journal Of Life Cycle Assessment*.
- Pasupuleti, R., Vadapalli, R., Mader, C. & Timothy, N. (2024). Popular LLM-Large Language Models in Enterprise Applications. In: *The 2nd International Conference on Foundation and Large Language Models (FLLM2024)*, 125-131.
- Peffer, K., Tuunanen, T., Rothenberger, M. A. & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. In: *Journal Of Management Information Systems*, 24(3), 45-77.

Rani, Y. A., Balaram, A., Sirisha, M. R., Nabi, S. A., Renuka, P. & Kiran, A. (2024). AI Enhanced Customer Service Chatbot. In: *2024 International Conference On Science Technology Engineering And Management (ICSTEM)*, 1-5.

SAP SE (2025). SAP AI Core. *SAP Help Portal*.

Shi, I., Li, Z., Wang, W., He, L., Yang, Y. & Shi, T. (2025). eSapiens: A Real-World NLP Framework for Multimodal Document Understanding and Enterprise Knowledge Processing. In: *arXiv (Cornell University)*.

Webster, J., Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. In: *MIS Quarterly* 26(2), xiii–xxiii.

Yu, H., Gan, A., Zhang, K., Tong, S., Liu, Q. & Liu, Z. (2025). Evaluation of Retrieval-Augmented Generation: A Survey. In: *Communications in computer and information science*, vol 2301, 102-120.

Zhao, H., Chen, H., Yang, F., Liu, N., Deng, H., Cai, H., Wang, S., Yin, D. & Du, M. (2024). Explainability for Large Language Models: A Survey. In: *ACM Transactions On Intelligent Systems And Technology*, 15(2), 1-38.

Tiny Models, Big Impact: A Federated Learning Approach to Monitoring Depression Treatment via Smartphones and Wearables

E. Kaaya & J. Gómez

The University of Oldenburg, Oldenburg, Germany

Abstract: This study proposes a federated learning framework for monitoring treatment response in patients with major depressive disorder using behavioral data collected from smartphones and wearable devices such as smartwatches. Unlike previous federated learning applications in the health domain that focus primarily on disease detection, diagnosis, risk classification and activity recognition, this framework addresses the underexplored challenge of monitoring treatment outcomes in real time while preserving user privacy. By leveraging tiny federated learning models optimized for low-power devices and time-series data, the framework supports decentralized training and on-device inference enabling scalable, privacy-conscious mental health monitoring. It incorporates model compression techniques to ensure feasibility in resource-constrained environments. While not yet empirically validated, this framework establishes a foundation for ethical and efficient digital health tools that can support timely clinical decision-making in mental health care.

Keywords: Major Depressive Disorder, Tiny Federated Learning, Behavioral Data

1. Introduction

Major Depressive Disorder (MDD) affects over 300 million people worldwide and stands as the leading cause of disability (Cui et al., 2024). Characterized by symptoms such as suicidal ideation, anhedonia (lack of previously enjoyable activities), fatigue, and weight loss, MDD profoundly impairs social, professional, and academic functioning (Marx et al., 2023). The World Health Organization predicts that by 2030, MDD will be the single largest contributor to the global burden of disease, underscoring the urgent need for more effective diagnostic, treatment, and monitoring strategies (Malhi & Mann, 2018).

Current approaches to diagnosing and monitoring MDD rely on self-reports and periodic questionnaires which fail to capture the disorder's dynamic course and often lead to delayed or suboptimal treatment adjustments (Koinis et al., 2022). Given that standard treatment follows a trial-and-error model, there is a pressing need for methods that can track a patient's evolving mental state in real time and alert clinicians to emerging non-response or relapse.

Smartphones and wearable devices (e.g., smartwatches) offer a non-invasive, continuous means of collecting behavioral and physiological metrics such as sleep patterns, physical activity, and phone usage that correlate with depressive symptoms (Betteridge et al., 2021; Kim et al., 2023). Machine Learning (ML) models trained on these large, passive data streams have shown promise in both emotion classification and depression detection (Singh et al., 2023). However, concerns over data privacy in mental health have limited centralized ML approaches, giving rise to Federated Learning (FL), which enables collaborative model training without sharing raw data (Shek et al., 2021).

To deploy FL on resource-limited devices, researchers have developed Tiny Machine Learning (TinyML) models that drastically reduce computational and memory demands (Lin et al.,

2023). Integrating these models with FL known as Tiny Federated Learning (TinyFL) allows smartphones and wearables to participate in decentralized training without sacrificing performance. Further, fully Decentralized Federated Learning (DFL) architectures, such as the GLow extension of the Flower framework, eliminate reliance on a central server by using Gossip Learning among neighboring devices (Belenguer et al., 2025). This paper proposes leveraging TinyFL within a DFL environment to monitor MDD treatment response via behavioral data, aiming to deliver scalable, privacy-preserving, real-time insights into patient progress across diverse populations.

1.1. Related Studies

Federated Learning in the Health Domain

Research in the healthcare domain is increasingly adopting FL to enable collaborative training among institutions (cross-silo FL) and devices (cross-device FL) in order to get insight on patient health. The widespread use of resource constrained devices such as smartphones and smartwatches has enabled the collection of passive health data (Betteridge et al., 2021). The collected data has been used to train ML models to provide insights necessary for disease diagnosis, treatment and monitoring. The following review organizes recent peer-reviewed contributions into thematic focus highlighting the methodologies used in developing FL systems in the health domain.

Privacy-Preserving Federated Learning on Wearable Devices

To enhance data privacy in FL systems, Can & Ersoy, 2021 introduced a privacy-preserving FL algorithm for stress detection using Photoplethysmography (PPG)-based heart activity signals collected from commercial smartwatches and smartbands such as Samsung Galaxy Gear S, Gear S2, and Empatica E4. The model architecture utilized a Multilayer Perceptron (MLP) with two hidden layers of 100 neurons each, employing Rectified Linear units (ReLU) activations, trained locally on clients and aggregated via federated averaging using TensorFlow Federated (TFF).

Wang et al., 2023 introduced FRESH, a FL framework for disease detection that leverages time-series physiological data such as blood pressure, ECG, physical activity, and heart rate collected from wearables and IoT devices. By running ML models directly on resource-constrained devices such as smartphones and tablets, FRESH preserves privacy using an optimized elliptic-curve ring signature scheme with batch verification to obscure client identities and reduce computational overhead during signature checks.

Furthermore, Song et al., 2024 enhanced FL for health risk prediction by securing gradient updates with Paillier homomorphic encryption and authenticating clients via a Schnorr zero-knowledge proof module. They trained models using the MNIST and CIFAR-10 datasets and further optimized performance with a gradient-filtering compression algorithm that discards non-essential updates and applies lightweight quantization to reduce computational overhead.

Model Compression & Resource Efficiency

Deployment of ML models on resource-constrained devices requires the use of model compression techniques to reduce computational and memory demands. The commonly used techniques are pruning, quantization and knowledge distillation. A study by **Qayyum et al., 2022** deployed COVID-19 diagnosis models on resource-constrained devices by combining

network pruning, 8-bit precision reduction, and knowledge distillation, using chest X-ray and ultrasound inputs. Their clustered FL approach outperformed standard FL and centralized ML, achieving a 16 percent F1-score gain on X-rays and an 11 percent gain on ultrasound by effectively handling multimodal data divergence.

Similarly, Aminifar et al., 2024 developed an epilepsy seizure detection framework using EEG data from wearable e-Glass systems (T7F7 and T8F8 electrodes). They applied knowledge distillation to compress a modular deep neural network by freezing the base model and training only the head resulting in a student model with 81–83 percent accuracy. The framework's scalability and real-world applicability were validated on the Amazon AWS cloud platform.

Advanced Neural Architectures for Edge Federated Learning

Khan et al., 2025 introduced Spiking neurons- Long Short-Term Memory (S-LSTM), a hybrid Human Activity Recognition (HAR) model that interleaves two 100-neuron LSTM layers with Leaky Integrate-and-Fire (LIF) spiking layers, uses surrogate gradients for spiking differentiability, and concludes with a 300-neuron dense LIF layer followed by a dropout and a final SoftMax. Tested on indoor and outdoor wearable data, S-LSTM outperformed standard LSTMs by up to 3.84 percent in accuracy and achieved a 32.3 percent increase in energy efficiency.

Standard Integration & Usability in Federated Learning

Akhmetov et al., 2025 integrated FL with Fast Healthcare Interoperability Resources (FHIR) to improve privacy and interoperability in wearable sensor data analysis. They evaluated ML models using Apple Watch data (Harvard Dataverse), Mi Band data (Kaggle), and Human Activity Recognition Trondheim (HARTH), and trained the ML model on the fmendes-DAT263x-demos dataset for activity classification. Automated Machine Learning (AutoML) was used to select and tune algorithms for non-expert end users. FL was simulated locally using Streamlit, PyCaret, and Flower. Results showed FL performance comparable to or slightly exceeding that of centralized ML.

Federated Learning for Multimodal Signals

A study conducted by Liu et al., 2025, developed a FL framework to classify three stress levels in nurses using multimodal wearable data such as accelerometer/gyroscope orientation, electrodermal activity, heart rate, and skin temperature. Their Neural Network featuring Rectified Linear Units (ReLU)-activated layers and a SoftMax output was trained with Stochastic Gradient Descent with a learning rate of 0.01, mini-batch size of 1024, and one local epoch per round across 100 rounds. It performed with an accuracy exceeding 90% and with precision, recall, and F1-scores above 85 percent, hence outperforming centralized learning.

The above studies address key concerns in FL for healthcare including data privacy, model compression, architectural design, multimodal data analysis and device interoperability. Despite these efforts, monitoring treatment response in MDD using behavioral data collected from smartphones and wearable devices remains unexplored. Moreover, fully decentralized TinyFL models capable optimised for real-time on-device inferences using time-series data are notably absent.

2. Proposed Framework

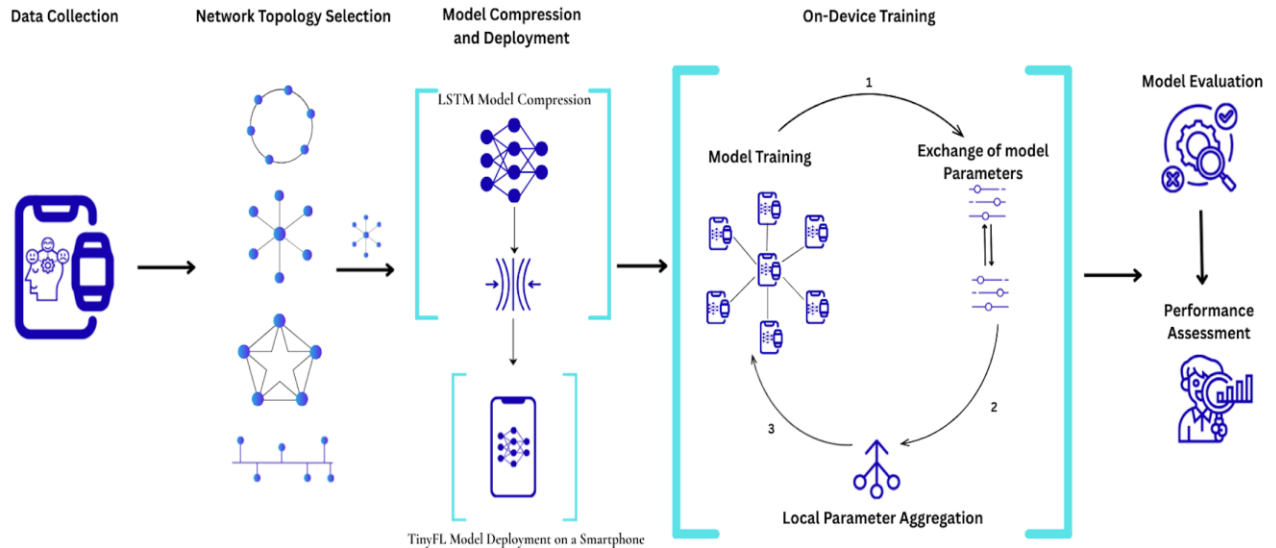


Fig. 1: Proposed Framework

2.1. Data Acquisition and Exploration

The framework leverages behavioral data collected from smartphones and wearable devices. A developed mobile application for data collection, the Data Collector (DC) together with wearable devices will collect behavioral and activity data such as sleep patterns, social media consumption, nutritional data, physical activity, GPS changes, smartphone usage, time spent at home, and sporting activity which are associated with MDD (Schröter et al., 2025). Post data acquisition, data cleaning will be carried out to remove all irrelevant and incomplete information, ensuring consistency. Following data cleaning, Exploratory Data Analysis (EDA) will be performed, this process involves analysing the data to identify patterns, correlations, and outliers. The Pandas library will be used for data analysis. Matplotlib and seaborn will be used for data visualization.

2.2. Network Topology and Client Behavior Customization

In Fig.1 above, a Topology Generator (TG) in the GLow framework is used to generate different network topologies. This allows experimentation to be carried out on different topologies to determine the ideal topology that enhances scalability, convergence, privacy, and model performance. Through the TG, multiple topologies ranging from fully disconnected to fully connected can be generated by linking each client to its closest n neighbors. Additionally, GLow has the unique ability to pre-define client behavior, together this will enable both the recognition of ideal network topologies and identification of optimal client configurations that support a DFL system capable of monitoring treatment response among MDD patients using behavioral data collected from smartphones and wearable devices.

2.3. Model Compression

As demonstrated in Fig.1, a Long Short-Term Memory (LSTM) model will be compressed into a TinyFL model suitable for deployment on smartphones and wearable devices. The choice of LSTM is influenced by its ability to process time-series data, making it suitable for analyzing changes in behavioral data of MDD patients which is an indicator of their treatment response (Walschots et al., 2024). The key compression techniques such as pruning, quantization and knowledge distillation will be applied depending on the model's complexity (Garcia et al., 2025). These techniques significantly reduce computational and communication overhead, enabling low-power devices, such as IoT devices and microcontrollers with limited resources, to leverage the capabilities of large models (Lin et al., 2023).

2.4. Model Deployment

Post-model compression, TinyFL model will be deployed on a smartphone. This deployment will be facilitated through the utilization of TensorFlow Lite for Microcontrollers (TFLite Micro), a lightweight version of TensorFlow Lite designed to run ML models on embedded systems and resource-constrained devices such as microcontrollers, wearables, and IoT devices (Manor & Greenberg, 2022).

2.5. Model Training, Local Parameter Exchange and Aggregation

As demonstrated above, an LSTM model will be trained on behavioural data collected from smartphones and wearable devices such as smartwatches, in an iterative manner. In the proposed framework, local training will begin in the smartphones, after the first round is complete, a single client is selected as the head using the round-robin algorithm. The selected head aggregates its local parameters with those of its neighbour by requesting their updated model parameters, this process is iterated until convergence is achieved. The proposed framework will be implemented using the Python programming language and Keras deep learning library. The TensorFlow library will be used for data pre-processing and visualization.

2.6. Model Evaluation

The final model will be assessed based on its accuracy, precision, recall, and F1-score. Furthermore, relevant performance indicators such as area under the ROC curve, mean squared error, and computational efficiency will be considered to ensure a thorough assessment. The rationale behind choosing these particular metrics, which encompass accuracy, precision, recall, and F1-score, is rooted in their capacity to offer a comprehensive evaluation of a model's performance across diverse contexts, with a specific emphasis on applications within the realms of machine learning and data science.

2.7. Performance Assessment

Key metrics such as model accuracy, convergence speed, memory footprint, latency, and energy usage are used to evaluate the system. This evaluation will be used to shed light on the technical affordances concerning trade-offs between model size and predictive performance as well as resource efficiency toward the long-term goal of deploying robust real-time monitoring of treatment response in patients with MDD.

3. Discussion

The proposed framework addresses a critical void in mental healthcare applications of FL by introducing a fully decentralized, TinyFL model specifically designed to monitor treatment response in MDD. Unlike previous FL studies focused on stress detection, disease diagnosis, or activity recognition this approach leverages behavioral data from smartphones and wearables such as sleep patterns, physical activity, GPS mobility, and smartphone usage to capture evolving, time-series indicators of treatment efficacy. It combines model-compression techniques (pruning, quantization, knowledge distillation), and resource-efficient architectures suitable for low-power devices. By eliminating the need for a central server via Gossip Learning in the GLow extension of the Flower framework, the system ensures seamless, scalable training and real-time, on-device inference thereby preserving patient privacy while delivering personalized assessment.

Although conceptually grounded in state-of-the-art methods, empirical validation remains an essential next step. Future work will involve deploying and evaluating the framework in real-world clinical settings to assess accuracy, scalability, and its ability to detect non-responsive treatment in diverse patient populations. Further enhancements such as incorporating personalized model adaptation and stronger privacy guarantees (e.g., differential privacy, homomorphic encryption) will be explored to reinforce trust, improve robustness, and optimize performance for practical, privacy-preserving mental health monitoring.

4. Conclusion

This study proposes a novel, FL framework designed to monitor treatment response in patients with MDD using behavioral data collected from smartphones and wearable devices. Unlike prior work in the intersection of FL and the health domain that focused primarily on disease diagnosis, detection, and activity recognition, this approach targets the underexplored area of treatment outcome monitoring offering the potential for real-time, on-device inference in a privacy-preserving manner. While empirical validation remains a future step, this work lays the groundwork for scalable and ethically sound mental health monitoring systems that can enhance clinical decision-making without compromising patient autonomy or data security.

5. References

- Akhmetov, A., Latif, Z., Tyler, B., & Yazici, A. (2025). Enhancing healthcare data privacy and interoperability with federated learning. *PeerJ Computer Science*, 11, e2870. <https://doi.org/10.7717/peerj-cs.2870>
- Aminifar, A., Shokri, M., & Aminifar, A. (2024). Privacy-preserving edge federated learning for intelligent mobile-health systems. *Future Generation Computer Systems*, 161, 625–637. <https://doi.org/10.1016/j.future.2024.07.035>
- Belenguer, A., Pascual, J. A., & Navaridas, J. (2025). *GLow—A Novel, Flower-Based Simulated Gossip Learning Strategy* (Version 1). arXiv. <https://doi.org/10.48550/ARXIV.2501.10463>
- Betteridge, C., Mobbs, R. J., & Ho, D. (2021). Proposed objective scoring algorithm for walking performance, based on relevant gait metrics: The Simplified Mobility Score

- (SMoSTM)—observational study. *Journal of Orthopaedic Surgery and Research*, 16(1), 419. <https://doi.org/10.1186/s13018-021-02546-8>
- Can, Y. S., & Ersoy, C. (2021). Privacy-preserving Federated Deep Learning for Wearable IoT-based Biomedical Monitoring. *ACM Transactions on Internet Technology*, 21(1), 1–17. <https://doi.org/10.1145/3428152>
- Cui, L., Li, S., Wang, S., Wu, X., Liu, Y., Yu, W., Wang, Y., Tang, Y., Xia, M., & Li, B. (2024). Major depressive disorder: Hypothesis, mechanism, prevention and treatment. *Signal Transduction and Targeted Therapy*, 9(1), 30. <https://doi.org/10.1038/s41392-024-01738-y>
- Ding, W., Liu, Y., Wang, Z., & Chu, Z. (2025). <https://doi.org/10.3390/a17090408>. *ICASSP 2025 - 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1–5. <https://doi.org/10.1109/ICASSP49660.2025.10888587>
- Garcia, M. E., Ochoa-Frongia, L., Neuhaus, J., Hong, J. C., Hinton, L., Livaudais-Toman, J., Feldman, M. D., Mangurian, C., Appelle, N., & Karliner, L. S. (2025). Depression Treatment After a Positive Depression Screen Result. *JAMA Internal Medicine*, 185(2), 221. <https://doi.org/10.1001/jamainternmed.2024.6211>
- Jerotic, S., Ignjatovic, N., Maric, N. P., Nesic, J., Jovandic, Z., Latas, M., Nestorovic, M., Jevtic, M., & Aftab, A. (2024). A Comparative Study on Mental Disorder Conceptualization: A Cross-Disciplinary Analysis. *Community Mental Health Journal*, 60(4), 813–825. <https://doi.org/10.1007/s10597-024-01240-3>
- Khan, S., Khan, M., Khan, M. A., Wang, L., & Wu, K. (2025). Advancing Medical Innovation through Blockchain-Secured Federated Learning for Smart Health. *IEEE Journal of Biomedical and Health Informatics*, 1–14. <https://doi.org/10.1109/JBHI.2025.3532976>
- Kim, J. S., Wang, B., Kim, M., Lee, J., Kim, H., Roh, D., Lee, K. H., Hong, S.-B., Lim, J. S., Kim, J.-W., & Ryan, N. (2023). Prediction of Diagnosis and Treatment Response in Adolescents With Depression by Using a Smartphone App and Deep Learning Approaches: Usability Study. *JMIR Formative Research*, 7, e45991. <https://doi.org/10.2196/45991>
- Koinis, L., Mobbs, R. J., Fonseka, R. D., & Natarajan, P. (2022). A commentary on the potential of smartphones and other wearable devices to be used in the identification and monitoring of mental illness. *Annals of Translational Medicine*, 10(24), 1420–1420. <https://doi.org/10.21037/atm-21-6016>
- Lin, J., Zhu, L., Chen, W.-M., Wang, W.-C., & Han, S. (2023). Tiny Machine Learning: Progress and Futures [Feature]. *IEEE Circuits and Systems Magazine*, 23(3), 8–34. <https://doi.org/10.1109/MCAS.2023.3302182>
- Liu, K., Xue, W., & Hou, D. (2025). Federated learning for nurse stress prediction using wearable sensors: Integrating biomechanical data. *Molecular & Cellular Biomechanics*, 22(5), 1699. <https://doi.org/10.62617/mcb1699>
- Malhi, G. S., & Mann, J. J. (2018). Depression. *The Lancet*, 392(10161), 2299–2312. [https://doi.org/10.1016/S0140-6736\(18\)31948-2](https://doi.org/10.1016/S0140-6736(18)31948-2)
- Manor, E., & Greenberg, S. (2022). Custom Hardware Inference Accelerator for TensorFlow Lite for Microcontrollers. *IEEE Access*, 10, 73484–73493. <https://doi.org/10.1109/ACCESS.2022.3189776>

- Marx, W., Penninx, B. W. J. H., Solmi, M., Furukawa, T. A., Firth, J., Carvalho, A. F., & Berk, M. (2023). Major depressive disorder. *Nature Reviews Disease Primers*, 9(1), 44. <https://doi.org/10.1038/s41572-023-00454-1>
- Qayyum, A., Ahmad, K., Ahsan, M. A., Al-Fuqaha, A., & Qadir, J. (2022). Collaborative Federated Learning for Healthcare: Multi-Modal COVID-19 Diagnosis at the Edge. *IEEE Open Journal of the Computer Society*, 3, 172–184. <https://doi.org/10.1109/ojcs.2022.3206407>
- Schröter, E., Klein, F., Elfert, P., Bredehorn, F., Räker, J., Aschwege, F., & Hein, A. (2025). Mobile Data Collection for Depression Analysis: An App Framework for Monitoring Mood and Depression Using Smartphone and Wearable Data: *Proceedings of the 18th International Joint Conference on Biomedical Engineering Systems and Technologies*, 781–788. <https://doi.org/10.5220/0013298700003911>
- Shek, A. C., Biondi, A., Ballard, D., Wykes, T., & Simblett, S. K. (2021). Technology-based interventions for mental health support after stroke: A systematic review of their acceptability and feasibility. *Neuropsychological Rehabilitation*, 31(3), 432–452. <https://doi.org/10.1080/09602011.2019.1701501>
- Singh, A. K., Kumar, A., & Choi, B. J. (2023). Privacy-Preserving Digital Intervention for Mental Health Using Federated Learning. In H. Zaynudinov, M. Singh, U. S. Tiwary, & D. Singh (Eds), *Intelligent Human Computer Interaction* (Vol. 13741, pp. 213–224). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-27199-1_22
- Song, C., Wang, Z., Peng, W., & Yang, N. (2024). Secure and Efficient Federated Learning Schemes for Healthcare Systems. *Electronics*, 13(13), 2620. <https://doi.org/10.3390/electronics13132620>
- Walschots, Q., Zarchev, M., Unkel, M., & Kamperman, A. (2024). Using Wearable Technology to Detect, Monitor, and Predict Major Depressive Disorder—A Scoping Review and Introductory Text for Clinical Professionals. *Algorithms*, 17(9), 408. <https://doi.org/10.3390/a17090408>
- Wang, W., Li, X., Qiu, X., Zhang, X., Brusic, V., & Zhao, J. (2023). A privacy preserving framework for federated learning in smart healthcare systems. *Information Processing & Management*, 60(1), 103167. <https://doi.org/10.1016/j.ipm.2022.103167>

Operating an automated SAP Helpdesk: Running a BERT-based Support Ticket Classifier on IBM Power10 Hardware

Simon Fuchs, Vinzenz Senftl, Holger Wittges, Helmut Krcmar
TUM School of Computation, Information and Technology, Garching, Germany

Abstract: The growing demand for Natural Language Processing (NLP) technologies combined with the growing data privacy requirement has flagged the critical role of hardware systems for hosting these NLP models. IBM's current Power10 processor architecture offers the MatrixMath Accelerator (MMA) hardware feature as a viable alternative to traditional GPU-based systems for specific Machine Learning (ML) workloads. Testing the use case of support ticket classification with BERT as an NLP model, this paper investigates the Power10 CPU architecture's performance across various metrics, including accuracy, training time, latency, resource utilization, and cost-effectiveness. We compare the performance of our BERT-based Support Ticket Classifier on Power10 CPUs with MMA to its performance on classical GPU Servers. Our results indicate that while the Power10 architecture excels in inference tasks and energy efficiency, GPU-based systems have a significant performance advantage in training BERT. We conclude that the Power10 offers a practical solution for less time-sensitive training scenarios and environments with idle CPU resources.

Keywords: BERT, Machine Learning, Support Ticket Classification, Natural Language Processing, Power10

1. Introduction

The rapid growth of deep learning (DL), accelerating the push towards business-ready artificial intelligence (AI) solutions, has created opportunities and challenges for organizations across various sectors (Baier et al. 2021; Subbarao et al. 2023; Raschka/Mirjalili 2019; Nielsen 2015). With the advent of Large Language Model (LLM) chatbots like ChatGPT, the DL transformer model architecture has become the new status quo in Natural Language Processing (NLP) (Sun et al. 2022; Zhao et al. 2023; Euronews 2023). However, this trend is also challenging due to the enormous rise in computing resource demand (Preuveneers et al. 2020; Yao/Xiao 2024), requirement for specialized hardware (Tulbure et al. 2024), and the privacy concerns connected to large LLMs running on cloud servers outside of Europe (Longpre et al. 2024; Kaczmarczyk 2024).

Against this context, IBM has developed the MatrixMath Accelerator (MMA) hardware feature integrated into IBM's Power10 processor architecture. MMA is designed to accelerate complex matrix computations, which are fundamental to DL workload, especially transformers (Venkat 2022). MMA enables high-speed matrix operations without offloading computational load to external GPUs or FPGAs. This allows for more efficient and faster processing of dense linear algebra tasks, such as matrix multiplications and convolutions (Venkat 2022).

With MMA, normal IBM Power10 Servers might become a valid asset for AI hosting across various industries (Carvalho et al. 2022), providing additional on-premise resources for on-premise AI workload, while already being present in the computing centers of various companies (TheirStack 2025). MMA might become an especially valuable asset for companies already using IBM Power servers, allowing idle CPU workload for AI workload without additional acquisition costs for dedicated AI server hardware.

Although there are multiple publications about the theoretical potential and inner machinations of the MMA hardware feature (Kuzma et al. 2023; Ferrari/Araujo 2024; Tukanov et al. 2022), we currently detect a lack of real-world evaluation of the Power10’s MMA hardware feature, measuring its performance under real-world business conditions and comparing its performance to GPU-based workload. For this reason, we decided to evaluate Power10’s MMA hardware feature in the real-world NLP use case of support ticket classification in the SAP UCC Munich’s support ticket system – a well-researched and well-understood NLP use case due to multiple previous studies (Fuchs et al. 2022a; Fuchs et al. 2022b; Fuchs et al. 2023; Fuchs et al. 2024; Fuchs et al. 2025). With this, we are in the distinct position to evaluate the Power10’s MMA hardware features capabilities under real-world conditions with a well-understood, real-world business NLP use case.

Accordingly, in this paper, we first roughly describe the classification problem. Afterward, we describe the hardware setups to compare in this paper. Then, we describe our NLP evaluation results from training and inference, and afterward discuss our findings.

2. NLP Use Case of Support Ticket Classification & Classifier Artifact

As described, we chose the support ticket automation project of the UCC Munich as a test use case for our evaluation of IBM Power10’s real-world NLP capabilities. We chose this use case due to our extensive knowledge and insight. This section roughly describes the use case and its underlying NLP classification problem.

The UCC Munich is a small company in Germany that operates as an education-as-a-service provider of SAP software products. Its core business is to host, provide, maintain, and support curricula for different SAP software and dedicated SAP systems, mainly SAP S/4HANA Application Server Systems. These curricula and SAP systems are provided to partners of the SAP University Alliance program (SAP 2024): universities, universities of applied sciences, schools, and so-called German “Berufsschulen.”

In the helpdesk, assistant workers work as part-time first-level support agents processing the more straightforward first-level requests according to extensive documentation in a Wiki. The full-time employees work as Product Owners, organized in three departments, besides regular assignments. They answer the more complex second-level requests that often involve deep product knowledge or need special authorization.

The goal of the automation project here is to replace the human support agents who manually distribute the support tickets to the respective support workers with three ML models. This support ticket classification pipeline is depicted in Figure 1.

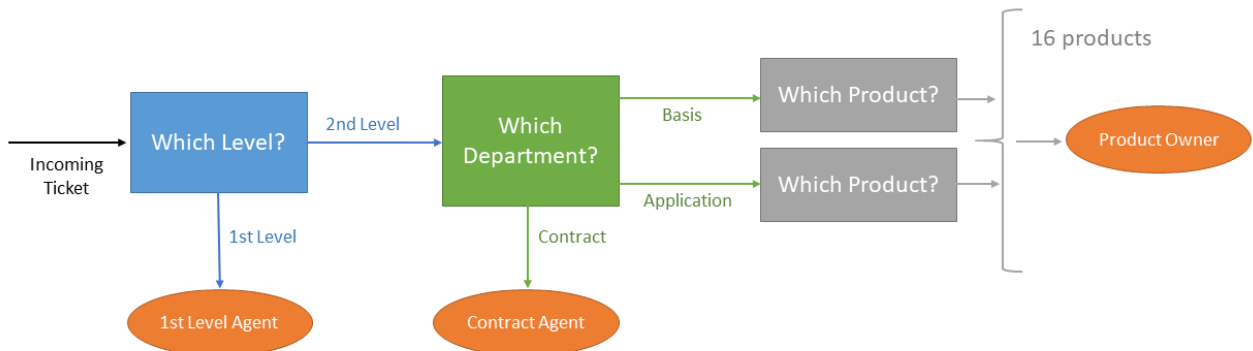


Figure 1: Support Ticket Classification Problem at the UCC (Fuchs et al. 2022b).

When a support ticket is created, it must first be decided whether it contains a first-level or second-level problem. Since every first-level agent should be able to solve all first-level requests, first-level tickets are directly distributed to the next free first-level agent. As mentioned above, the second level of support is separated into three departments: The contract department consists of only one person and is therefore not further subdivided. The department's technical and application staff (mainly responsible for requesting educational services) consists of multiple product owners, each expert in their specific product/service. Hence, a ticket classified as technical or application staff must be assigned a product label selected from sixteen distinct product classes.

To solve this classification problem, we developed multiple potent support ticket classification artifacts (Fuchs et al. 2022b; Fuchs et al. 2025) across multiple design iterations (Hevner et al. 2004), from which a context-aware BERT-based classifier (Devlin et al. 2018) showed the best classification accuracy in Level and Department classification. Further, it is the second best classifier in Product classification (Fuchs 2025). Accordingly, we chose this artifact for our evaluation use case. BERT's contextual embeddings make it particularly suitable for NLP classification tasks. The well-established model architecture, pre-trained embeddings, and fine-tuning approach allow the focus to remain on comparing hardware performance rather than model development (Liu et al. 2023).

3. Research Design

We obtained, filtered, and preprocessed about 7,500 support tickets for our experiments. We removed missing entries and balanced the dataset by combining downsampling and upsampling techniques (Yap et al. 2014). Each ticket includes categories, subcategories, and descriptive text. The preprocessing ensures that all categories are adequately represented for robust model evaluation.

As mentioned above, we compare ML training and inference of the (same) BERT-based support ticket classifier on two IBM Power-based hardware setups: (1) a Podman Container located on an IBM AC922³ with 20 vCPUs allocated, 128 GB RAM, and one NVIDIA Tesla V100 GPU, equipped with 32 GB of GPU memory, 5120 CUDA cores, and 640 Tensor Cores; and (2) an

³ <https://www.redbooks.ibm.com/redpapers/pdfs/redp5494.pdf>

LPAR located on an IBM E1050⁴ with 20 vCPUs allocated with MMA and 128 GB RAM allocated. The hardware specifications of both LPARs are listed in Table 1.

Feature	Power 10 (E1050 LPAR)	Power 9 + V100 GPU
vCPU	20	20
RAM	128 GB	128 GB
GPU	None (MMA acceleration on CPU)	NVIDIA Tesla V100 (32 GB GPU memory)
OS	RHEL 9.2 (ppcle)	Ubuntu 18.04. (ppcle)
CUDA Support	N/A (No GPU)	CUDA 10.2, cuDNN 8
Virtualization	LPAR on IBM pHyp, uncapped mode	Podman Container

Table 1: Hardware features of the two set-ups to compare in this study.

As far as possible, we tried to use the same libraries and packages for both setups. We trained and inferred on both setups with the same ten random 10%/10%/80% train/inference/test sets to ensure optimal comparability.

4. Results

In this section, we present the evaluation results of both hardware setups using the evaluation metrics presented in the previous section.

4.1. Model Accuracy

The classification accuracies of both setups for all three steps of the UCC use case are presented in Figure 2. The average values across all ten train/inference/test splits.

⁴ <https://www.redbooks.ibm.com/redpapers/pdfs/redp5684.pdf>

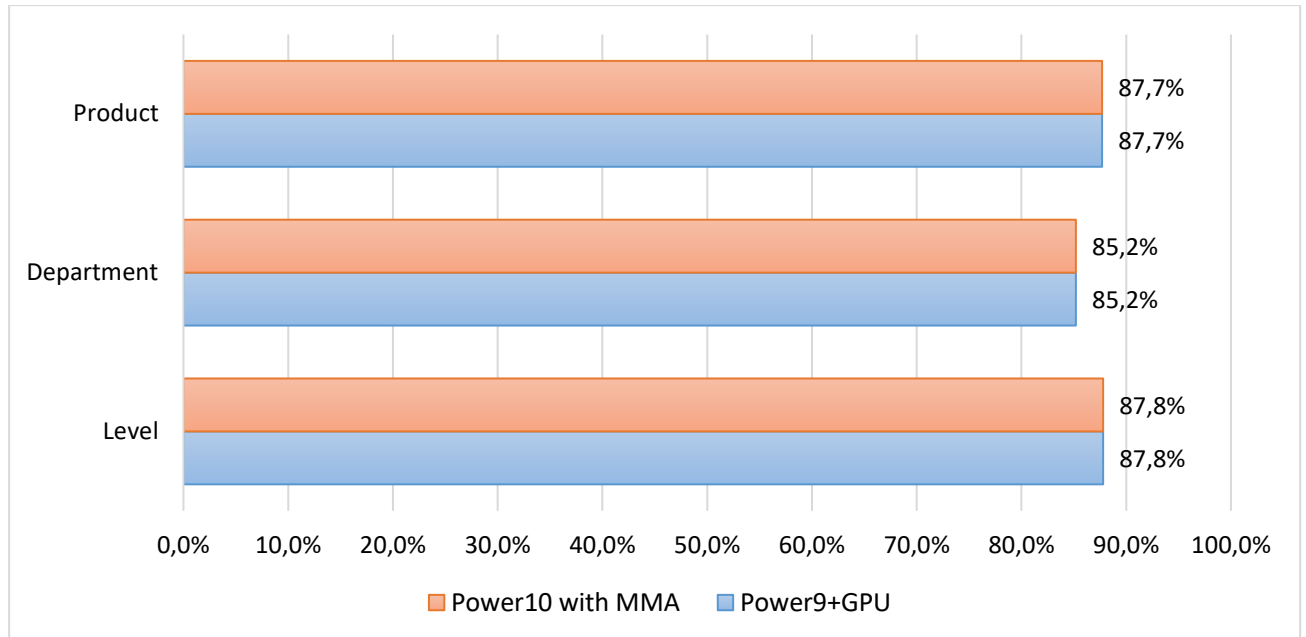


Figure 2: Support Ticket Classification Accuracy of the BERT-based Classifier run on the Power9+GPU and the Power10 with MMA Hardware Set-ups

The results were the same for both hardware setups, indicating that the underlying hardware does not impact the model's accuracy. These results suggest that the models' accuracy primarily depends on the algorithms and datasets used rather than the computational resources provided by the hardware, which is undoubtedly good news.

4.2. Model Training Time

From the hardware-related literature, we already expected that model training on CPUs would be far greater than on GPUs (Hayashi et al. 2015; Mittal et al. 2021; Yu 2024). Still, we were interested in determining the effect of MMA on the training time. For this reason, we trained the BERT classifier on the Power9 container without using the GPU and forcing the operating system to only use the Power9 CPUs. That way, we see the difference between Power9 (no MMA) and Power10 with MMA. The results of these evaluations are depicted in Figure 3. As depicted, the training time on the GPU is by far the smallest. However, the MMA feature significantly reduces the training time compared to the Power9 CPU-only setup.

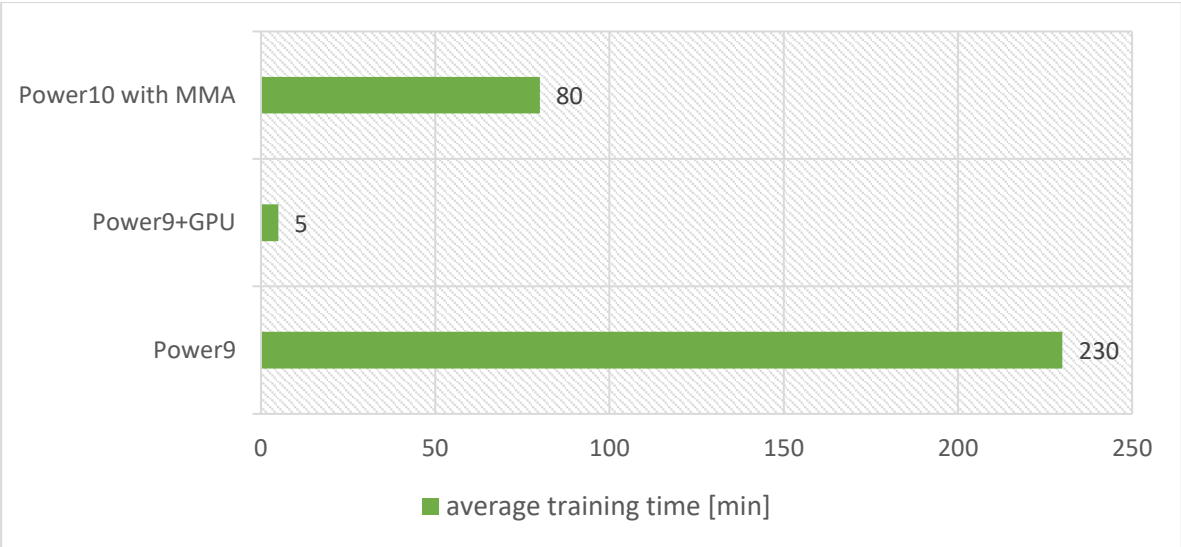
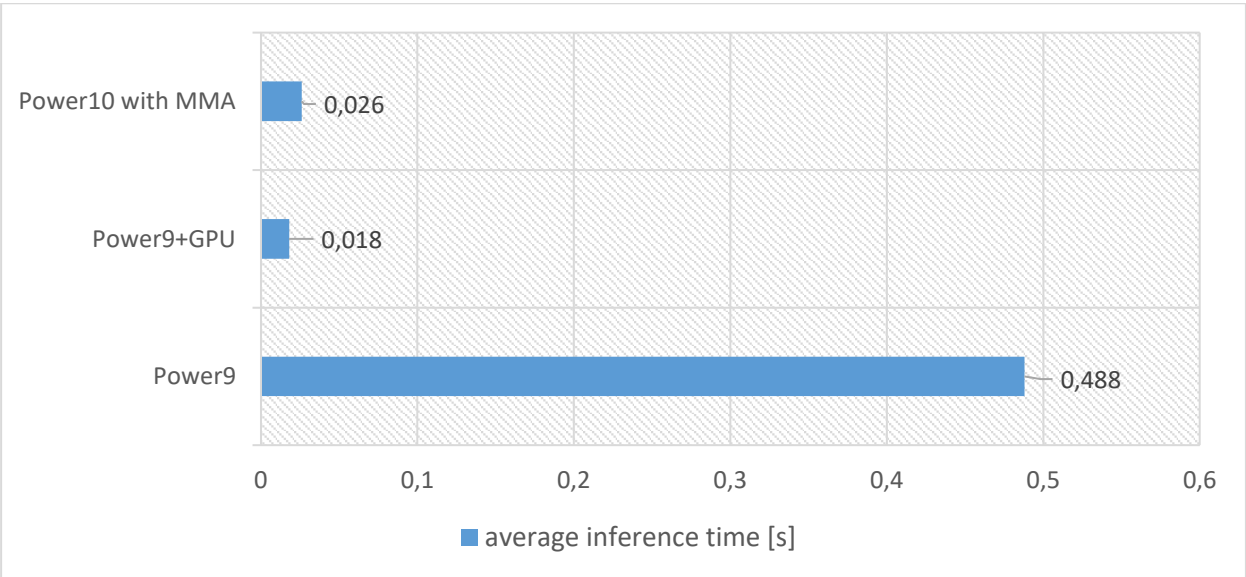


Figure 3: Average Model Training Time of the BERT-based Classifier run on Power9 CPUs only, Power9 CPUs+GPU, and the Power10 CPUs with MMA

4.3. Model Inference Time

Again, we were interested in determining the effect of MMA on the inference time. Accordingly, we evaluated a Power9 CPU only set up to compare the results to Power10 with MMA. The results of these evaluations are depicted in Figure 4.



As depicted, Power10+MMAs is nearly as fast in inference as the GPU. As expected, Power9 CPU-only is significantly slower in inference. These results demonstrate that Power10 is competitive in real-time application scenarios, where faster inference directly contributes to improved system responsiveness and the volume of requests a model can process.

4.4. Resource Utilization

The CPU and GPU utilization rates for Power9 and Power10 highlight each system's contrasting strategies for workload management. During training and inference, the Power9 system relied

heavily on its GPU, while the Power10 system used its CPU much more. The utilization rates were as follows:

- Power10 CPU Utilization: 72%
- Power9 CPU Utilization: 2.7% while Power9 GPU Utilization: 83%

The reliance on CPU resources in the Power10 system meant that it was tasked with handling all significant computational aspects of training, which led to a greater strain on its CPU cores. The Power9 system, with its dedicated GPU, was able to offload these tasks, leading to much lower CPU activity and more efficient processing. Figure 15 provides insights into the CPU utilization during a training run on the Power10 system. This visualization illustrates the stark contrast between the CPU workloads on Power10 and Power9, where the Power10 system ran close to its capacity while Power9 maintained minimal CPU involvement.

5. Discussion

Our results revealed key trade-offs between the IBM Power9 and Power10 systems. Both achieved similar accuracy, indicating that model performance depends more on architecture and data than hardware. This result validates the primary goal of ensuring that slight hardware differences do not compromise computational integrity, a foundational requirement for deployment in critical industries. Training times varied significantly, while Power10 with MMA showed competitive inference times. This considerable contrast highlights GPU acceleration's critical role in efficiently training deep learning models. The Power9 system leveraged a dedicated GPU and utilized the entire instance for parallel processing, resulting in much faster training times. In contrast, the Power10 system relied solely on its CPU within an LPAR virtualized environment, which may have constrained its computational performance.

The findings provide actionable insights for determining the appropriate hardware for different ML workloads. The decision depends on model complexity and time sensitivity- Deep learning models with extensive data requirements, such as transformer-based architectures, benefit from GPUs optimized for parallel computations. Power10 offers sufficient capabilities for simpler or smaller models. Training times can vary significantly based on hardware. Power10 is suitable for environments where time is less critical, such as weekend batch training in enterprise systems with otherwise idle CPU resources. Power10 excels at inference-based tasks, a key performance area highlighted by IBM in its marketing efforts. Industries such as insurance and banking, where real-time processing is crucial (e.g., for fraud detection), stand to benefit from Power10's efficiency. These use cases align with IBM's customer base and validate the processor's suitability for operational environments requiring low latency and high throughput.

One of the original motivations for this study was to address CPU underutilization in enterprise environments. The results suggest that Power10 could effectively leverage idle CPU resources for ML tasks, particularly in non-peak periods (e.g., weekends). For example, training smaller models over two idle days could significantly improve system efficiency without incurring additional hardware costs. This capability positions Power10 as a cost-effective solution for enterprises that want to maximize their existing infrastructure without investing heavily upfront in external GPUs.

The study underscores the evolving role of CPUs in ML. While GPUs dominate high-performance training, CPUs with accelerators like MMA can carve out a niche in less demanding yet cost-sensitive applications. However, Power10's broader adoption will hinge on

software optimizations that allow more profound integration of its MMA capabilities into popular ML frameworks. The findings of this study are particularly significant for the UCC, offering valuable insights for optimizing the utilization of its IBM Power10 systems. These results provide a roadmap for integrating ML tasks within existing infrastructures and highlight opportunities to enhance students' educational experience while minimizing operational costs. The tested use case, support ticket classification, demonstrates how Power10 systems can effectively handle specific ML tasks without requiring external GPUs. This workload is an ideal candidate for Power10 due to its non-time-critical nature, allowing for resource allocation flexibility.

6. Conclusion

This paper investigates the IBM Power10 processor's Matrix-Math Accelerator (MMA) and its role in enhancing machine learning (ML) workloads, with a particular focus on support ticket classification. The MMA is designed to accelerate matrix operations, a core component of ML, through specialized instructions that deliver up to 20 times higher throughput compared to CPU-only execution. Beyond performance, a major advantage of the MMA is its energy efficiency. Its in-core design reduces data movement, cutting both latency and power consumption. Benchmarks show that the MMA achieves energy-efficient processing with results comparable to GPU-based systems, reaching near-parity with NVIDIA V100 GPUs in inference tasks. This demonstrates its potential as a cost-effective and power-conscious alternative for enterprises that would otherwise rely heavily on specialized GPUs. The MMA also supports scalability and seamless infrastructure integration. Leveraging the Power10's high-bandwidth memory, it facilitates smooth data handling and efficient operation in virtualized environments, making it suitable for large-scale and flexible AI deployments. For enterprises in sectors like healthcare, finance, and customer service, the MMA provides practical benefits. Its efficiency with small- to mid-sized models matches industry needs for rapid decision-making and low-latency operations.

7. Acknowledgments & Disclaimers

We want to thank the IBM managers Alexander Loibl and Thomas Heinlein, who significantly provided the UCC's Power10 servers. Further, we want to thank the IBM employees, Felix Eberhardt, Michael Perzl, and Stephen Lutz, who helped set up the server landscape and contributed variously with their noticeable expertise.

This paper is primarily based on the unpublished master's thesis of Vinzenz Senftl, the second author of this paper. We used Grammarly to correct our Spelling, Style, and Grammar. No further Generative AI was used to write this paper's text.

8. References

- Baier, L.; Kühl, N.; Schüritz, R.; Satzger, G. (2021): Will the customers be happy? Identifying unsatisfied customers from service encounter data. In: *Journal of Service Management*, Vol. 32 (2021) No. 2, pp. 265-288.

- Carvalho, J.P.L.d.; Moreira, J.E.; Amaral, J.N. (2022): Compiling for the IBM Matrix Engine for Enterprise Workloads. In: IEEE Micro, Vol. 42 (2022) No. 5, pp. 34-40.
- Devlin, J.; Chang, M.-W.; Lee, K.; Toutanova, K. (2018): Bert: Pre-training of deep bidirectional transformers for language understanding. In: arXiv preprint arXiv:1810.04805, (2018).
- Euronews (2023): ChatGPT a year on: 3 ways the AI chatbot has completely changed the world in 12 months <https://www.euronews.com/next/2023/11/30/chatgpt-a-year-on-3-ways-the-ai-chatbot-has-completely-changed-the-world-in-12-months>, accessed at 30.12.2024.2024.
- Ferrari, V.; Araujo, G. (2024): Improving Direct Convolution through Tensor Slicing, Vectorized Packing and ISA Extensions. Paper presented at the Concurso de Teses e Dissertações (CTD), pp. 148-157.
- Fuchs, S. (2025): Dissertation Manuscript - Context Data-driven Automation of an SAP Support Ticket System using Machine Learning on a Small Training Data Set (V1). In TUM (Ed.), (pp. 240). in publication.
- Fuchs, S.; Böllhoff, J.; Andraichuk, V.; Wittges, H.; Krcmar, H. (2023): Towards an automated SAP service desk - Alignment of technical projects with a business perspective. SAP Academic Community Conference 2023 (pp. 72-85). Munich: Technical University Munich (TUM).
- Fuchs, S.; Drieschner, C.; Wittges, H. (2022a): Improving Support Ticket Systems Using Machine Learning: A Literature Review. Hawaii International Conference on System Sciences 2022 (pp. 10).
- Fuchs, S.; Schnellbach, J.; Schmidt, L.; Wittges, H. (2024): Data Annotation for Support Ticket Data: A Literature Review. 57th Hawaii International Conference on System Sciences 2024 (pp. 10). Oahu, Hawaii.
- Fuchs, S.; Wiehl, N.; Wittges, H.; Krcmar, H. (2022b): Towards an Automated SAP Service Desk -Design and Implementation of a Prototype Machine Learning Classifier for Support Tickets on a Small Data Set. Conference: SAP Academic Community Conference 2022 DACH. Magdeburg, Germany.
- Fuchs, S.; Wiehl, N.; Wittges, H.; Krcmar, H. (2025): Customizing Context: Discovering the Optimal Integration of Context Data to Elevate ML-Driven Automated Support Ticket Classification. 58th Hawaii International Conference on System Sciences (HICSS). Kona, Hawaii: scholarspace.manoa.hawaii.edu.
- Hayashi, A.; Ishizaki, K.; Koblenz, G.; Sarkar, V. (2015): Machine-learning-based performance heuristics for runtime cpu/gpu selection. In: Proceedings of the principles and practices of programming on the Java platform. Eds. 2015, pp. 27-36.
- Hevner, A.R.; March, S.T.; Park, J.; Ram, S. (2004): Design science in information systems research. In: MIS Q., Vol. 28 (2004) No. 1, pp. 75–105.
- Kaczmarczyk, B. (2024): Is OpenAI Safe? - A Practical Look at OpenAI Data Security. <https://www.softkraft.co/openai-data-security/#conclusion---is-openai-safe>, accessed at 02.03.2025.2025.
- Kuzma, B.; Korostelev, I.; De Carvalho, J.P.; Moreira, J.E.; Barton, C.; Araujo, G.; Amaral, J.N. (2023): Fast matrix multiplication via compiler-only layered data reorganization and

- intrinsic lowering. In: *Software: Practice and Experience*, Vol. 53 (2023) No. 9, pp. 1793-1814.
- Liu, Z.; Benghe, C.; Jiang, S. (2023): Ticket-bert: Labeling incident management tickets with language models. In: *arXiv preprint arXiv:2307.00108*, (2023).
- Longpre, S.; Mahari, R.; Chen, A.; Obeng-Marnu, N.; Sileo, D.; Brannon, W.; Muennighoff, N.; Khazam, N.; Kabbara, J.; Perisetla, K.; Wu, X.; Shippole, E.; Bollacker, K.; Wu, T.; Villa, L.; Pentland, S.; Hooker, S. (2024): A large-scale audit of dataset licensing and attribution in AI. In: *Nature Machine Intelligence*, Vol. 6 (2024) No. 8, pp. 975-987.
- Mittal, S.; Rajput, P.; Subramoney, S. (2021): A survey of deep learning on CPUs: opportunities and co-optimizations. In: *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 33 (2021) No. 10, pp. 5095-5115.
- Nielsen, A. (2015): *Neural networks and deep learning*. Determination Press.
- Preuveneers, D.; Tsingenopoulos, I.; Joosen, W. (2020): Resource Usage and Performance Trade-offs for Machine Learning Models in Smart Environments. In: *Sensors*, Vol. 20 (2020) No. 4, pp. 1176.
- Raschka, S.; Mirjalili, V. (2019): *Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2*. (3 ed.), Packt Publishing Ltd. 2019.
- SAP (2024): Education and research with SAP Next-Gen.
<https://www.sap.com/about/company/innovation/next-gen-innovation-platform/university-alliances.html>, accessed at 24.12.2024.2024.
- Subbarao, M.; Venkatarao, K.; Chittineni, S.; Kompella, S. (2023): Utilizing deep learning, feature ranking, and selection strategies to classify diverse information technology ticketing data effectively. In: *IAES International Journal of Artificial Intelligence (IJ-AI)*, Vol. 12 (2023), pp. 1985.
- Sun, T.-X.; Liu, X.-Y.; Qiu, X.-P.; Huang, X.-J. (2022): Paradigm Shift in Natural Language Processing. In: *Machine Intelligence Research*, Vol. 19 (2022) No. 3, pp. 169-183.
- TheirStack (2025): Companies that use IBM POWER Systems.
<https://theirstack.com/en/technology/ibm-power-systems>, accessed at 01.07.2025.2025.
- Tukanov, N.; Srinivasaraghavan, R.; Moreira, J.E.; Low, T.M. (2022): Modeling matrix engines for portability and performance. Paper presented at the 2022 IEEE International Parallel and Distributed Processing Symposium (IPDPS), pp. 1173-1183.
- Tulbure, A.; Ioan, S.; Danciu, D.; Adrian Alexandru, T.; Dulf, E. (2024): Training cost analysis on a large computer vision defect detection model for ceramics 2024.
- Venkat, S. (2022): MMA in IBM Power10 processor. IBM. In:
<https://community.ibm.com/community/user/blogs/sridhar-venkat1/2022/08/29/power10-mma>, accessed at 01.07.2025.2025.
- Yao, Y.; Xiao, Y. (2024): Trade-Off between Performance and Power Consumption in Deep Learning Training. Paper presented at the 2024 5th International Symposium on Computer Engineering and Intelligent Communications (ISCEIC), pp. 104-109.
- Yap, B.W.; Rani, K.A.; Rahman, H.A.A.; Fong, S.; Khairudin, Z.; Abdullah, N.N. (2014): An application of oversampling, undersampling, bagging and boosting in handling

imbalanced datasets. Paper presented at the Proceedings of the first international conference on advanced data and information engineering (DaEng-2013), pp. 13-22.

Yu, R. (2024): Application of CPU in AI and Machine Learning. Paper presented at the Proceedings of the International Conference on Decision Science & Management, pp. 221-224.

Zhao, W.X.; Zhou, K.; Li, J.; Tang, T.; Wang, X.; Hou, Y.; Min, Y.; Zhang, B.; Zhang, J.; Dong, Z. (2023): A survey of large language models. In: arXiv preprint arXiv:2303.18223, (2023).

Visuelle Gestaltung eines GenAI-Assistenzsystems für die SAP Build-Umgebung: Ein Design-Science-Research-Ansatz

E. E. D. Zuferi¹, K. Bauer¹, M. Fischer¹ & V. Micelli-Schmidt²

¹Hochschule Heilbronn, Deutschland

²SAP SE

Abstract: Die Entwicklung künstlicher Intelligenz (KI)-gestützter Assistenten stellt eine bedeutende Innovation im Bereich der Softwareentwicklung dar. Universelle KI-Assistenten stoßen jedoch insbesondere bei domänenspezifischen Anfragen häufig an ihre Grenzen und liefern nicht selten ungenaue oder wenig hilfreiche Antworten. Genau mit diesen Herausforderungen sieht sich auch das Unternehmen SAP konfrontiert. Ziel ist es, durch spezialisierte, kontextuell angepasste KI-Lösungen sowohl die Mitarbeitenden zu entlasten als auch die Effizienz in der Entwicklung innerhalb der SAP-Build-Umgebung zu steigern. Benötigt wird hierfür ein kontextualisierter generativer Artificial Intelligence-Chatbot (GenAI-Chatbot), welcher die Nutzer in einem speziellen Teilbereich unterstützt, bei dem der selbstentwickelte Assistent von SAP namens Joule nicht entsprechend funktional ist. Eine erhebliche Zeiteffizienz durch Zentralisierung der Informationen bei Interaktion mit der GenAI, sowie die Beseitigung technischer Probleme in der Entwicklung wird realisiert. Mithilfe von Designprinzipien aus der Literatur wird der Prototyp auf Basis von Use Cases im Design- und Prototyping-Tool Figma umgesetzt. Die Use Cases sind "Backend developing", „Frontend developing", "General links or informations", Explaining an existing application". Diese Use Cases werden in einer Reihe von Experteninterviews ermittelt, aus welcher der erste Prototyp erfolgt. Durch die anschließende Evaluierungsphase ergibt sich die finale Prototyp Version. Mit einem abschließenden Fazit und Limitation ist die Basis für zukünftige Forschungen gegeben.

Schlüsselwörter: SAP Build, Developer Experience, GenAI, Softwareentwicklung, Chatbot

1. Einführung

Die Entwicklung künstlicher Intelligenz-Assistenten (KI-Assistenten) stellt eine signifikante Innovation im Bereich der Softwareentwicklung dar (Pinto et al., 2024, S. 81). Ihre Relevanz ergibt sich aus der Komplexität und dem zeitlichen Aufwand softwarebezogener Prozesse, die eine effektive Entlastung der Entwickler erfordern (Dehaerne et al., 2022, S. 82434). Während regelbasierte Chatbots lediglich auf vorgefertigten Skripten basieren, nutzen moderne KI-Assistenten Technologien wie natürliche Sprachverarbeitung und maschinelles Lernen zur Generierung kontextsensitiver Antworten (Kortas, 2024).

Trotz des technischen Fortschritts bestehen bei universellen KI-Assistenten wesentliche Limitationen hinsichtlich der Genauigkeit, insbesondere bei domänenspezifischen Fragestellungen (Pinto et al., 2024, S. 81). Hinzu kommen Herausforderungen im Design solcher Systeme, da die offene Sprache der Nutzer differenzierte Anforderungen an die Interaktionsgestaltung stellt (Weisz et al., 2024, S. 4). Auch die Optimierung der Mensch-KI-Kommunikation ist entscheidend, um eine hohe Effektivität im Arbeitsprozess zu gewährleisten (Lestarini & Raflesia, 2024, S. 1). Gerade in der SAP Build-Umgebung zeigt sich ein gesteigertes Bedürfnis nach strukturierten Assistenzsystemen, die nicht nur Wissenslücken im

Bereich der SAP S/4HANA Erweiterung schließen, sondern auch Prozessabbrüche durch externe Recherche vermeiden (Brandt et al., 2009, S. 1589). Vor diesem Hintergrund entsteht die Notwendigkeit, einen kontextualisierten GenAI-Chatbot zu entwickeln, der speziell auf die Anforderungen von Entwicklern in der SAP Build-Umgebung zugeschnitten ist (Pinto et al., 2024, S. 81). Die Integration eines solchen Systems verspricht eine erhebliche Zeiteffizienz durch zentrale Informationsbereitstellung, erweiterte Dokumentationsverweise sowie konkrete Hilfestellungen bei technischen Problemen (Bahrman, 2023; Pinto et al., 2024, S. 82). Gleichzeitig wird durch eine Verbesserung der Developer Experience (DX) ein ganzheitlicher Nutzen erzeugt, der sowohl auf individueller als auch organisatorischer Ebene wirkt (Aslina & Nugraha, I Gusti Bagus Baskara, 2024, S. 144; Mahendra, 2023, S. 218; Tiedemann & Bach, 2021). Die Schaffung einer positiven DX unterstützt die Entwickler bei einer effizienten und effektiven Arbeit, welche der gesamten Organisation zugutekommt (Augsten, 2023). Die vorliegende Forschungsarbeit verfolgt das Ziel, einen generativen GenAI-Chatbot-Prototyp für die SAP Build-Plattform zu konzipieren. Die grafische Umsetzung erfolgt mittels des Design- und Prototyping-Tools Figma. Auf Basis qualitativer Experteninterviews werden relevante Use Cases erhoben, die exemplarisch durch Konversationsszenarien im Prototyp abgebildet werden. Die Gestaltung berücksichtigt dabei die sechs Designprinzipien für GenAI-Anwendungen nach Weisz et al. (2024). Daraus ergibt sich die folgende Forschungsfrage (F) mitsamt Unterfragen (UF):

F1: “Wie lässt sich ein GenAI-Chatbot-Assistent in der Entwicklungsumgebung von SAP Build grafisch implementieren?”

UF1: Welche Use Cases der Entwickler sollen vom GenAI-Chatbot abgedeckt werden?

UF2: Wie sollte die Design-Lösung in der SAP Build Umgebung konkret ausgestaltet sein?

Zur Beantwortung dieser Fragestellungen wird ein methodisches Vorgehen gewählt, das sowohl theoriegeleitete Grundlagen als auch empirisch erhobene Anforderungen aus der Praxis systematisch miteinander verknüpft.

2. Theoretischer Hintergrund: GenAI und SAP Build

Generative KI-Modelle haben das Potenzial, die Softwareentwicklung grundlegend zu verändern. Im Unterschied zu regelbasierten Systemen generieren GenAI-Modelle kontextabhängige Inhalte, etwa Code, Tests oder Erklärungen, basierend auf natürlicher Spracheingabe (Donvir et al., 2024, S. 1; Rumondor et al., 2024, S. 43). Damit begleiten sie Entwickler durch verschiedene Phasen des Softwareentwicklungsprozesses und tragen zur Reduktion manueller Aufwände sowie zur Beschleunigung iterativer Zyklen bei (Raghi et al., 2024, S. 1). Gleichzeitig ist der Einsatz solcher Systeme nicht frei von Herausforderungen – insbesondere hinsichtlich Genauigkeit, Zuverlässigkeit und Benutzerinteraktion (Lestarini & Raflesia, 2024, S. 1; Pinto et al., 2024, S. 81). Durch die Betrachtung der Developer Experience (DX) wird nach Aslina & Nugraha (2024, S. 144) ein ganzheitliches Betrachtungsmaß geschaffen, wie Entwickler bei ihrer Arbeit mit Tools, Prozessen und der Umgebung interagieren. Nach Tiedemann & Bach (2021) werden unter der DX alle Aspekte der Interaktion von Entwicklern eines Unternehmens mit dessen Tools und Systemen umfasst. Eine positive DX kann nicht nur Produktivität und Qualität steigern, sondern auch die langfristige Akzeptanz neuer Werkzeuge fördern (Mahendra, 2023, S. 218). Als Plattform für die Umsetzung des GenAI-Chatbots dient SAP Build – ein modulares Entwicklungstool zur Erstellung von Applikationen, automatisierten Prozessen und digitalen Workspaces im SAP-Kontext. Die

Plattform unterstützt unterschiedliche technologische Pfade, unter anderem ABAP-basierte on-stack-Entwicklungen und JavaScript-orientierte side-by-side-Architekturen mit Low-Code und Pro-Code Entwicklungsansätzen (Sourceforge, 2025).

3. Methodologie

Zur wissenschaftlichen Strukturierung des Forschungsprozesses wird das Design Science Research Methodology (DSRM) nach Peffers et al. (2007) herangezogen. Ziel ist die systematische Entwicklung und Evaluation eines GenAI-Chatbot-Prototyps für SAP Build, der an konkreten Anforderungen aus der Praxis ausgerichtet ist. Ergänzend wird eine strukturierte Literaturarbeit nach Webster und Watson (2002) sowie die gestaltungsorientierte Forschungssystematik nach Fettke (2006) verwendet. In Phase eins wird das Problem der begrenzten Developer Experience in SAP Build identifiziert. Phase zwei konkretisierte die Zielsetzung, einen kontextsensitiven Chatbot zur Unterstützung von Entwicklern zu entwerfen. Die Gewinnung der Use-Cases in Phase drei erfolgt durch qualitative Experteninterviews mit Entwicklern. Grundlage bildet ein semi-strukturierter Leitfaden gemäß der qualitativen Interviewmethodik nach Kaiser (2021). Die Auswertung erfolgt thematisch anhand inhaltlicher Cluster, etwa zur Informationssuche, Tool-Nutzung und erwarteten Funktionalitäten. Die daraus abgeleiteten Use-Cases fließt in die Gestaltung des Prototyps ein, wobei die sechs Designprinzipien (DP) nach Weisz et al. (2024) als methodischer Rahmen dienen. So werden z. B. „Generative Variability“ (1) durch Mehrfachantworten, „Trust & Reliance“ (2) durch Quellennachweise und „Co-Creation“ (3) durch anpassbare Eingaben im Freitextfeld umgesetzt. Weitere Prinzipien wie „Imperfection“ (4), „Mental Models“ (5) und „Responsibility“ (6) werden durch Rückmeldemechanismen (Daumen hoch bzw. runter), UI-Hinweise zur Datenverarbeitung und Use-Case-Ableitung berücksichtigt. In den Abbildungen eins-drei sind diese zu sehen. In Phase vier wird der Prototyp in Figma visualisiert und in Phase fünf mit den ursprünglichen Experten evaluiert. Die Rückmeldungen werden qualitativ analysiert – mit Fokus auf wahrgenommene Mehrwerte, Kritikpunkte und Erweiterungspotenziale. Phase sechs umfasst die wissenschaftliche Kommunikation der Ergebnisse im Rahmen dieses Beitrags, mit dem Ziel, die Erkenntnisse zur Gestaltung GenAI-gestützter Assistenzsysteme im SAP-Kontext für Forschung und Praxis zugänglich zu machen.

4. Ergebnisse

Im Folgenden werden die Ergebnisse, welche in die Konzeption des Chatbots münden, veranschaulicht.

4.1. Erhebung nutzerzentrierter Use-Cases

Im Rahmen der ersten Interviewphase wurden qualitative Expertengespräche mit drei langjährig erfahrenen SAP-Entwicklern und Projektverantwortlichen durchgeführt. Ziel ist es, potenzielle Anwendungsfälle für den Einsatz eines GenAI-Chatbots im SAP Build-Kontext zu identifizieren. Die inhaltsanalytische Auswertung der Interviews ermöglicht eine Strukturierung der Erkenntnisse entlang von vier funktionalen Pfaden, die verschiedene Entwicklungsszenarien abbilden.

Ein zentrales Themenfeld stellt die Unterstützung im Bereich der Backend-Entwicklung (Pfad eins) dar, insbesondere im Umgang mit ABAP und ABAP-Cloud. Alle drei Interviewpartner

heben hervor, dass ABAP-Entwickler durch die Einführung von SAP Fiori zunehmend auf das Backend reduziert wurden, da die UI-Komponenten heute oft durch spezialisierte UX-Developer erstellt werden. Diese Fragmentierung erschwert die End-to-End-Verantwortung von Full-Stack-Entwicklern. Zudem herrscht oft Unsicherheit bei der Auswahl geeigneter Architekturen (z. B. RAP vs. CAP) sowie beim Design des Datenmodells und der Geschäftslogik. Der GenAI-Chatbot soll hier Orientierung bieten, etwa durch Entscheidungshilfen zu Architekturpfaden, Erläuterungen zu Datenflüssen oder Vorlagen für typische CRUD-Funktionen. Auch das automatisierte Generieren von Basisdateien und Templates wird als wünschenswerte Unterstützung genannt.

Im Bereich der Frontend-Entwicklung (Pfad zwei) wird von zwei Experten vor allem der Wunsch nach gestalterischer Unterstützung geäußert. Entwickler benötigen Hilfestellung bei der Auswahl passender SAP Fiori-Templates oder beim Aufbau einer neuen UI-Struktur. Die Experten schlagen vor, dass der Chatbot entweder existierende UI-Templates vorschlagen oder diese generativ neu erstellen kann – angepasst an Zielgruppe, Anwendungsbereich und technische Vorgaben. Auch die Entscheidung, ob in ABAP oder JavaScript entwickelt werden soll, wird als potenzieller Use Case gesehen. Durch gezielte Ja/Nein-Fragen zur geplanten Architektur, zur Nutzerinteraktion und zu projektspezifischen Anforderungen soll der Chatbot eine kontextsensitive Empfehlung abgeben können.

„General links and information“ (Pfad drei) ergibt sich aus dem Bedarf nach aggregierten Informationen. Alle drei Experten beklagen die Verteilung relevanter Inhalte auf zahlreiche externe Quellen wie SAP Discovery Center, GitHub oder YouTube. Ein strukturierter Zugang über den Chatbot zu verifizierten Tutorials, Prototypen, Architekturleitlinien und Best Practices ist daher als hohes Potenzial identifiziert. Ziel ist es, den Chatbot als kuratierten Einstiegspunkt zu nutzen, der zielgerichtet auf relevante Ressourcen verweist – idealerweise direkt verlinkt und thematisch einordnet. Auch für neue Entwickler oder Quereinsteiger kann dies die Einarbeitung erheblich erleichtern.

“Explaining an existing application” (Pfad vier) stellt neben der Neuentwicklung auch Use Cases im Kontext bestehender Anwendungen dar. In der Praxis tritt häufig der Fall auf, dass eine Applikation nicht neu erstellt, sondern lediglich erweitert oder angepasst werden soll. Der Chatbot soll hier in der Lage sein, auf Basis weniger Angaben (z. B. Anwendungstyp, Fachbereich, gewünschte Erweiterung) bestehende Projekte zu identifizieren, eine Beschreibung des Status quo zu liefern und generative Unterstützung bei der Modifikation anzubieten – etwa in Form einer Schritt-für-Schritt-Anleitung oder durch das Vorschlagen konkreter Änderungsschritte. Idealerweise kann sogar ein geklonter Entwicklungsstand geladen werden, um die Erweiterung unmittelbar vorzunehmen. Die identifizierten Use Cases entlang dieser vier Pfade bilden die inhaltliche Grundlage für die anschließende Konzeption und Gestaltung des Chatbot-Prototyps. Sie verdeutlichen, dass der Bedarf an GenAI-Unterstützung im SAP Build-Kontext nicht nur in der Codegenerierung liegt, sondern insbesondere in der intelligenten Strukturierung, Auswahl und Einbettung von Informationen und Entscheidungsprozessen im Entwicklungsalltag. Die gewonnenen Use Cases werden durch Prototyping im Designtool Figma grafisch aufbereitet. Der Prototyp ist die Basis für die anschließende Evaluierung in Experteninterviews.

4.2. Qualitative Bewertung des GenAI-Chatbots

Zur Evaluation der ersten Version des GenAI-Chatbots wurden fünf Experteninterviews durchgeführt, davon drei mit bereits in Phase eins involvierten Personen sowie zwei mit neuen

Fachexperten. Ziel ist es, strukturiertes Feedback zur Benutzerführung, inhaltlichen Relevanz und funktionalen Gestaltung entlang der vier definierten Pfade zu erheben. Der Chatbot ist als PDF-Export in ein öffentliches Gitlab-Repository abgelegt. Dieser ist hier erreichbar. Alle nachfolgenden Seitenverweise in Kapitel fünf referenzieren auf diesen PDF-Export.

Pfad eins - Backend developing: Der prototypisch modellierte Konversationsfluss im Pfad „Backend developing“ wird von allen fünf Experten überwiegend positiv bewertet. Es wird empfohlen, mit der Interaktionsart (S. 5) zu beginnen, gefolgt von Template (S. 41), Datenmodell-Verhalten (S. 23) und schließlich Programmiersprache (S. 50). Die Abfrage der Eigenschaften wird von drei Experten als relevante Unterstützung empfunden. Positiv ist ebenfalls der Entscheidungspfad „Recommend me one“ (S. 3), der über gezielte Fragen zur Zielgruppe (S. 9), Infrastruktur (S. 17), Aufzeichnungs- und Abwicklungssystem (S. 25) sowie zu Key Usern (S. 34) eine datengestützte Empfehlung zwischen ABAP und JavaScript ermöglicht. In der zweiten Hälfte dieses Pfades folgen weitere Rückfragen – u. a. zu Interaktionsart (S. 16), CRUD-Funktionalität (S. 24), Programmiersprache (S. 33) und Template (S. 36, 61). Großes Lob von allen Experten erhält der ergänzende Pfad „Give me some input“, bei dem der Nutzer erklärende Inhalte zu komplexen Architekturentscheidungen erhält – etwa zu „ABAP (RAP) vs. JavaScript (CAP)“ (S. 51), „On stack vs. side-by-side-Extension“ (S. 31), „User vs. Two-System Interaction“ (S. 15) und „Data Model Behaviour Patterns“ (S. 22). Diese Hilfen werden als wertvoll zur Förderung von Entscheidungssicherheit bewertet. Der Dialog schließt mit der Option, eine passende Development Environment im Hintergrund direkt zu öffnen sowie Feedback zu geben (S. 44, 52, 57, 58, 64, 65, 67, 68). Diese Funktionen werden als konsequente Abrundung des Nutzerpfades von vier Experten positiv aufgenommen.

Pfad zwei - Frontend developing: Im Pfad „Frontend developing“ bietet der Prototyp die Möglichkeit, entweder ein bestehendes Template für die spätere Entwicklung auszuwählen oder mithilfe von SAP Fiori Elements ein eigenes Template von Grund auf zu erstellen (S. 6). Diese Option wird von vier Experten insgesamt als praxisrelevant und realitätsnah bewertet, da sie gängige Einstiegsszenarien in der SAP Build-Frontend-Entwicklung abbildet. Von einem Experten wird hingegen das Wording einzelner Antwortmöglichkeiten kritisch hinterfragt. So wird vorgeschlagen, die Formulierung „Match the skills, start application“ durch die präzisere und technisch passendere Variante „Match skills. start SAP Fiori Elements“ zu ersetzen (S. 10). Diese Anpassung soll die Verständlichkeit verbessern und gleichzeitig den korrekten Toolbegriff in den Vordergrund stellen. Auch im Bereich der visuellen Orientierung wird Optimierungspotenzial gesehen. Zwei Rückmeldungen empfehlen, die Dialogführung stärker an der Benutzerführung der SAP Build Lobby zu orientieren. Ein Experte merkt an, dass die visuelle Referenz auf den bereits existierenden „Create“-Button aus der SAP Build-Umgebung fehle, obwohl dieser für Nutzerinnen und Nutzer ein vertrauter Ausgangspunkt ist. Positiv hervorgehoben wird der Übergang vom Frontend- in den Backend-Pfad, wie er im späteren Verlauf des Prototyps realisiert wird (S. 30). Diese Verbindung wird von allen Experten als logisch nachvollziehbar und funktional stimmig empfunden – insbesondere vor dem Hintergrund, dass in der praktischen Umsetzung UI- und Backend-Komponenten meist eng miteinander verzahnt sind.

Pfad drei - General links or informations: Die Möglichkeit, gezielt auf externe Ressourcen wie Dokumentationen oder Referenzsysteme im SAP Discovery Center zuzugreifen, adressiert insbesondere Informationsbedarfe erfahrener Entwicklerinnen und Entwickler, die über klassische Entwicklungsfragen hinausgehen (S. 7). Von allen fünf Experten wird die Integration von vier thematisch differenzierten Katalogen positiv hervorgehoben: „Missions catalogue“ (S.

18), „Architecture catalogue“ (S. 19), „Service catalogue“ (S. 20) und „AI catalogue“ (S. 21). Diese Auswahl deckt zentrale Wissensbereiche der SAP Build-Plattform ab und bietet damit eine effektive Ausgangsbasis für weiterführende Recherchen. Als Verbesserungspotenzial zweier Experten wird angemerkt, dass vor der Weiterleitung in einen Katalog eine kurze inhaltliche Beschreibung im Chatfenster sinnvoll ist. Dadurch lässt sich die Zielsetzung des jeweiligen Katalogs bereits im Vorfeld vermitteln und so auch weniger erfahrenen Nutzerinnen und Nutzern eine informierte Auswahl ermöglichen. Neben einer verbesserten Nutzerführung würde dies auch das Vertrauen in die Qualität der verlinkten Inhalte stärken.

Pfad vier - Explaining an existing application: Diese Funktion wird von vier Experten als besonders praxisnah und relevant bewertet, da er reale Szenarien abbildet, in denen bestehende Applikationen analysiert und angepasst werden müssen. Besonders positiv hervorgehoben wird die Option, automatisiert einen „Step-by-Step“-Guide als PDF zu generieren (S. 39). Dieses Feature wird als effektive Unterstützung für die eigenständige Einarbeitung in eine vorhandene Anwendung verstanden – insbesondere für Entwicklerinnen und Entwickler, die bislang nicht am ursprünglichen Projekt beteiligt waren. Ergänzend wird die Funktion „Let AI implement it“ (S. 40) sehr begrüßt, da sie es erlaubt, vorgeschlagene Anpassungen direkt in eine Entwicklungsumgebung zu überführen. Dadurch kann der manuelle Konfigurationsaufwand reduziert und ein schnellerer Projektstart ermöglicht werden. Es soll die gewünschte Hauptfunktion der Applikation abgefragt werden (S. 14), gefolgt von einer konkreten Use-Case-Beschreibung der gewünschten Erweiterung (S. 12–14). Erst nach diesen Eingaben soll die Bestätigung über eine passende bestehende Anwendung erfolgen (S. 29). Darüber hinaus wird das verwendete Wording in einzelnen Antwortoptionen von zwei Experten als verbesserungswürdig eingestuft.

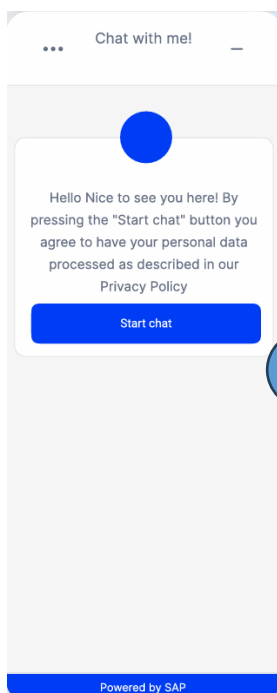


Abbildung 1:
Startbildschirm

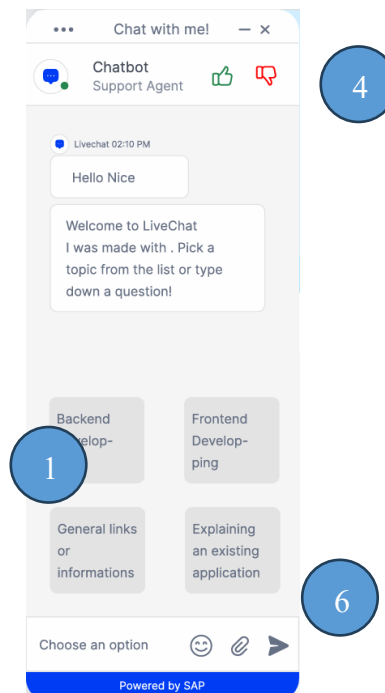


Abbildung 2:
Auswahl der vier
Pfade (Use Cases)

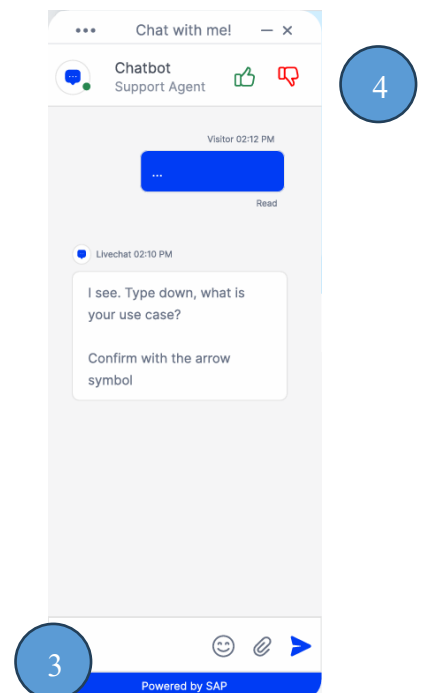


Abbildung 3: Pfad
vier - Explaining an
existing application

5. Fazit

Die vorliegende Arbeit verfolgte das Ziel, einen prototypischen GenAI-Chatbot zur Unterstützung von Entwicklerinnen und Entwicklern in der SAP Build-Umgebung im Kontext der SAP S/4HANA Erweiterungen und organisations-spezifischen Innovationen zu konzipieren, zu gestalten und zu evaluieren. Auf Basis realer Entwicklungsbedarfe sowie einer systematischen Interviewauswertung wird ein dialogbasiertes Assistenzsystem geschaffen, das sich durch kontextsensitive Interaktionspfade und praxisnahe Funktionalitäten auszeichnet.

Trotz positiver Rückmeldungen der Expertinnen und Experten zeigen sich auch Einschränkungen, die bei der Interpretation und Weiterentwicklung des Prototyps berücksichtigt werden sollten. Zum einen basiert die Gestaltung des Chatbots auf einer begrenzten Anzahl qualitativer Interviews, wodurch keine statistisch generalisierbaren Aussagen getroffen werden können. Auch wenn die Auswahl der Teilnehmenden gezielt auf Rollenvielfalt und Erfahrungsbreite ausgerichtet sind, besteht dennoch die Möglichkeit, dass bestimmte Nutzungsperspektiven nicht erfasst wurden. Zudem beschränkt sich der funktionale Umfang des Chatbots auf die Entwicklung von Applikationen in SAP Build – andere SAP Build-Komponenten wie SAP Build Process Automation oder SAP Build Work Zone bleiben außen vor. Diese Fokussierung ist zwar forschungsökonomisch begründet, limitiert jedoch die Übertragbarkeit auf andere Arbeitskontexte. Gleiches gilt für die visuelle Umsetzung des Prototyps in Figma, die den technischen Integrationsaspekt unberücksichtigt lässt. Ein weiterer Aspekt betrifft das Spannungsfeld zwischen Automatisierung und Kontrolle: Während einige Experten eine stärkere Entscheidungsunterstützung begrüßen, wird zugleich betont, dass nachvollziehbare Informationspfade und transparente Quellenangaben essenziell sind, um Vertrauen in das System aufzubauen und Fehlnutzung zu vermeiden. Für eine Weiterentwicklung des Chatbots ergeben sich verschiedene potenzielle Forschungs- und Gestaltungsfelder. Der nächste logische Schritt wäre die technische Umsetzung als interaktiver Prototyp in einer cloudbasierten Sandbox-Umgebung. Damit könnte die tatsächliche Interaktion mit der Zielgruppe in realen Szenarien validiert und durch Usability-Tests ergänzt werden. Langfristig stellt sich die zentrale Frage, wie sich GenAI-gestützte Assistenzsysteme sinnvoll in bestehende Arbeitsabläufe integrieren lassen und welche Governance-Strukturen erforderlich sind, um eine verantwortungsvolle Nutzung, die Sicherstellung der Datenintegrität sowie die langfristige Wartbarkeit dieser Systeme zu gewährleisten.

6. Literaturverzeichnis

- Aslina, Y. R. & Nugraha, I Gusti Bagus Baskara. (2024). Exploring Potential AI Use Cases in Internal Developer Portals: A Path to Enhanced Developer Experience. In 2024 IEEE International Conference on Data (S. 143–148).
<https://doi.org/10.1109/ICoDSE63307.2024.10829893>
- Augsten, S. (2023). Optimierung der Developer Experience in der Softwareentwicklung.
<https://www.datacenter-insider.de/optimierung-der-developer-experience-in-der-softwareentwicklung-a-6cf1ba2adc9a57c506c2c2c10e2dbe4e/>
- Bahrman, S. (2023). Game Changer im UX Design: Von Chatbots bis Predictive UX.
<https://www.it-p.de/blog/ux-game-changer/#dipipopup-25909502>
- Brandt, J., Guo, P. J., Lewenstein, J., Dontcheva, M. & Klemmer, S. R. (2009). Two studies of opportunistic programming: interleaving web foraging, learning, and writing code. In

- Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09) (S. 1589–1598). Association for Computing Machinery.
<https://doi.org/10.1145/1518701.1518944>
- Dehaerne, E., Dey, B., Halder, S., Gendt, S. de & Meert, W. (2022). Code Generation Using Machine Learning: A Systematic Review. In IEEE Access (Volume: 10) (Bd. 10, S. 82434–82455). <https://doi.org/10.1109/ACCESS.2022.3196347>
- Donvir, A., Panyam, S., Paliwal, G. & Gujar, P. (2024). The Role of Generative AI Tools in Application Development: A Comprehensive Review of Current Technologies and Practices. In 2024 International Conference on Engineering Management of Communication and Technology (EMCTECH) (S. 1–9).
<https://doi.org/10.1109/EMCTECH63049.2024.10741797>
- Kaiser, R. (2021). Qualitative Experteninterviews: Konzeptionelle Grundlagen und praktische Durchführung. Springer VS. <https://doi.org/10.1007/978-3-658-30255-9>
- Kortas, T. (2024). KI-Chatbot: Vorteile und Tipps für Unternehmen.
<https://otrs.com/de/blog/ki-ml/ki-chatbot/>
- Lestarini, D. & Raflesia, S. P. (2024). Towards AI Chatbot Effectiveness: Crucial Elements Impacting User Engagement and Satisfaction. In 2024 International Conference of Adisutjipto 2024 (S. 1–5). <https://doi.org/10.1109/ICAAEEI63658.2024.10899167>
- Mahendra, A. (2023). AI Startup Strategy: A Blueprint to Building Successful Artificial Intelligence Products from Inception to Exit. Apress Media. <https://doi.org/10.1007/978-1-4842-9502-1>
- Peffer, K., Tuunanen, T., Rothenberger, M. & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. In Journal of Management Information Systems, Volume 24, Issue 3 (Bd. 24, S. 45–77).
<https://doi.org/10.2753/MIS0742-1222240302>
- Pinto, G., Souza, C. de, Rocha, T., Steinmacher, I., Souza, A. de & Monteiro, E. (2024). Developer Experiences with a Contextualized AI Coding Assistant: Usability, Expectations, and Outcomes. In 2024 IEEE/ACM 3rd International Conference on AI Engineering – Software Engineering for AI (CAIN) (S. 81–91).
- Raghi, K. R., Sudha, K., A. M., S. & Joshua S, S. (2024). Software Development Automation Using Generative AI. In 2024 International Conference on Emerging Research in Computational Science (ICERCS) (S. 1–6).
<https://doi.org/10.1109/ICERCS63125.2024.10894980>
- Rumondor, A. D., Abimanyu, A., Abiyu'Ammaar, M., Achmad, S. & Sutoyo, R. (2024). Improving Programmer Work Quality with ChatGPT Assistance. In 2024 10th International HCI and UX Conference in Indonesia (CHIuXiD) (S. 42–47).
<https://doi.org/10.1109/CHIuXiD64022.2024.10860622>
- Sourceforge. (2025). ABAP vs. JavaScript Comparison Chart.
<https://sourceforge.net/software/compare/ABAP-vs-JavaScript/>
- Tiedemann, J. B. & Bach, T. (2021). Blogs Banner Warum Sie in gute Developer Experience investieren sollten. <https://www.thoughtworks.com/de-de/insights/blog/experience-design/why-you-should-invest-good-developer-experience-today>

Weisz, J. D., He, J., Muller, M., Hoefer, G., Miles, R. & Geyer, W. (2024). Design Principles for Generative AI Applications. In Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24) (S. 1–22). Association for Computing Machinery. <https://doi.org/10.1145/3613904.3642466>

TRACK 3: DIGITALE TRANSFORMATION/MANAGEMENT

Track Chair: Prof. Dr. Peter Langkafel

Leveraging AI and Micro-Credentials in Education Insights from the EduAId Use Case Review

Salma.Sassi¹, Richard Chbeir¹, Sasapu,Venkatesh², Judit Horgas³, Smiljana ukic'in Vuc'kovic'⁴
& Marx Go'mez, Jorge²

Universite' de Pau et des Pays de l'Adour Angle, France¹

University of Oldenburg Oldenburg, Germany²

European School Heads Association Utrecht, Netherlands³

Minds Europe (ME) Novi Sad, Serbia⁴

Abstract: This paper offers a systematic and multidimensional review of real-world applications of Artificial Intelligence (AI) in education, conducted within the framework of the Erasmus+ Ed- uAId project. Drawing on more than 20 validated use cases across primary, secondary, vocational, and higher education settings, we map the current landscape of AI integration along five analytical dimensions: pedagogical integration, ethical and regulatory align- ment, personalization and learner agency, institutional readiness, and micro-credentialing mechanisms. Our analysis reveals a vibrant yet fragmented ecosystem, with significant disparities in implementation maturity, ethical governance, and alignment of AI-generated learning data with verifiable skill recognition. AI- though several initiatives demonstrate promising outcomes such as improved student participation, teacher support, and adaptive learning, challenges persist regarding institutional embedding, explainability, and credential portability.

This study aims to consolidate theoretical foundations, stake- holder insights, and regulatory analyses to inform a sound and inclusive integration of AI in education. The paper identifies key gaps and proposes a strategic agenda to support the development of interoperable, ethically grounded, and pedagogically mean- ingful micro-credentialing architectures. By situating EduAId's contribution within this evolving landscape, we advocate for a co-designed, policy-aligned, and research-informed model for the sustainable integration of AI in education.

Index Terms—Artificial Intelligence in Education, Use Cases, AI Ethics, Micro-Credentials, Digital Competence

1. Introduction

The integration of Artificial Intelligence (AI) into education is rapidly expanding across primary, secondary, vocational, and higher education systems. AI-powered tools now offer unprecedented opportunities for personalized learning, automated assessment, predictive analytics, and intelligent tutoring. How- ever, this technological momentum is accompanied by deep structural and pedagogical challenges. Current deployments of AI in education are often fragmented, disconnected from national curriculum, and lacking consistent ethical safeguards. Many systems operate as isolated pilot initiatives without long-term institutional integration. Furthermore, AI-supported learning remains poorly recognized in formal qualification.

frameworks: while learners may engage with adaptive tools and receive algorithmically generated feedback, there is rarely a mechanism for formally validating the acquired competencies. The lack of interoperable micro-credentialing systems further exacerbates this gap, reducing the potential of AI to contribute meaningfully to lifelong learning and work- force readiness. From a governance perspective, the situation is equally fragmented. Despite the presence of guidelines like the European AI Act or UNESCO's AI in Education recommendations, most real-world use cases show uneven attention to ethical transparency, learner agency, and data protection. Educational actors (e.g., teachers, school leaders, policymakers) face growing pressure to adopt AI without clear models for pedagogical validity, explainability, or evidence- based evaluation. In this context, the very legitimacy of AI in education remains contested.

The EduAId project, funded by Erasmus+, was designed to respond to these challenges. As a research-led and system- level innovation initiative, EduAId seeks to build a sustainable and inclusive model for integrating AI into education. Its objectives are threefold:

- 1) Document and analyze real-world AI use cases in edu- cation across Europe and beyond, identifying practices, gaps, and enablers.
- 2) Propose coherent ethical, scalable and pedagogically frameworks for using AI as a support to teaching, learning, and skill recognition.
- 3) Design and validate a federated micro-credentialing ar- chitecture that ensures the formal recognition of AI- mediated learning outcomes, aligned with European qualification frameworks (e.g., EQF, ESCO). This paper contributes to these goals by presenting a system- atic review and structured synthesis of more than twenty rep- resentative use cases in education. These cases were selected, documented, and validated as part of EduAId's WP2 activities between April and June 2025. A standardized five-criterion

analysis framework was used to assess each case: (C1) ped- agogical integration, (C2) ethical and regulatory alignment, (C3) personalization and learner agency, (C4) institutional readiness, and (C5) micro-credentialing mechanisms.

The rest of this paper is organized as follows. Section 2 presents the research methodology and collaborative valida- tion process used to construct the case database. Section 3 analyzes the landscape of educational use cases, highlighting typologies, institutional differences, and critical gaps. Section 4 proposes strategic research and policy directions based on cross-case synthesis. Finally, Section 5 situates EduAId's unique contribution and long-term vision within the broader European agenda on ethical and inclusive AI in education.

2. Systematic Methodology for Identifying and Validating Use Cases

As mentioned previously, this review was conducted as part of the EduAId project's desk research activities and was designed to ensure both methodological rigor and collaborative engagement across all participating institutions. The entire process was implemented between April and June 2025 and followed the Preferred Reporting Items for Systematic Re- views and Meta-Analyses (PRISMA) guidelines to enhance transparency, replicability, and consistency in evidence collec- tion.

The methodological workflow consisted of six interdepen- dent phases:

2.1. Definition of validated sources and keywords.

A shared corpus of databases (scientific, institutional, and practitioner-oriented) and a validated list of multilingual keywords were collaboratively defined and approved by all project partners. This ensured that the case identification process was both exhaustive and aligned with the project's research questions, particularly the intersection between AI adoption, pedagogical integration, and micro-credentialing mechanisms.

2.2. Standardized drafting of use cases.

Each partner contributed a minimum of four use cases, documented using a common structured template. This template was co-designed during the project's kickoff and focused on capturing five core dimensions: pedagogical integration, ethical and regulatory alignment, personalization and learner agency, institutional readiness, and micro-credentialing mechanisms. The sources of these cases included peer-reviewed publications, policy reports, project deliverables, and technical documentation from AI platforms deployed in education.

2.3. Structured cross-validation and peer review.

To ensure analytical coherence and reduce subjectivity, each use case was independently validated by two actors: (i) a designated validator responsible for structural consistency and thematic relevance, and (ii) a reviewer tasked with assessing completeness, clarity, and alignment with the five analytical dimensions. Feedback was logged directly in shared documents, and validation progress was tracked using a central Excel matrix.

2.4. Final quality assurance and harmonization.

A final harmonization phase, coordinated by the Work package leader, ensured uniformity in formatting, vocabulary, and conceptual framing. No content was rewritten but use cases were tagged "ready for inclusion" only if they met predefined quality thresholds.

2.5. PRISMA-based traceability and transparency.

All reference materials and source justifications were systematically recorded using a dedicated PRISMA tracking log. This enabled transparent traceability for each use case, strengthened the replicability of the review, and facilitated future updates or extensions of the dataset.

2.6. Meta-synthesis and synthesis matrix development.

Once the individual cases were validated, they were cross compared using a shared synthesis table. Each case was mapped to the selected five analytical criteria (C1–C5), enabling a comparative overview of trends, gaps, and innovations in the landscape of AI applications and micro-credentialing in education.

This multi-phase methodology not only ensured high-quality evidence collection but also fostered alignment between partners and created a shared analytical lens. The result is a robust, empirical foundation that supports both case-level insights and system-level conclusions regarding the integration of ethical AI and micro-credentials in educational ecosystems.

3. AI and Micro-Credentials in Education

The integration of AI into education is not solely a matter of technological advancement, it also raises fundamental questions about pedagogical alignment, ethical governance, and the formal recognition of learning outcomes. Recent years have witnessed a proliferation of AI-driven tools across all levels of education, from intelligent tutoring systems and curriculum support to data-driven personalization and institutional automation. However, these innovations vary widely in their educational validity, ethical robustness, and capacity to support verifiable learning pathways. In particular, the emergence of micro-credentials digital attestations of discrete competencies offers a promising avenue for formalizing AI-enhanced learning. Yet, the integration of micro-credentials into AI-supported educational ecosystems remains fragmented. Many initiatives prioritize personalization or automation but neglect recognition and certification, limiting their systemic value. This section provides a structured review of representative use cases, emphasizing the intersection between AI applications and micro-credentialing practices. We classify and analyse each initiative according to five analytical dimensions: (C1) pedagogical integration, (C2) ethical and regulatory alignment, (C3) personalization and learner agency, (C4) institutional readiness, and (C5) credentialing mechanisms. We begin by presenting the landscape of real-world cases, followed by a critical synthesis that identifies trends, gaps, and implications for policy and research.

4. Use case landscape and micro-credential integration

A growing body of initiatives and research projects across Europe and beyond have sought to operate AI in educational contexts, with varying degrees of technological maturity, ethical oversight, and credential integration. This section provides a critical mapping of these initiatives with particular attention to the role of micro-credentials, focusing on five analytical dimensions: (a) pedagogical integration, (b) ethical and regulatory alignment, (c) personalization and learner agency, (d) institutional readiness, and (e) credentialing mechanisms.

a) Teacher Training and Institutional Preparedness: The AI4T initiative [1], [2] represents a structured effort to build AI literacy among professional educators through a multilingual MOOC deployed in five European countries. The project introduced a seven-layer template for AI tool characterization and emphasized ethical understanding and pedagogical applicability. Despite increased teacher confidence, challenges such as uneven digital infrastructure and emotional resistance (e.g., "AI anxiety") persisted. Building on this, the AI Pioneers project [7] targeted adult and vocational education, mapping AI competencies onto an extended DigCompEdu framework. Over twenty tools were piloted, with a strong emphasis on ethical auditing (privacy, bias, environmental impact) and the institutionalization of micro-credentialing systems to formally recognize AI-related teaching skills.

b) Learner-Centred Intelligent Assistance: Several plat-

forms have explored AI to support personalized learning. Brainly [3], [4] combines community answers with NLP and computer vision (e.g., Snap to Solve), offering fast assistance but raising concerns around over-reliance and pedagogical misalignment. In contrast, CENTURY Tech [10] delivers adaptive learning experiences using AI-driven analytics, showing measurable improvement in learner outcomes and teacher efficiency. ThirdSpace Learning [11] supplements tutoring with real-time AI prompts, improving math proficiency while maintaining

ethical safeguards. However, none of these systems currently offers formal micro-credentials, indicating a gap between technological personalization and institutional recognition.

c) Collaborative and Critical Literacies: Social annotation

tools like Perusall [8], [9], deployed in Dutch schools, support collaborative reading through NLP clustering and feedback mechanisms, and integrate GDPR-compliant architectures. In higher education, a study across Chinese universities evaluated the use of ChatGPT (GreatGPT) in critical writing [13], [14], highlighting students' reflective engagement despite mixed outcomes. The University of Hohenheim [13] advanced this model by integrating a GPT-4-based assistant for academic writing, issuing Open Badges v2.0 for iterative feedback cycles. This is one of the few examples where micro-credentialing was natively embedded within the AI-supported learning process.

d) Equity of Access and Open Educational Resources:

The X5GON project [5], [6] uses AI to personalize OER discovery across linguistic and accessibility boundaries. It incorporated ethical data practices and interoperable standards but did not explore formal recognition mechanisms. A contrasting case comes from UC3M [19], where the OpenLearning AI Coach tailors digital skill pathways and awards micro-credentials aligned with EQF and ESCO. Learner motivation improved, and dropout rates declined in AI-augmented cohorts. Similarly, IBM SkillsBuild with Polytechnique Montr al [20], [21] provides AI-validated, blockchain-secured badges in cybersecurity and AI, which have demonstrable labour market value.

e) Creative and Contextual Applications: In Serbia, flipped classrooms used AI-supported tools like InShot and Genially [12] to build asynchronous geography modules and gamified escape rooms. These showed pedagogical promise but also revealed access inequities and high teacher preparation time. At the University of Sydney [16], AI tools such as ChatGPT and DALL E were integrated into multimedia capstone projects. While ethical instincts were strong among students, the absence of institutional micro-credentialing frameworks limited the formalization of these emerging competencies.

f) AI for Curriculum Design and Institutional Tools: Catholic CoPilot [15] demonstrated how generative AI could be embedded in faith-based K–12 curricula to automate assessment and improve teacher efficiency. MagicSchool.ai [17] was trialed for daily teacher support tasks. These tools, while impactful, operate largely outside formal credentialing structures. By contrast, the University of Jyv skyl  [18] developed an NLP-based tool to assess student-teacher portfolios against the DigCompEdu framework, combining multimodal input with human-validated AI scoring.

5. Synthesis and discussion

Across the examined initiatives, Table I reveals significant disparities in how AI is pedagogically integrated and how it supports learner agency, institutional scalability, and formal skill recognition through micro-credentials. While a majority of projects exhibit moderate to high levels of pedagogical integration (C1), particularly in structured educational environments (e.g., AI4T, AI Pioneers, CENTURY Tech), there remains a notable gap in aligning these pedagogical innovations with robust, standardized credentialing systems (C5). Projects such as AI4T and AI Pioneers stand out for their dual emphasis on professional

educator preparedness and the integration of formalized micro-credentialing—leveraging Open Badges and DigCompEdu-aligned frameworks. These initiatives not only address skill acquisition but also contribute to institutional capacity building (C4), showing that well-aligned digital strategies can scale across contexts. In contrast, several platforms with high personalization potential (C3) such as Brainly, CENTURY Tech, and ThirdSpace Learning lack mechanisms to formally acknowledge learner progress or digital competencies. This reveals a fragmentation in the educational value chain: while these tools personalize learning experiences, they often operate outside formal educational structures and validation protocols. From an ethical and regulatory standpoint (C2), most projects achieve medium to high alignment, particularly those originating within EU-funded or institutional frameworks. Notably, the UC3M OpenLearning Coach and IBM SkillsBuild illustrate how ethical compliance and labor market relevance can be jointly addressed by integrating AI-based evaluation and blockchain-backed credentials. These projects align closely with EQF and ESCO frameworks, offering interoperability and recognition across borders. Micro-credentialing (C5) emerges as the least uniformly addressed dimension. Only six out of sixteen initiatives implement verifiable credentials, with three (AI4T, Hohenheim, UC3M) embedding them directly into instructional workflows. The rest either omit this dimension or treat it as an afterthought, despite using AI for tasks that could meaningfully generate evidence of learning or competence (e.g., flipped classrooms, MagicSchool.ai, Perusall). The analysis also highlights a geographic and institutional imbalance: large-scale or university-backed initiatives are more likely to address all five dimensions systematically, while start-up or tool-centric solutions often prioritize C3 (learner personalization) at the expense of C1 (curricular coherence) and C5 (recognition). This calls for a more holistic framework that bridges edtech innovation with formal accreditation pathways. Ultimately, the diversity captured in Table I suggests that while the ecosystem of AI in education is maturing, the consolidation of ethical, personalized, and credentialed learning remains an open challenge. Future directions must focus on designing interoperable micro-credential systems that are not only technically sound but pedagogically meaningful and institutionally sustainable. The following section outlines such pathways, addressing both research and policy implications.

6. Future Directions

The analysis of current AI-driven educational initiatives reveals a vibrant yet fragmented ecosystem, characterized by a diverse range of pedagogical objectives, technological maturity levels, and degrees of integration of micro-credentialing. While many projects demonstrate promising outcomes in terms of learner engagement, adaptive support, and teacher facilitation, these benefits are often isolated and lack systemic anchoring within institutional and policy frameworks.

6.1. Key Challenges

Based on the comparative mapping and synthesis, three transversal challenges can be identified:

Weak institutional embedding and pedagogical coherence A substantial number of AI implementations remain disconnected from institutional strategies and national educational agendas. Although some projects, such as AI4T or UC3M's AI Coach, align with frameworks such as DigCompEdu or EQF, most are confined to pilot studies or local classroom practices without a long-term vision for sustainability, professional development or curricular

integration. This gap limits both the scalability and legitimacy of these innovations within formal education systems.

Fragmented ethical governance: Although compliance with GDPR and basic transparency measures is frequently reported, few initiatives embed holistic ethical governance. Notable exceptions (e.g., AI Pioneers, IBM SkillsBuild) introduce structured auditing protocols and environmental safeguards, yet most systems fall short on ensuring explainability, inclusiveness, and student agency. As AI increasingly mediates access to feedback, learning paths, and assessment, ethics must shift from compliance-oriented checklists to embedded design principles, guided by anticipatory governance models.

Underdeveloped micro-credentialing ecosystems: Micro-credentials have emerged as a strategic enabler for recognizing granular, practice-based, and AI-supported learning. Yet, as shown in Table I, only a minority of initiatives offer verifiable credentialing mechanisms, and even fewer integrate them meaningfully into pedagogical workflows. A lack of common ontologies, quality standards, and formal interoperability with public qualification systems (e.g., ESCO, EQF) threatens the credibility and uptake of such credentials.

6.2. Strategic directions

To overcome these limitations and pave the way toward inclusive and robust AI-enabled education systems, the following strategic directions are proposed:

Institutional alignment and scalable pedagogical models: Future initiatives should embed AI tools within broader institutional development plans, linking them explicitly to teaching and learning objectives, curriculum standards, and teacher training pathways. This includes the co-development of AI-readiness rubrics and integration into accreditation and evaluation systems.

Ethics-by-design and multi-level governance: Beyond compliance, ethics must be built into the lifecycle of AI tools from data sourcing and algorithmic design to user interaction and decision-making autonomy. National education strategies should promote shared ethical frameworks co-developed with professional educators, students, developers, and policymakers.

Federated and interoperable micro-credentialing architectures: A federated micro-credentialing model should be developed to ensure alignment between AI-supported learning activities and formal qualification frameworks. This requires standardized metadata schemas (e.g., Open Badges 2.0), verifiable digital infrastructures (e.g., blockchain, digital wallets), and partnerships with industry stakeholders to enhance recognition and employability.

Participatory design and evidence-based evaluation: Stakeholders including teachers, learners, school leaders, parents, students and policymakers must be involved in the iterative design, testing, and refinement of AI tools. Impact assessments should go beyond usability and focus on learning outcomes, equity, trust, and long-term engagement.

Tab.1: Comparative Analysis of AI Use Cases in Education

Use Case / Project	C1	C2	C3	C4	C5
AI4T [1], [2]	High	High	Medium	Medium	Yes (Open Badges)
AI Pioneers [7]	High	High	Medium	High	Yes (DigCompEdu-aligned)
Brainly [3], [4]	Low	Medium	High	Low	No
CENTURY Tech [10]	High	Medium	High	Medium	No
ThirdSpace Learning [11]	Medium	High	High	High	No
Social Annotation (Perusall) [8], [9]	Medium	High	Medium	Medium	Proposed
GreatGPT (China) [13], [14]	Medium	Medium	Medium	Low	No
Hohenheim Writing Assistant [13]	High	High	Medium	High	Yes (Open Badges 2.0)
X5GON [5], [6]	Medium	High	High	Medium	No
UC3M OpenLearning Coach [19]	High	High	High	High	Yes (EQF/ESCO)
IBM SkillsBuild [20], [21]	Medium	High	Medium	High	Yes (Blockchain)
Serbian Flipped Classrooms [12]	Medium	Medium	Medium	Low	No
Sydney Capstone Projects [16]	Medium	Medium	Medium	Medium	No
Catholic CoPilot [15]	Medium	High	Low	High	No
MagicSchool.ai [17]	Medium	Medium	Medium	Medium	No
Jyva"skylä" DigCompEdu Assessment [18]	High	High	Medium	High	Yes (Competency-Based)

C1: Pedagogical integration C2: Ethical & regulatory alignment C3: Personalization & learner agency C4: Institutional readiness
C5: Micro-credentialing mechanism

7. EduAId’s strategic contribution to the AI and micro-credentialing ecosystem

As both a research-driven initiative and a system-level innovation project, EduAId occupies a unique position at the intersection of ethical AI deployment, pedagogical co-design, and micro-credentialing infrastructure. Unlike isolated tech- nological interventions, EduAId adopts an integrated, inter- disciplinary approach that treats these domains as interdepen- dent pillars for building inclusive, learner-centered educational ecosystems.

7.1. Ethical AI and pedagogical co-design as foundational principles

One of EduAId’s key innovations lies in embedding ethical governance frameworks directly into the architecture of its AI-enhanced micro-credentialing processes. The project operationalizes principles such as transparency, fairness, and learner agency by aligning its system design with European and international guidelines (e.g., the EAIFT framework [23]) and integrating human-in-the-loop mechanisms across its pipeline. Rather than treating AI as a black-box optimization tool, EduAId positions it as a co-creative partner in the learning process. Teachers and students are actively involved in the co-design of AI-supported pathways, ensuring that adaptivity and automation are pedagogically meaningful, contextually appropriate, and trustworthy. This reinforces the legitimacy of the learning evidence used in micro-credentialing and preserves the centrality of human educational agency.

7.2. Use-Case driven and policy-aligned implementation

EduAId’s implementation strategy is explicitly use-case driven. Pilot experiments and demonstrators are selected to tackle critical bottlenecks in European education, including teacher workload, curriculum differentiation, equity in access, and the fragmentation of credentialing systems. The project aligns with the European Commission’s Digital Education Ac- tion Plan (2021–2027), the European Education Area (EEA), and principles outlined in the AI Act, thereby responding to policy imperatives for trustworthy and inclusive digital education. Through this lens, EduAId contributes to both pedagogical transformation and systemic policy innovation. Its work on micro-credentials is particularly noteworthy, seeking

to enhance their portability, transparency, and alignment with frameworks such as EQF and ESCO.

7.3. Toward a scalable and sustainable european model

EduAid's long-term vision goes beyond immediate deliverables. It aims to develop a replicable European model for AI-supported education that foregrounds sustainability, explainability, and inclusion. This vision encompasses:

The design of a pedagogically grounded AI framework adaptable by ministries of education and teacher training institutions;

The development of ethically certified, open-source AI modules for formative assessment and adaptive learning;

The construction of a federated, interoperable repository of micro-credentials mapped to European and national qualification frameworks.

In doing so, EduAid does not aim to replace professional educators, but to enhance their capacities and reduce systemic friction in competency recognition. Its approach supports the emergence of a digitally sovereign, ethically grounded education infrastructure in Europe. By integrating AI, micro-credentials, and participatory co-design, EduAid offers a forward-looking yet actionable pathway to rethinking educational validation systems in an age of rapid technological transformation.

8. Conclusion and Future Work

This review has highlighted the growing diversity—but also the persistent fragmentation—of AI applications in education. While a number of use cases demonstrate pedagogical promise, enhanced personalization, and ethical awareness, most implementations remain limited in scope and systemic alignment. In particular, we observe that many initiatives are:

- Technically opaque, lacking transparent or explainable AI mechanisms;
- Ethically inconsistent, with variable adherence to principles of data protection, non-discrimination, and user agency;
- Pedagogically limited, focusing on automation over meaningful learning transformation;
- Credential-wise disconnected, with few examples of verifiable, portable, and interoperable micro-credentials.

Future work will extend this analysis through a broader and more inclusive systematic review framework, building on the PRISMA-based methodology currently developed within the EduAid project. This forthcoming review will explore a wider array of educational levels, geographic regions, technological architectures, and inclusion criteria, to better map the landscape of AI use in education.

In parallel, EduAid will also design and deploy targeted stakeholder questionnaires to gather first-hand insights from professional educators, school leaders, policymakers, and learners. These instruments will help surface expectations, perceived risks, and unmet needs, providing a grounded perspective to guide the co-design of ethical and pedagogically sound AI systems.

Taken together, these complementary research strands aim to advance a more integrated, evidence-informed, and future-ready ecosystem for AI and micro-credentialing in education.

9. Acknowledgements

This work was supported by the Erasmus+ project EduAId with project number KA220-NI-24-36-253480, co-funded by the European Union. The authors would like to thank all project partners for their collaboration and valuable feedback throughout the use case analysis. Appreciation is also extended to the educators, trainers, and learners who informed the development of the micro-credential design and evaluation activities.

10. References

- [1] AI4T Project. (2024). *Template for AI characterization*. <https://www.ai4t.eu/>
- [2] AI4T Project. (2024). *Recommendations on AI in education*. <https://ec.europa.eu/...>
- [3] Brainly. (2023). *Honor code and AI use policy*. <https://brainly.com/honor-code>
- [4] Trustpilot & Reddit. (2024). *User reviews on Brainly*. <https://www.trustpilot.com/review/brainly.com>
- [5] X5GON Project. (2023). *Final report on piloting*. <https://cordis.europa.eu/project/id/761758>
- [6] X5GON Discovery Platform. (2023). *X5GON discovery platform*. <https://discovery.x5gon.org/>
- [7] AI Pioneers Project. (2024). *Evaluation schema*. <https://aipioneers.org/>
- [8] Pieter, M., & van der Meer, J. (2024). *The value of social annotation in AI-enhanced classrooms*. <https://...>
- [9] Kennisnet. (2023). *Digitale geletterdheid*. <https://www.kennisnet.nl/...>
- [10] CENTURY Tech. (2024). *AI in education*. <https://www.century.tech/...>
- [11] ThirdSpace Learning. (2024). *AI maths tutor*. <https://thirdspacelearning.com/...>
- [12] Stajić, S., Ristić, L., & Živanović, M. (2025). AI-supported flipped classrooms in Serbia: Lessons from asynchronous gamification. *Sustainability*, 17(6), 2464. <https://www.mdpi.com/2071-1050/17/6/2464>
- [13] Gimpel, H., et al. (2023). *Unlocking generative AI for higher education: A case study from the University of Hohenheim*. <https://www.econstor.eu/...>
- [14] Yang, Y., Zhang, X., & Li, M. (2024). From surface to deep learning with generative AI: A study of Chinese universities. *Studies in Higher Education*, 49(5).
- [15] Microsoft. (2024). *BCE to deploy Microsoft 365 Copilot to enhance curriculum automation*. <https://news.microsoft.com/...>
- [16] Charles, K., Nguyen, T., & Rizvi, A. (2025). AI in action: Creative integration in capstone multimedia projects. *European Journal of Pharmacology*, 998.
- [17] MagicSchool AI. (2024). *Official website*. <https://www.magicschool.ai/>

- [18] University of Jyväskylä. (2024). *AI portfolio analysis tool for DigCompEdu alignment*.
- [19] Universidad Carlos III de Madrid. (2024). *OpenLearning AI coach report*.
<https://www.uc3m.es/openlearning>
- [20] IBM SkillsBuild. (2024). *IBM SkillsBuild program overview*.
<https://skillsbuild.org/>
- [21] OpenCerts. (2024). *Blockchain-based credential verification platform*.
<https://www.opencerts.io/>
- [22] Joint Research Centre. (2024). *DigCompEdu: European framework for the digital competence of educators*. <https://joint-research-centre.ec.europa.eu/digcompedu>
- [23] Ejjami, F. (2024). EAI FT: A framework for trustworthy AI integration in education. *Journal of Ethical AI in Education*, 2(1), 45–63.

Wie betriebliche Weiterbildung skalierbar wird: Ein Zielbild digitaler SAP-Schulungen

Sebastian Junghans, Massimiliano Perini, Laura Schladitz & Martin Trommer
Westfälische Hochschule Zwickau, Deutschland

Abstract: Der steigende Weiterbildungsbedarf in Unternehmen trifft auf einen wachsenden Mangel an Schulungspersonal sowie zunehmend heterogene Lerngruppen. Besonders im Kontext großvolumiger Transformationsprojekte wie der SAP-S/4HANA-Einführung stoßen herkömmliche Schulungsformate an organisatorische und didaktische Grenzen. Der vorliegende Beitrag beschreibt ein skalierbares Konzept zur Digitalisierung betrieblicher Weiterbildung auf Basis des Learning Cycle, eines KI-gestützten Frameworks zur automatisierten Aufbereitung und adaptiven Bereitstellung von Schulungsinhalten. Bestehende Materialien wie Präsentationen oder Tonspuren werden systematisch erfasst, strukturiert und in personalisierte Lernpfade überführt, die zeitlich flexibel und ortsunabhängig im Lernmanagementsystem bereitgestellt werden. Anhand eines realen SAP-Rollouts wird aufgezeigt, wie sich damit sowohl Schulungskapazitäten entlasten als auch Lernerfolge sichern lassen. Der Beitrag versteht sich als konzeptioneller Impuls für eine zukunftsfähige, skalierbare Weiterbildungsstrategie unter Bedingungen von Fachkräftemangel, Digitalisierung und organisationaler Komplexität.

Keywords: Adaptive Lernpfade, Betriebliche Weiterbildung, Learning-Cycle, KI in der Erwachsenenbildung

1. Einleitung

Die digitale Transformation verändert Arbeitsprozesse und Qualifikationsanforderungen in nahezu allen Branchen. Besonders im IT-Bereich, für den ein starkes Stellenwachstum erwartet wird (World Economic Forum 2025), steigt der Bedarf an betrieblicher Weiterbildung. Gleichzeitig fehlt es an qualifizierten Fachkräften. Im Jahr 2024 waren laut (Bitkom e.V. 2024) rund 149.000 IT-Stellen unbesetzt. Auch im Bereich der betrieblichen Weiterbildung zeigt sich ein zunehmender Mangel an geeignetem Schulungspersonal.

Demografische Entwicklungen verschärfen diese Situation zusätzlich. Die Erwerbsbevölkerung schrumpft und altert, wodurch eine höhere Zuwanderung notwendig ist (Deschermeier 2024). Zugleich steigt die Zahl an Quereinsteigenden (Anger et al. 2025), was die Heterogenität innerhalb der Schulungsgruppen weiter erhöht. Unterschiedliche fachliche Vorkenntnisse, digitale Kompetenzen und sprachliche Voraussetzungen erschweren die Gestaltung wirksamer Weiterbildungsformate – insbesondere unter Bedingungen begrenzter personeller Ressourcen.

Für das Schulungspersonal ergeben sich daraus hohe didaktische Anforderungen. Lernende mit Unterstützungsbedarf benötigen zusätzliche Betreuung, während fortgeschrittene Teilnehmende nur eingeschränkt gefördert werden können. Eine weitergehende Individualisierung ließe sich nur durch kleinere Gruppen und einen erhöhten Personaleinsatz realisieren, was in der Praxis häufig an Kosten und Kapazitäten scheitert. Herkömmliche Formate geraten dadurch zunehmend in ein ungünstiges Verhältnis von Aufwand und Wirkung.

Gefragt sind daher Schulungskonzepte, die sich skalieren lassen und gleichzeitig differenzierte Lernwege ermöglichen, auch unter Bedingungen begrenzter personeller Ressourcen. Digitale Lernsysteme, die vorhandene Schulungsinhalte effizient strukturieren und adaptive Lernpfade bereitstellen, eröffnen hierfür ein vielversprechendes Potenzial. Kapitel 2 stellt mit dem Learning Cycle ein solches Konzept vor.

2. Learning Cycle: Das Konzept

Der Learning Cycle ist ein KI-gestütztes Multi-Agenten-Framework zur systematischen Erfassung, Strukturierung und didaktischen Aufbereitung von betrieblichem Wissen. Es verarbeitet sowohl strukturierte als auch unstrukturierte Quellen, darunter Präsentationen, Textdokumente, Bildschirmaufnahmen und Tonspuren. Auf diese Weise lassen sich wertvolle Wissensbestände sichern, standardisieren und in digital gestützte Lernprozesse überführen. (Junghans 2025)

Ein zentraler Bestandteil des Frameworks ist die automatische Erstellung adaptiver Lernpfade, die auf die Vorkenntnisse, Lernpräferenzen und den individuellen Fortschritt der Lernenden reagieren. Die Bereitstellung erfolgt über ein Lernmanagement-System, das sowohl die Kursorganisation als auch die Benutzerverwaltung und Lernstandserhebung zentral steuert. Die Potenziale adaptiver Lernpfade werden auch im Hochschulkontext zunehmend diskutiert, etwa im Rahmen des *WHZ* (Westfälische Hochschule Zwickau) *LearnXP-Projekts* (Trommer et al. 2024). Die dort entwickelten Ansätze lassen sich grundsätzlich auch auf betriebliche Weiterbildungsszenarien übertragen.

Der Learning Cycle kam bislang bei zwei Projekten der WHZ in prototypischen Anwendungsfällen zum Einsatz: Zum einen zur Sicherung impliziten Erfahrungswissens langjähriger Fachkräfte vor dem Renteneintritt, zum anderen zur Unterstützung ausgewählter Onboarding-Prozesse. In beiden Fällen wurden reale Arbeitsplatzdialoge verarbeitet und in strukturierte, digital gestützte Lernpfade überführt. Die entsprechenden Veröffentlichungen befinden sich derzeit in Vorbereitung.

3. Digitale Schulung im Großrollout mit dem Learning Cycle

Aufbauend auf einem realen, aus Vertraulichkeitsgründen anonymisierten Rollout-Projekt zur Einführung von SAP S/4HANA in einem deutschen Großunternehmen wird der Learning Cycle in diesem Beitrag als Lösung zur Digitalisierung und Steuerung komplexer, großvolumiger Schulungsprojekte weitergedacht.

Im Zuge dieses Projekts sollen mehrere tausend Mitarbeitende ohne Vorerfahrung im Umgang mit ERP-Systemen geschult werden. Bereits in der Pilotphase, der Qualifizierung zentraler Rollen wie Projekt- und Teilbereichsleitungen, zeigen sich strukturelle Herausforderungen: Die Teilnehmenden sind bundesweit verteilt und Präsenzformate daher organisatorisch kaum umsetzbar. Auch gestreckte Online-Schulungen erfordern einen hohen personellen Aufwand und sind mit organisatorischen Unsicherheiten verbunden, etwa wenn Termine aufgrund geringer Zahl an Teilnehmenden nicht zustande kommen. Eine Skalierung auf die gesamte Belegschaft würde die verfügbaren Schulungskapazitäten deutlich übersteigen.

Ziel ist es daher, die in Kapitel 2 beschriebenen Mechanismen des Learning Cycle auf vollständige Schulungseinheiten anzuwenden und so die bislang manuell organisierten,

verteilten Online-Formate in ein systematisch strukturiertes, digital gestütztes Schulungskonzept zu überführen. Auf diese Weise wird es möglich, Mitarbeitende dezentral, zeitlich flexibel und bedarfsangepasst zu qualifizieren, unabhängig von Standort, Gruppenzugehörigkeit oder Verfügbarkeit einzelner Schulungspersonen. Dies reduziert nicht nur den Koordinationsaufwand und Ausfallzeiten, sondern kann auch die Akzeptanz der Schulungsformate erhöhen. Formative Tests ermöglichen eine kontinuierliche Überprüfung des Lernerfolgs. Ergänzend lassen sich Feedbackdaten zu einzelnen Lerneinheiten zentral sammeln, analysieren und perspektivisch auch KI-gestützt auswerten, um Inhalte kontinuierlich weiterzuentwickeln und gezielt nachzusteuern.

Voraussichtlich wird auch im Zielbild ein hybrider Modus beibehalten, jedoch in anderer Ausprägung: Das Schulungspersonal übernimmt stärker moderierende Aufgaben, etwa durch die Betreuung asynchroner Lernphasen in Frageforen oder als Ansprechperson für individuelle Rückfragen. Zusätzlich werden feste Gruppenkonsultationen etabliert, in denen vertiefende Inhalte besprochen oder typische Fehlerbilder adressiert werden können. Einzelgespräche lassen sich nach Bedarf individuell vereinbaren. Die zeitliche Entlastung durch die digitale Vorstrukturierung schafft somit Freiräume für qualitativ hochwertige, gezielt eingesetzte Betreuung.

4. Diskussion und Anschlussfähigkeit

Das skizzierte Konzept zur Nutzung des Learning Cycle eröffnet vielversprechende Perspektiven für die Skalierung betrieblicher Schulungsprozesse. Insbesondere in Kontexten mit hoher Standardisierung und gleichzeitig wachsender Heterogenität der Zielgruppen, wie etwa im SAP-Schulungsumfeld, lassen sich durch eine automatisierte, KI-gestützte Aufbereitung bestehender Inhalte signifikante Effizienzgewinne erzielen. Die Verlagerung repetitiver Wissensvermittlung in digitale Formate entlastet das Schulungspersonal und schafft Freiräume für individualisierte Betreuung und gezieltes Coaching. Gleichzeitig erhöht die adaptive Steuerung der Inhalte die Passgenauigkeit der Lernangebote, was sich positiv auf Lernmotivation und nachhaltigen Lernerfolg auswirken kann.

Die Umsetzung des Konzepts ist jedoch an bestimmte Voraussetzungen gebunden. So erfordert die inhaltliche Verarbeitung vorhandener Materialien eine gewisse Mindestqualität hinsichtlich Struktur, Sprache und Medienformat. Auch eine geeignete technische Infrastruktur, etwa in Form eines leistungsfähigen Lernmanagement-Systems, muss vorhanden sein oder in Einklang mit der bestehenden IT-Landschaft des jeweiligen Unternehmens gebracht werden. Hinzu kommt, dass die initiale Aufbereitung und Systemkonfiguration mit einem nicht zu unterschätzenden Implementierungsaufwand verbunden ist, insbesondere wenn zusätzliche KI-Komponenten zur Inhaltsanalyse oder Testgenerierung zum Einsatz kommen. Für die Qualitätssicherung ist bislang ein menschlicher Abnahmeschritt („Human in the Loop“) notwendig. Ein weiterer Erfolgsfaktor liegt in der Akzeptanz durch Schulungspersonal und Lernende, die frühzeitig eingebunden und bei der Einführung begleitet werden sollten.

Für die weitere Entwicklung besteht Forschungsbedarf, vor allem hinsichtlich der empirischen Evaluation der Lerneffekte adaptiver Lernpfade im Vergleich zu traditionellen Schulungsformaten. Auch Fragen der Integration in bestehende betriebliche Weiterbildungsstrategien und Prozesslandschaften bedürfen vertiefter Betrachtung. Perspektivisch erscheint zudem eine Übertragung auf andere Schulungsbereiche denkbar, etwa im Kontext Arbeitssicherheit oder Compliance-Schulungen. Der vorgestellte Ansatz liefert

somit einen konzeptionellen Impuls für eine systematische Weiterentwicklung betrieblicher Bildung unter den Bedingungen von Fachkräftemangel, Digitalisierung und organisatorischer Komplexität.

5. Quellenverzeichnis

Anger, S. et al. (2025): Zentrale Befunde zu aktuellen Arbeitsmarkt-Themen 2025. Institut für Arbeitsmarkt und Berufsforschung. <https://iab.de/publikation/?id=14679002> [Stand 2025-07-21].

Bitkom e.V. (2024): Mangel an IT-Fachkräften droht sich dramatisch zu verschärfen.

Deschermeier, P. (2024): IW-Bevölkerungsprognose 2024. IW-Trends 3, 67.

Junghans, S. (2025): Learning Cycle (internes Dokument, AIIS – Arbeitskreis Integrierte Informationssysteme).

Trommer, M. et al. (2024): Enhancing Educational Outcomes Through Learning Pathways and Artificial Intelligence-Driven Learning Analytics. In: Smart Technologies for a Sustainable Future. Cham: Springer Nature Switzerland.
https://link.springer.com/10.1007/978-3-031-61891-8_27 [Stand 2025-07-22].

World Economic Forum (2025): The Future of Jobs Report 2025.
<https://www.weforum.org/publications/the-future-of-jobs-report-2025/>.

Framing Labour Inclusion Through Micro-Credentials: A Theoretical Synthesis at the Intersection of Sustainability, Digital Transformation, and Human Capability

Musharu, Timothy, Sasapu, Venkatesh, Gomez, Jorge Marx, Cremering, Insa
Dept of Computer Science University of Oldenburg Oldenburg, Germany

Abstract: Micro-credentials are increasingly promoted as flexible instruments for upskilling and reskilling within economies shaped by digital transformation and sustainability imperatives. Yet definitions and purposes remain fragmented across education, sustainability, and informatics literatures. This study undertakes an interpretive conceptual synthesis of selected Literature to map these conceptualisations and explain how micro-credentials can expand human capability for equitable labour market inclusion, and to specify the conditions under which effects are strengthened or weakened. Using a transparent search and screening protocol with documented reliability procedures, manually coded the literature deductively against four core domains and inductively for emergent themes. The analysis yields a four-component framework in which Micro-Credential Design and Recognition, supported by robust recognition of prior learning, enhances Capability Expansion that in turn promotes Labour Market Inclusion. Two contextual forces condition these pathways. Digital transformation, including the design and ethical governance of AI enabled platforms, can amplify capability gains through personalisation and verifiable records, or can constrain them where transparency and bias safeguards are absent. Sectoral demand for green skills increases the employability premium associated with sustainability aligned competences. From this model we derive a set of testable propositions and operational indicators that can support programme evaluation and future empirical studies. Policy and practice implications include alignment of qualification frameworks with modular credentials, capability weighted funding for inclusive upskilling in green sectors, and 4C design principles that foreground competence specificity, credit transparency, cumulative stackability, and contextual relevance. The framework offers a clear, capability centered basis for building credential ecosystems that are innovative, ethical, and socially inclusive.

Index Terms: Micro-credentials, Capability Expansion, Labour market inclusion.

Funded By the EU Erasmus+ programme

1. Introduction

1.1. A. Background

Micro-credentials have rapidly emerged on the global education and training landscape as flexible, bite-sized certifications of skills and competencies. They are increasingly offered by universities, professional bodies, and online platforms to meet the needs of lifelong learners and evolving labor markets [1]. In recent years, international organizations and governments have promoted micro-credentials as a means to upskill and reskill the workforce in the face of disruptive economic and technological changes [1]. This global rise of micro-credentials is coupled with strong policy momentum. In Europe, for example, the European Commission's

Digital Education Action Plan (2021–2027) explicitly supports micro-credentials to foster digital skills and innovation in education [2]. Similarly, the European Green Deal highlights the urgent need for training in green skills to support the transition to a sustainable economy, an agenda in which micro-credentials are seen as a key tool for rapid upskilling [3]. These policy initiatives signal high-level recognition of micro-credentials as important instruments for addressing the digital transformation and sustainability challenges of the coming decades.

1) *Fragmented Conceptualisation Across Education, Sustainability, and Informatics Literatures*: Despite the global enthusiasm, the concept of “micro-credentials” is far from uniformly defined or understood across different scholarly communities. In the education literature, micro-credentials are often conceptualised as short, competency-based courses or digital badges that certify specific skills, frequently aligned with formal qualifications frameworks or professional development pathways [4]. By contrast, in the information technology and informatics literature, micro-credentials may be discussed in terms of digital badges, blockchain-enabled certificates, and platform-based credential ecosystems [5]. Meanwhile, nascent discussions in sustainability and environmental education contexts frame micro-credentials as a means to quickly equip learners with green competencies and sustainability-related skills (e.g., renewable energy management or sustainable practices) to meet emerging labour market needs [6]. These divergent perspectives have led to a fragmented understanding of what constitutes a micro-credential, with convergent themes (such as focus on specific skills and flexibility) but also divergent emphases and terminologies across domains [4][5]. This fragmentation points to the need for an integrative conceptual framework that can bridge education, sustainability, and informatics perspectives on micro-credentials. To address the above gap, this paper proposes a theoretical advance by articulating a four-component integrative model. In this model, micro-credentials serve as a catalyst for expanding learners’ capability sets that is, the range of skills, competencies, and opportunities they can effectively utilise for gainful employment and personal development. We posit that obtaining micro-credentials can enhance an individual’s human capital and agency, thereby increasing their capability to pursue meaningful work and lifelong learning. However, the extent of this enhancement is hypothesised to be moderated by the broader context of digital transformation and shaped by sustainability demands. The rapid digital transformation of industries (often termed the “Fourth Industrial Revolution”) is changing the nature of work and skills demand [7][8], thus influencing how and which micro-credentials translate into real opportunities. Likewise, the imperative of sustainable development and the green transition exerts pressure on education and training systems to produce “green-skilled” workers [3][6]. Our integrative model interlinks these components – micro-credential uptake, human capability expansion, digital transformation, and sustainability context – providing a novel framework to explain how micro-credentials may foster inclusive labor-market outcomes amid contemporary global challenges [9].

2) *Research Aim and Guiding Questions*: Building on the trends and gaps identified above, the aim of this research is to synthesise the literature and develop a unifying framework for understanding the role of micro-credentials in enhancing human capabilities for inclusive and sustainable employment. It also considers how emerging educational technologies, such as AI-driven learning platforms, might mediate the effectiveness and equity of micro-credentials [10].

Specifically, the paper is guided by the following research questions:

- RQ1: How are micro-credentials currently conceptualised across education, sustainability, and informatics literatures, and what convergent or divergent themes emerge?
- RQ2: Through which mechanisms do micro-credentials expand learners' capability sets to foster labour-market inclusion, and how are these mechanisms conditioned by digital transformation and green-skills demand?
- RQ3: What integrative theoretical model and testable propositions emerge from this synthesis to guide future empirical research?
- RQ4: How does the design and ethical governance of AI-powered learning platforms mediate or moderate micro-credential outcomes, especially for vulnerable groups?

The remainder of this article is structured as follows. Section

2 explains the methodology, including the literature search strategy and analytical approach used to address the research questions. Section 3 presents a synthesis of how micro-credentials are defined and conceptualised in different bodies of literature (addressing RQ1). Section 4 examines the mechanisms by which micro-credentials may enhance learners' capabilities and promote inclusive employment, with attention to the influences of digital transformation and green skill demands (addressing RQ2). Section 5 integrates these insights to propose an integrative theoretical model and a set of testable propositions (addressing RQ3). Section 6 discusses the role of AI-driven learning platforms in shaping micro-credential outcomes and considers ethical governance issues, particularly regarding impacts on vulnerable learners (addressing RQ4). Finally, Section 7 concludes the paper, highlighting key findings, implications for policy and practice, and directions for future research.

2. Methodology

The analysis of this paper uses the ICS (Interpretive Conceptual Synthesis) approach to enable them to construct a theoretically rigorous view of how micro-credentials can contribute to labour market and inclusion, within their broader context of green and digital transitions. ICS is well suited to areas suffering from limited conceptual unity or coherence, as it allows for an integrative, theory-building, rather than merely aggregative interpretation [11], [12].

2.1. Search Strategy

A systematic literature search was conducted across five academic databases: Scopus, Web of Science, ERIC, Education Source (EBSCO), and IEEE Xplore. The search covered the time period from January 2012 to March 2025, capturing over a decade of developments following the Bologna Process updates and the rise of micro-credentials in EU policy. Search strings combined Boolean operators and controlled vocabulary terms. Examples included: *"micro-credentials" OR "digital badges" OR "alternative credentials" AND "labour inclusion" OR "employability" OR "skills recognition" AND "green skills" OR "sustainability competencies" AND "AI in education" OR "learning analytics"* In parallel, a grey-literature sweep targeted high-level policy documents from the European Commission, United Nations Educational, Scientific and Cultural Organisation (UNESCO), Organisation for Economic Co-operation and Development OECD, and national skills agencies. Project websites such as Digital Competence Framework for Citizens (DIGCOMP), and MICROHE were manually screened to capture project-level insights and practices not yet published in peer-reviewed outlets.

2.2. Screening Criteria

A two-step screening process was applied to ensure methodological rigour:

- 1) Inclusion criteria: English-language articles; publications addressing at least one of the three domains (education, sustainability, informatics); papers engaging substantively with micro-credentials, not merely referencing them.
- 2) Exclusion criteria: Purely technical AI papers without pedagogical relevance; opinion pieces lacking empirical or conceptual depth; articles focused solely on traditional credentialing mechanisms.

The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework guided documentation of the screening process. After deduplication and initial abstract review ($n = 1,256$), full-text screening yielded a final set of 57 documents for conceptual coding as shown in Figure 1 below.

B. Coding Framework

Data was coded using a hybrid coding framework within manual coding.

- 1) Deductive domains were derived from the theoretical framework and policy background:

- Micro-credential design and delivery
- Skills recognition and RPL (recognition of prior learning)
- Human capability development (agency, self-efficacy)
- Labour market inclusion (green and decent employment)

- 2) Inductive categories emerged through iterative analysis, including:

- Ethical tensions in AI-enabled credentialing
- Platform governance and learner agency
- Fragmentation in terminology (e.g., badges vs nano-degrees)
- Green skills taxonomies and their institutional up-take

Emergent themes were documented in analytic memos, then compared with

GREEN Credentials project documentation to ensure the synthesis remained aligned with practice. To enhance interpretive validity, several bias mitigation strategies were employed:

- 1) **Intercoder Reliability:** A second coder independently coded 15% of the sample, achieving a Cohen's kappa (K) of 0.79, indicating substantial agreement [13].
- 2) **Reflexive memos:** were maintained throughout the process to document interpretive decisions, theoretical tensions, and coding shifts.
- 3) **Triangulation:** was achieved by cross-referencing emergent themes against key EU policy texts (such as European Skills Agenda, Digital Education Action Plan) and

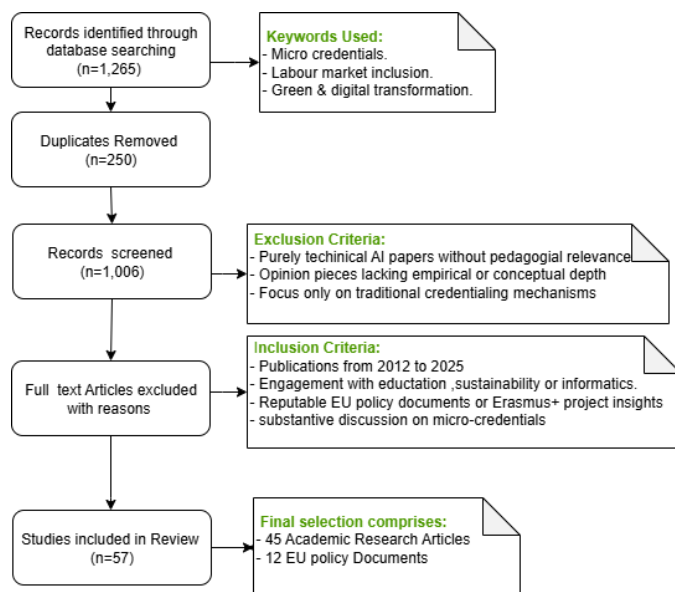


Fig 1: PRISMA flow diagram illustrating the systematic review and conceptual synthesis process. The model outlines five sequential stages: database identification, duplicate removal, abstract screening, full-text eligibility assessment, and final inclusion. Inclusion and exclusion criteria are specified, reflecting the focus on micro-credentials, labour-market inclusion, and green and digital transitions. The final corpus comprises 57 sources, including academic articles, EU policy documents, and grey literature. Green callouts highlight conceptual parameters used for thematic alignment

practitioner documents from previous Erasmus+ projects. Additionally, the review team reflected on positionality and potential policy entanglements, ensuring the synthesis was not unduly shaped by any single normative agenda.

2.3. Scope and Limitations

This review adopts a conceptual synthesis model and does not attempt meta-analysis or effect size aggregation due to the wide diversity of methodological approaches and outcome variables in the literature. The aim is not to quantify impact, but to develop a robust, integrative framework that can guide future empirical validation and programme design. Further, while the review includes EU-centric policy and project data, its conclusions may require contextual recalibration for Global South contexts, where credential infrastructure and digital equity differ substantially.

3. Sustainability and Digital Transformation Contexts

Micro-credentials do not exist in a vacuum; their relevance and design are shaped by broader transformations—particularly the dual forces of sustainability imperatives and digital innovation. This section situates the GREEN-Credentials framework within these two structural contexts, analysing how sector-specific green transitions and technological shifts influence the skills landscape, and by extension, the role of micro-credentials in advancing inclusive labour market participation.

3.1. Environmental Informatics in Sustainable Sectors

- 1) *Skills Needs in Agriculture and Forestry:* The agriculture and forestry sectors are undergoing a rapid transformation, driven by the twin imperatives of climate resilience and digital innovation. To meet sustainability goals, emerging competency frameworks highlight the need for integrated skill sets that address environmental challenges while embracing technological solutions. One critical area is climate-resilient crop production, which equips workers with the knowledge to adapt agricultural practices to increasingly unpredictable weather patterns, ensuring food security and ecosystem stability. Similarly, precision farming and drone-assisted monitoring enable real-time data collection and efficient resource use, reducing environmental impact while improving yields. Equally important is expertise in soil biodiversity and carbon sequestration techniques. These practices not only preserve long-term soil health but also contribute directly to climate mitigation by capturing atmospheric carbon. In the forestry domain, ecosystem-based management is essential to align timber production with the protection of biodiversity and the provision of carbon sinks. Yet, these specialised competencies are often siloed within advanced academic or vocational tracks, creating access barriers for many—particularly migrants, low-qualified individuals, or rural populations. Micro-credentials, when designed as short, stackable, and job-relevant learning units, provide flexible entry points into these fields. They validate discrete capabilities, empower learners to build toward larger qualifications, and support inclusive participation in the green transition Trienekens et al. [14] note that while these areas are rich in innovation potential, the associated knowledge domains are often siloed within academic or vocational pathways that are inaccessible to marginalised groups such as migrants or low-qualified workers. Micro-credentials—particularly when designed as modular, stackable learning units can reduce entry barriers into such knowledge domains, allowing learners to certify discrete

Micro-credentials for Sustainable Agro-forestry Skills

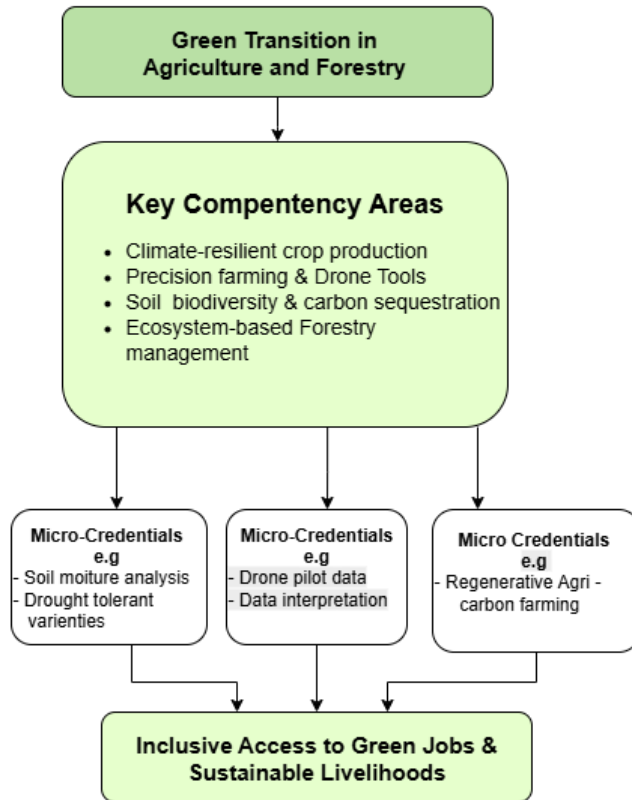


Fig. 2. Micro-Credentials for Sustainable Agri-Forestry Skills

competencies relevant to sustainable sectoral work. As shown in Figure 2, the diagram maps the green transition in agriculture and forestry to four competence areas, links each area to an example micro credential that certifies practical skills, and shows how these units support inclusive access to green jobs and sustainable livelihoods.

2) *Environmental Informatics and Credential Tracking Platforms:* Environmental informatics which has the integration of environmental science, data analysis, and digital infrastructure is increasingly used to monitor and guide the development of green skills[15]. Projects like the EU Agri- Food Skills Observatory and the Bioeconomy Skills Monitor are pioneering data platforms that identify gaps in labour market supply and demand for sustainability-related skills [15]. These platforms have three key implications for micro- credential design:

1) Evidence-Based Curriculum

Development: By mapping granular skill needs in real-time, they support the agile development of learning modules tailored to labour market demands.

- 2) **Credential Interoperability:** Data standardisation across EU member states allows micro-credentials to be portable and comparable, which is critical for migrant and mobile workers.
- 3) **Feedback Loops for RPL Systems:** The integration of micro-credential achievements into sectoral databases facilitates the recognition of informal and non-formal learning, enhancing RPL effectiveness.

Environmental informatics thus acts as a bridge between

green policy ambitions and credential infrastructure, aligning learning systems with sustainability transitions while also supporting equitable access to these new economic opportunities.

3.2. Artificial Intelligence and Digital Credentialing

Parallel to the sustainability transition is the rise of Artificial Intelligence (AI) and broader digital transformation, often referred to as the Fourth Industrial Revolution (4IR). The automation of routine tasks, algorithmic decision-making, and datafication of learning are changing how skills are acquired, measured, and recognised [16]. This has major implications

for the micro-credential ecosystem—both as an enabler of personalised, just-in-time learning and as a potential vector for exclusion. AI technologies are disrupting job profiles in both low-skilled and high-skilled domains. For example, while manual agricultural labour may decline with increased mechanisation, new roles emerge in agri-tech operations, data interpretation, and equipment maintenance. However, these roles often require hybrid skillsets that combine domain knowledge with digital fluency—precisely the space where micro-credentials can offer rapid upskilling or reskilling. The challenge lies in ensuring that vulnerable groups are not left behind as legacy skills become obsolete. This is particularly acute in rural or under-resourced regions where digital infrastructure is limited, and where workers may not have access to devices or the digital literacy required to engage with online learning platforms [17].

3) *Digital Badges, Blockchain, and Learning Analytics*: Emerging digital credentialing technologies promise greater trust, transparency, and learner control. For instance:

- Digital badges represent verifiable achievements that can be shared across platforms like LinkedIn or Europass.
- Blockchain-secured micro-credentials enable tamper-proof storage and decentralised verification, which can be critical for migrants lacking traditional documentation.
- Learning analytics and AI-based assessment allow platforms to track learner engagement, recommend next steps, and provide feedback in real time.

These features contribute to what the GREEN-Credentials platform refers to as a “smart credentialing infrastructure” which is a digital learning ecosystem in which credential pathways are modular, transparent, and responsive to learner needs. However, these innovations also raise critical governance questions. As micro-credentials become increasingly mediated by data and algorithms, concerns around surveillance, bias, and learner autonomy emerge [18].

4) *Risks of Digital Exclusion and Principles for Inclusive Design*: The risk is that the same systems designed to promote flexibility and access may replicate or deepen existing inequalities. Key challenges identified in literature include:

- 1) Algorithmic bias: AI-driven assessments may underperform with non-native language users or neurodiverse learners.
- 2) Data privacy: Learners may not fully understand how their behavioural data is collected, used, or monetised.
- 3) Credential fragmentation: A proliferation of platforms and standards can create confusion rather than empowerment, particularly for disadvantaged learners.

To mitigate these risks, inclusive design principles must be embedded at every level of digital credentialing:

- 1) Multilingual and culturally sensitive interfaces
- 2) Mobile-first design for low-connectivity settings
- 3) Transparency in AI decision-making processes
- 4) User control over data sharing and portability Furthermore, alignment with frameworks like the EU Artificial Intelligence Act and UNESCO’s Ethics of AI in Education

guidelines is essential to ensure micro-credential systems are

ethically grounded and socially equitable [19]. Sustainability and digital transformation are not merely background trends, they are structural forces reshaping the meaning and accessibility of work, skills, and credentials. As shown in this section, the intersection of environmental

informatics and AI- driven learning platforms presents both an opportunity and a challenge. On one hand, it allows for the agile, data-informed deployment of micro-credentials that are tightly coupled with real labour market needs. On the other hand, it risks rein- forcing exclusion unless explicit design attention is given to issues of access, control, and equity. For micro-credentials to truly serve as tools of empowerment especially for migrants, refugees, and low-qualified workers they must be embedded in a digital ecosystem that is not only smart and scalable, but also fair and human-centred. This requires governance frameworks that go beyond technological feasibility to consider ethical accountability, learner agency, and systemic inclusion.

4. Integrated Theoretical Framework and Propositions

Figure 2 depicts a moderated-process model comprising four primary constructs namely 1) Micro-Credential Design & Recognition (independent variable), 2) Capability Expansion (mediator), 3) Labour-Market Inclusion (outcome), and two contextual moderators: Digital-Transformation Intensity & AI- Governance Quality and Green-Skills Demand. Recognition of prior learning (RPL) is nested within the micro-credential box to emphasise its facilitating role, while dashed arrows indicate moderation effects. To translate the graphical logic of Figure 1 into a testable analytical framework, each node and pathway must be anchored in precise, literature-derived constructs. Es- tablishing clear boundaries and operational definitions not only ensures conceptual coherence across education, sustainability, and informatics perspectives, but also provides a reproducible coding scheme for the forthcoming systematic review. Accord- ingly, the study delineates five core constructs—three causal components and two contextual moderator, drawing on the most commonly accepted definitions in the micro-credential, capability, and green-skills literatures [3], [4], [6], [7], [9], [10]. These constructs are summarised below.

- 1) Micro-Credential Design & Recognition: Short, stack- able, competency-based certifications issued or verified through digital platforms; includes robustness of RPL and alignment with qualification frameworks [4], [9].
- 2) Capability Expansion: Growth in learners’ knowledge, agency, and self-efficacy that enlarges their set of feasi- ble opportunities, consistent with capability theory.

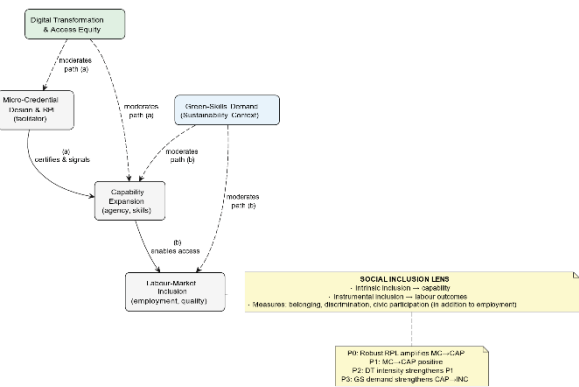


Fig. 3. Moderated-process model and theoretical framework comprising (i) Micro-Credential Design and Recognition (independent variable), (ii) Capa- bility

- 3) Labour-Market Inclusion: Entry into, and progression within, decent and sustainable employment, encompass- ing job quality and social-inclusion indicators [7].
- 4) Digital-Transformation Intensity & AI- Governance Quality: Sectoral prevalence of AI/automation plus the extent to which learning platforms follow transparent, bias- mitigating governance principles [10].
- 5) Green-Skills Demand: Labour-market demand for com- petencies aligned with the European Green Deal and related sustainability agendas [3], [6].

4.1. Directional Propositions

Grounded in the construct definitions above and aligned with Research Questions 2–4, the model yields three directional propositions that specify how micro-credentials are expected to translate into equitable labour-market outcomes. Proposition P1 articulates the core causal mechanism: capability expansion mediates the influence of micro-credentials on inclusion. Propositions P2 and P3 introduce contingent factors—digital-transformation intensity, AI-governance quality, and green-skills demand—that theory and policy analyses identify as critical amplifiers or dampeners of credential value in contemporary labour markets [3], [6], [7], [10]. Together, these propositions convert the conceptual model into a falsifiable programme for empirical testing and policy evaluation.

- **P1** Micro-credentials improve labour-market inclusion when they demonstrably expand learners' capability sets.
- **P2** The Micro-credential → Capability link (P1) is strengthened in sectors undergoing rapid digital transformation and governed by high-quality, ethical AI practices.
- **P3** Rising demand for green skills amplifies the Capability → Labour-Market-Inclusion pathway, making sustainability-aligned competences more valuable.

4.2. Mediating and Moderating Pathways

The model posits a two-stage causal chain. First, well-designed micro-credentials, especially those embedded in robust RPL systems enhance capability by certifying existing skills, signalling new competencies, and improving learner agency. Second, expanded capability translates into concrete labour-market outcomes when employers recognise these credentials and when learners can leverage new skills in growing sectors. Two contextual forces condition these relationships which form proposition 0 and 4:

- 1) **P0** Digital-Transformation Intensity & AI-Governance Quality moderate Path (a) (Micro-credential → Capability). In digitally advanced environments, AI-powered platforms can personalise learning and credential verification, thereby magnifying capability gains; conversely, poor governance or algorithmic bias may restrict benefits for vulnerable groups [10].
- 2) **P4** Green-Skills Demand moderates Path (b) (Capability → Inclusion). Where sustainability agendas drive job growth, capability gains that include green competences yield higher employability premiums [3], [6].

TABLE I CONSTRUCTS AND ILLUSTRATIVE INDICATORS

Construct / Path	Illustrative Indicators (Literature Based)
Micro Credential Design and Recognition	Stackability index; credit value; presence of formal RPL procedure; interoperability with national qualification frameworks [4], [9]
Capability Expansion	Self efficacy scale scores; breadth of digital portfolio; validated competence assessments
Labour Market Inclusion	Six month employment status; wage progression; job quality index; perceived social belonging scale
Digital Transformation Intensity	Sectoral automation index; AI adoption rate; availability of adaptive learning analytics
AI Governance Quality	Existence of bias audits, transparency policies, and learner data protection measures [10]
Green Skills Demand	Sectoral employment growth in EU Taxonomy “green” occupations; frequency of green skills vacancies [3], [6]

4.3. Operational Indicators for Empirical Testing

Table 1 shows indicators provide a clear roadmap for coding empirical studies in the forthcoming systematic review and enable quantitative tests of Propositions P1–P3.

5. Implications and Contributions

5.1. Strategic (Policy-Level) Implications

- 1) Alignment of qualification frameworks. The Council Recommendation on Micro-Credentials encourages Member States to embed short, stackable certifications within national qualifications frameworks (NQFs) and to map them transparently to the European Qualifications Framework (EQF) [9]. Our findings reinforce this directive by showing that capability gains translate into measurable inclusion only when employers and public employment services recognise micro-credential metadata. Policymakers should therefore (i) create EQF-referenced descriptors for levels 3–6 that accommodate micro-learning outcomes, and (ii) formalise recognition of prior learning (RPL) within NQF statutes, thereby amplifying the RPL facilitator pathway identified in Proposition P0.
- 2) Funding levers for inclusive green upskilling. The European Green Deal and Skills Agenda earmark substantial funds under ESF+, Erasmus+, and the Just Transition Mechanism for green-skills development [3], [7]. To avoid allocative bias toward already-skilled workers, these instruments should adopt capability-weighted funding formulas: higher reimbursement rates for providers that demonstrably enrol migrants, refugees, or low-qualified adults. Such conditional funding would operationalise our model's equity mandate by linking public investment to demonstrable inclusion outcomes rather than enrolment volumes alone. International agencies (e.g., ILO) can adapt similar formulas for global South contexts where micro-credentials are emerging as rapid-response tools for green transitions [6].

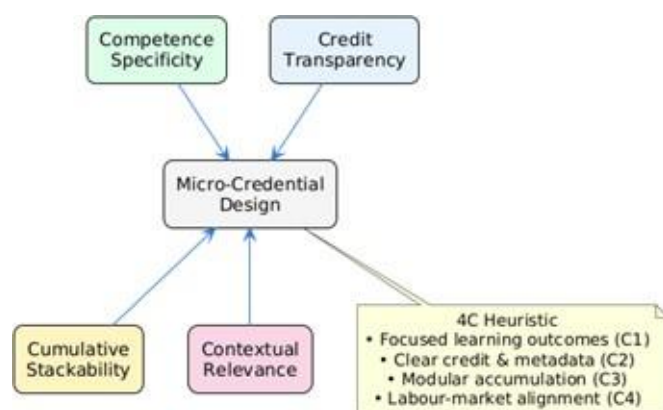
5.2. Operational (Practitioner-Level) Implications

Design principles for credential structure for assessment, and recognition. Micro-credential providers should adhere to the “4 C” design heuristic distilled from the literature that is (competence specificity, credit transparency, cumulative stackability, and contextual relevance) as shown in Figure 3 [4], [5]. Assessment must privilege authentic performance tasks over multiple-choice recall to ensure that issued credentials genuinely reflect capability expansion. Embedding machine-readable metadata (e.g., EQF level, learning hours, assessment rubric) supports cross-border recognition and feeds directly into environmental-informatics skills observatories [21]. AI-driven analytics for personalised pathways. Adaptive learning engines, when ethically governed, can diagnose individual knowledge gaps, recommend micro-credential sequences, and trigger RPL assessments, thereby accelerating the Micro-credential → Capability link in P1. However, inclusion gains hinge on algorithmic transparency and bias auditing, as emphasised by UNESCO's AI ethics framework [10] and recent empirical work on algorithmic decision-making in education [24]. Practitioners should:

- Implement multilingual interfaces and accessibility standards to mitigate digital-exclusion risks highlighted by Selwyn [23].
- Publish explainability dashboards that disclose how recommendation algorithms weight learner data [22].

- Offer “offline-sync” or mobile-first delivery modes to support low-bandwidth regions, particularly in rural green-economy contexts.

5.3. Theoretical Contributions and Positioning



The proposed moderated-process model extends prior micro-credential scholarship in three ways:

1) Capability-centric mediation. Whereas most extant models adopt a human-capital lens that emphasises earnings returns, our framework situates capability expansion as the mediating engine linking credentials to inclusion, thereby incorporating intrinsic dimensions of social participation absent

from human-capital or technology- acceptance models.

2) Dual contextual moderators. By integrating digital- transformation intensity (with an explicit AI-governancequality component) and green-skills demand as moderators, the model

Fig. 4. The 4C Design heuristics for Micro-Credentials

operationalises macro-structural forces often treated only descriptively in prior studies [8], [20]. This adds analytical

precision, enabling falsifiable interaction hypotheses (P2, P3).

3) Ethics-aware digital ecosystem perspective. The framework positions ethical AI governance not as an ex-post concern but as a constitutive factor shaping capability gains, thereby bridging emerging ed-tech critiques [22] with inclusion-oriented credential research.

By foregrounding capability expansion, modelling dual contextual moderators, and embedding AI-ethics considerations, the present framework extends existing micro-credential theories and offers a more comprehensive, equity-focused lens for future empirical work and policy design as shown in Table

2. Collectively, these contributions deliver a more holistic explanation of how micro-credentials function within digitally mediated, sustainability-driven labour markets and provide a robust scaffold for subsequent quantitative tests and policy evaluations.

6. Limitations, Reproducibility, and Future Research

This review draws exclusively on sources contained in the Basque International Research Association (BIRA) literature worksheet, which is dominated by English-language publications and European policy documents. Although this focus aligns with the GREEN-Credentials context, it limits generalisability to non-European jurisdictions and non-English discourses where micro-credential ecosystems may differ markedly. Moreover, within the sustainability domain the corpus is weighted toward agriculture and forestry, leaving sectors such as renewable energy or circular construction under-represented. Findings should therefore be interpreted as sector-specific rather than universally applicable. The study’s interpretive conceptual synthesis prioritises theoretical integration over effect-size aggregation; there was no conduct of statistical meta-analysis. While rigorous inclusion criteria and intercoder

agreement (0.79) mitigate subjective bias, the deductive coding frame may still have privileged constructs that fit the proposed model, inadvertently overlooking alternative explanatory variables. In addition, grey literature of uncertain quality was excluded, potentially omitting innovative but unpublished practices. To enhance transparency, full search strings, and a PRISMA flow diagram are provided in the paper. These assets enable replication of search and screening procedures, facilitate re-analysis with updated datasets, and support extensions that incorporate additional languages or sectors. Future research can look into the following avenues:

- 1) Quantitative testing of propositions. The five propositions (P0–P4) can be examined through multi-level regression or structural equation modelling using survey or administrative data that link micro-credential attainment to employment outcomes.
- 2) Sectoral transferability. Replicating the framework in energy, construction, or health-care sectors will reveal

TABLE II POSITIONING OF THE PRESENT FRAMEWORK VIS A VIS DOMINANT MICRO CREDENTIAL MODELS

Analytical dimension	Dominant models in the literature (illustrative sources)	Present capability centric framework	Significance for future research and policy
Outcome construct	Human capital lens, micro credentials valued primarily for wage returns or credit accumulation [4], [5].	Capability expansion mediates inclusion; accounts for agency, self efficacy and social participation.	Broadens evaluative metrics beyond earnings to intrinsic inclusion; supports SDG aligned policy objectives.
Contextual variables	Macro forces such as digitalisation and the green transition acknowledged descriptively but not modelled causally [1], [3], [6].	Two explicit moderators: (i) digital transformation intensity and AI governance quality, (ii) green skills demand.	Enables falsifiable interaction hypotheses (P2, P3); informs context sensitive programme design and funding formulas.
Digital ethics perspective	EdTech studies highlight risks from datafication but treat them as downstream concerns [13].	Ethical AI governance embedded as a constitutive factor shaping capability gains and equity outcomes.	Integrates UNESCO AI ethics norms [16] into the theoretical core; provides design criteria for inclusive platforms.
Equity orientation	Many models assume general adult learners; marginalised cohorts discussed separately [7].	Migrants, refugees and low qualified workers targeted as primary beneficiaries; RPL positioned as a facilitator (P0).	Aligns micro credential funding and quality assurance with social inclusion mandates; operationalises capability weighted funding levers.
Credential artefact	Predominantly university issued digital badges; limited focus on interoperability [4].	Emphasises interoperable, metadata rich artefacts aligned to NQF/EQF and to green skills taxonomies [9], [12].	Supports cross border portability and integration into environmental informatics skills observatories.
Empirical testability	Prior conceptual papers offer narrative propositions or single context case studies.	Provides five numbered propositions (P0 to P4) with operational indicators in Section 5.5.	Establishes a roadmap for mixed methods validation and comparative sectoral analyses.

whether the moderating effects of digital transformation and green-skills demand vary across sustainability domains.

- 3) Longitudinal analyses. Panel studies tracking learners before and after credential completion would clarify temporal dynamics, including whether capability gains persist and translate into long-term labour-market mobility.
- 4) AI-governance efficacy. Experimental or quasi-experimental designs comparing platforms with and without bias-mitigating features could quantify the equity impact posited in Proposition P2.
- 5) Non-European contexts. Extending the study to Global South settings will test the framework's robustness where digital infrastructure and policy environments differ, and may necessitate additional moderators such as connectivity or informal-economy prevalence.

Recognising these limitations and delineating concrete research pathways, positions the framework as a cumulative and testable contribution rather than a definitive account. Contin-

uous empirical validation across contexts and time frames is essential to refine the model and to inform inclusive, evidence based policy on micro credentials.

7. Conclusion

This subsection summarises the theoretical advance by answering each research question directly. We first show how the three literatures conceptualise micro credentials and where their themes converge or diverge, which resolves the initial fragmentation and sets the scope for synthesis. We then explain the mechanism by which well designed micro credentials, supported by recognition of prior learning, expand capability and how this mechanism is shaped by digital transformation and by demand for green skills. Next, we present the integrated model and its testable propositions that follow logically from this evidence. Finally, we consider how the design and ethical governance of AI enabled learning platforms mediate or moderate outcomes, with particular attention to migrants, refugees, and other low qualified learners. Taken together, these elements establish a clear, capability centred account of how micro credentials can support equitable labour market inclusion in green and digitally evolving sectors.

- **RQ1. Conceptualisations across fields.** The review shows clear areas of convergence and divergence in how micro credentials are understood across education, sustainability, and informatics. All three literatures value short, competency based learning that is portable and stackable. Education emphasises alignment with qualification frameworks and assessment quality. Sustainability prioritises sector specific green competences and rapid workforce conversion. Informatics focuses on verifiable digital artefacts and platform interoperability. This mapping resolves the fragmentation noted in the introduction by showing where definitions overlap and where purposes differ, which provides a coherent basis for synthesis.
- **RQ2. Mechanisms and conditioning factors.** The evidence supports a capability centred mechanism. Well designed micro credentials, especially when supported by recognition of prior learning, expand knowledge, agency, and confidence. These capability gains increase the likelihood of decent and sustainable employment. The strength of this pathway is not constant. It is shaped by the intensity of digital transformation and by the level of demand for green skills in the relevant sector. Where digital adoption is high and where green demand is rising, capability gains are more likely to translate into inclusion.
- **RQ3. Integrative model and propositions.** On the basis of these findings the paper sets out a four component model that links micro credential design and recognition to labour market inclusion through capability expansion, with digital transformation and green skills demand as moderators. The model is expressed through a small set of testable propositions, including the facilitating role of recognition of prior learning, the positive effect of micro credentials on capability, and the two contextual moderators. This provides a clear programme for empirical testing and a structure for future evidence synthesis.
- **RQ4. Role of AI powered platforms and ethical governance.** Design and governance choices in AI supported learning environments mediate or moderate outcomes. Personalisation, diagnostic feedback, and verifiable digital records can strengthen capability gains and improve signalling to employers. Without transparent algorithms,

bias auditing, and meaningful learner control of data, these same systems can widen exclusion, especially for migrants and other vulnerable groups. Ethical governance is therefore not an afterthought but a necessary condition for equitable outcomes.

Conceptually, the framework extends existing models by:

- Elevating capability expansion—rather than wage gain—as the pivotal mediator, thereby aligning micro-credential research with twenty-first-century inclusion discourse;
- Operationalising digital and green transitions as testable moderators, moving them from narrative background to analytic foreground; and
- Integrating AI-ethics principles as causal conditions, bridging critical ed-tech studies with credential theory [10], [22].

For practitioners and policymakers, the study distils a “4 C” design heuristic—Competence specificity, Credit transparency, Cumulative stackability, and Contextual relevance and links it to operational indicators that can populate skills observatories and NQF dashboards. Additionally, the paper proposes capability-weighted funding formulas that tie public investment to demonstrable inclusion outcomes, offering an actionable road map for instruments such as The European Social Fund Plus (ESF+) and EU programme for education, training, youth and sport (Erasmus+) [3].

Sustainability imperatives and rapid digitalisation are not merely parallel trends; they are intertwined forces redefining what counts as work, skill, and learning. Micro-credentials, when ethically governed and linked to robust RPL systems, can serve as bridges between displaced workers and emergent green-digital occupations, particularly for migrants, refugees, and other low-qualified adults. Yet the same AI-driven platforms that personalise learning can entrench inequities if algorithmic transparency and data-justice safeguards are absent. Future credential ecosystems must therefore be designed as human-centred socio-technical systems that couple agile, data-informed learning pathways with strong protections for learner agency and privacy. Empirical validation of the five propositions set out in this paper, through longitudinal, multi-sectoral, and cross-regional studies will be essential to refine the model and to guide evidence-based policy formulation. In doing so, the research community can ensure that micro-credentials fulfil their promise: which is to unlock human capability, power inclusive (green) transitions, and render digital transformation an engine of opportunity rather than exclusion.

8. Acknowledgment

This paper is part of the GREEN Credentials project, a three-year Erasmus+ initiative that advances vocational education and training in sustainability, agriculture, and forest management. The authors thank the GREEN Credentials consortium of ten partners across eight countries for collaboration, technical input, and constructive feedback on the framework and pilots. The University of Oldenburg contributed expertise in artificial intelligence, data analytics, and sustainability and led the design and implementation of smart learning and assessment tools. Appreciation is also extended to the VET providers, trainers, and learners who informed the development of the micro-credential design and evaluation activities.

9. References

- [1] Abdel Fattah et al., “The potential skilling, upskilling, and reskilling opportunities for the migration and mobility of workers, with a specific focus on gender aspects of workers in the countries of origin,” 2024, doi: 10.5281/ZENODO.14523918.
- [2] Anslinger et al., “Dimensions of validation of prior learning in Eu- rope: Empirical insights from Denmark, Poland, Turkey and Germany,” 2022. [Online]. Available: [https://www.pedocs.de/volltexte/2024/30015/pdf/Anslinger et al 2022 Dimensions of validation.pdf](https://www.pedocs.de/volltexte/2024/30015/pdf/Anslinger_et_al_2022_Dimensions_of_validation.pdf)
- [3] M. J. I. Ayuso, “Decoding micro-credentials: Analysis of initial in- tentions and early implementations in the Erasmus+ framework: From theory to practise,” *J. Learn. Develop.*, vol. 12, no. 1, 2025, doi: 10.56059/jl4d.v12i1.1355.
- [4] Bourke et al., “‘Teaching is a learning experience’: Exploring faculty engagement with low-income adult learners in a college-community partnership program,” *Can. J. Educ.*, 2020. [Online]. Available: <https://www.jstor.org/stable/26954690>
- [5] U. Brandi et al., “The European year of skills 2023: Skills for now and in the future?,” *Int. J. Lifelong Educ.*, 2023, doi: 10.1080/02601370.2023.2212424.
- [6] V. Cvetkoska et al., “Skill matching and migration links in the North Macedonia, Turkiye, Ukraine and Ethiopia,” 2025. [Online]. Available: [https://skills4justice.eu/wp-content/uploads/2025/02/Working-Paper-Synthesis-Report-from-WP3-January-2025-UKIM-final-2. pdf](https://skills4justice.eu/wp-content/uploads/2025/02/Working-Paper-Synthesis-Report-from-WP3-January-2025-UKIM-final-2.pdf)
- [7] Dalipi, M. Ferati, and A. Kurti, “Integrating MOOCs in regular higher education: Challenges and opportunities from a Scandinavian perspective,” in *Learning and Collaboration Technologies: Design, Development and Technological Innovation*, 2018, doi: 10.1007/978-3- 319-91743-6 15.
- [8] F. Dunkerley et al., “Labour market and skills demand horizon scanning and future scenarios,” Research Report, Department for Education, 2022. [Online]. Available: [https://assets.publishing.service.gov.uk/media/628ca1fce90e071f6af1462a/Labour Market and Skills Demand Horizon Scanning and Future Scenarios FINAL.pdf](https://assets.publishing.service.gov.uk/media/628ca1fce90e071f6af1462a/Labour_Market_and_Skills_Demand_Horizon_Scanning_and_Future_Scenarios_FINAL.pdf) . Farkas, “University lifelong learning in Hungary – The role and potential of universities to support the development of lifelong learning,” *Eur. J. Univ. Lifelong Learn.*, 2022, doi: 10.53807/0602glhf.
- [9] O. Glumac and F. Pallonetto, “Education needs and ūpskilling opportunities in the renewable energy sec- tor,” Deliverable 3.4, 2023. [Online]. Available: <https://www.res4city.eu/wp-content/uploads/2024/11/RES4CITY-D2.3-A-multidisciplinary-microcredential-basket-for-the-sustainable-transition-in-cities PU v1.0.pdf>
- [10] Dixon-Woods et al., “Conducting a critical interpretive synthesis of the literature on access to healthcare by vulnerable groups,” *BMC Med. Res. Methodol.*, vol. 6, p. 35, Jul. 2006.
- [11] Boell and D. Cecez-Kecmanovic, “A hermeneutic approach for conducting literature reviews and literature searches,” *Commun. Assoc. Inf. Syst.*, vol. 34, no. 1, pp. 257–286, 2014.

- [12] R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," *Biometrics*, vol. 33, no. 1, pp. 159–174, Mar. 1977.
- [13] Trienekens, L. Klerkx, and M. Bacco, "Skills and competencies for sustainable agri-food systems: A review," *Sustainability*, vol. 14, no. 3, pp. 1123–1141, 2022.
- [14] EU Agri-Food Skills Observatory, "Skills Intelligence Reports," [On- line]. Available: <https://skills4agrifood.eu>
- [15] Williamson and R. Hogan, "The datafication of education: A critical approach to emerging edtech regimes," *Learn. Media Technol.*, vol. 46, no. 1, pp. 7–20, 2021.
- [16] S. Selwyn, "Digital exclusion in education: Addressing the challenge,"
- [17] *Br. J. Sociol. Educ.*, vol. 41, no. 6, pp. 756–770, 2020.
- [18] M. Prinsloo and N. Slade, "Algorithmic decision-making in education: Ethical implications for micro-credentials," *J. Learn. Analytics*, vol. 10, no. 1, pp. 65–80, 2024. UNESCO, Recommendation on the Ethics of Artificial Intelligence. Paris, 2021.
- [19] S. Huckstep and H. Dempster, "Meeting skill needs for the global green transition: A role for labour migration?," *Policy Papers 318*, Centre for Global Development, 2024. [Online]. Available: <https://ideas.repec.org/p/cgd/ppaper/318.html>
- [20] F. Iniesto et al., "Introducing a reflective framework for the assessment and recognition of microcredentials," *Open/Technol. Educ., Soc., Schol- arsh. Assoc. J.*, vol. 2, no. 2, 2022, doi: 10.18357/otessaj.2022.2.2.37.
- [21] Kemisso, "The vulnerable graduates? Exploring the post-bachelor's degree un/employment experiences of Somali and Eritrean refugees in Ethiopia," *Ethiop. J. Educ.*, 2023. [Online]. Available: <https://www.jstor.org/stable/48815833>
- [22] R. C. Levin, "Online learning & the transformation of global higher education," *Daedalus*, 2024. [Online]. Available: <https://www.jstor.org/stable/48774451>
- [23] M. Macauley, R. Duvekot, and Y. J. Berthier, "Non-formal and in- formal learning: A gateway to lifelong learning for all. The case of migrants and refugees," in *Third Int. Handbook Lifelong Learn.*, 2022. [Online]. Available: https://link.springer.com/referenceworkentry/10.1007/978-3-030-67930-9_65-1
- [24] M. Pavlova and M. Singh, "Recognizing green skills through non- formal learning: A comparative study in Asia," *Educ. Sustain.*, 2022, doi: 10.1007/978-981-19-2072-1.

This work forms part of the GREEN Credentials project and is co-funded by the Erasmus+ Programme of the European

Union (Key Action 2, Cooperation Partnerships in VET;). The views expressed are those of the authors and do not necessarily reflect the position of the European Union or the European Commission.

Social Architecture - The near future of distributed acting and learning

Otto Schell

IGDCR - Institute for Global Digital Creativity and Relevance pw by OS-Enterprise GmbH,
Deutschland

Within the new digital environment we have different potential to introduce new business models, if we keep in mind that a connected world will use intelligent sensors and intelligent infrastructures supported by self-driving software through any kind of agents. Anticipate intelligence networks, we can also run different models on co-opetition human-machine-human. Therefore the next phase will have not only human consumers participating in the market but also intelligent assets of companies will interact with intelligent assets from many other companies, and these assets will all be consumers based on intelligent services. Having addressed this learning will get also federated dimensions.

The main difference between running digital today and in the 1980s when the concept of CIM (Computer Integrated Manufacturing) was the latest disruptor is as simple as the availability of realtime technology because real-time technology gives us the potential to connect everything with everything else. In the early days of CIM, ideation around E2E processes already took place, but later in the 90s, with the rise of ERP, companies started aiming to “clean house” and began investing more and more in process standardization and optimization.

Nevertheless, when observing companies now, one can see that they are still in the old mindset where they want to “clean house and optimize” their own businesses instead of looking for opportunities to connect their companies to others. If you ask yourself, in your current business environment, to what extent your blueprints are aligned with those from your network, you should not be ashamed if the answer is zero (even though we all know that the IoT will give us more opportunities to work in a cross-silo and cross-functional way). When we speak about social architectures, we are discussing the combination of processes, data, and values related to a consumer. First how to integrate customers in all aspects, and then how to use devices to move consumers from the current business model and attract them to new ones.

Doing these things requires different kinds of E2E processes, different ways of relating to consumers, different ways of acting towards competitors along with knowing that potential consumers in the IoT world are also visible to all other participants in the market. Capabilities such as AI or Blockchain, quantum computing, open borders in IoT will make different business models and kinds of collaboration possible.

As companies journey toward a future where data security and actionable insights coexist harmoniously, these technologies stand as essential enablers of progress and will be the base for those new business models are possible. This will require also fast adoptable skill sets.

Looking in to todays challenges about project lead times in ERP or similar projects, we could anticipate timing / performance different and furthermore also strategical collaborating with business partners/networks. Having the possibility to work in an anonymized way with data, we are also able to project business models in metaverses as one visioning approach how we have to transform models.

The software landscape in companies has often grown historically and is heterogeneously shaped by decentralized decisions for different systems as well as by individual in-house developments or company mergers. For example, ERP or CRM systems from different vendors are often used in different divisions or on the basis of different, outdated release levels. In order to compete with tomorrow's digital companies, a consistent inventory of your business situation is required – which analyses the following basic dimensions:

With highly platform and modular driven system landscape, significant cost reductions (synergies in operation, licenses, maintenance, further development, etc.) can be generated – also by replacing individual systems with standard software.

One of the biggest challenges in the day-to-day operations of most companies is an ever increasing reliance on manual processes and customer-specific tools and systems, often supported by classic office applications as a bridging technology. Companies that continue to operate without intelligent automation of their processes will have to face these challenges even more in the future. The dynamic environment of companies encourages all decision-makers to work less administratively and instead to invest more time in strategic considerations and exchanges with colleagues and customers about new IT or business models.

In addition to the above-mentioned aspects and solutions, the terms cloud and cloud computing are on everyone's doorstep. While technology is not the only major digital technology, it provides central access to other technologies such as artificial intelligence, the Internet of Things (IoT) or blockchain. The cloud is considered the heart of digital transformation. But the move to the decentralized IT infrastructure wants to be well structured in order to avoid a manufacturer dependency, in this case from the respective cloud provider. If you thought with open source software and cloud solutions, the vendor lock would be a thing of the past.

Here "architects" will take the role of orchestration / mentoring / anticipation. Another set of skills to be addressed

Störungen in Produktionssystemen Eine strukturierte Literaturanalyse zur Resilienzsteigerung

J. Kallisch & J. Marx Gómez

OFFIS Institut für Informatik, Oldenburg, Deutschland

Carl von Ossietzky Universität Oldenburg, Oldenburg, Deutschland

Abstract: Produktionssysteme sind zunehmend anfällig für interne und externe Störungen, deren Auswirkungen durch globale Just-in-time-Lieferketten zusätzlich verstärkt werden. Diese Ereignisse führen regelmäßig zu erheblichen wirtschaftlichen Schäden. Folglich wird die Resilienz, verstanden als dynamische Anpassungsfähigkeit eines Produktionssystems, zum entscheidenden Wettbewerbsfaktor. Ziel dieser Arbeit ist es, den aktuellen Forschungsstand zu Störungen und Strategien der Resilienzsteigerung systematisch zu analysieren. Durch eine strukturierte Literaturanalyse wurden relevante Veröffentlichungen aus den Bereichen Produktionsmanagement und Operations Research seit 2010 ausgewertet. Die Ergebnisse zeigen, dass Resilienz deutlich über reine Robustheit hinausgeht und proaktive Anpassungsfähigkeiten wie Flexibilität, Agilität und kontinuierliches Lernen verlangt. Für Unternehmen bedeutet dies, dass nur eine ganzheitliche Kombination aus Redundanzen, flexiblen Prozessen, vorbeugender Instandhaltung sowie reaktivem und lernorientiertem Störungsmanagement langfristige Wettbewerbsvorteile bietet. Die Studie unterstreicht, dass Produktionsunternehmen eine gezielte und strategisch abgestimmte Integration dieser Resilienzmaßnahmen anstreben sollten.

Keywords: Resiliente Produktionssysteme, Produktionsplanung und -steuerung, Produktionsstörungen, Resilienzstrategien

1. Einleitung

In dynamischen, vernetzten Produktionssystemen gehören unvorhergesehene Störungen zur Normalität. Externe Schocks – von Naturkatastrophen über Pandemien bis hin zu geopolitischen Ereignissen – treffen auf global just-in-time ausgerichtete Lieferketten und können tiefgreifende, langfristige und negative Veränderungen auslösen. Bereits kurzfristige Unterbrechungen können nachhaltige Auswirkungen entlang ganzer Wertschöpfungsketten haben.

Eine aktuelle Studie zeigte, dass etwa jede fünfte Lieferkettenstörung Schäden in Millionenhöhe verursacht (Bundesverband Materialwirtschaft Einkauf und Logistik 2019). Folglich rückt die Resilienz produzierender Unternehmen – verstanden als Fähigkeit, trotz unerwarteter Einflüsse stabil zu bleiben und sich anzupassen – in den Fokus. Resiliente Unternehmen können externe Schocks standhalten und sich schnell an neue Bedingungen anpassen. Im Unterschied zur reinen Robustheit, bei der starre Strukturen trotz Änderungen unverändert funktionieren, impliziert Resilienz eine dynamische Anpassungsfähigkeit. Dies

Diese Studie wurde im Rahmen des Projektes "Proaktive Steigerung der Anpassungsfähigkeit produzierender Unternehmen durch agiles Resilienz-Management" erstellt und wurde durch das Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR) – unter dem Förderkennzeichen 02J23C022 gefördert.

bedeutet, dass Prozesse und Strukturen bei Bedarf rasch umgestaltet werden können, um handlungsfähig zu bleiben (Heinicke 2014). Idealerweise geht ein resilientes System gestärkt aus der Krise hervor und agiert anschließend sogar agiler und innovativer als zuvor (Johnston et al. 2011).

Der wirtschaftliche Schaden durch das Fehlen solcher Resilienz ist dabei erheblich, kann aber branchenspezifisch unterschiedlich gelegt sein. In der Automobilindustrie können bereits kurze Produktionsstopps enorme Kosten verursachen – Schätzungen gehen von bis zu 50.000 Euro pro Minute Stillstand aus. (TPC Training 2025) Das Fehlen eines einzigen günstigen Bauteils kann eine ganze Fertigungslinie stoppen und millionenschwere Verluste nach sich ziehen (TPC Training 2025). Die Halbleiter-Knappheit 2021 führte global zu rund 210 Milliarden USD Umsatzverlust und einer Produktionslücke von 7,7 Millionen Fahrzeugen (Deloitte 2023). Gleichzeitig hat die Automobilbranche durch Lean-Production und Just-in-Time-Logistik besonders schlanke Puffer, was die Verletzbarkeit erhöht, wenn ein Glied ausfällt, steht oft weltweit die Produktion still (Dudenhöffer 2016).

Solche Beispiele illustrieren, dass Resilienz – also robuste und anpassungsfähige Systeme – zum wettbewerbsentscheidenden Faktor geworden ist. Unternehmen integrieren daher zunehmend ein systematisches Störungsmanagement in ihre Strategie, um präventiv vorbereitet zu sein und reaktiv schnell handeln zu können.

2. Forschungsfragen und Methodik

Vor dem Hintergrund der Auswirkungen fehlender Resilienz von Produktionssystemen zielt dieser Beitrag darauf ab, den Stand der Forschung zu Störungen und Resilienz in Produktionssystemen systematisch auszuwerten. Die Untersuchung wird durch drei Forschungsfragen geleitet:

- Wie kann die Resilienz von Produktionssystemen beschrieben und gemessen werden?
- Wie lassen sich Störungen in Produktionssystemen voneinander abgrenzen und strukturieren?
- Welcher Forschungsstand liegt bezüglich Strategien und Maßnahmen zum Umgang mit Produktionsstörungen und zur Steigerung der Resilienz vor?

Zur Beantwortung der Forschungsfragen wurde eine strukturierte Literaturanalyse durchgeführt. Ausgangspunkt war eine umfassende Suche nach peer-reviewten Publikationen ab 2010 in einschlägigen Datenbanken. Diese umfassten das Web of Science, Scopus und um technische Aspekte der Fragen zu umfassen IEEE Xplore.

Verwendete Stichwörter kombinierten die Begriffe production disturbances, manufacturing resilience, production simulation, digital twin und disruption management. Die Suche wurde durch Vorwärts- und Rückwärts-Suche ergänzt, um auch häufig zitierte Werke und aktuelle Veröffentlichungen einzubeziehen. Als Auswahlkriterien für die Literatur dienten neben der zeitlichen Begrenzung der Veröffentlichung von 2010 bis 2025 vor allem Praxisrelevanz, namentlich den konkreten Bezug zu realen Produktionssystemen, und wissenschaftliche Qualität, die durch die Nutzung von Peer-Review Verfahren gesichert werden sollte.

Insgesamt wurden 59 Arbeiten identifiziert, die den definierten Kriterien entsprachen. Diese umfassen überwiegend Journal- und Konferenzbeiträge aus den Bereichen Produktionsmanagement, Industrial Engineering und Operations Research, ergänzt um wenige

empirische Studien und Übersichtsarbeiten. Für diese Arbeit wurden 16 der 59 Quellen direkt verwendet, während die anderen zur Validierung der Aussagen genutzt werden konnten.

3. Begriff Resilienz

Resilienz bezeichnet allgemein die Fähigkeit eines Systems, die Auswirkungen von Störungen oder Krisen zu überstehen und zügig wieder funktionsfähig zu werden (Kneifel 2025). Übertragen auf Produktionssysteme bedeutet dies, dass die Produktion trotz unerwarteter Unterbrechungen weiter handlungsfähig bleibt und schnell zur ursprünglichen Leistungsfähigkeit zurückkehrt (Siellaff et al. 2023). Wichtig ist dabei, dass Resilienz mehr umfasst als klassische Robustheit eines Produktionssystems. Daher starten wir zunächst mit der Abgrenzung der beiden Begriffe.

Robustheit bezeichnet, dass ein System Veränderungen begegnen kann, ohne seine Struktur anzupassen (Heinicke 2014). Genauer bezeichnet es die Fähigkeit Störungen Stand zu halten und weitgehend unbeeinflusst weiterzuarbeiten. Robustheit erfordert dementsprechend weder Planung noch Handlungen, sondern stellt eine passive Eigenschaft eines Produktionssystems dar (Scholz et al. 2012).

Resilienz hingegen beinhaltet eine Anpassungsfähigkeit des Systems, inkludiert aber auch Robustheit als Teilbereich. Ein resilientes System nimmt bei Bedarf Änderungen an Struktur oder Prozessen vor, um handlungsfähig zu bleiben (Leng et al. 2025). Im Unterschied zu einem robusten System, dass nach einer Störung wieder in den Ausgangszustand zurückkehrt, muss ein resilientes System nicht zwingend den Originalzustand anstreben. Stattdessen kann es in einem neuen Gleichgewicht weiterarbeiten, welches den neuen Optimalzustand darstellt (Okoh und Haugen 2015). In der Praxis kann somit auch ein robustes System als resilient angesehen werden, da es Störungen ohne Leistungsverlust übersteht, jedoch umfasst Resilienz mehr als Robustheit. Es umfasst die Fähigkeit zur schnellen Erholung und Anpassung an veränderte Bedingungen (Heinicke 2014).

Resilienz kann nach verschiedenen Quellen als Kombination verschiedener Eigenschaften gesehen werden. Zu diesen zählen unter anderem Antizipationsvermögen, Agilität, Flexibilität, Adaptivität, Robustheit sowie Regenerations- und Lernfähigkeit (Leng et al. 2025).

4. Störungen in Produktionssystemen

Störungen in Produktionssystemen umfassen alle Ereignisse, die den geplanten Produktionsablauf unterbrechen oder beeinträchtigen. Im engeren Sinne wird oft an Maschinenstillstände gedacht, doch in einem erweiterten Verständnis zählt alles dazu, was eine Fabrik daran hindert, optimal zu arbeiten (Gronau et al. 2019). Solche Störungen können zu erheblichen Verlusten führen und sind nicht auf das einzelne Unternehmen begrenzt, sondern können sich über Wertschöpfungsketten hinweg auswirken (Sawik 2013). Eine Umfrage des BME ergab etwa, dass jede fünfte Lieferkettenunterbrechung Schaden in Millionenhöhe verursacht. Entsprechend ist ein effektives Störungsmanagement seit langem von hoher Relevanz (Gronau et al. 2019).

In der Literatur werden Störungen häufig nach ihrer Ursache oder Herkunft systematisiert. Eine gebräuchliche Unterteilung unterscheidet interne vs. externe Störfaktoren (Ihlenfeldt et al. 2021). Interne Störungen gehen vom Unternehmen selbst aus, beispielsweise technische

Defekte, menschliches Versagen oder organisatorische Fehlleistungen. Externe Störungen werden durch das Umfeld ausgelöst, etwa Probleme in der Zulieferkette, Nachfrageschwankungen oder Naturereignisse. Tab. 1 gibt einen Überblick über typische Störungskategorien in Produktionssystemen.

Tab. 1: Übersicht bekannter Störungen in Anlehnung an Gronau et al. (2019), Ihlenfeldt et al. (2021), Kneifel (2025) und Muhammad et al. (2023).

Kategorie	Unterkategorie	Beispiel
Interne Störungen (innerbetrieblich)	<i>Technische Störungen</i>	Maschinen- oder Anlagenausfall, Werkzeugbruch, Qualitätsprobleme
	<i>Personelle Störungen</i>	Bedienfehler, krankheitsbedingte Personalausfälle, hohe Mitarbeiterfluktuation
	<i>Prozessuale/Organisatorische Störungen</i>	Planungsfehler, Abstimmungsprobleme, fehlende Materialien durch interne Logistikfehler
Externe Störungen (unternehmensextern)	<i>Lieferkettenstörungen</i>	Lieferverzögerungen, Rohmaterialmangel, Zuliefererausfall
	<i>Marktbedingte Störungen</i>	Nachfrageschwankungen, Absatzkrisen, Preisvolatilität
	<i>Umweltbedingte Störungen</i>	Naturkatastrophen, Pandemien, politische Regulierung (z. B. Exportbeschränkungen)
	<i>Infrastruktur/IT-Störungen</i>	Energieausfälle, Cyberangriffe, IT-Systemausfall

Wie Tab. 1 verdeutlicht, decken Störungen ein breites Spektrum von technischen, sozio-technischen und umfeldbedingten Ursachen ab. Produktionsstandorte sind heute über Liefernetzwerke und Informationssysteme derart verflochten, dass lokale Störungen Kettenreaktionen auslösen können (Muhammad et al. 2023). Ein Maschinenausfall in der Fabrik führt nicht nur intern zu Produktionsausfall, sondern kann z. B. auch Lieferverpflichtungen gefährden; umgekehrt wirken externe Schocks (wie Materialengpässe oder plötzliche Nachfrageeinbrüche) direkt auf die Fabrikleistung (Kneifel 2025).

5. Strategien zur Erhöhung der Resilienz

Zur Erhöhung der Resilienz von Produktionssystemen steht Unternehmen ein breites Spektrum an Strategien zur Verfügung. In der wissenschaftlichen Literatur werden verschiedene Handlungsfelder und Maßnahmen diskutiert, die sich hinsichtlich ihrer Wirkungsweisen ergänzen und eine ganzheitliche Betrachtung der Resilienz erlauben. Die in der aktuellen Forschung besonders hervorgehobenen Strategien lassen sich in fünf Kernansätze unterteilen: (1) Redundanz schaffen, (2) Flexibilität und Anpassungsfähigkeit erhöhen, (3) Prävention durch Instandhaltung und Überwachung, (4) Schnelle Reaktion und Störungsmanagement-Prozesse sowie (5) Wissensmanagement und Lernen. Die nachfolgende Tab. 2 fasst die wesentlichen Merkmale, Maßnahmen und die jeweilige Wirkungsweise dieser Strategien übersichtlich zusammen.

Tab. 2: Übersicht bekannter Resilienzstrategien in Anlehnung an Heinicke (2014), Okoh & Haugen (2015), Ihlenfeldt et al. (2021) und Leng et al. (2025).

Resilienzstrategie	Beschreibung	Beispiele /Maßnahmen	Wirkungsweise
Redundanz schaffen	Kritische Ressourcen mehrfach vorhalten, um bei Ausfall auf Alternativen umzuschalten. Ziel ist es, das Produktionssystem trotz Störungen funktionsfähig zu halten und robuste Prozesse sicherzustellen.	Parallelmaschinen, zusätzliche Betriebsmittel, Ersatzteillager, mehrere Lieferanten	Erhöht Robustheit (Ausfallsicherheit)
Flexibilität und Anpassungsfähigkeit erhöhen	Fähigkeit, schnell auf Veränderungen und Störungen zu reagieren, indem Abläufe kurzfristig angepasst werden. Technische und operative Flexibilität erlauben schnelle Umstrukturierungen und Anpassungen, um negative Auswirkungen zu minimieren oder zu verhindern.	Modulare Anlagen, multifunktionale Maschinen, agile Produktionsplanung, variabler Personaleinsatz, adaptive Fabriken, flexible Prozessrouten	Erhöht Agilität und Anpassungsfähigkeit
Prävention durch Instandhaltung und Überwachung	Ziel ist, Störungen durch vorausschauende Wartung und Zustandsüberwachung frühzeitig zu erkennen und zu verhindern. Moderne Technologien erlauben es, schleichende Probleme zu antizipieren, bevor sie akut werden und Produktionsstörungen verursachen.	Vorausschauende Wartung (Preventive Maintenance), prädiktive Wartung (Predictive Maintenance), Condition-Monitoring-Systeme	Erhöht Antizipationsfähigkeit, reduziert Störowahrscheinlichkeit
Schnelle Reaktion und Störungsmanagement-Prozesse	Vorbereitung auf unvermeidbare Störungen durch klar definierte, eingeübte Abläufe, um schnell reagieren und den Schaden begrenzen zu können. Ziel ist es, die Erholzeit nach Störungen möglichst kurz zu halten und Ausbreitung zu verhindern.	Alarmierungssysteme, Notfallpläne, eingespielte Diagnoseteams, automatische Umschaltmechanismen, Pufferkapazitäten	Verringert Schaden und Erholzeit
Wissensmanagement und Lernen	Systematisches Lernen aus Störungen, um diese langfristig zu reduzieren. Ziel ist es, Erfahrungswissen intern zu teilen und kontinuierlich zu verbessern, um zukünftige Störungen besser bewältigen oder ganz vermeiden zu können.	Ursachenanalysen (Root Cause Analysis), Lessons-Learned-Workshops, Standardisierung, Schulungen	Erhöht langfristige Adaptivität und Lernfähigkeit

Aus der Zusammenstellung von Tab 2. wird deutlich, dass resiliente Produktionssysteme auf einer ausgewogenen Kombination verschiedener Maßnahmen beruhen. Während Redundanzmaßnahmen vor allem die Robustheit und Stabilität gegen akute Störungen erhöhen, sorgen flexible und adaptive Strategien dafür, dass Unternehmen bei unerwarteten Ereignissen schnell reagieren und Prozesse kurzfristig umplanen können. Präventive Strategien wie vorausschauende Instandhaltung wirken dabei als erste Verteidigungslinie, indem sie potenzielle Ausfälle antizipieren und verhindern. Gleichzeitig spielen reaktive Prozesse und schnelle Reaktionen eine entscheidende Rolle, um Schäden möglichst gering zu halten und Erholungszeiten zu verkürzen. Nicht zuletzt ermöglicht ein konsequentes Wissensmanagement langfristige Lern- und Anpassungseffekte, durch die Unternehmen zukünftig besser mit ähnlichen Störungen umgehen können.

6. Messbarkeit von Resilienz

Durch die durchgeführte Literaturanalyse wurde eine Reihe geeigneter Key Performance Indicators (KPIs) identifiziert, mit denen sich die Resilienz von Produktionssystemen systematisch und differenziert bewerten lässt (Raoufi et al. 2020) (Dudenhöffer 2016) (Sielaff et al. 2023). Die in der Analyse identifizierten Indikatoren sind in Tab. 3 zusammengefasst.

Tab. 3: Übersicht bekannter Resilienzkenzzahlen in Anlehnung an Raoufi et al. (2020), Dudenhöffer (2016) und Sielaff et al. (2023).

KPI	Zielgröße	Berechnungsmethodik
Wiederanlaufzeit	<i>Schnelligkeit der Wiederaufnahme der Produktion (Erholungsfähigkeit)</i>	Zeit von Behebung der Störursache bis Erreichen des Soll-Leistungsniveaus. Direkt zeitlich messbar
Störungsausmaß	<i>Tiefe des Leistungseinbruchs (Robustheit)</i>	Differenz zwischen Referenzleistung und minimaler Leistung während der Störung.
Störungsverluste	<i>Gesamtausfallmenge / -dauer (Schadenumfang)</i>	Integral der Leistungsminderung über die Störungsdauer. Fläche zwischen Soll- und Ist-Leistungskurve, oft normiert auf Soll-Produktion
Störungseintrittszeit	<i>Frühzeitigkeit bzw. Verzögerung des Leistungseinbruchs (Warnsignal/Trägheit)</i>	Zeitspanne vom Störungsbeginn (z.B. Auslöser) bis zum spürbaren Leistungsabfall; definiert z.B. als Dauer bis Erreichen eines bestimmten Performanceabfalls
Lösungsfindungszeit	<i>Reaktions- und Diagnosefähigkeit (Agilität in der Störungsbehebung)</i>	Diagnosefähigkeit (Agilität in der Störungsbehebung) Zeit von Störungserkennung bis zur Implementierung einer Lösung; entspricht der Dauer der Fehlerdiagnose und Entscheidungsfindung
Adaptivitätsmaß	<i>Fähigkeit zur Umstrukturierung bei Veränderungen (Wandelbarkeit)</i>	Rekonfigurationszeit für alternative Produktionsrouten; Flexibilitätsindex (Anzahl Produktvarianten oder Volumenschwankungen, die ohne Leistungsabfall handhabbar sind); allgemeiner: „Grad der Selbst-Reorganisation“ des Systems

7. Fazit und Implikationen

Die für diese Arbeit durchgeführte strukturierte Literaturanalyse verdeutlicht, dass Produktionssysteme in einer zunehmend vernetzten und globalisierten Welt besonders anfällig für eine Vielzahl an internen und externen Störungen sind. Externe Schocks wie Naturkatastrophen, Pandemien oder geopolitische Spannungen sowie interne Probleme wie technische Ausfälle oder organisatorische Defizite führen regelmäßig zu erheblichen wirtschaftlichen Schäden entlang der Wertschöpfungskette. Durch diese Verkettung ist die Resilienz zu einem entscheidenden Faktor für die Wettbewerbsfähigkeit produzierender Unternehmen geworden.

Die Analyse zeigt deutlich, dass Resilienz weit mehr umfasst als reine Robustheit. Es geht nicht nur um das passive „Aushalten“ von Störungen, sondern um proaktives und adaptives Handeln, um in kritischen Situationen handlungsfähig zu bleiben und anschließend gestärkt daraus hervorzugehen. Die betrachteten Resilienzstrategien bilden dabei ein umfangreiches Spektrum an Maßnahmen ab, die Unternehmen zur Verfügung stehen.

Eine zentrale Implikation aus den Ergebnissen ist, dass eine hohe Resilienz von Produktionssystemen nicht durch einzelne Maßnahmen erreicht werden kann, sondern vielmehr auf einer integrierten und systematisch abgestimmten Kombination verschiedener Ansätze

beruhen muss. Unternehmen sollten daher genau analysieren, welche Störungstypen in ihrem spezifischen Kontext besonders relevant sind und ihre Resilienzstrategien entsprechend gezielt ausrichten. Hierbei spielt auch die Berücksichtigung wirtschaftlicher Gesichtspunkte eine zentrale Rolle, da übermäßige Redundanzen zwar die Robustheit steigern, zugleich jedoch Kosten und Komplexität erhöhen.

8. Literaturverzeichnis

- Bundesverband Materialwirtschaft Einkauf und Logistik (2019): Unternehmen brauchen bessere Risikoversorge. Online verfügbar unter <https://chemanager-online.com/de/themen/bme-unternehmen-brauchen-bessere-risikoversorge>, zuletzt geprüft am 18.07.2025.
- Deloitte (2023): Halbleiterindustrie im Umbruch. In: Deloitte, 20.06.2023. Online verfügbar unter <https://www.deloitte.com/de/de/Industries/tmt/analysis/halbleiterindustrie-im-umbruch.html>, zuletzt geprüft am 18.07.2025.
- Dudenhöffer, Ferdinand (2016): Wer kriegt die Kurve? Zeitenwende in der Autoindustrie. Frankfurt, New York: Campus Verlag (Business 2016).
- Gronau, Norbert; Kern, Eva-Maria; Jonitz, Hendrik (2019): Herausforderungen im Umgang mit Produktionsstörungen. In: I40M 2019 (6), S. 33–36. DOI: 10.30844/I40M_19-6_S29-32.
- Heinicke, Matthias (2014): Framework for Resilient Production Systems. In: Eduardo Bayro-Corrochano und Edwin Hancock (Hg.): Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications, Bd. 8827. Cham: Springer International Publishing (Lecture Notes in Computer Science), S. 200–207.
- Ihlenfeldt, Steffen; Wunderlich, Tim; Süße, Marian; Hellmich, Arvid; Schenke, Christer-Clifford; Wenzel, Ken; Mater, Sarah (2021): Increasing Resilience of Production Systems by Integrated Design. In: Applied Sciences 11 (18), S. 8457. DOI: 10.3390/app11188457.
- Johnston, Matthew; Lee, Hyang-Won; Modiano, Eytan (2011): A robust optimization approach to backup network design with random failures. In: 2011 Proceedings IEEE INFOCOM: IEEE, S. 1512–1520.
- Kneifel, Dena (2025): Resiliente Fabriken in bewegten Zeiten - WGP. Online verfügbar unter <https://wgp.de/de/resiliente-fabriken-in-bewegten-zeiten/>, zuletzt geprüft am 18.07.2025.
- Leng, Jiewu; Xie, Junxing; Li, Rongjie; Zhou, Xueliang; Gu, Xi; Liu, Qiang et al. (2025): Resilient manufacturing: A review of disruptions, assessment, and pathways. In: Journal of Manufacturing Systems 79, S. 563–583. DOI: 10.1016/j.jmsy.2025.02.006.
- Muhammad, Abdulgaffar; Idris, Mohammed Bello; Ishaq, Aisha Ahmad; Abdullah, Auwal Kabir (2023): The Butterfly Effect and its Implications for Resilience in Complex Socio-Ecological Systems. In: JESCAE 2 (2), S. 38–49. DOI: 10.56556/jescae.v2i2.533.
- Okoh, Peter; Haugen, Stein (2015): Improving the robustness and resilience properties of maintenance. In: Process Safety and Environmental Protection 94, S. 212–226. DOI: 10.1016/j.psep.2014.06.014.

Raoufi, Habibollah; Vahidinasab, Vahid; Mehran, Kamyar (2020): Power Systems Resilience Metrics: A Comprehensive Review of Challenges and Outlook. In: Sustainability 12 (22), S. 9698. DOI: 10.3390/su12229698.

Sawik, Tadeusz (2013): Selection of resilient supply portfolio under disruption risks. In: Omega 41 (2), S. 259–269. DOI: 10.1016/j.omega.2012.05.003.

Scholz, Roland W.; Blumer, Yann B.; Brand, Fridolin S. (2012): Risk, vulnerability, robustness, and resilience from a decision-theoretic perspective. In: Journal of Risk Research 15 (3), S. 313–330. DOI: 10.1080/13669877.2011.634522.

Sielaß, Lennard; Adolf, Thomas; Lucke, Dominik; Friedmann, Andreas (2023): Resilienzmetriken für Produktionssysteme/Resilience Metrics for Production Systems. In: wt 113 (06), S. 266–270. DOI: 10.37544/1436-4980-2023-06-58.

TPC Training (2025): Troubleshooting Thursday | True Cost of Factory Downtime: How Downtime Affects Productivity. Online verfügbar unter <https://www.tpctraining.com/blogs/news/the-true-cost-of-downtime-what-you-dont-know-about-how-downtime-affects-your-productivity>, zuletzt geprüft am 18.07.2025.

Disrupting the Chicken-and-Egg Barrier in Green Hydrogen: A Distributed-Ledger Subsidy Scheme with Purpose-Bound Tokens

Ali Amin Rezaei, ali.amin.rezaei@uol.de, Carl von Ossietzky Universität Oldenburg
Jorge Marx Gómez, jorge.marx.gomez@uol.de, Carl von Ossietzky Universität Oldenburg

Abstract: Europe's green-hydrogen rollout is trapped in a chicken-and-egg loop: high costs depress demand, and weak demand deters the investment needed to cut costs. We propose a Distributed Ledger Technology (DLT) architecture in which government-minted, purpose-bound tokens travel with each kilogram of green H₂ through production, storage, logistics and retail. Because only the final consumer can redeem the token at the time of purchase, the full subsidy is reflected in a lower retail price, eliminating upstream political or lobbying influences. The immutable ledger simultaneously gives regulators real-time proof of origin and price pass-through, ensuring that public funds stimulate consumption and crowd-in private capital. By breaking the cost-demand deadlock, the scheme would support the EU's goal of 20 Mt of renewable hydrogen by 2030.

Keywords: Tokenization, Painted Token, Supply Chain, Blockchain

1. Introduction

Was that first the egg or chicken? This problem can manifest itself in many domains including allocating subsidies to green hydrogen.

The compass of the market is toward profit. Therefore, higher demand in a niche turns it into a profitable niche and therefore attract more investments. However, generating higher demand has also correlation with the price. When there are other providers which providing the same product with much lower price, the demand would automatically go toward that direction.

Hydrogen is poised to play a significant role in the future of Europe's energy system as a key component of the continent's strategy to achieve net-zero emissions and energy transition goals (Kountouris *et al.*, 2024). Currently, hydrogen's role in Europe's energy consumption is minimal, accounting for less than 2% in 2022 (*Hydrogen* 2025). This hydrogen is primarily "gray hydrogen," produced from fossil fuels like natural gas, which results in significant carbon dioxide emissions. However, the European Union has set ambitious targets to change this. The REPowerEU strategy, updated in 2022, aims for Europe to produce 10 million tons of renewable hydrogen and import another 10 million tons by 2030 (fgiacco, 2023).

The hydrogen generated from renewable energy sources like sun and wind called green hydrogen. As producing green hydrogen generates zero-carbon footprint; it's the goal of hydrogen production. Meanwhile the gray hydrogen considered only as a transition plan. However, since the cost of green hydrogen is significantly higher than gray hydrogen (fgiacco 2023). There is a concern that because of such price difference the complete transition won't take place (Hussain *et al.*, 2020). Since the price of green hydrogen is so high, therefore, it doesn't generate enough demand to attract more private investment. Here comes chicken and egg problem; Because of the high price the demand naturally goes toward gray hydrogen and since there are little demand makes it less attractive for the private investors and with lack

enough investment and competitors the price won't come down. These relationships create a vain loop which needs to be broken in some point. This paper is about to propose a solution based on Distributed Ledger technology, to address this problem and get out of the loop.

2. Proposed Approach

Distributed Ledger Technology (DLT) (Antal *et al.*, 2021) can create a fabric of trust between participants. The immutability of the DLTs makes data verifiable. One of the use-cases is to track and trace the supply chain. Distributed Ledger Technology (DLT) can be used to track and trace a supply chain by creating a decentralized, transparent, and immutable record of all transactions and movements of goods or states of their origin to the final consumer (Herbe *et al.*, 2024).

By recording transactions and the movement of goods on a shared digital ledger, DLT can securely link all participants of an entire supply chain. This creates a credible and transparent record of a product's or services which changing the state to the final user (Pearson *et al.*, 2019).

Tracking the changes of the state in supply chain can provide verifiability and authenticity for the green hydrogen; besides providing a clear picture for the regulators and politicians for decision making.

The role of distributed ledger in track and trace the supply chain have been researched in previous studies (Kumar *et al.*, 2025); however, the aim of this paper is propose a solution, by using the purpose-bound token in DLT based supply chain effectively distribute the subsidies across the participants and give an end to vein cycle mentioned. For this, an investment injection needs to take place to decrease the price and derive more demand. The question comes, where should the investment come? from, public sector or private?

The private sector needs guarantee on return on investment to become motivated. Apparently this guarantee and motivation at this stage doesn't exists in hydrogen domain. Then we need investment from public sector to trigger the market and make the domain more attractive for more investment. However, public sector needs indication that the subsidies eventually reducing the price for the end-user and attracting more investment and competition from private sector to the point that the market without subsidies vibrant enough.

Governments usually direct subsidies to producers at the start of the supply chain. A per-unit subsidy, lowers producers' marginal costs, shifts the supply curve outward, and—at least in standard economic models—leads to a lower market price and a higher quantity sold. However, in practice numerous academic studies shows, energy subsidies and public funds often fail to lower end-user prices or boost consumption due to supply-chain inefficiencies, profit-maximizing behaviors by private firms, and poor policy implementation (Höckner, Voswinkel and Weber, 2020; Hussain *et al.*, 2020; Shao, Yang and Bai, 2021; Yuan, Li and Su, 2022). A detailed examination of these obstacles is beyond the scope of this short paper. Instead, the paper's purpose is to outline an alternative approach that avoids these potential pitfalls.

To grasp the function of the token, picture it as a digital coupon. A coupon has no intrinsic worth; it is **purpose-bound**. It can be redeemed only for a specific good or service, at a designated location, and only because an authorised issuer validates it. In the same way, a token's value derives solely from the right it grants, not from any material substance.

This paper introduces a purpose-bound token that is issued at the start of the supply chain and redeemed by the end-user at the point of purchase. Before detailing the mechanics of this token flow, we first explain the fundamentals of a DLT-enabled supply chain, as this architecture determines how the token can be created, transferred, and ultimately collected.

Distributed Ledger Technology (DLT) enhances supply chain operations by creating a decentralized and transparent digital record of transactions. DLT distributes and synchronizes information across a network of computers, which removes the risk of a single point of failure. This structure provides a secure and shared ledger where all participants in the supply chain can view and record information (Antal *et al.*, 2021).

In our proposed system for tracking and tracing the origin of green hydrogen, the government initiates the supply chain by issuing purpose-bound subsidy tokens digitally on a distributed ledger technology (DLT) platform (Islam *et al.*, 2024; Odenweller and Ueckerdt, 2025). This is followed by renewable electricity producers, who generate data for each unit of electricity produced—such as proof of renewable sourcing—and record it immutably on the DLT, linking it to an existing subsidy token (Reda *et al.*, 2024). The process continues sequentially through subsequent stages, including hydrogen manufacturers (who electrolyze water using the renewable electricity to produce green hydrogen), storage facilities, logistics partners, and finally end-users, with each participant appending verified data to the shared ledger. Unlike centralized systems, this structure is collaboratively managed by the network participants rather than a single authority, ensuring decentralization and reducing risks like tampering or single points of failure.

Key participants include the government's unit (as token issuer), electricity producers, hydrogen manufacturers, storage operators, logistics providers, and end-seller. Overall, the DLT network interconnects these stakeholders, providing them with access to a single, shared, and immutable record of the green hydrogen's entire journey from production to consumption, thereby enhancing transparency, traceability, and the effective delivery of subsidies for sustainable energy goals.

Here is the step-by-step workflow of the proposed DLT-Enabled supply chain:

1. **Issuance at the Supply Chain Start:** The government or authorizing entity issues the purpose-bound tokens digitally on a blockchain network. These tokens are "purpose-bound," meaning they can only be used for specific subsidies (e.g., reducing energy costs for end-users) and are tied to particular products or services. Blockchain's decentralized and immutable nature ensures tokens are created transparently and cannot be tampered with, providing a complete, tamperproof record from the outset.
2. **Integration and Flow Through the Supply Chain:** As the product moves through the supply chain—from raw materials to manufacturing, logistics, and distribution—the token is transferred digitally alongside the physical goods. DLT enables traceability, allowing each stakeholder (e.g., suppliers, manufacturers, distributors) to verify and pass the token without altering it (*Using Blockchain to Drive Supply Chain Transparency and Innovation* | Deloitte US, no date). This creates a transparent chain of custody, reducing risks like fraud or diversion of subsidies (which notes how tokens can incentivize sustainable practices and provide anti-counterfeit measures). For instance, in energy subsidies, tokens could be attached to subsidized fuel or materials, ensuring they follow the supply chain's path.

3. **Redemption at End-User Purchase:** The end-user collects or redeems the token at the point of purchase, unlocking the subsidy benefit (e.g., a discounted price). This could involve scanning a QR code or using a digital wallet to validate the token on the blockchain, confirming it hasn't been misused. By connecting on- and off-chain assets, tokens facilitate value transfer between blockchain systems and traditional finance, enabling efficient subsidy delivery (PricewaterhouseCoopers, no date). This direct-to-consumer mechanism boosts consumption by ensuring lower end-user prices, bypassing upstream inefficiencies like profit skimming by intermediaries.

3. Conclusion

We have sketched a way to link green-hydrogen subsidies to an auditable token that moves with every kilogram of fuel along a distributed ledger. This token should let the full subsidy reach the final buyer, prove the hydrogen's origin, and give policymakers real-time market data. The idea is still at the drawing-board stage. It must undergo strict technical testing—scalability, data privacy, cyber-security—and clear legal work on token status, consumer protection, and cross-border rules. We therefore invite engineers, economists, and regulators to refine, pilot, and standardize the framework so that it can reliably cut prices, raise demand, and unlock private investment in green hydrogen.

4. References

- Antal, C. *et al.* (2021) 'Distributed Ledger Technology Review and Decentralized Applications Development Guidelines', *Future Internet*, 13(3), p. 62. Available at: <https://doi.org/10.3390/fi13030062>.
- fgiacco (2023) 'Green Hydrogen to Undercut Gray Sibling by End of Decade', *BloombergNEF*, 9 August. Available at: <https://about.bnef.com/insights/clean-energy/green-hydrogen-to-undercut-gray-sibling-by-end-of-decade/> (Accessed: 8 July 2025).
- Herbe, A. *et al.* (2024) 'How to effectively use distributed ledger technology in supply chain management?', *International Journal of Production Research*, 62(7), pp. 2522–2547. Available at: <https://doi.org/10.1080/00207543.2023.2218947>.
- Höckner, J., Voswinkel, S. and Weber, C. (2020) 'Market distortions in flexibility markets caused by renewable subsidies – The case for side payments', *Energy Policy*, 137, p. 111135. Available at: <https://doi.org/10.1016/j.enpol.2019.111135>.
- Hussain, J. *et al.* (2020) 'Pricing behavior of monopoly market with the implementation of green technology decision under emission reduction subsidy policy', *Science of The Total Environment*, 709, p. 136110. Available at: <https://doi.org/10.1016/j.scitotenv.2019.136110>.
- Hydrogen* (no date). Available at: https://energy.ec.europa.eu/topics/eus-energy-system/hydrogen_en (Accessed: 8 July 2025).
- Islam, A. *et al.* (2024) 'Accelerating the green hydrogen revolution: A comprehensive analysis of technological advancements and policy interventions', *International Journal of*

Hydrogen Energy, 67, pp. 458–486. Available at:
<https://doi.org/10.1016/j.ijhydene.2024.04.142>.

Kountouris, I. *et al.* (2024) ‘A unified European hydrogen infrastructure planning to support the rapid scale-up of hydrogen production’, *Nature Communications*, 15(1), p. 5517. Available at: <https://doi.org/10.1038/s41467-024-49867-w>.

Kumar, N. *et al.* (2025) ‘Blockchain technology in supply chain management: Innovations, applications, and challenges’, *Telematics and Informatics Reports*, 18, p. 100204. Available at: <https://doi.org/10.1016/j.teler.2025.100204>.

Odenweller, A. and Ueckerdt, F. (2025) ‘The green hydrogen ambition and implementation gap’, *Nature Energy*, 10(1), pp. 110–123. Available at: <https://doi.org/10.1038/s41560-024-01684-7>.

Pearson, S. *et al.* (2019) ‘Are Distributed Ledger Technologies the panacea for food traceability?’, *Global Food Security*, 20, pp. 145–149. Available at: <https://doi.org/10.1016/j.gfs.2019.02.002>.

PricewaterhouseCoopers (no date) *Tokenization in financial services: Delivering value and transformation*, PwC. Available at: <https://www.pwc.com/us/en/tech-effect/emerging-tech/tokenization-in-financial-services.html> (Accessed: 17 July 2025).

Reda, B. *et al.* (2024) ‘Green hydrogen as a source of renewable energy: a step towards sustainability, an overview’, *Environment, Development and Sustainability* [Preprint]. Available at: <https://doi.org/10.1007/s10668-024-04892-z>.

Shao, W., Yang, K. and Bai, X. (2021) ‘Impact of financial subsidies on the R&D intensity of new energy vehicles: A case study of 88 listed enterprises in China’, *Energy Strategy Reviews*, 33, p. 100580. Available at: <https://doi.org/10.1016/j.esr.2020.100580>.

Using Blockchain to Drive Supply Chain Transparency and Innovation | Deloitte US (no date). Available at: <https://www.deloitte.com/us/en/services/consulting/articles/blockchain-supply-chain-innovation.html> (Accessed: 17 July 2025).

Yuan, S., Li, J. and Su, X. (2022) ‘Impact of Government Subsidy Strategies on Supply Chains Considering Carbon Emission Reduction and Marketing Efforts’, *Sustainability*, 14(5), p. 3111. Available at: <https://doi.org/10.3390/su14053111>.

Emerging Technologies and Artificial Intelligence in Business Software

F.Salma, JM. Gómez
University of Oldenburg, Germany

Abstract: The rapid embracing of AI is revolutionizing the way in which companies produce, deliver and use business software. Besides scientific research, AI-driven applications are already being integrated into the enterprise software platforms such as customer relationship management (CRM) systems, software development environments, and workflow automation systems. In This presentation, we demonstrate the actual examples of AI application in business software based on the recent GitHub Copilot, ServiceNow, and Salesforce Einstein deployments.

We show ways where AI can bring about more developer efficiency, IT and HR process automation, and customer engagement in real-time through predictive analytics and natural language processing. However, organizations are also dealing with a number of obstacles such as compatibility with old systems, data privacy, and the risk of non-compliance.

The presentation, supported by case evidence, offers practice-oriented advice to the practitioners: AI implementation in business processes, finding the right balance between automation and human control, as well as the use of responsible AI. Next, the presentation outlines the upcoming generation's pragmatic views of AI-native applications, among them energy-efficient designs, explainable AI systems, and composable model marketplaces.

By stressing on the ground truth, the presentation provides technology leaders and software vendors a clear path to the use of AI for the purpose of achieving efficiency, innovation, and sustainable digital transformation.

Keywords: AI, Business Software, Digital Transformation, Practical Applications

Künstliche Intelligenz in der Internen Revision – Eine Studie gegenwärtiger Methoden

Eguchi, Alexander

Carl von Ossietzky Universität Oldenburg, Oldenburg, Deutschland

Abstract: Die fortschreitende Digitalisierung und die damit verbundene Datenflut stellen Unternehmen – insbesondere die interne Revision – vor neue Herausforderungen. Eine zentrale Aufgabe der internen Revision ist die risikoorientierte Prüfungsplanung, deren Effizienz und Qualität zunehmend durch den Einsatz Künstlicher Intelligenz (KI) verbessert werden kann. Ziel dieser Arbeit ist es, den Stand von Wissenschaft und Technik beim Einsatz von KI in der internen Revision systematisch zu untersuchen. Zur Beantwortung dieser Frage wurde ein systematisches Literaturreview nach dem Ansatz von Webster und Watson durchgeführt. Insgesamt wurden 18 relevante Arbeiten ausgewertet. Die Analyse zeigt, dass der Einsatz von KI in der Risikobewertung bislang überwiegend auf strukturierte Daten fokussiert ist. Methoden wie Machine Learning zur Anomalie- oder Betrugserkennung sind erprobt, jedoch fehlt es an Anwendungen, die unstrukturierte Daten in die Risikoanalyse einbeziehen.

Keywords: AI, Internal Audit, Systematic Literature Review, Machine Learning, Audit Planning

1. Einleitung

Die stetige Entwicklung der Digitalisierung und der Informationsgenerierung in den unterschiedlichsten Bereichen erzeugt riesige Datenmengen (Berwanger & Hahn 2020). Datenanalysen ermöglichen Einblicke und Erkenntnisse zu erhalten, die Entscheidungsträgern eine bessere Entscheidungsgrundlage bieten (Roissy & Perron 2018).

Die interne Revision als Organisation umfasst die objektive Prüfung und Unterstützung interner Kontroll- und Überwachungsprozesse, um Schäden unterschiedlicher Arten vom Unternehmen abzuwenden und Risiken zu minimieren. Zu ihren Kernaufgaben gehören die Prüfungsplanung und -durchführung (Berwanger & Kullmann 2012). Die Datenanalyse gehört zu den fundamentalen Aufgaben der internen Revision (Cauers & Scharr 2017). Eine Hürde bei der Datenanalyse sind die riesigen Datenmengen, die erzeugt und nicht mehr händisch ausgewertet werden können (Berwanger & Kullmann 2012).

Das Ziel dieses Literaturreviews ist es, den aktuellen Forschungsstand zum Einsatz von KI-Methoden in der Prüfungsplanung der internen Revision, ohne geografischer oder branchenspezifischer Grenzen systematisch zu analysieren und zu synthetisieren. Der Schwerpunkt dieser Arbeit liegt auf der internen Revision, die als Instrument zur Unterstützung der Überwachungs- und Kontrollaufgaben der Geschäftsleitung etabliert wird. Die Identifizierung von Anwendungen, Chancen, Herausforderungen sowie zentrale Forschungslücken werden herausgearbeitet. Durch die strukturierte Aufarbeitung der relevanten Literatur wird ein umfassender Überblick über den Stand der Wissenschaft und Technik geschaffen, der sowohl theoretische als auch praxisrelevante Implikationen bietet. Die Forschungsfrage, die mit diesem Literaturreview beantwortet wird, ist: Welche KI-Methoden werden zur Risikobewertung in der internen Revision eingesetzt?

Diese Arbeit wird nach der „Preferred Reporting Items for Systematic Reviews and Meta-Analyses“ (PRISMA) Methode durchgeführt.

Nach dieser Einleitung wird in Kapitel das Literaturreview erläutert, einschließlich der Methodik und damit auch die Auswahl der relevanten Literatur, der Suchstrategie und der angewandten Kriterien zur Analyse der Studien. In Kapitel folgen die Analyse und Synthese der identifizierten Literatur, wobei die Ergebnisse thematisch geordnet und in einen konzeptionellen Rahmen eingeordnet werden. Kapitel diskutiert, identifizierte Forschungslücken und leitet zukünftige Forschungsrichtungen ab. Abschließend fasst Kapitel die wichtigsten Erkenntnisse zusammen und reflektiert die Bedeutung der Ergebnisse für Theorie und Praxis.

2. Literaturrecherche

2.1. Methodik

Die Quellen für die Literaturrecherche sind die wissenschaftlichen Datenbanken IEEE Xplore, SpringerLink, ScienceDirect und die Suchmaschine BASE. Diese Datenbanken bieten ein breites Spektrum an qualitativ hochwertigen Arbeiten aus den Bereichen der Wirtschaftsinformatik und Informatik. BASE ist eine der weltweit größten Suchmaschinen für wissenschaftliche Dokumente. Der Index enthält über 400 Mio. Nachweise und ca. 60% der indexierten Nachweise sind in Volltexten frei zugänglich. Der Fokus dieser Arbeit liegt in auf der technischen Implementierung und Evaluation, sodass wirtschaftswissenschaftliche Datenbanken unbeachtet bleiben.

Die Suchstrategie basierte auf der gezielten Auswahl und Kombination relevanter Schlüsselbegriffe. Die Finalisierung des Suchstrings umfasst mehrere Iterationen. Auf dieser Grundlage sind die Suchbegriffe systematisch entlang dreier inhaltlicher Bereiche aufgebaut.

Der erste Bereich betrifft den kontextuellen Rahmen der Untersuchung, konkret die interne Revision in Aktiengesellschaften. Zur eindeutigen thematischen Eingrenzung werden die Begriffe „Internal Audit“ und „Internal Auditing“ verwendet. Diese Fokussierung ist erforderlich, da Risikobewertungsverfahren auch in fachfremden Disziplinen – etwa in der medizinischen Diagnostik, der Projektplanung oder im Betrieb sicherheitskritischer Infrastrukturen Anwendung finden. Um eine thematisch konsistente Literatursammlung sicherzustellen, ist eine klare Abgrenzung zum Untersuchungsfeld der internen Revision notwendig.

Der zweite Bereich adressiert den inhaltlichen Schwerpunkt der Untersuchung: die Risikobewertung im Kontext der internen Revision. Hier sind Begriffe wie „Risk Assessment“, „Risk Management“ und „Audit Planning“ herangezogen worden, da sie zentrale Aspekte der risikoorientierten Prüfungsplanung abbilden.

Keywords OR Title OR Abstract=: ((„Internal Audit“ OR „Audit*“) AND („Risk Assessment“ OR „Risk Management“ OR „Audit Planning“) AND („Machine Learning“ OR „Deep Learning“ OR „Large Language Models“ OR „LLM“ OR „Natural Language Processing“ OR „NLP“))*

Publicity Year :=>2018

Algorithmus 1: Suchstring für die Suche in den Datenbanken.

Der dritte Bereich bildet die methodische Komponente ab – konkret den Einsatz von künstlicher Intelligenz (KI). Hierbei liegt der Fokus sowohl auf allgemeinen KI-Verfahren als auch auf spezifischen Methoden des Natural Language Processing (NLP). Entsprechend wurden Begriffe wie „Machine Learning“, „Deep Learning“, „Natural Language Processing“, „Large Language Model“ sowie einschlägige Abkürzungen berücksichtigt. Der so entstandene Suchstring wird in Algorithmus 1 dargestellt.

Zur ersten inhaltlichen Eingrenzung sind ausschließlich wissenschaftliche Fachartikel und Konferenzbeiträge berücksichtigt, die im Volltext frei zugänglich und in deutscher oder englischer Sprache verfasst sind.

Ein weiteres zentrales Auswahlkriterium bildet der Veröffentlichungszeitraum: Es sind nur Arbeiten einbezogen, die im Zeitraum von 2018 bis Februar 2025 erschienen sind.

Abb. 1 zeigt den PRISMA-Auswahlprozess der Literaturrecherche sowie die jeweiligen Teilergebnisse inklusive der Anzahl an Arbeiten nach jeder Phase. Insgesamt wurden 96 wissenschaftliche Arbeiten identifiziert. Die Datenbank IEEE Xplore enthält 46 Arbeiten. Über ScienceDirect wurden 33 relevante Arbeiten identifiziert, weitere 17 über BASE und SpringerLink. In der Vorauswahl schloss man insgesamt 78 Arbeiten aus. Zunächst wurden Titel, Schlüsselbegriffe und Abstracts gesichtet, um eine erste inhaltliche Relevanzbewertung vorzunehmen. Es wurden 58 Arbeiten ausgeschlossen, die nicht im Fokus dieser Arbeit lagen, aber im Bereich der internen Revision. Weitere 10 Arbeiten wurden ausgeschlossen, da sie ebenfalls systematische Literaturreviews umfassen. Eine Publikation wurde in Spanisch verfasst. Die letzten neun Arbeiten wurden aufgrund des fehlenden Bezugs zur internen Revision ausgeschlossen. Anschließend wurde eine Volltextanalyse durchgeführt, um die Arbeiten abschließend qualitativ zu bewerten. In dieser Phase wurden keine weiteren Arbeiten ausgeschlossen.

Die auf diese Weise als relevant eingestuften Beiträge bildeten die Grundlage für eine anschließende Vorwärts- und Rückwärtssuche. Diese wurden ebenfalls durch einer Eignungsprüfung unterzogen. Es wurden 18 Arbeiten für die Auswertung genutzt.

Eine Inhaltsanalyse erfolgt anhand der relevanten Arbeiten unter Anwendung vordefinierter Kategorien, die eine systematische Auswertung und die gezielte Beantwortung der Forschungsfrage ermöglichen. Insgesamt wurden acht Merkmale definiert ergänzt um weitere drei Merkmale. Hervorzuheben sind die Merkmale ‚Ziel der Arbeit‘, die ‚Methoden‘ sowie der ‚Anwendungsdomäne‘. Der Forschungs-fokus wurde nachträglich hinzugefügt, da sich im Zuge der Sichtung der Arbeiten eine detailliertere Unterteilung innerhalb des Bereichs der internen Revision als sinnvoll herausstellte. Das primäre Ziel dieser Untersuchung ist die Analyse der Methoden im Risk Assessment und in der Prüfungsplanung innerhalb der internen Revision. Aufgrund der begrenzten Anzahl relevanter Arbeiten in diesem Bereich wurde zudem der vorgeschlagene Einsatzbereich einer genaueren Betrachtung unterzogen.

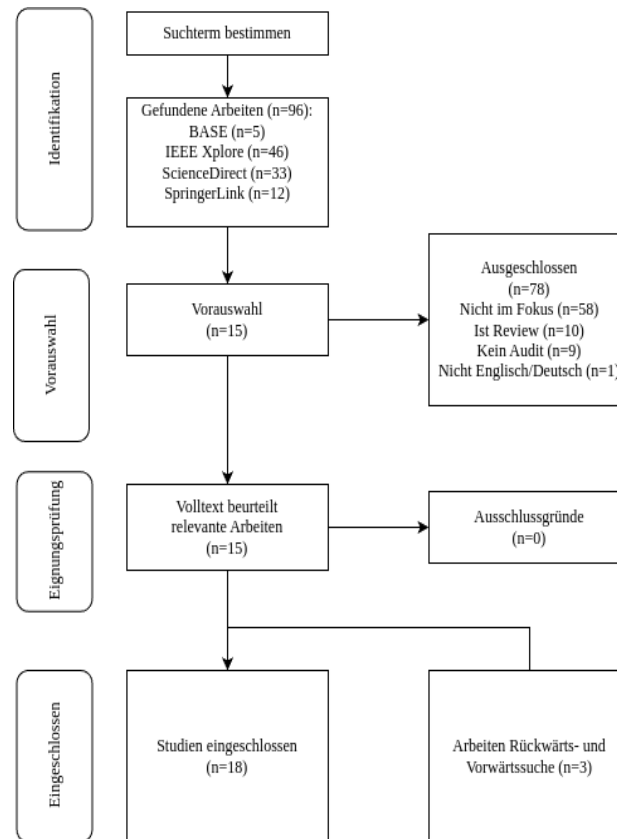


Abb. 1 Vorgehen bei der Identifikation von relevanten Arbeiten nach PRISMA

Das Merkmal Forschungsfokus unterscheidet in konzeptionelle, praxisorientierte und gemischter Fokus. Die konzeptionellen Arbeiten konzentrieren sich auf die Entwicklung von Modellen und theoretischen Rahmenwerken und basieren primär auf literaturgestützten abschließend qualitativ zu bewerten. In dieser Phase wurden keine Arbeiten ausgeschlossen.

Die auf diese Weise als relevant eingestuft Beiträge bilden die Grundlage für eine anschließende Vorwärts- und Rückwärtssuche. Diese wurden ebenfalls durch einer Eignungsprüfung unterzogen.

Der Prozess der qualitativen Inhaltsanalyse erfolgte anhand der relevanten Arbeiten unter Anwendung vordefinierter Kategorien, die eine systematische Auswertung und die gezielte Beantwortung der Forschungsfrage ermöglichten. Insgesamt wurden acht Merkmale definiert ergänzt um weitere drei Merkmale. Hervorzuheben sind die Merkmale ‚Ziel der Arbeit‘, die ‚Methoden‘ sowie der ‚Anwendungsdomäne‘. Der Forschungs-fokus wurde nachträglich hinzugefügt, da sich im Zuge der Sichtung der Arbeiten eine detailliertere Unterteilung innerhalb des Bereichs der Internen Revision als sinnvoll herausstellte. Das primäre Ziel dieser Untersuchung ist die Analyse der Methoden im Risk Assessment und in der Prüfungsplanung innerhalb der Internen Revision. Aufgrund der begrenzten Anzahl relevanter Arbeiten in diesem Bereich wurde zudem der vorgeschlagene Einsatzbereich einer genaueren Betrachtung unterzogen.

Das Merkmal Forschungsfokus unterscheidet in konzeptionelle, praxisorientierte und gemischter Fokus. Die konzeptionellen Arbeiten konzentrieren sich auf die Entwicklung von Modellen und theoretischen Rahmenwerken und basieren primär auf literaturgestützten

Argumentationen, ohne empirische Untersuchungen durchzuführen. Dazu gehören theoretisch fundierte Beiträge, die durch illustrative Beispiele ergänzt werden.

Die Ausprägung praxisorientierte Studien rücken den Praxisbezug in den Vordergrund. Hier werden theoretische Grundlagen zwar herangezogen, um die praktische Umsetzung in realen Kontexten einzuordnen. Solche Arbeiten dokumentieren meist die Umsetzung von Konzepten in konkreten Settings und evaluieren diese praxisnah. Oft stehen dabei praktische Erfahrungen und Beobachtungen im Vordergrund, ohne tiefere theoretische Fundierung. Ziel ist vor allem der Wissenstransfer durch nachvollziehbare Darstellung von Implementierungsprozessen.

Die Kombination konzeptioneller Studien und praxisorientierten Studien, die gleichermaßen Theorie und Praxis aufweisen, umfassen die dritte Ausprägung. Diese Arbeiten kombinieren theoretische Fundierung mit empirischen Untersuchungen, wobei die theoretischen Modelle sowohl entwickelt als auch systematisch durch Datenvalidierung überprüft werden.

3. Analyse und Synthese der Literatur

Nach der systematischen Identifikation und Kategorisierung relevanter Literatur erfolgt in diesem Abschnitt die Analyse und Synthese der identifizierten Arbeiten zum Einsatz von KI-Methoden in der Internen Revision. In diesem Abschnitt sollen zentrale Erkenntnisse gebündelt und thematische Schwerpunkte identifiziert werden. Dieses strukturierte Vorgehen während der Analyse und Synthese soll ein fundierter Überblick über den aktuellen Forschungsstand schaffen, um daraus Implikationen für Wissenschaft und Praxis abzuleiten und zukünftige Forschungsrichtungen aufzuzeigen.

3.1. Analyse

Die tatsächlich relevanten Arbeiten wurden zwischen 2020 und 2025 veröffentlicht. Die Abb. 2 zeigt den zeitlichen Verlauf der Arbeiten über die Jahre 2018-2025. Ein signifikanter Trend ist seit 2021 zu erkennen. Zwischen 2020 und 2022 wurden zusammen fünf Arbeiten veröffentlicht.

Im Jahr 2023 kamen 2 Arbeiten hinzu auf insgesamt 4 Arbeiten. Im darauffolgenden Jahr gab es eine Steigerung auf 7 Arbeiten. Bis Feb. 2025 sind zwei Arbeiten veröffentlicht worden. Die Analyse der Arbeiten nach Anwendungsdomäne wird in der Abb. 3 gezeigt, es kann deutlich eine Konzentration auf die Prüfungsdurchführung beobachtet werden, die mit sechs Beiträgen den größten Anteil ausmacht. Es folgt die interne Revision als organisationale Funktion mit fünf Veröffentlichungen, was die zentrale Rolle interner Kontrollmechanismen unterstreicht. Die interne Revision als Funktion beschreibt die Überwachungs- und Prüfungsaufgaben. Die Studien gehen nicht auf eine spezifische Aufgabe innerhalb der internen Revision ein.

Sowohl das Risk Assessment als auch die interne Revision als Funktion sind mit jeweils vier deutlich geringer ist die Zahl der Arbeiten zu dem Thema external Audits, das nur einen Beitrag verzeichnet. Die Arbeiten (Qiuju 2020, Nonnenmacher *et al.* 2021, Dr. Antje Heinen *et al.* 2023, F. Yang *et al.* 2023, Wang 2024, *The Global Industry Classification Standard (GICS®)* 2025) erproben Methoden in der Prüfungsdurchführung. Diese Arbeiten haben einen vor allem einen praxisorientierten Forschungsfokus. Die Ziele umfassen eine effizienteren Prüfprozess. Dieser soll durch datengetriebene Modelle ermöglicht werden. Darüber hinaus wird eine objektive Gestaltung der Prüfungen angestrebt. Bei allen Arbeiten ist eine Datenorientierte Prüfung zentraler Bestandteil des Prüfprozesses. In (Nonnenmacher *et al.* 2021) wurde mithilfe

eines Autoencoder die Anomalieerkennung für Produktionsbezogene Daten untersucht. Die Auswirkung auf die Prüfungsdurchführung ist eine verbesserte Auswahl des Prüfungsbereiche. Wohingegen (X. Liu *et al.* 2024) eine Echtzeitüberwachung mittels einer KI digitalen Prüfplattform vorschlägt. In der Anwendungsdomäne Internen Revision als Funktion umfassen die Arbeiten von (M. S. Thomas & J. Mathew 2022, Adamyk *et al.* 2023, Zhang 2023, Zhang *et al.* 2024). Einen konzeptionellen Fokus werden in (Adamyk *et al.* 2023, Zhang 2023) verfolgt. Es werden keine expliziten Methoden genannt, die genutzt werden können. Vielmehr wird auf die Möglichkeiten eingegangen die unterschiedlichen Methoden ermöglichen. So wird in (Adamyk *et al.* 2023) beschrieben, dass NLP genutzt werden kann, um die Prüfungsabschlussberichte zu erzeugen. Darüber hinaus wird erörtert, dass unterschiedliche Texte analysiert und aufbereitet werden können. Dies ermöglicht Einblicke und Stimmung aus den Texten zu extrahieren. In (Roussy & Perron 2018, Oluwatosin Ilori *et al.* 2024, Huang *et al.* 2025) wurden Methoden in der Risikobewertung erprobt. Die Arbeiten (Zhenzheng Yan 2024, Ou & Zhang 2025) sind praxisorientiert, wohingegen (Oluwatosin Ilori *et al.* 2024, Huang *et al.* 2025) konzeptionelle Studien sind. Die Publikation (Huang *et al.* 2025) entwickelte ein Konzept bzw. erstellte ein Datenset um bestehend aus Mitarbeiterbewertungen potenzielle Kontrollprobleme im Rechnungswesen festzustellen. Eine Textanalyse und Text Mining soll später die Bewertungen auswerten. Die genutzten Methoden in (Zhenzheng Yan 2024, Ou & Zhang 2025) umfassen den Einsatz von CNN-LSTM und einer Kombination aus unterschiedlichen MLP zur Feature Extraction und zweier Klassifikatoren. Dieses Konstrukt als Multimodales Neuronales Netz wird dann mittels zweier Loss-funktionen Rückwärtspropagiert trainiert.

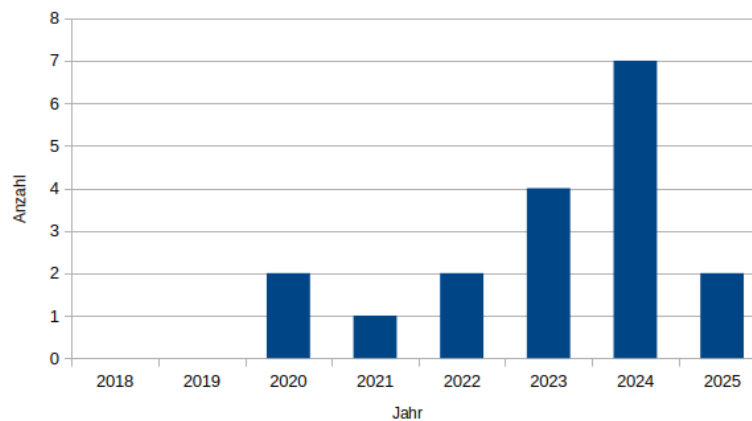


Abb. 2 Anzahl der Arbeiten von Jan. 2018 - Feb. 2025

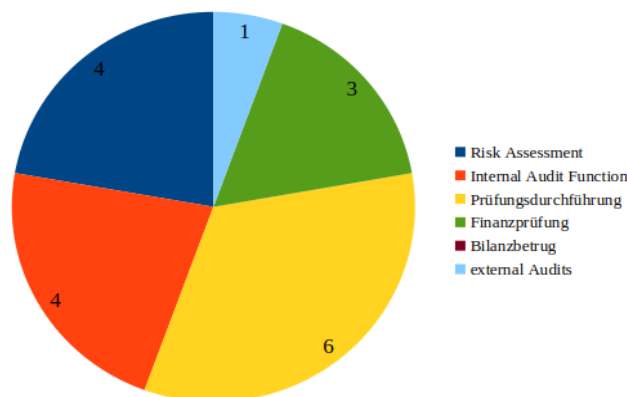


Abb. 3 Verteilung der Arbeiten nach Anwendungsbereich.

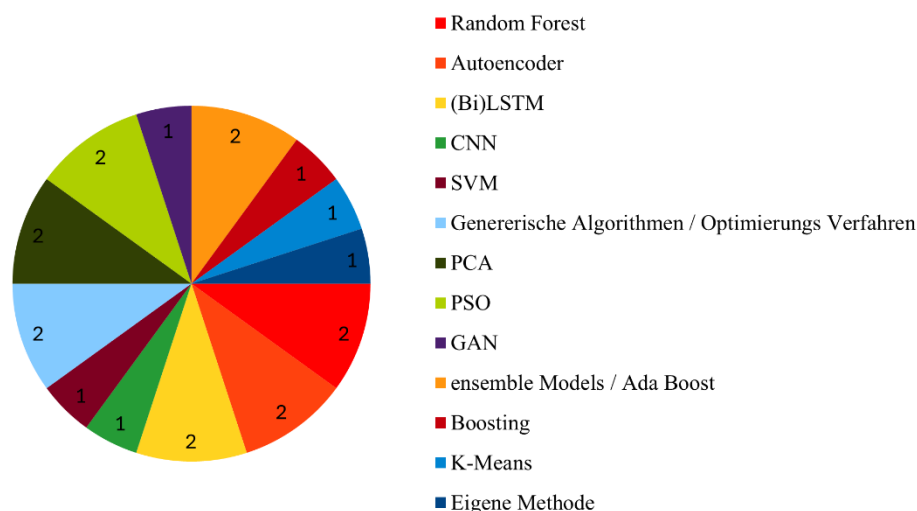


Abb. 4 Anzahl der in den Arbeiten genutzten KI-Methoden

Die Abb. 4 zeigt die genutzten KI-Methoden, die in den Arbeiten erprobt wurden. Die Summe der Methoden entspricht nicht der Anzahl der Arbeiten, da in einigen Arbeiten eine Mehrzahl von Methoden genutzt wurde.

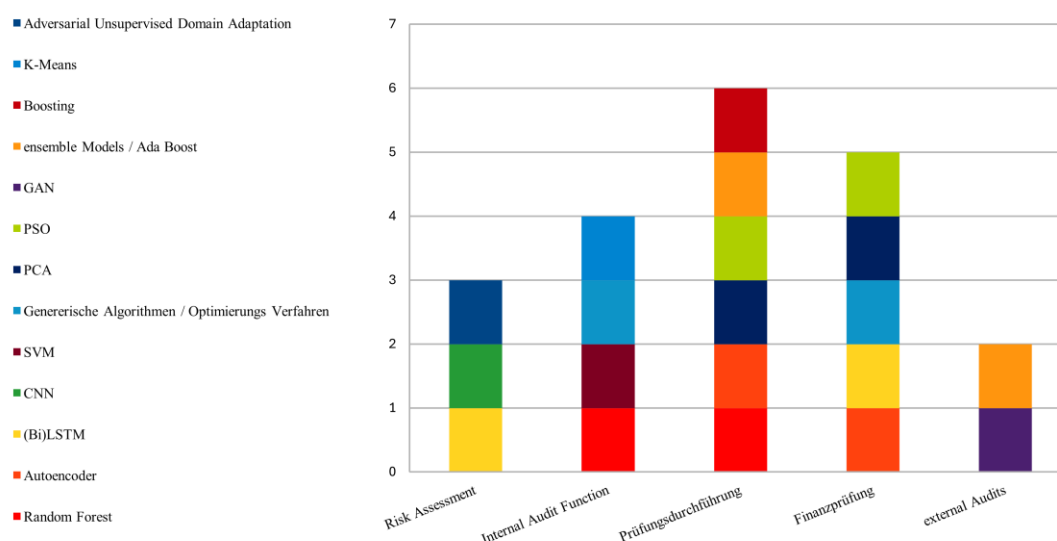


Abb. 5 Die in den Arbeiten genutzten Methoden gruppiert nach Anwendungsbereich

Die Abb. 5 zeigt die Anwendungsbereiche Risk Assessment, internal Audit Funktion, die Prüfungsdurchführung, Finanzprüfung und external Audits. Dazu werden die genutzten KI-Methoden und die Anzahl in diesen Bereichen dargestellt. In der Prüfungsdurchführung sind die meisten Methoden eingesetzt worden, gefolgt von der Finanzprüfung mit fünf zugeordneten Methoden. Genauer sind die Methoden Random Forest, Autoencoder, PCA, PSO, Ensemble Methoden, Boosting der Prüfungsdurchführung zugeordnet.

3.2. Synthese

Die Analyse zeigt, dass einige Anstrengungen unternommen werden, um KI in der internen Revision zu realisieren. Ein Großteil der identifizierten Studien veröffentlichte in den letzten Jahren, was auf ein zunehmendes Forschungsinteresse auf dieses Thema hindeutet. Die Analyse der Erscheinungsjahre zeigt eine deutliche Zunahme der Arbeiten ab dem Jahr 2022. Die Ursache kann nicht identifiziert werden. Insgesamt spiegelt die Verteilung der Studien auf die Anwendungsdomäne, die starke Ausrichtung der Forschung auf praktische Prüfungsaufgaben sowie interne Prüfungsstrukturen wider, während externe Prüfungen und andere Bereiche eher randständig behandelt werden. Die Analyse zeigt auch, dass dem Bereich Risk Assessment und dem Bereich interne Revision als Funktion mit vier Beiträgen eine mittlere Bedeutung zukommt. Betrachtet man die Zuordnung Methoden zu Bereich erkennt man eine breite Verteilung der Methoden ohne einen Schwerpunkt auf einen Bereich. So enthält die Prüfungsdurchführung Methoden des Machine Learning, Deep Learning, Optimierungsverfahren und Boosting-Ansätzen. Innerhalb des Risikobewertung sind Deep Learning-Methoden vorherrschend.

Angesichts der zunehmenden Komplexität unternehmerischer Risiken könnte eine intensivere wissenschaftliche Auseinandersetzung mit diesem Themenfeld künftig an Bedeutung gewinnen. So zeigt sich, dass in allen Bereichen keine Dokumentenanalysen genutzt werden und daher noch Forschungslücken aufweisen.

Die Auswertung der genutzten Methoden zeigt keinen klaren Schwerpunktsetzung. Die detaillierte Analyse der Arbeiten ergibt ein deutlich praxisorientiertes Gesamtbild: 12 der 18 Arbeiten fokussieren sich auf anwendungsnahe Umsetzungen, während 6 Arbeiten konzeptionelle Modelle oder theoretische Rahmenwerke entwickeln. Diese Verteilung zeigt, dass der Forschungsschwerpunkt derzeit klar auf der praktischen Anwendbarkeit moderner Technologien im Bereich der Internen Revision liegt. Dies deutet darauf hin, dass die Integration von Künstlicher Intelligenz und Machine Learning in Audits nicht mehr nur ein theoretisches Konzept ist, sondern zunehmend konkrete Systementwicklungen und Prozessoptimierungen beeinflusst. Die vergleichsweise geringe Anzahl rein konzeptioneller Arbeiten unterstreicht, dass die Forschung zunehmend auf direkte Nutzbarkeit und technologische Umsetzbarkeit abzielt, was den Trend zur Digitalisierung und Automatisierung von Prüfprozessen weiter beschleunigt.

4. Fazit

Das Ziel dieser Arbeit ist es den Stand von Wissenschaft und Technik bei dem Einsatz von KI in der Internen Revision aufzuzeigen. Die Forschungsfrage lautet: Welche KI-Methoden werden zur Risikobewertung in der internen Revision eingesetzt?

Um diese Frage zu beantworten, wurde ein Systematisches Literaturreview nach Webster & Watson durchgeführt. Es wurden 18 relevante Arbeiten aus insgesamt 3 Datenbanken und einem Suchindex für wissenschaftliche Arbeiten genutzt.

Die Analyse und Synthese zeigen, dass der Einsatz von KI in der internen Revision, vor allem in der Risikobewertung unbeachtet ist und noch mit offenen Forschungsfragen verbunden ist.

Zukünftige Forschungsarbeiten sollten sich verstärkt auf die Analyse unstrukturierter Daten konzentrieren. Insbesondere im Bereich der Dokumenten- und Textanalyseverfahren besteht

erhebliches Potenzial, um tiefergehende Erkenntnisse zu gewinnen. Die vorliegende Literaturrecherche wurde nach dem PRISMA-Schema durchgeführt, unterliegt jedoch gewissen Einschränkungen. Die Auswahl der Studien war auf bestimmte Datenbanken sowie auf Publikationen in Deutsch und Englisch beschränkt. Dadurch besteht die Möglichkeit, dass relevante Arbeiten in anderen Sprachen oder aus nicht indexierten Quellen unberücksichtigt geblieben sind. Zudem wurden ausschließlich veröffentlichte Studien berücksichtigt, sodass ein Publikationsbias nicht ausgeschlossen werden kann. Die Auswahl und Bewertung der eingeschlossenen Literatur erfolgte nach definierten Kriterien, enthält jedoch trotz systematischem Vorgehen ein gewisses Maß an Subjektivität.

5. Literaturverweise

- Adamyk, O. et al. (2023) ‘Does Artificial Intelligence Help Reduce Audit Risks?’, in 2023 13th International Conference on Advanced Computer Information Technologies (ACIT). 2023 13th International Conference on Advanced Computer Information Technologies (ACIT), Wrocław, Poland: IEEE, pp. 294–298. Available at: <https://doi.org/10.1109/ACIT58437.2023.10275661>.
- Aftabi, S.Z., Ahmadi, A. & Farzi, S. (2023) ‘Fraud detection in financial statements using data mining and GAN models’, *Expert Systems with Applications*, 227, p. 120144. Available at: <https://doi.org/10.1016/j.eswa.2023.120144>.
- Berwanger, J. & Hahn, U. (2020) *„i und Compliance: Operative Grundlagen und Recht*. Wiesbaden: Springer Fachmedien Wiesbaden. Available at: <https://doi.org/10.1007/978-3-658-31807-9>.
- Berwanger, J. & Kullmann, S. (2012) *Interne Revision: Funktion, Rechtsgrundlagen und Compliance*. Wiesbaden: Springer Fachmedien Wiesbaden. Available at: <https://doi.org/10.1007/978-3-8349-3879-4>.
- Cauers, L. & Scharr, C. (2017) ‘Quality Assessments nach dem neuen DIIR Revisionsstandard Nr. 3’, *Zeitschrift Interne Revision* [Preprint], (2). Available at: <https://doi.org/10.37307/j.1868-7814.2017.02.03>.
- Ding, R. (2022) ‘Enterprise Intelligent Audit Model by Using Deep Learning Approach’, *Computational Economics*, 59(4), pp. 1335–1354. Available at: <https://doi.org/10.1007/s10614-021-10192-9>.
- Dr. Antje Heinen et al. (2023) ‘Generative KI in der Internen Revision: Chancen und Herausforderungen’. Available at: https://www.diir.de/content/uploads/2023/11/Fachbeitrag_Nr._4_ChatGPT.pdf (Accessed: 15 January 2025).
- F. Yang, J. Deng, & X. Song (2023) ‘Research on Audit Model Method based on Ensemble Learning’, in 2023 6th International Conference on Data Science and Information Technology (DSIT). 2023 6th International Conference on Data Science and Information Technology (DSIT), pp. 265–271. Available at: <https://doi.org/10.1109/DSIT60026.2023.00047>.
- H. Zhang (2024) ‘Application of Hybrid Intelligent Optimization Algorithm in Internal Audit Informatization Construction’, in. 2024 International Conference on Telecommunications

- and Power Electronics (TELEPE), pp. 170–175. Available at: <https://doi.org/10.1109/TELEPE64216.2024.00036>.
- Huang, L. et al. (2025) ‘Identifying Accounting Control Issues from Online Employee Reviews’, *Data Science and Management*, p. S2666764925000074. Available at: <https://doi.org/10.1016/j.dsm.2025.02.001>.
- M. S. Thomas & J. Mathew (2022) ‘Supervised Machine Learning Model for Automating Continuous Internal Audit Workflow’, in. 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI), pp. 1200–1206. Available at: <https://doi.org/10.1109/ICOEI53556.2022.9776888>.
- Nonnenmacher, J. et al. (2021) ‘Using Autoencoders for Data-Driven Analysis in Internal Auditing’. Available at: <http://hdl.handle.net/10125/71317>.
- Oluwatosin Ilori, Nelly Tochi Nwosu, & JOUR (2024) ‘Advanced data analytics in internal audits: A conceptual framework for comprehensive risk assessment and fraud detection’. Available at: <http://www.fepbl.com/index.php/farj/article/view/1213>.
- Ou, Z. & Zhang, Z. (2025) ‘AUDAN: An adversarial unsupervised domain adaptation network informed decision-making in audit risk assessment’, *Knowledge-Based Systems*, 314, p. 113171. Available at: <https://doi.org/10.1016/j.knosys.2025.113171>.
- Qiuju, H. (2020) ‘Influence of Big Data on Modern Risk-oriented Audit and Countermeasures’, in 2020 2nd International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI). 2020 2nd International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI), Taiyuan, China: IEEE, pp. 319–322. Available at: <https://doi.org/10.1109/MLBDBI51377.2020.00068>.
- Roussy, M. & Perron, A. (2018) ‘New Perspectives in Internal Audit Research: A Structured Literature Review’, *Accounting Perspectives*, 17(3), pp. 345–385. Available at: <https://doi.org/10.1111/1911-3838.12180>.
- The Global Industry Classification Standard (GICS®) (2025). Available at: <https://www.msci.com/indexes/index-resources/gics> (Accessed: 13 June 2025). Wang, C. (2024) ‘Research on Audit Risk Prediction and Management Based on Machine Learning Algorithm’, in 2024 6th International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI). 2024 6th International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI), Hangzhou, China: IEEE, pp. 189–192. Available at: <https://doi.org/10.1109/MLBDBI63974.2024.10823844>.
- Wu, Y. (2024) ‘Enterprise financial sharing and risk identification model combining recurrent neural networks with transformer model supported by blockchain’, *Heliyon*, 10(12), p. e32639. Available at: <https://doi.org/10.1016/j.heliyon.2024.e32639>.
- X. Liu et al. (2024) ‘Artificial Intelligence Digital Audit System Under Machine Learning Technology’, in. 2024 3rd International Conference on Artificial Intelligence and Autonomous Robot Systems (AIARS), pp. 739–743. Available at: <https://doi.org/10.1109/AIARS63200.2024.00139>.
- Zhang, J., Su, X. & Su, R. (2024) ‘Research on the Application of Digital Audit in Power Grid Investment Based on Big Data and Neural Networks’, in 2024 3rd International Conference on Energy, Power and Electrical Technology (ICEPET). 2024 3rd International Conference on Energy, Power and Electrical Technology (ICEPET),

Chengdu, China: IEEE, pp. 666–669. Available at:
<https://doi.org/10.1109/ICEPET61938.2024.10626242>.

Zhang, L. (2023) ‘Intelligent Internal Audit Platform Architecture Based on Machine Learning’, in 2023 2nd International Conference on Artificial Intelligence and Autonomous Robot Systems (AIARS). 2023 2nd International Conference on Artificial Intelligence and Autonomous Robot Systems (AIARS), Bristol, United Kingdom: IEEE, pp. 531–534. Available at: <https://doi.org/10.1109/AIARS59518.2023.00112>.

Zhenzheng Yan (2024) ‘Audit Risk Assessment Model Construction by Incorporating Deep Learning Techniques’. Available at: <https://doi.org/10.2478/amns-2024-2138>.

ERPFinder: Self-Service-Plattform zur Erstellung von ERP-Spezifikationen für KMU

Sébastien Gard, Benjamin Morel, Rainer Fux, Niklas Maier
HES-SO Valais-Wallis, University of Applied Sciences of Western Switzerland

Zusammenfassung: Weniger als ein Drittel der kleinen und mittleren Unternehmen (KMU) mit 10 bis 49 Mitarbeitenden haben bislang ein ERP-System implementiert. Ziel dieser Studie ist es, eine Lösung zu finden, um die Einführung eines solchen Systems zu erleichtern. Zu diesem Zweck wurde eine neutrale und systemunabhängige Self-Service-Plattform entwickelt, die Führungskräfte bei der Definition ihrer Bedürfnisse unterstützt und als Abschluss ein Pflichtenheft generiert. In zwei Testphasen wurde die Relevanz dieses Ansatzes überprüft, wobei zentrale Fragen im Fokus standen: Kann ein Self-Service-Tool Hindernisse bei der ERP-Einführung abbauen? Bietet es eine Analysequalität, die mit der eines Experten vergleichbar ist, und hilft es, versteckte Bedürfnisse zu identifizieren? Wird es von Führungskräften als sinnvoll wahrgenommen? Die Rückmeldungen zeigen: Auch wenn die Mehrheit der Befragten das generierte Pflichtenheft als relevant einstuft, reicht das Tool allein nicht aus, um die Angst vor Fehlentscheidungen zu beseitigen. Eine menschliche Begleitung bleibt notwendig. Die Ergebnisse unterstreichen den Nutzen eines hybriden Ansatzes, bei dem Self-Service-Elemente mit Expertenberatung kombiniert werden, um die Effizienz und Akzeptanz von ERP-Projekten in KMU zu maximieren.

Schlüsselwörter ERP, KMU, BPM, Geschäftsprozessmanagement, Digitalisierung

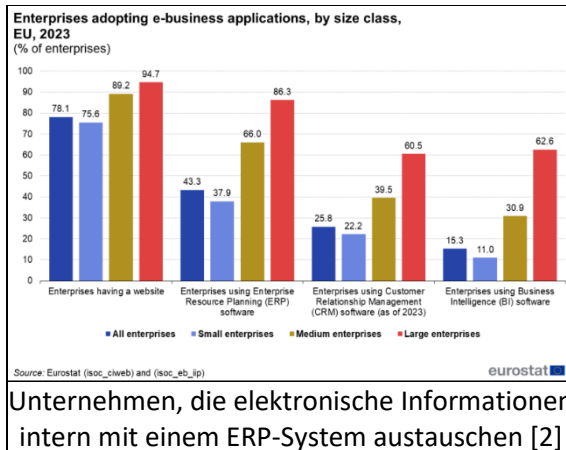
1. Einleitung

Im Rahmen der digitalen Transformation eines Unternehmens stellt die Einführung eines ERP-Systems eine grundlegende und unverzichtbare Maßnahme dar. Ein ERP (Enterprise Resource Planning) ist ein integriertes Softwaresystem, das die konsistente Verwaltung aller Unternehmensdaten und Geschäftsprozesse ermöglicht. Es bildet eine essenzielle Grundlage für die Wettbewerbsfähigkeit von Unternehmen, indem es eine effiziente Prozessabwicklung unterstützt. Aufgrund seiner strukturierenden Wirkung und seiner zentralen Rolle im Daten- und Aktivitätsmanagement ist ein ERP zudem Voraussetzung für den Einsatz fortschrittlicher Technologien wie Künstliche Intelligenz (KI), Maschinelles Lernen (ML), Datenanalytik oder Industrie 4.0.

Trotz zahlreicher Initiativen und den politischen Zielsetzungen für den digitalen Wandel hinkt die Digitalisierung kleiner und mittlerer Unternehmen (KMU) in Europa weiterhin hinter den großen Unternehmen und den politischen Zielvorgaben her. Laut dem Digital Economy and Society Index (DESI) sowie dem von Eurostat veröffentlichten Bericht 2022 verfügen nur 55 % der KMU in der EU über ein grundlegendes Digitalisierungsniveau – das Ziel liegt jedoch bei mindestens 90 % bis zum Jahr 2030 [1].

Ein genauerer Blick auf die DESI-Statistik – insbesondere auf die Integration von Geschäftsprozessen mittels ERP – zeigt ein noch ernüchterndes Bild: Nur 37,9 % der kleinen Unternehmen mit 10 bis 49 Mitarbeitenden nutzen ein ERP, und dieser Anteil ist seit 2015 weitgehend unverändert geblieben.

Während man annehmen könnte, dass ERP-Systeme längst Standard sind, zeigt sich, dass die Digitalisierungsziele für KMU ohne gezielte Maßnahmen in diesem Bereich nicht erreichbar sind.



Die Schwierigkeiten KMU bei der Einführung von ERP-Systemen sind bekannt und wurden in zahlreichen Studien thematisiert. Diese Untersuchungen haben einerseits die Motivationen zur ERP-Einführung beleuchtet [3], den Digitalisierungsgrad von KMU in bestimmten Regionen wie Brandenburg [4] oder der Schweiz [5] erfasst und andererseits zentrale Hürden identifiziert [6][7]. Zu diesen Hemmnissen zählen unter anderem:

- Begrenzte Budgets der KMU und die als hoch wahrgenommenen Implementierungskosten
- Fehlende IT-Ressourcen und -Kompetenzen
- Mangelnde Projektmanagement- und Change-Management-Erfahrung
- Unternehmenskultur, Veränderungsresistenz und digitales Reifegradniveau
- Der Glaube, ein ERP sei nur für große Unternehmen geeignet und für KMU zu komplex
- Die Überzeugung, über spezielle Anforderungen zu verfügen, verbunden mit einer Überschätzung der betrieblichen Besonderheiten im Vergleich zu ERP-Standards (z. B. Konfigurations- und Entwicklungsbedarf)
- Geringes Wissen über moderne ERP-Funktionalitäten (z.B. Cloud-Integration, Internet der Dinge (IoT) oder Nutzung künstlicher Intelligenz bzw. AI), deren Flexibilität und Modularität sowie über bestehende KMU-geeignete Lösungen
- Zeitmangel in der Geschäftsführung zur Vorbereitung der Implementierung (Bedarfsanalyse)

Trotz dieser Hindernisse besteht ein echtes Bedürfnis seitens der KMU, ERP-Systeme als Standardlösung zur Digitalisierung ihrer Prozesse zu nutzen – verbunden mit einem Bewusstsein für deren strategische Bedeutung und Nutzen.

In diesem Kontext lassen sich zwei zentrale Forschungsrichtungen unterscheiden, die sich diesen Herausforderungen widmen:

1. Methodisch-orientierte Ansätze: Dieser Forschungsstrang deckt hauptsächlich die Methodik ab, indem er Ansätze, Arbeitsblätter und Grundsätze für die Formalisierung des Bedarfsausdrucks sowie für die Bewertung und Umsetzung von Lösungen bereitstellt. Dazu zählen praxiserprobte Maßnahmen für die digitale Transformation [3], Modelle zur Abschätzung des erforderlichen Anpassungsgrads eines ERP [8], Methoden zur Analyse bestehender Anforderungen [9], Ansätze zur Formalisierung von Anforderungen bei der ERP-Einführung [10], sowie Leitfäden zur Bedarfsdefinition [11] und zur Auswahl von ERP-Systemen [12].

Diese Ansätze strukturieren den Prozess der Bedarfsermittlung und der ERP-Auswahl, setzen jedoch eine manuelle Analyse durch Fachleute mittels Interviews und Dokumentenstudien voraus. Sie bieten KMU keine konkrete Unterstützung, um diese Phase eigenständig und kostengünstiger durchzuführen.

2. Technisch-orientierte Ansätze zur Automatisierung: Der zweite Forschungsstrang konzentriert sich auf technische Lösungen und Automatisierung, um die ERP-Einführung zu erleichtern. Diese Arbeiten zielen hauptsächlich auf die automatische Implementierung von ERP-Systemen auf Basis bereits vorhandener und formalisierter Spezifikationen. Ein Beispiel ist das Projekt „NEXT“ [13], das die vollständige Generierung maßgeschneiderter ERP-Systeme anhand ontologischer Unternehmensmodelle anstrebt.

Andere Forschungsarbeiten zielen darauf ab, die Erstellung von Spezifikationen zu automatisieren, indem sie sich auf Softwaremodelle stützen oder die vom Benutzer in natürlicher Sprache bereitgestellten Informationen interpretieren [14].

Diese Ansätze beschränken sich jedoch auf die Generierung von Spezifikationen und bieten keine Vergleichsmöglichkeiten mit Standardprozessen oder Funktionalitäten – auch begleiten sie die Unternehmen nicht aktiv in der Analysephase. Nach aktuellem Kenntnisstand des Projektteams existiert kein unabhängiges, agnostisches und vollständig Self-Service-basiertes Tool, das Unternehmen bei der Anforderungsdefinition und Pflichtenhefterstellung für ein ERP unterstützt.

Bezüglich der Bibliothek an ERP-Systemen und Integratoren wurde eine weitere Plattform identifiziert (<https://www.erpfuehrer.ch/>). Diese ist jedoch sehr einfach gehalten, umfasst nur 10 Fragen und verfolgt primär kommerzielle Ziele.

Im Rahmen dieses Projekts schlagen wir daher eine neutrale, agnostische und kostenfrei nutzbare Self-Service-Plattform vor, die auf einer Bibliothek standardisierter ERP-Prozesse und -Funktionen sowie auf einem Verzeichnis von Lösungen und Anbietern basiert. Diese Plattform ermöglicht zunächst die Erstellung eines ersten Pflichtenhefts für die ERP-Einführung und reduziert somit den Zeitaufwand und die Anfangsinvestitionen in der Analysephase. Im Anschluss unterstützt sie Entscheidungsträger bei der Auswahl geeigneter ERP-Systeme und Dienstleister.

Ziel des Projekts ist es zu prüfen, ob eine solche Plattform den Einstieg in die ERP-Einführung erleichtert und spezifische Herausforderungen, wie von Geschäftsführern beschrieben, adressieren kann. Im Zentrum stehen dabei folgende Forschungsfragen:

- Lassen sich identifizierte Probleme (z. B. Überschätzung der Individualität, mangelnde Verfügbarkeit, fehlendes Wissen, Ressourcen- und Budgetrestriktionen) durch ein Self-Service-Analysetool verringern?
- Liefert ein solches Tool Analyseergebnisse von vergleichbarer Qualität wie eine kostenintensive manuelle Expertenanalyse?
- Kann das Tool verborgene, nicht explizit geäußerte Anforderungen ebenso identifizieren wie ein erfahrener Berater?
- Wird ein solches Tool von KMU-Führungskräften als nützlich und zielführend eingeschätzt?

Zu beachten ist, dass sich dieses Projekt auf in der Schweiz ansässige Industrie-KMU beschränkt. Prozesse, Funktionen und Lösungen aus anderen Branchen wurden im initialen Rahmen nicht berücksichtigt. Ebenso blieben Themen wie Projekt- und Change-Management-Kompetenz, Unternehmenskultur oder digitale Reife unberücksichtigt, da sie bereits ausführlich in der Literatur behandelt und in anderen Projekten in der Schweiz und Europa untersucht wurden.

2. Methodologie

Zur Entwicklung der Plattform wurde zunächst nach einer geeigneten Bibliothek für Geschäftsprozesse gesucht. Eine Bestandsaufnahme des aktuellen Stands der Technik führte zur Identifikation von acht in der Industrie weit verbreiteten Frameworks. Diese wurden anschließend anhand von zehn quantitativen und qualitativen Kriterien (Umfang/Anzahl der Prozesse, Preis/Lizenzmodell, Anwendungsbereich, Aktualität/Frequenz der Updates, Anpassbarkeit/Verfügbarkeit, Detaillierungsgrad, Abdeckung/Vollständigkeit, Qualität der Beschreibungen) bewertet, um ihre Relevanz für unser Projekt zu bestimmen und eines davon auszuwählen.

Die zweite zentrale Komponente der Plattform ist die ERP-Bibliothek. Hierfür wurden die relevanten Metadaten definiert, die für die Auswahl und Empfehlung von ERP-Systemen nach der Generierung eines Pflichtenhefts erforderlich sind. Es folgte eine systematische Recherche zur Erhebung aller in der Schweiz verfügbaren ERP-Systeme und deren aktiven Integratoren.

Basierend auf diesen Vorarbeiten wurde die Plattform entwickelt und mit zwei aufeinanderfolgenden Unternehmensgruppen getestet.

Zur Bewertung der Nützlichkeit der Plattform und zur Beantwortung der Forschungsfragen wurde ein Fragebogen mit 9 geschlossenen Ja/Nein-Fragen und 12 Fragen mit Bewertungsskala erstellt. Dieser richtete sich an KMU-Verantwortliche, die an den Tests teilnahmen.

Die erste Testgruppe bestand aus fünf Unternehmen. Diese Phase diente in erster Linie der Behebung technischer Probleme (Bugs) im Tool. Anschließend wurde eine zweite Testphase mit zwölf weiteren Unternehmen durchgeführt.

Neben den Fragebögen wurden zur Erfassung der Eindrücke und Rückmeldungen auch unstrukturierte Interviews mit allen Nutzerinnen und Nutzern geführt.

3. Ergebnisse

Im Rahmen dieses Projekts wurden acht in der Industrie weit verbreitete Frameworks identifiziert und evaluiert, um als Grundlage für unsere Prozessbibliothek zu dienen. Diese Frameworks waren:

- APQC
- COBIT
- EAS EssentialViewer
- eTOM
- SAP Best Practices
- Microsoft Dynamics 365 Process Library
- MIT Process Handbook
- SCOR

Letztlich fiel die Wahl auf das Productivity & Quality Center (APQC) des American Productivity & Quality Center (APQC) [15], das wir inhaltlich angepasst und sprachlich überarbeitet haben, um die Verständlichkeit für Geschäftsführende ohne IT-Hintergrund zu erhöhen. Die Inhalte wurden vollständig ins Französische und Deutsche übersetzt. Unsere Bibliothek umfasst 13 funktionale Module (z. B. Finanzen, Produktion), 46 Hauptfunktionen (z. B. Einkauf, Personalrekrutierung), 184 Prozesse und 673 detaillierte Aktivitäten (z. B. „Erstellung eines Bewerberdossiers“, „Berechnung von Beschaffungskosten“). Die Kriterien zur Auswahl des zu nutzenden Frameworks waren u.A.: **Größe** (Größe des Frameworks im Verhältnis zur Anzahl der Prozesse und Funktionen), **Preis** (z.B. Lizenzgebühren für die Nutzung des Frameworks), **Bereich** (Sektor, für den das Framework erstellt wurde), **Aktualität** (letztes Update des Frameworks), **Hierarchiegrad** (Detaillierungsgrad der

Hierarchie). **Vollständigkeit** (Vollständigkeit des Frameworks), **Granularität** (Detailgrad in Bezug auf Funktionen, Prozesse und Aktivitäten) sowie **Qualität / Präzision** (persönliche Einschätzung zur Qualität des Frameworks).

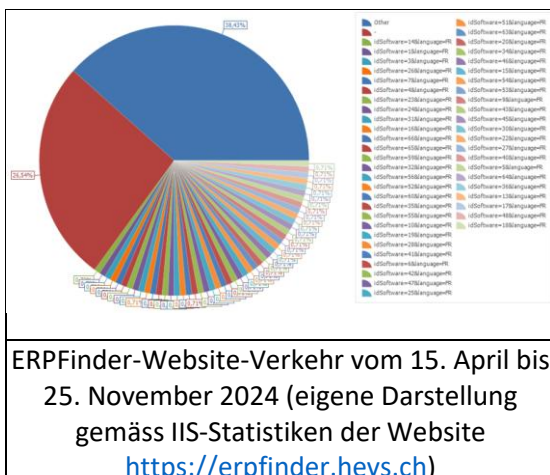
Des Weiteren wurde eine umfassende Datenbank mit 75 ERP-Systemen und 241 aktiven Integratoren in der Schweiz aufgebaut. Die entwickelte Plattform ist als klassische Webanwendung in .NET nach dem MVC-Modell mit Microsoft Entity Framework und einer SQL-Server-Datenbank realisiert und unter <https://erpfinder.hevs.ch> öffentlich zugänglich. Die Plattform verfügt über zwei Hauptfunktionen:

1. Generierung eines Pflichtenhefts durch eine geführte Auswahl relevanter Prozesse und Aktivitäten, mit Ausgabe eines formatierten Word-Dokuments.
2. Empfehlung passender ERP-Systeme und Integratoren anhand des erstellten Pflichtenhefts. Die erste Version arbeitete mit festen Gewichtungen; später wurde ein Modul integriert, das Nutzer*innen erlaubt, Gewichtungen (z. B. geografische Nähe) individuell anzupassen.

Nach Abschluss der Implementierungsphasen wurden die Rückmeldungen der 17 beteiligten Unternehmen durch standardisierte Fragebögen und nicht-strukturierte Interviews gesammelt. Nicht alle Unternehmen haben die Tests vollständig abgeschlossen oder Feedback gegeben. Die wichtigsten Ergebnisse im Hinblick auf die ursprünglich formulierten Forschungsfragen lauten wie folgt:

- Reduktion von Einführungsbarrieren: 70 % der Befragten sind der Meinung, dass das Self-Service-Werkzeug allein nicht ausreicht, um die Haupthindernisse für die Einführung eines ERP-Systems signifikant zu senken.
- Qualität der automatisierten Analyse: 88 % der Teilnehmenden bewerteten das generierte Pflichtenheft als relevant und qualitativ hochwertig. Jedoch gaben 35 % an, dass die Ergebnisse nicht präzise genug seien und die spezifischen Gegebenheiten ihres Unternehmens nicht vollständig abbilden.
- Nur wenige Nutzer berichteten, dass sie die erweiterten Funktionen des Tools genutzt haben, um Bedürfnisse aufzulisten, die in der Standardbibliothek nicht spezifiziert sind.
- Akzeptanz durch KMU: 100 % der Befragten lobten die Benutzerfreundlichkeit, Klarheit und Struktur des Tools. Keine Geschäftsführer*in sprach sich grundsätzlich gegen den Einsatz eines solchen Werkzeugs aus.

Zusätzlich wurden Zugriffszahlen der Plattform ausgewertet: Nur 26,5 % der Besucher*innen begannen die Generierung eines Pflichtenhefts, 38,4 % verweilten lediglich auf der Startseite und 35 % riefen die ERP-Bibliothek auf.



4. Diskussion

Entgegen unseren ursprünglichen Erwartungen und Hypothesen konnte das entwickelte Self-Service-Werkzeug zur Generierung von Pflichtenheften die wesentlichen Hürden bei der Einführung eines ERP-Systems nicht vollständig beseitigen. Die wahrgenommene Komplexität sowie der Mangel an Fachkenntnissen seitens der Geschäftsleitungen bleiben bedeutende Hindernisse. Trotz einer schrittweisen Benutzerführung und der Nutzung eines standardisierten Referenzmodells berichten viele Nutzerinnen von einer anhaltenden „Angst, etwas falsch zu machen“ und ziehen es vor, die Analyse spezialisierten Beraterinnen zu überlassen. Diese Abhängigkeit unterstreicht, dass der menschliche Faktor in der Voranalyse weiterhin zentral bleibt und das Tool kein vollständiger Ersatz für die Expertise eines Fachberaters sein kann.

Die Qualität der automatisierten Analyse wird verschieden wahrgenommen: Während einige Nutzer*innen die standardisierte Basis als solide und zuverlässig einstufen, erkennen andere ihre betriebliche Realität in den generischen Beschreibungen nicht wieder. Diese Divergenz ist teilweise der notwendigen Granularität geschuldet, um die Prozessbibliothek handhabbar zu halten, zeigt jedoch zugleich die Bedeutung einer stärker individualisierbaren Ausgestaltung der Inhalte auf.

Die Identifikation versteckter Anforderungen stellt insbesondere für weniger erfahrene KMU eine Herausforderung dar: Nur fortgeschrittene Nutzerinnen konnten zusätzliche, nicht antizipierte Bedürfnisse über die standardisierten Prozesse identifizieren. Den übrigen fehlt häufig der notwendige Abstand, um die abstrakten Beschreibungen mit ihren konkreten Geschäftsprozessen in Verbindung zu bringen. Dies bestätigt, dass menschliche Begleitung eine zentrale Rolle spielt, um die Autonomie der Unternehmen in der Analysephase zu stärken. Insbesondere Führungskräfte benötigen Beratung, Begleitung und Vertrauen in ihre Implementierungspartner. Diese Beobachtung steht im Einklang mit den Rückmeldungen zur ERP-Bibliothek: Die Nutzerinnen sind in erster Linie auf der Suche nach einem verlässlichen Partner zur Projektumsetzung.

Trotz dieser Einschränkungen wird das Tool durchweg für seine Klarheit, Lesbarkeit und Struktur geschätzt. Es besteht keine grundsätzliche Ablehnung gegenüber der Nutzung eines digitalen Instruments zur Bedarfserhebung. Vielmehr wird sein Potenzial zur Vereinfachung der Analysephase anerkannt, und das Bewusstsein für die Notwendigkeit der Digitalisierung ist bei den Unternehmen vorhanden – auch wenn ein gewisser Bedarf an menschlicher Unterstützung weiterhin besteht.

Die größte Wirkung unserer Plattform liegt in der ERP- und Integratorenbibliothek: Die Bereitstellung einer aktuellen und vollständigen Übersicht über Lösungen und Anbieter deckt ein drängendes Bedürfnis der KMU. Sie durchbricht das bisher dominierende Muster der informellen Informationsbeschaffung über persönliche Netzwerke oder sporadische Online-Recherchen. Diese Funktion erzeugt starkes Nutzerengagement und strukturiert den Auswahlprozess auf faktenbasierter Grundlage. Dies steht im Kontrast zur vergleichsweise geringeren Nutzung des Pflichtenheftgenerators und zeigt, dass viele Unternehmen dazu neigen, direkt zur Lösungsauswahl überzugehen, ohne vorher eine systematische Bedarfsanalyse durchzuführen. Dieser Effekt offenbart eine gewisse Diskrepanz zwischen den Zielen des Projekts und den tatsächlichen Verhaltensweisen der Zielgruppe.

Trotz der positiven Ergebnisse ist das Projekt nicht frei von Beschränkungen:

- Die Teststichprobe ist klein, dies schränkt die Verallgemeinerbarkeit der Ergebnisse ein.
- Keine Langzeitstudie: Der langfristige Einfluss des Werkzeugs auf die tatsächliche ERP-Einführung wurde nicht gemessen.
- Kompetenzbias: Erfahrene Nutzer*innen konnten das Tool effektiver einsetzen, was die Einschätzung zur Erkennung verborgener Anforderungen verzerrt.

Zur Vertiefung der bisherigen Ergebnisse und zur Überwindung dieser Einschränkungen erscheint es vielversprechend, Möglichkeiten zur Stärkung der Autonomie weniger erfahrener Entscheidungsträger

zu untersuchen. Interaktive Tutorials oder geführte Szenarien könnten dabei helfen, das notwendige Abstraktionsvermögen für die Prozessanalyse zu entwickeln. Ebenso wäre es sinnvoll, den Grad der Prozessindividualisierung durch die Nutzer*innen steuerbar zu machen, um zwischen einfacher Bedienbarkeit und inhaltlicher Präzision wählen zu können.

Letztlich sollten weiterführende Tests mit größeren und diversifizierten Unternehmensgruppen erfolgen, ergänzt durch langfristige Untersuchungen, um den tatsächlichen Einfluss auf die ERP-Einführungsquote valide zu erfassen. Im weiteren Sinne stellt sich die grundsätzliche Frage nach dem optimalen Verhältnis zwischen digitalen Werkzeugen und menschlichem Coaching bei der Analyse von Geschäftsprozessen – eine Thematik, die vertieft erforscht werden sollte.

5. Schlussfolgerung

Zur Unterstützung kleiner und mittlerer Unternehmen (KMU) bei der Einleitung ihrer digitalen Transformation und der Einführung eines ERP-Systems wurde ein neutrales, anbieterunabhängiges Self-Service-Werkzeug entwickelt, das die Erfassung betrieblicher Anforderungen erleichtert und die Erstellung eines ersten ERP-Pflichtenhefts ermöglicht. Ergänzt wird die Plattform durch eine umfassende Bibliothek zu verfügbaren ERP-Systemen und deren Integratoren in der Schweiz.

Die Tests zeigen, dass die Anwendung hinsichtlich Struktur und Bedienbarkeit positiv aufgenommen wurde. Dennoch bleibt eine persönliche Begleitung durch Expert*innen unerlässlich, insbesondere um die verbreitete Unsicherheit („Angst, etwas falsch zu machen“) abzubauen und eine passgenaue Abbildung der unternehmensspezifischen Prozesse sicherzustellen.

Ein wesentliches Potenzial liegt in der ERP-Bibliothek: Die strukturierte Übersicht über Systeme und Anbieter trifft auf starkes Interesse bei den Unternehmen und bietet eine datenbasierte Alternative zur üblichen informellen Informationsbeschaffung. Dennoch zeigen die Ergebnisse auch die Grenzen einer rein automatisierten Vorgehensweise auf: Die Identifikation verdeckter Anforderungen und eine tiefgreifende Individualisierung erfordern weiterhin menschliche Expertise. Zudem bleibt der langfristige Einfluss auf die tatsächliche ERP-Adoption bislang unbewertet, was vor allem dem kleinen Stichprobenumfang und dem Fehlen einer Längsschnittstudie geschuldet ist.

Zusammenfassend lässt sich festhalten: Die erfolgsversprechendste Strategie zur Steigerung der ERP-Adoptionsrate in KMU liegt in einer **hybriden Lösung**, welche digitale Werkzeuge mit gezieltem Coaching kombiniert. Diese Herangehensweise bietet insbesondere unerfahrenen Entscheidungsträger*innen Orientierung und stärkt ihre Selbstständigkeit in der Analyse und Auswahl geeigneter ERP-Lösungen.

6. Referenzen

- [1] Commission Européenne. (2022, July 28). Indice relatif à l'économie et à la société numériques pour 2022 (Press release No. IP/22/4560).
https://ec.europa.eu/commission/presscorner/detail/fr/ip_22_4560
- [2] Eurostat (2023, August 19), E-business integration,
https://ec.europa.eu/eurostat/statistics-explained/index.php?title=E-business_integration
- [3] Sticha, V., Zeller, V., Hicking, J., & Kraut, A. (2020). Measures for a successful digital transformation of SMEs. <https://doi.org/10.1016/j.procir.2020.03.023>
- [4] Kilimis, P., Zou, W., Lehmann, M., & Berger, U. (2019). A survey on digitalization for SMEs in Brandenburg, Germany. <http://doi.org/10.1016/j.ifacol.2019.11.522>

- [5] Institut suisse pour les PME et l'entrepreneuriat, Université de Saint-Gall & OBT. (2022, April 19). Etude PME 2022. <https://www.obt.ch/de/infoboard/kmu-studie-2022/>
- [6] Matt, D. T., Modrák, V., & Zsifkovits, H. (2020). Industry 4.0 for SMEs: Challenges, opportunities and requirements. *Procedia Manufacturing*, 45, 447–455. <https://doi.org/10.1007/978-3-030-25425-4>
- [7] Agrawal, P., Narain, R., & Ullah, I. (2019). Analysis of barriers in implementation of digital transformation of supply chain using interpretive structural modelling approach. *Journal of Modelling in Management*, 14(3), 739–760. <http://doi.org/10.1108/JM2-03-2019-0066>
- [8] Deltour, F., Farrajallah, M., & Lethiais, V. (2014). L'équipement des PME en systèmes ERP : une adoption guidée par les priorités stratégiques ? *Management International*, 18(2), 155–168. <http://doi.org/10.7202/1024200ar>
- [9] Parthasarathy, S., & Daneva, M. (2016). An approach to estimation of degree of customization for ERP projects using prioritized requirements. *Information and Software Technology*, 77, 9–24. <http://doi.org/10.1016/j.jss.2016.04.006>
- [10] Khaleel, Y., Abuhamdah, A., Abu Sara, M., & Al-tamimi, B. (2016). Components and analysis method of enterprise resource planning requirements in small and medium enterprises. *International Journal of Business Information Systems*, 22(2), 216–236. <http://doi.org/10.11591/ijece.v6i2.pp682-689>
- [11] Asgar, T. S., & King, T. M. (2014). Formalising requirements in ERP software implementations. *Requirements Engineering Journal*, 19(4), 367–380. <http://doi.org/10.7763/LNSE.2016.V4.220>
- [12] Burgues Illa, X., Franch, X., & Pastor, J. A. (2000). Formalising ERP selection criteria. *Computers in Industry*, 43(2), 201–213. <https://dl.acm.org/doi/proceedings/10.5555/857171>
- [13] van der Schuur, H., van de Ven, E., de Jong, R., Schunselaar, D., Reijers, H. A., Overeem, M., de Graaf, M., Jansen, S., & Brinkkemper, S. (2017). NEXT: Generating tailored ERP applications from ontological enterprise models. *Information Systems*, 64, 245–266. https://doi.org/10.1007/978-3-319-70241-4_19
- [14] Nicolás, J., & Toval, A. (2009). On the generation of requirements specifications from software engineering models: A systematic literature review. *Journal of Systems and Software*, 82(12), 2121–2136. <http://doi.org/10.1016/j.infsof.2009.04.001>
- [15] American Productivity & Quality Center. (2024). APQC Process Classification Framework (PCF) – Cross-Industry, <https://www.apqc.org/resource-library/resource-listing/apqc-process-classification-framework-pcf-cross-industry-excel-11>

Inclusive Pedagogy for Learners with Hearing Impairments: A Comparative Study of Uganda and Germany

Keti Kajumba

Carl von Ossietzky Universität Oldenburg, Deutschland

Abstract: Inclusive education remains limited for learners with hearing impairments (LHIs), particularly in secondary schools. This study compares inclusive pedagogy practices and challenges in Uganda and Germany through interviews and classroom observations with 19 teachers and 7 students across four schools. Guided by Universal Design for Learning and inclusive special education theory, the analysis highlights varied pedagogical strategies, including differentiated instruction, group work, and multi-dimensional assessments. Germany further emphasizes parent-teacher engagement, collaborative teaching, and strong teacher-student relationships.

Challenges more pronounced in Uganda include undifferentiated curricula, inadequate teacher training, low motivation, and limited parental support. The study recommends greater government investment, specialized teacher training, adapted assessments, and strengthened parent-teacher collaboration. Findings underscore the importance of context-sensitive policies to enhance inclusion for LHIs in secondary education.

TRACK 4: DIDAKTIK

Track Chair: Prof. Dr. Tim Neumann

Hochschulübergreifende Unternehmenssimulation mit ERPsim: Konzeption, Herausforderungen und Lessons-learned

Agnes Pechmann

Hochschule Emden/Leer, Deutschland

Dirk Frosch-Wilke

Fachhochschule Kiel, Deutschland

Abstract: In der heutigen globalisierten Welt sind internationale Kompetenzen und die Fähigkeit, in virtuellen Teams effektiv zu kommunizieren und zusammen zu arbeiten, wichtige Anforderungen im Berufsleben. Um Hochschulstudierenden die Möglichkeit zu bieten, mit Studierenden aus unterschiedlichen Ländern in virtuellen Projektteams zu arbeiten, stellen wir im Folgenden ein hochschulübergreifendes Projekt vor. In diesem Projekt haben internationale Masterstudierende der Hochschule Emden/Leer mit Masterstudierenden der Fachhochschule Kiel in gemischten, virtuellen Teams am Unternehmensplanspiel ERPsim teilgenommen. In dieser Studie stellen wir die wesentlichen Erkenntnisse dieses Projektes bzgl. notwendiger organisatorisch-technischer Maßnahmen, der Teambildung und -reife und der Zusammenarbeit der studentischen Teams vor. Damit wollen wir anderen Hochschulen die Möglichkeiten zur hochschulübergreifenden Durchführung von Unternehmensplanspielen aufzeigen und Handlungsempfehlungen geben.

Keywords: Virtuelle Teamarbeit, Unternehmens-ERP-Planspiel, Gruppenbildung/-reife, Hochschulübergreifende Lehrveranstaltungen

1. Einleitung

Die Hochschulausbildung soll Studierende auf eine erfolgreiche berufliche Karriere vorbereiten, weshalb innovative Lehrmethoden erforderlich sind [Hughes et al. 2008, Beranic, Hericko 2022]. Ein vielversprechender Ansatz ist das digitale spielbasierte Lernen, das interaktive Unterhaltung und ernsthaftes Lernen durch digitale Spiele kombiniert [Anastasiadis et al 2018]. Pädagogische digitale Lernspiele, wie Educational Serious Games, können auf die Anforderungen und Bedürfnisse der Studierenden eingehen und umfassen Unternehmensplanspiele, theoriebasierte Spiele und Trainings in Fertigungs- und Lieferprozessen [z.B. Zarte, Pechmann 2019].

Unternehmensplanspiele haben sich als wirksames Instrument für den Erwerb von Wirtschafts- und Managementkompetenzen bei Hochschulstudenten erwiesen [Huang et al. 2022]. Durch die Teilnahme an Unternehmensplanspielen können Studierende ihre Fähigkeiten im Bereich der Teamarbeit, Kommunikation und Entscheidungsfindung entwickeln [Goi 2019]. Die Integration von virtueller und geografisch-verteilter Teamarbeit in Unternehmensplanspiele ist von besonderer Bedeutung, da Unternehmen aufgrund der Globalisierung und des technologischen Fortschritts Mitarbeitende benötigen, die in virtueller Teamarbeit geschult sind [Jony, Serradell-López 2021].

Um diese Kompetenzen auszubilden und zu trainieren, haben die AutorInnen dieses Berichts im Sommer 2023 eine hochschulübergreifende Durchführung von Planspielen geplant. Das Ziel dieses Beitrages ist es, Erfahrungen mit der hochschulübergreifenden Durchführung des Planspiels ERPsim aufzuarbeiten und darauf basierend Handlungsempfehlungen herzuleiten. Insbesondere wird untersucht, ob Planspiele eine effektive Methode zur Förderung von Teamfähigkeiten sind.

Der Artikel ist wie folgt aufgebaut: Im ersten Teil wird der theoretische Hintergrund des Projekts vorgestellt, im zweiten Teil die Methodik des Projekts beschrieben, im dritten Teil die Ergebnisse präsentiert und im vierten Teil die Ergebnisse diskutiert und Handlungsempfehlungen gegeben.

2. Methodik

2.1. Kontext der Studie

Die Daten für diese Studie wurden durch ein gemeinsames Unternehmensplanspiel im Rahmen von Lehrveranstaltungen im Sommersemester 2025 von zwei unterschiedlichen Hochschulen gewonnen. Die Vorbereitung hierfür wurden im Herbst 2024 gestartet, nachdem ein vorheriger Anlauf wegen zu geringem Planungsvorlauf abgebrochen wurde.

Von der Fachhochschule Kiel nahmen sechs Studierende des deutschsprachigen Masterstudiengangs Wirtschaftsingenieurwesen im Rahmen des Moduls Informationsmanagement teil. Dieses Modul ist curricular dem ersten Semester des Studiengangs zugeordnet. Die Durchführung des Unternehmensplanspiels ERPSim ist üblicherweise nicht Inhalt des Moduls wurde aber im Sommersemester 2025 zum Zweck der Durchführung der Studie ermöglicht.

Von der Hochschule Emden/Leer waren 10 Studierende der englischsprachigen Masterstudiengänge Technical Management sowie Business Intelligence and Data Analytics im Rahmen des Moduls Project ERPSim Challenge beteiligt. Die internationalen Studierenden kamen zum großen Teil aus Indien. Das Modul ist ein ergänzendes, fakultatives Angebot zum Modul ERP- Systems im Umfang von 3 ECTS. Es wurde zum zweiten Mal durchgeführt.

Die Studierenden wurden in fünf verschiedene Teams aufgeteilt. Dabei wurde sichergestellt, dass in jedem der Teams Studierende unterschiedlicher Hochschulen beteiligt waren.

Das Unternehmensplanspiel wurde an der Hochschule Emden/Leer in einem Veranstaltungsraum durchgeführt und die Studierenden der Fachhochschule Kiel nahmen daran virtuell mittels der Videokonferenzsysteme Zoom bzw. Big Blue Button teil. In der ersten Simulationsrunde trafen sich die Kieler Studierenden noch in einem gemeinsamen Raum an der Fachhochschule. In weiteren Simulationsrunde schalteten diese sich jedoch von unterschiedlichen Orten zu.

2.2. Das verwendete Planspiel

Als Unternehmensplanspiel kam ERPSim mit den Spielen Manufacturing Intro bzw. Sustainability zum Einsatz. Die ERPSim Manufacturing Planspielreihe des ERPSim Lab an der HEC Montréal simuliert betriebswirtschaftliche Entscheidungsprozesse in einem realitätsnahen Produktionsumfeld auf Basis von SAP S/4HANA. Teilnehmende übernehmen die Steuerung eines virtuellen Unternehmens der Müsliproduktion und treffen unter Zeitdruck strategische Entscheidungen zu Einkauf, Produktion und Vertrieb. In jedem Spiel können bis zu 12 Runden a 20 Tagen (steps) gespielt werden. Jeder Tag läuft in der Regel binnen einer Minute ab. Im Manufacturing Introduction Game wird der Umfang des Geschäftsprozesses pro Runde erhöht, so dass ab der 3. Runde der volle Geschäftsprozess dieser Version gespielt wird. Dies dient der Einführung der Studierenden in das Spiel, ohne sie mit dem Prozess und der Handhabung des SAP S/4 HANA Systems zu stark zu überfordern.

Das Sustainability Manufacturing Game erweitert dieses Szenario um Aspekte der Logistik und Produktionsoptimierung sowie der Nachhaltigkeit. Die Nachhaltigkeit wird über die CO₂-Kennzahlen gesteuert insbesondere über den CO₂-Ausstoß, nachhaltige Investitionen, CO₂-Bilanzen und gesetzliche Rahmenbedingungen, die in die Spielmechanik integriert sind. [ERPSim Lab 2025]

2.3. Der Lehrplan der gemeinsamen Veranstaltungen

Das Ziel der Veranstaltung war die Studierenden zum einen im Bereich der virtuellen Teamarbeit mit verteilen, in zunächst unbekannten und kulturell inhomogenen Teams zu schulen. Zum anderen sollten sie praktische Erfahrungen im Bereich der Entscheidungsfindung von vereinfachten, aber realitätsnahen Geschäftsprozesse machen und die Bedeutung von ERP-Systemen und Datenanalysen für die Entscheidungsfindung erkennen. Die gemeinsamen Veranstaltungen sollten sich in die übergeordneten Modul-Lehrplan einordnen und die Erkenntnisse und Erfahrungen dort integriert werden.

Es wurden acht Veranstaltungen a 1 ½ h wöchentlich im Zeitraum von Anfang April bis Ende Mai 2025 in englischer Sprache durchgeführt. Nach einer Einführung in die Veranstaltung, der Einführung in das Spiel sowie der Gruppeneinteilung, sind vier Planspiele mit ansteigendem Komplexitätsgrad durchgeführt worden. Zudem sind

zwei Veranstaltungen dem Thema Datenanalyse gewidmet worden, eine unter Nutzung von Business Builders <https://erpsim.hec.ca/en/businessbuilders> mit SAP Analytics Cloud und einmal eine Einführung in KMIME <https://www.knime.com/downloads> mit exportierten Planspieldaten via ODATA nach Excel. Als Prüfungsform für die Studierenden aus Kiel wurde sich für die unbenotete, erfolgreiche Teilnahme an der Veranstaltung entschieden und die Benotung erfolgte mit einer Prüfungsleistung im übergeordneten Modul. Im Rahmen des eigenständigen, fakultativen Projekt-Moduls der Emder Studierenden als Studienleistung, hatten diese neben der erfolgreichen Teilnahme eine kurze Analyse des letzten gemeinsamen Planspiels abzugeben oder alternativ, am Planspielwettbewerb auf EMEA-Ebene teilzunehmen.

2.4. Vorbereitungen für die hochschulübergreifenden Veranstaltungen

Die Durchführung der gemeinsamen Veranstaltungen erforderte eine sorgfältige Planung und Organisation. Um sicherzustellen, dass die Veranstaltungen erfolgreich durchgeführt werden konnten, werden zunächst die organisatorisch administrative Aufgaben in Tabelle 1 aufgeführt und dann die organisatorisch technischen Aufgaben in Tabelle 2. Zeitlich gesehen liefen die Durchführung dieser Aufgaben z.T. parallel ab.

In Tabelle 1 werden die durchgeführten Aufgaben stichpunktartig dargestellt und können als eine Art "Checkliste" verwendet werden.

Tabelle 1 - Organisatorisch-administrative Aufgaben

Willenserklärung der interessierten Professoren	Alle beteiligten Professoren haben ihre Bereitschaft zur Teilnahme an der Veranstaltung erklärt und ihre generelle Unterstützung für die Durchführung der gemeinsamen Veranstaltungen zugesagt.
Klärung der Kosten	Kosten für die Planspiellizenzen und deren Übernahme sowie ggfls weiterer anfallender Kosten.
Bestimmung des gemeinsamen Veranstaltungszeitraums und Anzahl der Veranstaltungen	Der gemeinsame, d.h. überlappende Vorlesungszeitraum wurde mit den beteiligten Professoren abgestimmt und die Anzahl der Veranstaltungen wurde festgelegt. Dies war wichtig, da die gemeinsamen Planspielveranstaltungen in das Gesamtkonzept des Studiengangs bzw. der beteiligten Module integriert sein sollten.
Abstimmung des Stundenplans	Der Stundenplan wurde mit den beteiligten Professoren abgestimmt und den Stundenplankoordinatoren der beteiligten Hochschulen abgestimmt, um sicherzustellen, dass die Veranstaltungen ohne Konflikte für die Studierenden durchgeführt werden konnten.
Wahl der verwendeten Lernplattform	Die Lernplattform wurde ausgewählt und konfiguriert, um sicherzustellen, dass alle Studierenden den gleichen Zugang zu den Materialien und der Veranstaltung hatten. Zudem diente sie der unidirektionalen Kommunikation (Professoren zu Studierenden).
Abstimmung bzw. Abgleich der groben Vorlesungsinhalte der beteiligten Module der Hochschulen	Die Professoren tauschten sich über die groben Vorlesungsinhalte, um sicherzustellen, dass die Veranstaltungen den Anforderungen der Studiengangsmodule entsprachen.
Abstimmung der Lehrziele	Die Lehrziele für die gemeinsamen Veranstaltungen wurden mit den beteiligten Professoren abgestimmt.
Erstellen eines Vorlesungsplans mit Inhalten	Ein Vorlesungsplan wurde erstellt, der die Inhalte der Veranstaltungen enthielt.
Abstimmung der begleitenden	Die begleitenden Analysemethoden und -werkzeuge wurden mit den beteiligten Professoren abgestimmt und Lizenzfragen geklärt.

Analysemethoden im Planspielkontext	
Auswahl der zu verwendeten Planspielversionen, Gruppengröße und -zusammensetzung	Die Planspielversionen, Gruppengröße und -zusammensetzung wurden ausgewählt, um sicherzustellen, dass sie zu den Lernzielen der Professoren passten.
Aufteilung der Lehrinhalte auf die beteiligten Lehrenden	Die Lehrinhalte wurden auf die beteiligten Lehrenden aufgeteilt, um sicherzustellen, dass diese inhaltlich und zeitlich durchgeführt werden konnten.
Anzahl der teilnehmenden Studierenden	Die Zahl der Teilnehmenden für die Klärung der Planspiellizenzen wurde benötigt.
Identifikation der teilnehmenden Studierenden und ihrer Email-Adressen	Die Email-Adressen waren notwendig für die Freischaltung der Studierenden zur Lernplattform.
Methode und Werkzeug zur Einteilung der Teams	Die Einteilung der dezentral verteilten, und sich zum Teil unbekannten Teammitglieder musste festgelegt werden und die Bekanntgabe der Teammitglieder organisiert werden. Dies erfolgte z.T. online nach dem Zufallsprinzip sowie nachfolgend mit Eintragung in vorbereitete Teams im Moodle-Kurs zum Austausch der Email-Adresse und Weitergabe der persönlichen Kontaktmöglichkeiten durch die Studierenden.

Tabelle 2 - Organisatorisch-technische Aufgaben

Klärung der technischen Möglichkeiten einer gemeinsamen Lernplattform	Die technischen Zugangsmöglichkeiten einer gemeinsamen Lernplattform wurden geklärt, um sicherzustellen, dass alle Studierenden Zugang zu den Veranstaltungen hatten.
Einrichten und Vorbereiten eines Moodle-Kurses (Lernplattform)	Ein Moodle-Kurs wurde eingerichtet und die Zugriffsmöglichkeiten für die beteiligten Professoren sichergestellt.
Einschreiben der externen Studierenden auf die gewählte Lernplattform	Die externen Studierenden wurden auf die gewählte Lernplattform vom Rechenzentrum der hostenden Hochschule eingeschrieben und über die Zugangscredentials informiert.
Wahl der Video-konferenzplattform und Kommunikation von Alternativen	Die Video-konferenzplattform wurde ausgewählt und die Alternative wurden kommuniziert
Wahl des Kommunikationsmediums für die Studierendenteams (Angebot und freie Wahl)	Eine Kommunikationsmöglichkeit über die Lernplattform wurde angeboten, und deren Nutzung oder einer Alternative freigestellt.
Einbindung der Interviews	Die Umfragen wurden eingebunden, um sicherzustellen, dass alle Studierenden einen einfachen Zugang zu diesen zu gewährleisten.
Bestimmung der Hardware für die Durchführung der Planspiele mit Studierenden in Präsenz und virtuell (Gruppe) sowie einzeln, verteilt	Die Hardware wurde bestimmt, um sicherzustellen, dass die Veranstaltungen erfolgreich in Präsenz durchgeführt werden konnten. In diesem Fall wurde ein vorhandenes Konferenzsystem zur Verbindung mit dem Dozenten-Rechner gewählt.

Installation des Mini-Konferenzsystems im Pool-Raum	Das Mini-Konferenzsystem wurde im Pool-Raum installiert, um sicherzustellen, dass die Veranstaltungen erfolgreich durchgeführt werden konnten.
Stummschalten der Studierenden Mikrofone und Lautsprecher während der gemeinsamen Session	Die Studierenden Mikrofone und Lautsprecher wurden stummgeschaltet, um sicherzustellen, dass die Veranstaltungen ohne Rückkopplungen durchgeführt werden konnten.

2.5. Datensammlung und -analyse

Die Daten wurden mittels eines Mixed-Methods-Ansatzes (vgl. [Creswell, Clark 2011]) erhoben. Als quantitative Methode wurden zwei Umfragen durchgeführt. Beobachtung und Interview dienten der qualitativen Datenerhebung.

Die Studierenden wurden vor der ersten Spielsitzung im Rahmen einer anonymen Umfrage zu folgenden Aspekten befragt: Enthusiasmus bzgl. der Teilnahme am Unternehmensplanspiel und der Zugehörigkeit zu einem studentischen Team; Erwartung zum Wissenserwerb; Bedenken bzgl. der Remote-Teamarbeit; Bedeutung des Unternehmensplanspiels im Vergleich zu anderen Modulen im selben Semester.

Die zweite Umfrage wurde nach drei Planspielrunden durchgeführt. Bei dieser Umfrage ging es primär darum den Teamentwicklungsstand zu erheben. Dieser Zeitpunkt wurde deshalb gewählt, weil Gersick [Gersick 1988, 1989] herausfand, dass Gruppen, die an einer spezifischen Aufgabe mit einem vorgegebenen Zeitlimit arbeiteten, signifikante Übergänge (d. h. Verschiebungen) in den Ansätzen ihrer Arbeit machten, und zwar zur oder fast zur Mitte der Gruppenarbeit.

Insbesondere in der ersten Spielrunde wurde die Methode der Fremdbeobachtung eingesetzt. Hierbei wurden das Verhalten der Studierenden durch die Dozierenden erfasst. Nach Abschluss des Unternehmensplanspiels wurden noch Interviews mit einzelnen Studierenden durchgeführt. Diese dienten der Reflexion des Planspiels bzgl. der Aspekte Lernziele, Teamentwicklung, verwendete Kommunikationsmittel und Eignung von ERPsim als Planspiel im Kontext einer hochschulübergreifenden Durchführung. Dabei sollten die Studierenden auch Angaben zu ihren „lessons-learned“ und Verbesserungsmöglichkeiten machen.

3. Ergebnisse

3.1. Planspielergebnisse

Bei den Planspielen wurde zunächst die Version Manufacturing Introduction gespielt. Aufgrund der zeitlichen Restriktion einer Vorlesung (1,5 h) wurde zunächst nach der Einführung nur eine Runde gespielt. Die erste Runde besteht in der Regel nur aus dem Verkauf von bereits produzierten Fertigprodukten. Die Studierenden beherrschten die Transaktionen relativ gut, so dass bereits nach der ersten Hälfte der Runde 1 die Transaktionen zum Steuern der Produktion freigegeben wurden. Im Vergleich zum regulären Spielbetrieb der Introduction Version wurden also die Runde 1 und 2 zu einer Runde zusammengelegt. Dies spiegelt sich auch in den Ergebnissen dieser Planspielrunde wieder (s. Tabelle 3). Alle Teams haben einen sehr hohen Unternehmenswert (Company Valuation) von über 120 Mio € bei einer geringen Produktivität von maximal 40%. Da die Produktivität jedoch auf der Basis von 20

Tagen berechnet wird (=eine Runde) jedoch nur in der zweiten Hälfte, also für zehn Tage produziert wurde und zudem eine Rüstzeit von acht Stunden nach jedem Produktwechsel anfällt, ist die Produktivität als hoch zu betrachten.

Tabelle 3- Planspielergebnis Runde 1 Manufacturing Introduction

Leaderboard																			
Round	Team	Team	Credit ratings	Interest Rate (%)	Rank	Company valuation	Cumulative Net Income	Total sales	Gross Margin (%)	Net Margin (%)	ROE (%)	ROA (%)	D/E (%)	Mktg/S (%)	Round Net Income	Round Productivity	Round Sales per Team		
1	A	A	AAA+	6.000	2	189,702,772.80	1,580,856.44	3,718,269.58	67.157	42.516	7.325	7.143	2.556	0.015	1,580,856.44	39.999	3,718,269.58		
	B	B	AAA+	6.000	3	179,178,844.80	1,493,157.04	3,503,010.69	69.826	42.625	6.947	6.694	3.778	0.069	1,493,157.04	38.332	3,503,010.69		
	C	C	AAA+	6.000	1	189,913,173.60	1,582,609.78	3,640,555.42	67.260	43.472	7.333	7.242	1.251	0.001	1,582,609.78	39.999	3,640,555.42		
	D	D	AAA+	6.000	4	176,767,406.40	1,473,061.72	3,388,522.16	69.022	43.472	6.860	6.749	1.643	0.000	1,473,061.72	17.291	3,388,522.16		
	E	E	AAA+	6.000	5	122,132,130.00	1,017,767.75	2,980,350.88	68.540	34.149	4.842	4.573	5.887	0.203	1,017,767.75	40.000	2,980,350.88		

Bei den nachfolgenden Planspielen wurde auf die Version von Manufacturing Sustainability gewechselt.

Beim nächsten Spiel erfolgte zunächst die Einführung in die neue Spielversion. Es wurden zwei Runden an nacheinander folgenden Vorlesungen durchgeführt. Das Spiel wurde aufgrund fehlender Teammitglieder an beiden Vorlesungstagen mit vier Teams gespielt (Team B bis E). Die verbleibenden Teammitglieder von Team A wurden auf die anderen Gruppen aufgeteilt. Es kam somit in den beiden Runden zu einer unterschiedlichen Teamzusammensetzung. Zunächst wurde mit vereinfachter Logistik, ohne Lieferantenauswahl und Beachtung des CO₂-Ausstosses gespielt und diese Funktionalitäten dann sukzessive freigegeben. Das Ergebnis ist in Tabelle 4 dargestellt.

Tabelle 4- Spiel 2 – Manufacturing Sustainability

Leaderboard																			
Round	Team	Team	Credit ratings	Interest Rate (%)	Rank	Company valuation	Cumulative Net Income	Total sales	Gross Margin (%)	Net Margin (%)	ROE (%)	ROA (%)	D/E (%)	Mktg/S (%)	Round Net Income	Round Productivity	Round Sales per Team	CO ₂ e (kg)	CO ₂ e per Unit Sold (kg/unit)
1	B	B	A+	7.750	3	29,582,604.26	289,663.00	1,333,670.18	64.896	21.719	1.361	0.961	41.520	0.014	289,663.00	65.000	1,333,670.18	475,637.50	1.64
	C	C	AA-	7.250	4	18,054,866.13	169,264.37	1,196,143.79	61.568	14.151	0.800	0.567	41.074	0.000	169,264.37	63.333	1,196,143.79	396,443.25	1.47
	D	D	AA-	7.250	2	33,809,553.07	316,964.56	1,407,875.18	62.651	22.514	1.487	1.054	41.046	0.000	316,964.56	70.000	1,407,875.18	408,648.75	1.32
	E	E	A+	7.750	1	35,395,204.09	346,578.04	1,490,838.96	62.289	23.247	1.624	1.143	42.036	0.000	346,578.04	73.333	1,490,838.96	511,321.50	1.53
2	B	B	A-	8.250	2	107,327,982.20	2,137,112.97	9,215,883.90	68.287	23.189	9.237	6.512	41.851	0.044	1,847,449.97	68.474	7,882,213.72	4,518,509.72	2.51
	C	C	A+	7.750	4	70,622,457.70	1,383,023.13	7,349,717.33	62.844	18.817	6.179	4.411	40.066	0.003	1,213,758.76	80.000	6,153,573.54	2,971,371.00	1.75
	D	D	AA	7.000	3	95,705,029.64	1,754,592.21	6,970,713.14	65.511	25.171	7.711	5.650	36.478	0.054	1,437,627.65	55.000	5,562,837.96	1,568,630.00	1.06
	E	E	BB	10.000	1	119,833,936.00	2,787,791.84	9,180,594.36	69.310	30.366	11.719	7.054	66.148	0.027	2,441,213.80	73.332	7,689,755.40	1,993,278.80	1.21

Beim nachfolgenden Spiel (s. Tabelle 3) waren aus organisatorischen Gründen nur Emders Studierende bei der Veranstaltung und es wurden die fünf Gruppen zu zwei Gruppen (G, H) zusammengefasst. Inhaltlich wurde die Komplexität weiter erhöht, da Änderungen der Stückliste und Änderungen der Losgröße ermöglicht wurden. Die Gruppen haben von dieser Option jedoch keinen Gebrauch gemacht. Zudem wurden die Reports kostenpflichtig. Die Erfahrungen mit den neuen Inhalten mussten an die nicht anwesenden Teammitglieder aus Kiel für folgende Spiel weitergegeben werden.

Im abschließenden Spiel (s. Tabelle 5) wurden vier Runden in den Originalbesetzungen vom Anfang (Gruppen A bis E) gespielt. Die Spielergebnisse können als solide bis sehr gut (Team B) betrachtet werden. Die Produktivität ist bei allen Gruppen relativ hoch. Ausnahmen bilden hier nur die erste und letzte Runde. In der ersten Runde muss vor Produktionsbeginn das Material beschafft werden. Ist bekannt, dass die letzte Runde gespielt wird, lassen Team die Produktion z.T. auslaufen. Zudem ist bei allen Team der Umsatz (Round Sales per Team)

konstant über eine Mio €. Das Nettoeinkommen ist in Runde ein z.T. negativ (i.d.R. aufgrund von Investitionen in Rüstzeitoptimierung), und bei Gruppe E in Runde vier, der letzten, sehr niedrig. Hier wurde versucht das Lager durch Preisdumping zu leeren.

Tabelle 5 Spiel 4 (final) – Manufacturing Sustainability in der Anfangsbesetzung

Leaderboard																				
Round	Team	Team	Credit ratings	Interest Rate (%)	Rank	Company valuation	Cumulative Net Income	Total sales	Gross Margin (%)	Net Margin (%)	ROE (%)	ROA (%)	D/E	Mktg/S (%)	Round Net Income	Round Productivity	Round Sales per Team	CO2e (kg)	CO2e per Unit Sold (kg/unit)	CO2e per Sales (kg/€)
1	A	A	AA-	7.250	1	56,135,574.04	242,104.34	1,550,107.14	74.593	15.619	1.140	0.870	30.991	0.000	242,104.34	64.583	1,550,107.14	317,830.96	1.30	0.21
	B	B	A+	7.750	2	40,694,036.43	343,462.44	1,372,923.78	72.645	25.017	1.609	1.247	29.007	0.109	343,462.44	59.375	1,372,923.78	319,483.24	1.62	0.23
	C	C	BBB+	8.750	5	10,119,877.33	(620,857.03)	1,027,995.04	71.945	(60.395)	(3.047)	(2.187)	39.331	11.119	(620,857.03)	50.266	1,027,995.04	594,629.59	2.95	0.58
	D	D	A	8.000	4	15,194,007.67	(77,226.59)	1,010,684.28	74.935	(7.641)	(0.369)	(0.263)	40.512	0.000	(77,226.59)	56.667	1,010,684.28	284,994.00	1.62	0.28
	E	E	BBB+	8.750	3	22,247,156.39	(927,790.63)	1,515,909.07	63.124	(61.204)	(4.622)	(3.138)	47.315	0.000	(927,790.63)	72.500	1,515,909.07	423,294.60	1.23	0.28
2	A	A	AA-	7.250	2	71,101,203.38	1,074,814.23	3,542,726.06	73.348	30.339	4.869	3.877	25.581	0.000	832,709.89	76.458	1,992,618.92	717,481.76	1.23	0.20
	B	B	AA-	7.250	1	92,834,878.76	1,273,987.31	3,991,726.27	74.961	31.916	5.720	4.583	24.794	0.406	930,524.87	80.000	2,618,802.49	596,548.71	1.02	0.15
	C	C	A-	8.250	5	30,702,135.51	160,168.60	3,457,154.33	70.172	4.633	0.757	0.540	40.082	7.046	781,025.63	89.166	2,429,159.29	1,609,598.03	2.56	0.47
	D	D	AA-	7.250	3	47,594,572.09	425,731.56	3,021,946.59	72.511	14.088	1.987	1.466	35.525	0.165	502,958.15	66.667	2,011,262.31	567,857.13	1.02	0.19
	E	E	BBB	9.000	4	36,781,458.15	(261,401.74)	3,486,836.22	64.901	(7.497)	(1.260)	(0.833)	51.377	0.000	666,388.89	80.981	1,970,927.15	781,039.17	0.98	0.22
3	A	A	AA	7.000	2	85,881,317.82	2,129,236.24	5,793,520.41	73.498	36.752	9.206	7.715	19.322	0.117	1,054,422.01	73.542	2,250,794.35	969,721.70	1.02	0.17
	B	B	AA+	6.500	1	110,246,265.67	2,542,868.39	6,405,661.59	74.977	39.697	10.801	9.297	16.182	0.334	1,268,881.08	72.292	2,413,935.32	875,954.66	0.93	0.14
	C	C	A	8.000	4	40,760,009.33	802,800.28	5,791,057.99	71.264	13.863	3.682	2.780	32.467	7.289	642,631.68	86.666	2,333,903.66	2,672,172.64	2.58	0.46
	D	D	AA	7.000	3	59,393,214.91	1,213,313.41	4,892,623.21	73.227	24.799	5.462	4.229	29.167	0.102	787,581.85	67.292	1,870,676.62	767,558.60	0.92	0.16
	E	E	BB+	9.750	5	39,315,591.85	398,973.47	5,658,047.74	63.712	7.051	1.864	1.171	59.248	0.000	660,375.21	81.833	2,171,211.52	1,195,691.10	0.93	0.21
4	A	A	AA+	6.500	2	88,739,342.42	2,973,159.77	7,751,671.14	73.691	38.355	12.402	10.410	19.140	0.087	843,923.53	69.167	1,958,150.73	1,323,273.09	1.04	0.17
	B	B	AAA+	6.000	1	132,467,740.00	4,042,258.00	9,122,377.82	75.056	44.311	16.142	14.614	10.452	0.336	1,499,389.61	79.375	2,716,716.23	1,140,067.13	0.84	0.13
	C	C	AA	7.000	3	73,171,219.45	2,309,611.38	8,731,884.88	71.698	26.450	9.908	8.027	23.435	5.325	1,506,811.10	87.499	2,940,826.89	2,924,503.77	1.89	0.33
	D	D	AA-	7.250	4	67,324,372.36	2,151,330.63	6,897,334.79	74.171	31.191	9.292	7.515	23.659	0.072	938,017.22	62.800	2,004,711.58	960,324.36	0.84	0.14
	E	E	BB+	9.750	5	28,162,618.62	444,120.02	6,817,336.55	63.683	6.515	2.071	1.313	57.692	0.000	45,146.55	52.835	1,159,288.81	1,421,883.93	0.92	0.21

3.2. Studentische Erwartungshaltung vor Spielbeginn

Nach der Gruppeneinteilung und vor Beginn der ersten Spielrunde wurden die Studierenden mittels einer Umfrage zu ihrer Motivation, der Einstellung zur Teamzugehörigkeit, den Erwartungen an den eigenen Beitrag zur Teamleistung (Einbringen von Wissen und Fähigkeiten sowie die individuelle Bedeutung des Unternehmensplanspiels im Vergleich mit anderen Modulen des laufenden Semesters) und Befürchtungen bzgl. der Remote-Teamarbeit befragt.

Die jeweiligen Antworten erfolgten durch eine Einfachauswahl auf einer fünfstelligen Likert-Skala (Strongly agree, Agree, Neutral, Disagree, Strongly Disagree). An der Umfrage nahmen 18 Studierende (N=18).

Die Studierenden waren überwiegend sehr motiviert an dem Unternehmensplanspiel teilzunehmen. So stimmten über 83% der Teilnehmenden der Aussage „I am very enthusiastic to play the ERPsim game“ (stark) zu (vgl. Abb. 1)

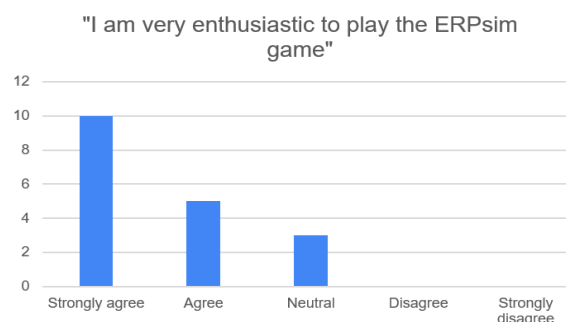


Abb. 1 Motivation der Studierenden (N=18)

Ein vergleichbares Ergebnis ergab sich auch bei der Zustimmung zu der Aussage „I am very enthusiastic to become a member of an ERPsims team“. Auch hier gab es keine Ablehnung dieser Aussage und lediglich 22% der Studierenden hatten eine neutrale Einstellung dazu.

Ein etwas gemischteres Ergebnis ergab sich beim individuellen Beitrag der Studierenden zur Teamleistung.

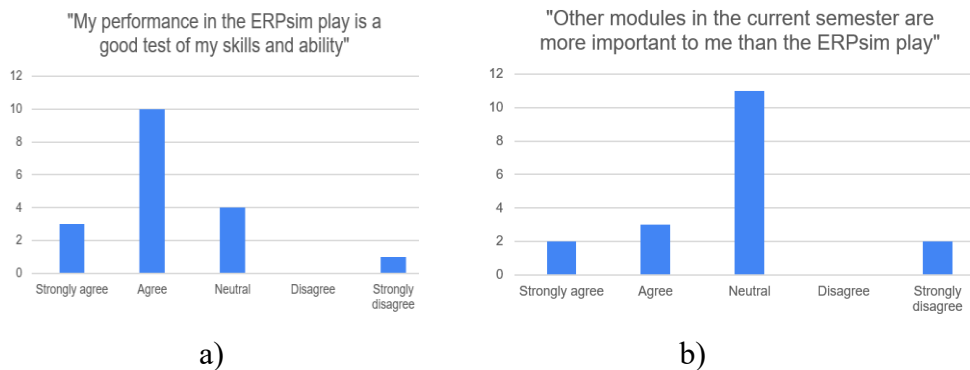


Abb. 2 Beitrag der Studierenden zur Teamleistung (N=18)

Hier zeigt sich, dass die Studierende das Einbringen der eigenen Leistung zur Teamarbeit noch überwiegend positiv einschätzen (vgl. Abb. 2 a), dass aber die Bedeutung des Unternehmensplanspiels im Vergleich zu anderen Modulen nur für ca. 28 % der Studierenden größer als die der anderen Module im selben Semester ist und für immerhin 11% der Studierenden sogar geringer ist (vgl. Abb. 2b).

Befürchtungen, dass Remote-Teamarbeit nicht funktionieren kann, äußerten lediglich 22% der Studierenden. Gleich viele Studierende vertraten die Ansicht, dass Sie keine Bedenken haben, während die Mehrheit der Studierenden eher eine neutrale Einstellung dazu besaß (vgl. Abb. 3)

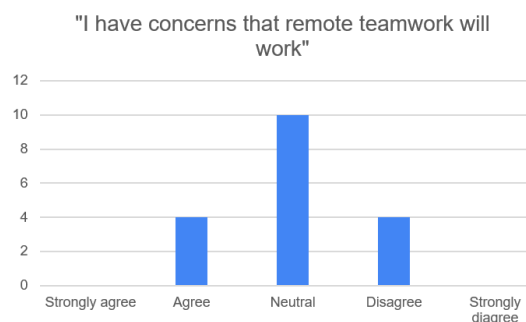


Abb. 3 Befürchtungen bzgl. Remote-Teamarbeit (N=18)

Fazit: Die Motivation der Studierenden an der Teilnahme des Unternehmensplanspiels ERPsims mit virtuellen Teams kann als hoch eingeschätzt werden. Auch die Bereitschaft sich mit den eigenen Fähigkeiten und Möglichkeiten einzubringen war gut ausgeprägt. Jedoch ist letzteres dadurch eingeschränkt, dass die Studierenden parallel auch noch Leistungen in anderen Modulen erbringen müssen und dem ERPsims Unternehmensplanspiel im Mittel keinen größeren Stellenwert bei der Verteilung ihres Workload einräumen.

3.3. Herausforderung der Teambildung und Kommunikation

Die virtuelle Teamarbeit bietet sowohl Herausforderungen als auch Möglichkeiten, die bei der persönlichen Zusammenarbeit nicht gegeben sind.

Dabei ist die virtuelle Teambildung ein wichtiges Instrument in einer digitalen kollaborativen Lernumgebung, wie es das Unternehmensplanspiel ERPsim im Kontext dieser Studie darstellt. Dabei handelt es sich um eine Reihe von Maßnahmen in der Lernumgebung die Vertrauen aufbauen, den Zusammenhalt im Team fördern, die Teamnormen klären, das Verständnis für die Arbeit der virtuellen Kolleg*innen fördern und effektive Sitzungen in einer virtuellen Umgebung ermöglichen sollen [Kuznetsova et al. 2023].

Obwohl personalisierte, unterrichtliche Interventionen bei der Teambildung das Verhalten, die Einstellung, die Motivation und die akademischen Leistungen der Studierenden in einer virtuellen Lernumgebung effektiv verbessern können [Branch, Dousay, 2015; Zhang et al., 2020], wurde bei dieser Studie darauf verzichtet. Dies ist u.a. damit begründet, dass die Motivation und das persönliche Engagement der Studierenden sich in die Teamarbeit einzubringen hoch war (vgl. Kap. 3.3) und gleichzeitig sollte beobachtet werden, wie eine nicht angeleitete Teambildung stattfindet.

Nach der Teameinteilung entschied sich jedes Team, jeweils eine eigene WhatsApp Gruppe einzurichten. Darüber sollte zunächst ein gegenseitiges Kennlernen und die Teambildung ermöglicht werden. Prinzipiell sind WhatsApp Gruppen ein geeignetes Instrumentarium, eine informelle und freundliche Atmosphäre unter den Teammitgliedern zu schaffen [Jailobaeva 2021]. Gleichzeitig schienen alle Studierenden mit diesem Instant Messaging Dienst vertraut zu sein und diesen für ihre sonstigen sozialen Aktivitäten nutzen, so dass es durchaus möglich ist, dass sie WhatsApp auch für die akademische Arbeit bequem und einfach finden. Nach dem Effizienzrahmen von [Nowak et al. 2005, 2009] neigen Menschen dazu, die Kommunikationsmedien zu nutzen, die sie für das Erreichen ihrer Ziele als am effektivsten erachten und die weniger kognitiven und verhaltensbezogenen Aufwand und weniger Zeit erfordern. Im Gegensatz zu persönlichen Gesprächen fehlt den textbasierten Kommunikationsmedien jedoch die Fähigkeit, nonverbale Hinweise zu übermitteln. Infolgedessen müssen sich die Menschen mehr Mühe geben, um einander besser kennenzulernen und zu verstehen [Watt et al., 2002].

Die durchgeführten Interviews ergaben, dass die Studierenden die WhatsApp Gruppe vor allem für eine kurze Vorstellung der eigenen Person und WhatsApp als Kommunikationsmedium in der ersten Planspielrunde nutzten. In weiteren Spielrunden wechselten die Studierenden dann auf Google Meet, weil hier umfangreichere Funktionalitäten zur Zusammenarbeit genutzt werden konnten. Insbesondere das Teilen des Bildschirms eines Teammitgliedes war für die Studierende eine nützliche Funktion. Dies ist vor allem darin begründet, dass bei ERPsim immer nur ein Teammitglied die Datenpflege für eine bestimmte Transaktion vornehmen kann und der Zugriff dann für andere Teammitglieder gesperrt ist.

Die Beziehungen zwischen den Menschen in Gruppen und Aspekte der Gruppenreife haben nachweislich große Auswirkungen auf die Effektivität der Gruppenarbeit, denn es hat sich gezeigt, dass reife Gruppen in einer Vielzahl von Bereichen besser abschneiden, z. B. Projekte schneller abschließen [Wheelan et al., 1998]. Daher war für diese Studie die Betrachtung des Reifegrads und somit die Gruppenentwicklung der studentischen Teams von besonderer Bedeutung.

In ihrer Übersicht über die Gruppenentwicklungsliteratur fanden Chidambaram und Bostrom [Chidambaram, Bostrom, 1996] eine Übereinstimmung zwischen verschiedenen Modellen bezüglich charakteristischer Verhaltensweisen, die zwischen gut und nicht so gut entwickelten Gruppen unterscheiden: (1) Wahrgenommene Zusammenhalt in der Gruppe (Zugehörigkeitsgefühl und soziale Emotionen), (2) Konfliktbewältigung, (3) Aufgabenbezogene und sozio-emotionale Bedürfnisse, (4) Kommunikation und (5) Teilnahme an Gruppenaufgaben. Jede dieser fünf Verhaltensweisen wurde mit separaten Skalen gemessen. Alle Skalen wurden aus der Literatur entnommen und wurden dort als valide und zuverlässig (Reliabilität) nachgewiesen.

Die sechs Punkte umfassende Perceived Cohesion Scale (PCS) [Chin et al., 1999] wurde entwickelt, um den wahrgenommenen Zusammenhalt in Gruppen zu messen, also wie verbunden sich die Mitglieder der Gruppe untereinander fühlen und wie stark sie sich zu der Gruppe hingezogen fühlen. Die Fähigkeit zur Konfliktbewältigung wurde anhand einer Skala gemessen, die von [Chidambaram, Bostrom, 1999], entwickelt und validiert wurde. Eine Skala zur Messung des Verhaltens von Gruppenmitgliedern wurde von [Gouran et al., 1978] entwickelt⁶. [Davidson, 1997] hat eine Skala entwickelt, um Aspekte der Kommunikation in einer Gruppe zu messen. [Hemphill, 1956] entwickelte eine zehnstufige Skala zur Beteiligung an Gruppenaktivitäten⁷.

Die oben genannten Skalen wurden im Rahmen einer Umfrage nach drei Spielrunden angewendet. An dieser Umfrage haben keine Studierende von Team A teilgenommen, so dass im Folgenden nur die Ergebnisse der Gruppen B – E vorgestellt werden.

Ein zusammengesetzter Wert für die Gesamtentwicklung der Gruppen wurde als Durchschnitt der fünf Skalenwerte berechnet. Dabei entspricht 5 der höchsten Gruppenentwicklungsreife und 1 der niedrigsten (vgl. Tabelle 6).

Tabelle 6 Gruppenentwicklungsreife

Team/Maße	B	C	D	E
Zusammenhalt	4,9	4,2	4,2	3,5
Konfliktbewältigung	4	5	4	3
Aufgabe/Sozial-Emotional	3,4	3,2	3,5	3,7
Kommunikation	3,9	4	5	3,8
Teilnahme	3,9	3,8	4	4,2
Gruppenentwicklungsreife:	4,02	4,04	4,14	3,64

Die Teams haben bzgl. der Kriterien der aufgabenbezogenen und sozio-emotionalen Bedürfnisse sowie der Teilnahmen an den Gruppenaufgaben einen vergleichbaren Entwicklungsstand erzielt. Unterschiede hier sind wahrscheinlich über marginal unterschiedliche Bewertungen und die geringe Zahl der befragten Studierenden erklärbar. Größere Unterschiede gibt es beim wahrgenommenen Zusammenhalt in der Gruppe und der Konfliktbewältigung. Letzteres ist etwas überraschend, denn zu mindestens bei den

⁶ Die ursprüngliche Skala besteht aus neun Items, die drei Arten von Verhaltensweisen erfassen sollen: aufgabenorientiertes, sozio-emotionales und verfahrenstechnisches Verhalten. Die sechs Items zur Messung des aufgabenbezogenen und sozio-emotionalen Verhaltens wurden in dieser Studie verwendet

⁷ Allerdings trafen nicht alle Items auf die vorliegende Situation zu. So beziehen sich beispielsweise einige der Items auf das Ansehen einer Gruppe oder darauf, ob „unproduktive“ Gruppenmitglieder aus der Gruppe ausgeschlossen werden (was während der Unternehmenssimulation nicht möglich war). Daher wurden nur die relevanten Items für diese Untersuchung ausgewählt.

Planspielrunden konnte keine Konflikte innerhalb der Gruppen durch die Dozierenden beobachtet werden.

Der Gruppenzusammenhalt wurde im Team B am besten und im Team E am schlechtesten bewertet, während die Teams C und D mit jeweils gleicher Punktzahl genau in der Mitte der beiden Extremwerte liegen. Vergleicht man dies mit dem Planspielergebnissen aus Tabelle 4, so zeigt sich dieselbe Reihenfolge auch bei der für die Erfolgsmessung des Planspiels maßgeblichen Kennzahl „Company Valuation“. Somit lässt sich aus diesem Ergebnis schließen, dass der Gruppenzusammenhalt einen Einfluss auf den Erfolg der Gruppe beim Unternehmensplanspiel ERPsim hat. Damit lässt sich in diesem Fall die Vermutung bestätigen, dass es eine positive Beziehung zwischen Zusammenhalt und Leistung in virtuellen Teams gibt [Langfred,1998].

3.4. Herausforderungen bei der Vorbereitung und Durchführung

Die Liste der organisatorisch administrativen und organisatorisch technischen Punkte (s. Kapitel 2.4) ist über das abgebrochene Vorhaben im Sommersemester 2024 und das aktuell durchgeführte Vorhaben zusammengetragen worden. Die Herausforderungen waren zum einen die Identifikation dieser Punkte und sofern möglich, deren vorausschauende Abarbeitung. Einige Punkte bleiben jedoch bis zum Beginn unbekannt, z.B. die Anzahl teilnehmender Studierender. Davon hängen die Zugänge zu der Lernplattform und die Lizenzbeschaffung für das Planspiel ab. Nachzügler, deren Zugangsdaten im Hochschulsystem noch definiert werden müssen sind dabei eine der Herausforderungen. Der Zeitaufwand für diesen und anderen Punkte sind ein anderer.

Neben dem Zeitaufwand kommt noch die Synchronisation der Aktivitäten zwischen den zwei Hochschulen hinzu. So musste im Vorfeld das Vorhaben mit drei Hochschulen durchzuführen aufgegeben werden. Die zeitliche Synchronisation der Vorlesungen war nicht möglich. Eine weitere Herausforderung besteht in der dezentralen, teils Präsenz, teils online Teilnahme an der Veranstaltung. Dies erfordert für die Professoren (und Studierenden) Multitasking-Arbeitsweisen. Neben der Einführung in das Planspiel sowie dessen Durchführung bzw. dessen Analyse, also der eigentlichen Lehre, muss die Medientechnik beherrscht und die Kommunikation durchgeführt werden. Hier ist eine ständige Rückkopplung bezüglich Bild- und Tonübertragung notwendig. Dies erschwert sowohl den Lehrenden als auch den Studierenden sich auf den Lehr- und Lerninhalt zu konzentrieren. Die für die Rückkopplung notwendige Zeit muss noch stärker eingeplant werden. Wünschenswert wäre hier eine unterstützende Kraft für das Medienhandling.

4. Fazit und Handlungsempfehlungen

Die vorliegende Studie untersuchte die Durchführung eines hochschulübergreifenden Planspiels mit heterogenen, räumlich verteilten und internationalen Teams. Die Forschungsfrage lautete, ob Planspiele eine effektive Methode zur Förderung von Teamfähigkeiten und zur Verbesserung der Kommunikation zwischen Studierenden sind. Die Ergebnisse der Studie zeigen, dass die Durchführung von gemeinsamen Veranstaltungen verschiedener Hochschulen eine aufwändige Vorbereitung erfordert, aber dass die Studierenden in der Lage sind, in verteilten, internationalen Teams zu arbeiten und Entscheidungen zu treffen.

Die wichtigsten Ergebnisse der Studie sind, dass die Planspielergebnisse in Zusammenhang mit der Umfrage darauf hindeuten, dass die Teams mit einem hohen Zusammenhalt mit einem besseren Ergebnis abschließen. Zwar bevorzugen Studierenden einfache, bereits bekannte Medien für die Kommunikation, die sich aber bei der Durchführung des Unternehmensplanspiels als weniger geeignet herausstellten.

Die Ergebnisse bestätigen den Nutzen von Planspielen für das Training von verteilter, internationaler Teamarbeit, aber auch, dass der Aufwand für die Durchführung solcher hochschulübergreifender Lehrveranstaltungen hoch ist. Hier können die in der Arbeit enthaltenen Auflistungen der vorbereitenden Maßnahmen als Handlungsempfehlung verstanden werden.

Für die weitere Forschung wäre es interessant, die Veranstaltung wiederholt mit einer größeren Anzahl Studierender durchzuführen und die Ergebnisse zu vergleichen. Dafür wäre wichtig, die Einbindung von unterstützenden Programmen, hochschulintern oder extern, wie z.B. Erasmus+ zu prüfen, um die Durchführung solcher Veranstaltungen zu unterstützen.

Insgesamt kann zusammengefasst werden, dass die vorliegende Studie einen wichtigen Beitrag über die Erkenntnisse zur Durchführung von hochschulübergreifenden Planspielen leistet.

5. References

- Anastasiadis, T.; Lampropoulos, G.; Siakas, K. Digital Game-based Learning and Serious Games in Education. *Int. J. Adv. Sci. Res. Eng.* 2018, 4, 139–144.
- Beranič, T.; Heričko, M. The Impact of Serious Games in Economic and Business Education: A Case of ERP Business Simulation. *Sustainability* 2022, 14, 683
- Branch, R. M., & Dousay, T. A. *Survey of Instructional Development Models* (5th ed.). Association for Educational Communications and Technology, 2015
- Chidambaram, L. and Bostrom, R.P. A longitudinal study of the impact of group decision support systems on group development. *Journal of Management Information Systems*, 7, 3, 1991, 7-25
- Chidambaram, L. & Bostrom, R.P.. Group development (I): A review and synthesis of development models. *Group Decision and Negotiation*, 6, 1996, 159-187.
- Chin, W. W., Salisbury, W. D., Pearson, A. W., & Stollak, M. J. Perceived cohesion in small groups: Adapting and testing the perceived cohesion scale in a small-group setting. *Small group research*, 30(6), 1999, 751-766.
- Creswell, J. W., Plano Clark, V. L. . *Designing and Conducting Mixed Methods Research*. Sage Publications, 2011
- Davison, R. An instrument for measuring meeting success. *Information and Management*, 32, 1997, 163-176.
- Gersick, C.J.G. Time and Transition in Work Teams: Toward a New Model of Group Development. *Academy of Management Journal*, 31, 1, 9-41, 1988
- ERPsims Lab – HEC Montréal. (2025). ERPsims Sustainability Challenge 2025. Abgerufen am 25. Juni 2025 von <https://erpsim.hec.ca/en/node/316>

- Gersick, C.J.G.. Marking Time: Predictable Transitions in Task Groups. *Academy of Management Journal*, 32, 2, 274-309, 1989
- Goi, C.L. The use of business simulation games in teaching and learning. *J. Educ. Bus.* 2019, 94, 342–349
- Gouran, D. S., Brown, C., and Henry, D. R. Behavioral correlates of perceptions of quality in decision-making discussions., *Communication Monographs*, 45, 1978, 51-63.
- Huang, Y.-M.; Silitonga, L.M.; Wu, T.-T. Applying a business simulation game in a flipped classroom to enhance engagement, learning achievement, and higher-order thinking skills. *Comput. Educ.* 2022, 183, 104494
- Hemphill, J.K. *Group Dimensions: A Manual for Their Measurement*, Research Monograph No. 87, Columbus: Ohio State University, Bureau of Business Research, 1956.
- Hughes, T.; Regan, N.; Wornham, D.; O'Regan, N. The credibility issue: Closing the academic/practitioner gap. *Strateg. Change.* 2008, 17, 215–233
- Jailobaev, T., Jailobaeva, K., Baialieva, M., Baialieva, G., & Asilbekova, G. WhatsApp groups in social research: new opportunities for fieldwork communication and management. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, 2021, 149(1), 60-82.
- Jony, A.I., Serradell-López, E. An Evaluation of Virtual Teamwork Model in Online Higher Education. In: Visvizi, A., Lytras, M.D., Aljohani, N.R. (eds) *Research and Innovation Forum 2020*, 2021
- Kuznetsova, M., Gura, D., Vorona-Slivinskaya, L. Virtual team building in an intelligent collaborative learning environment. *Journal of Information Technology Education: Research*, 22, 2023, 157-175
- Langfred, C.W. Is group cohesiveness a double-edged sword? An investigation on the effects of cohesiveness on performance. *Small Group Research*, 29, 1998, 124-143.
- Nowak, K-L., Watt, J.H., Walther, J.B. The influence of synchrony and sensory modality on the person perception process in computer-mediated groups. *Journal of Computer-Mediated Communication* 10(3), 2005.
- Nowak, K.L., Watt, J.H., Walther, J.B. Computer mediated teamwork and the efficiency framework: Exploring the influence of synchrony and cues on media satisfaction and outcome success. *Computers in Human Behavior* 25, 2009, 1108–1119
- Peterková, J., Repaská, Z., Prachařová, L. Best Practice of Using Digital Business Simulation Games in Business Education. *Sustainability*, 14(15), 2022, 8987
- Watt, J.H., Walther, J.B., Nowak, K.L. Asynchronous videoconferencing: A hybrid communication prototype. In: *The 35th Hawaii international conference on system sciences*, Big Island, USA, 2002, 7–10.
- Wheelan, S., Murphy, D., Tsumura, E., & Kline, S. F. . Member perceptions of internal group dynamics and productivity. *Small Group Research*, 29, 1998, 371–393
- Zarte, M., Pechmann, A., . Concept for Introducing Sustainable Product Service Systems through a Business Game, *Procedia CIRP*, Volume 83, 2019, Pages 44-49, ISSN 2212-8271, <https://doi.org/10.1016/j.procir.2019.02.139>

Zhang, J. H., Zou, L. C., Miao, J. J., Zhang, Y. X., Hwang, G. J., & Zhu, Y. An individualized intervention approach to improving university students' learning performance and interactive behaviors in a blended learning environment. *Interactive Learning Environments*, 28(2), 2020, 231–245.

Individuelle SAP-Ausbildung mit LEGO: Innovative Lehrmethoden

C. Albrecht, S. Wolf, T. Teich, S. Bergelt, A. Kreisel
Westfälische Hochschule Zwickau, Zwickau, Deutschland

Abstract: Komplexe SAP-Lehrveranstaltungen erfordern über längere Zeiträume hohe Aufmerksamkeit, die jedoch erfahrungsgemäß schnell abnimmt. Dieses Paper untersucht, wie kurze, spielerische Impulse - konkret der Einsatz von LEGO - helfen können, die Aufmerksamkeit zu erhöhen und den Lernerfolg zu fördern. Die Methode verbindet theoretische Inhalte mit haptischem Bauen und nutzt Konzepte wie Serious Play und den Kolb-Lernzyklus, um ERP-Prozesse greifbar zu machen. Anhand einer Fallstudie wird gezeigt, wie Teilnehmer eine simulierte Lieferkette inklusive Materialflüssen, Stücklisten und Arbeitsplätzen sowohl digital im SAP-System abbilden als auch physisch mit LEGO nachbauen. Ergebnisse aus Kursen belegen eine gesteigerte Motivation und ein besseres Verständnis komplexer Zusammenhänge. Herausforderungen bestehen vor allem in der logistischen Vorbereitung und der Skalierung. Ein Ausblick zeigt Potenziale den Transfer auf weitere Module und den Einsatz von KI-gestützten Modellen. Insgesamt liefert der Ansatz eine praxisnahe Antwort auf die Frage, wie spielerische Impulse in komplexe Lehrformate integriert werden können, um Aufmerksamkeit und Lernerfolg nachhaltig zu steigern.

Keywords: SAP-Ausbildung, Interaktive Lehrformate, Innovative Lehrmethoden, LEGO, Haptik

Zur besseren Lesbarkeit wird auf die Nennung aller Geschlechter verzichtet; alle Bezeichnungen gelten für alle Geschlechter.

1. Einleitung und zentrale Forschungsfrage

Komplexe Lehrveranstaltungen im SAP-Kontext erstrecken sich häufig über mehrere Stunden am Tag und erfordern ein hohes Maß an Aufmerksamkeit. Untersuchungen legen jedoch nahe, dass die Aufmerksamkeit in Vorlesungen bereits nach kurzer Zeit abnimmt (Open Collage 2025). Daher ist es von besonderer Relevanz, zu Beginn eines ausgedehnten Lernintervalls die zentralen Lernziele prägnant und nachvollziehbar zu kommunizieren sowie die Inhalte mittels diverser Modalitäten (visuell, auditiv, kinästhetisch) zu gestalten.

In Anbetracht dessen wird in diesem Paper die Implementierung einer spielerischen Impuls-Pause nach einer ausgedehnten Lerneinheit empfohlen. Diese kann als Aktivierungsschub fungieren, der dazu geeignet ist, mentale Ermüdung zu reduzieren und den Fokus wiederherzustellen, ohne dabei die inhaltliche Kohärenz zu unterbrechen. Daraus ergibt sich die Forschungsfrage:

Auf welche Art und Weise können kurze, spielerische Impulse in eine Lehrveranstaltung integriert werden, um die Aufmerksamkeit der Lernenden zu erhöhen und den Lernerfolg zu fördern?

Im Rahmen der SAP-Ausbildung erfolgte der Einsatz in der LEGO-Fabrik, mit dem Ziel, die Material- und Informationsflüsse praktisch zu veranschaulichen. Dieser Ansatz diente dazu, die Komplexität der Prozesse zu reduzieren und eine größere Greifbarkeit zu gewährleisten. Ein

Nachteil dieses Vorgehens besteht darin, dass den Teilnehmern die Möglichkeit verwehrt bleibt, ein eigenes Modell zu bauen. Aus diesem Grunde wurde nach der Wissensvermittlung sowie der Klickeinheit eine kurze LEGO-Pause implementiert, in welcher die Teilnehmer das zuvor vermittelte Inhalte als Modell nachbauen. Diese haptische Aktivität zeichnet sich durch eine besondere Effektivität aus, da sie die inhaltliche Verbindung stärkt und Motivation sowie Engagement merklich steigert.

2. Innovative Lehrmethoden

2.1. Aktuelle Herausforderungen in der Bildungslandschaft

Die Bildungslandschaft steht vor tiefgreifenden Veränderungen. Eine der zentralen Herausforderungen besteht darin, auf die veränderten Lerngewohnheiten der Generation Z einzugehen - einer Generation, die in einer visuell und digital geprägten Welt mit Smartphones, Tablets, Social Media, Kurz-Videos oder Storys aufgewachsen ist. Klassisch textbasierte Lernformen verlieren zunehmend an Wirksamkeit, da viele junge Menschen eher durch Bilder, Videos und interaktive Inhalte erreicht werden. Die Fähigkeit, längere Texte konzentriert zu lesen und zu verarbeiten, ist bei vielen Schülern weniger stark ausgeprägt als noch in früheren Generationen. Um dem entgegenzuwirken, braucht es neue, innovative Bildungsansätze, die sowohl die Lebensrealität der Lernenden berücksichtigen als auch zentrale Kompetenzen wie Teamfähigkeit, Problemlösung und Kreativität fördern. Durch den Einsatz von Legobausteinen können Lerninhalte spielerisch und haptisch vermittelt werden. Die Kursteilnehmer bauen im Team Modelle, lösen gemeinsam Aufgaben und entwickeln kreative Lösungen - alles in einem der Generation Z ansprechenden praxisnahen Lernumfeld sowie Lerngewohnheiten. LEGO wird zum verbindenden Element, das kognitive, soziale und emotionale Lernprozesse anstößt. (Parment, 2024)

2.2. Spielerische Impulse als innovative Lehrmethoden

Spielerische Ansätze bieten eine motivierende und abwechslungsreiche Möglichkeit, Lerninhalte nachhaltig zu vermitteln. Gerade bei der Generation Z, die Abwechslung und Interaktivität schätzt, können innovative Methoden helfen, Wissen praxisnah zu festigen und den Lernprozess aufzulockern. (Gabryelski 2022)

Viele Menschen verbinden mit LEGO positive Erinnerungen - fast jeder baut gern mit den bunten Steinen. Mit LEGO Serious Play wird diese spielerische Freude genutzt, um komplexe Inhalte greifbar zu machen, kreative Ideen zu entwickeln und den Austausch in der Gruppe zu fördern. Der Ansatz des Serious Play verfolgt das Ziel, spielerische Aktivitäten gezielt für ernsthafte Lern-, Reflexions- und Problemlösungsprozesse zu nutzen. Anders als herkömmliche Spiele zielt Serious Play nicht primär auf Unterhaltung, sondern auf kognitive Aktivierung, Kommunikation und Erkenntnisgewinn (Michael & Chen 2006). Die Kombination aus haptischer Interaktion, bildhafte Darstellung und narrativer Einbettung fördert eine tiefere Verarbeitung und stärkt den Transfer auf reale Anwendungskontexte. In der Hochschullehre ermöglicht Serious Play einen Zugang zu strategischem Denken und Teamprozessen und wird daher zunehmend als didaktisches Instrument komplexer Lernszenarien eingesetzt (Kristiansen & Rasmussen 2014). Weiterhin setzen die Trainer Auflockerungsspiele ein. Beim Format „*Walk and Talk*“ erhalten die Teilnehmer drei thematisch passende Fragen und werden dazu ermutigt, nach draußen an die frische Luft zu gehen. So wird Bewegung ganz bewusst in den

Lernprozess integriert. Dieser Ansatz kommt besonders gut bei der Generation Z an, die großen Wert auf körperliche und mentale Gesundheit legt (Nieżurawska-Zajac, 2023). Im Spiel „*Wer bin ich - Transaktion Edition*“ erraten die Teilnehmer verschiedene Transaktionen im SAP-System. So wird spielerisch Wissen gefestigt und das Systemverständnis auf unterhaltsame Weise vertieft. Nach dem bekannten Prinzip von „Stadt, Land, Fluss“ wird bei „*Stadt, Land, SAP*“ Wissen nachhaltig gesichert und motiviert die Teilnehmer, ihr Wissen kreativ abzurufen.

2.3. Der Kolb-Zyklus in der Praxis: Aktives Experimentieren als Lehrmethode

Der Lernzyklus nach David Kolb bietet eine praxisnahe Grundlage für moderne Lehr- und Lernkonzepte. Besonders die vierte Phase „Das aktive Experimentieren“ spielt dabei eine zentrale Rolle (Kolb 2015). Hier setzen die Lernenden ihre neu gewonnenen Erkenntnisse direkt in die Praxis um, wenden sie auf unterschiedliche Situationen an und überprüfen, ob die entwickelten Ideen tragfähig sind. Durch diese Phase wird Wissen nicht nur theoretisch erarbeitet, sondern unmittelbar beim Bauen von Lego erprobt und gefestigt. So entsteht ein lebendiger Lernprozess, in dem Erfahrungen reflektiert, Konzepte entwickelt und diese in einem kontinuierlichen Zyklus weiter verbessert werden. Aktives Experimentieren fördert die Eigenverantwortung, Motivation und Kreativität der Lernenden - und macht Lernen damit nicht nur effektiver, sondern auch nachhaltiger und praxisnaher.

2.4. Vorteile von Haptik in der Unterstützung des Verständnisses von ERP-Systemen

Der Einsatz haptischer Elemente kann das Verständnis komplexer ERP-Systeme wesentlich erleichtern. Durch das Anfassen, Bauen oder physische Visualisieren werden abstrakte Prozesse greifbar und besser nachvollziehbar. Haptische Methoden wie Bauen mit LEGO ermöglichen es, Strukturen, Abläufe und Zusammenhänge bildlich darzustellen, wodurch theoretisches Wissen mit praktischer Erfahrung verknüpft wird. Dies fördert nicht nur das Verständnis, sondern auch die Merkfähigkeit und die aktive Beteiligung der Lernenden. Besonders bei komplexen Themen wie Geschäftsprozessen, Datenflüssen oder Systemarchitekturen bietet die Haptik einen niedrigschwelligen Zugang, der Motivation schafft und Hemmschwellen abbaut. (Müller et. Al. 2022)

Darüber hinaus zeigen neurodidaktische Studien, dass multisensorische Lernprozesse - insbesondere solche mit haptischer Beteiligung - eine höhere kognitive Aktivierung bewirken und den Wissenstransfer deutlich verbessern (Shams & Seitz 2008). Haptische Interaktionen aktivieren dabei auch motorische Areale im Gehirn, was zu einer tieferen Verankerung im Langzeitgedächtnis führen kann (Gallese & Lakoff 2005). Zudem wirkt sich das sogenannte „Enactment“ - also das aktive Ausführen von Handlungen im Lernprozess - positiv auf das Behalten und die Strukturierung komplexer Informationen aus (Mayer 2009).

3. Praktische Umsetzung

3.1. Ablauf der Fallstudie

Im Zentrum der Fallstudie steht die gemeinsame Fertigung eines Fahrzeugs mit dem Namen Milo One. Die Teilnehmer übernehmen dabei eine von drei Rollen innerhalb einer simulierten Lieferkette: Supplier 1 - Blechverarbeitung (SP1), Supplier 2 - Karosseriebau (SP2) oder den Original Equipment Manufacturer - Kraftfahrzeugbau (OEM). Ziel ist es, den gesamten

Zu Beginn legen die Gruppen jeweils auf Basis einer Klickanleitung ihre benötigten Materialien im System an, erfassen Stücklisten und verknüpfen diese über alle drei Unternehmen hinweg. Zusätzlich werden für die geplanten Fertigungsschritte die zugehörigen Arbeitsplätze, Arbeitspläne sowie eine Fertigungsversion angelegt. Die Teilnehmer legen die Unternehmen als Geschäftspartner mit den Rollen Lieferant und Kunde an, um sich gegenseitig die Materialien zu beliefern.

Nach der Vermittlung betriebswirtschaftlicher Grundlagen und deren Umsetzung im SAP-System folgt eine Trainingseinheit, in der die Teilnehmer zentrale Funktionen wie Anlegen von Materialstamm und Stücklisten direkt im System anwenden. Im Anschluss daran wird eine kurze, aktivierende LEGO-Einheit als Impuls eingeplant, um das zuvor Gelernte haptisch zu verankern und für spätere integrative SAP-Module zu verinnerlichen.

The BOM tree for the Mikro-Controller board is structured as follows:

- Mikro-Controller** (ID: 0_0110)
 - (1) BG Polystyrol 1_0410**
 - (4) CG Wider 2_0210
 - (2) BG Aktoren 2_021E
 - (1) BG Steckgruppe 2_0213
 - (7) BG Stecker 1_0410
 - (1) BG Amplitudenstabilisator 1_0210
 - (1) BG SMD-Gesamt 2_0210
 - (6,3 mH) Rückstrom-Schutz 2_0210
 - (1,5 mH) Rückstrom-Schutz 2_0210
 - (1) BG Kunststoff 1_0420**
 - (1) BG Polystyrol 2_0110
 - (2) Koffergel hartem 0_0120
 - (1) Motorhaube 0_0140
 - (7) Koffergel weich 0_0110
 - (1) Kofferschutzhülse 0_0130
 - (1) Deckel 0_0108
 - (2) Fuß 0_0120
 - (7,5 mH) Rückstrom-Schutz 1,5 mm 1_0310
 - (5,05 mH) Rückstrom-Schutz 1 mm 1_0320

Legend: OCM, SWP, SP1

3.3. Beispiel Arbeitsplatz

225

Zur Förderung des Verständnisses für diese abstrahierten Systemelemente werden die im System angelegten Arbeitsplätze im Anschluss physisch nachgebaut. Hierfür erhalten die Teilnehmer LEGO-Bauanleitungen, anhand derer sie zentrale Stationen des Fertigungsprozesses modellhaft umsetzen. Abb. 2 zeigt die Umsetzung einer solchen Einheit: ein Auszug aus der Bauanleitung für die Lackierkabine und das gebaute Modell. Durch diese Verbindung von Systemarbeit und haptischer Umsetzung werden abstrakte Objekte wie Arbeitsplätze greifbar und funktional erfahrbare.

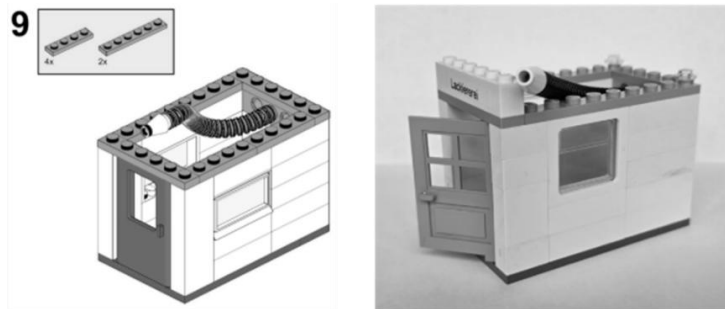


Abb. 2: LEGO-Bauanleitung des Arbeitsplatz Lackierkabine

3.4. Beispiel Umlagerungsbestellung

Ein zentraler Bestandteil der Fallstudie ist die Belieferung innerhalb der simulierten Supply Chain (s. Abb. 3). Hierbei nutzen die Teilnehmer im SAP-System den Prozess der Umlagerungsbestellung, um Materialien zwischen den drei beteiligten Unternehmen zu transferieren. (Teich, 2020)

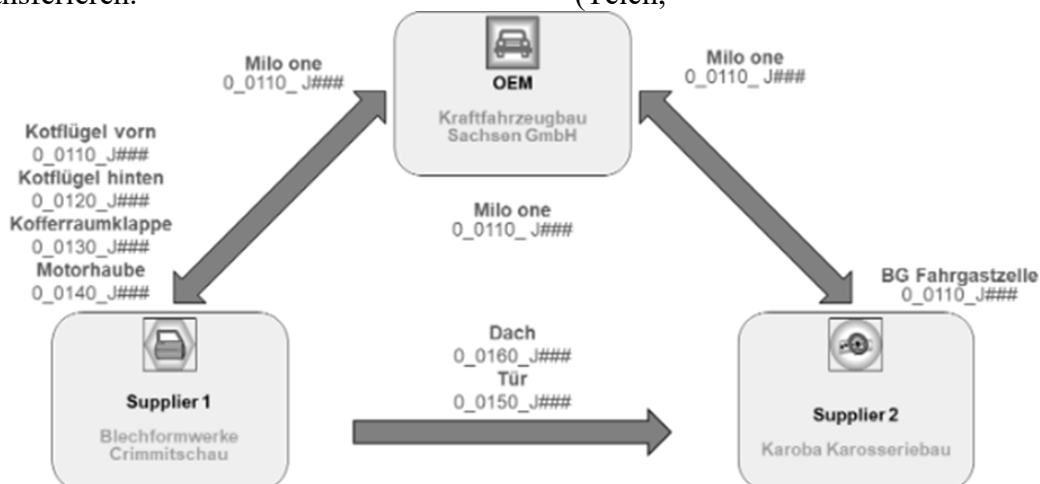


Abb. 3: Darstellung der Lieferantenbeziehungen in der Supply Chain

Der Prozess umfasst aufeinanderfolgende Schritte: Bestellung des Kunden, Auslieferung und Warenausgang des Lieferanten sowie Anlieferung und Wareneingang beim Kunden.

Ein in der Praxis häufig auftretender Fehler besteht darin, dass der empfangende Betrieb bereits einen Wareneingang bucht, bevor der sendende Betrieb die Auslieferung mit Warenausgang vollständig abgeschlossen hat. Der Fehler unterbricht den Ablauf und führt zu Systemfehlern.

Zur Reduktion solcher Fehler und zur Veranschaulichung der Abläufe wird der gesamte Prozess durch eine LEGO-basierte Darstellung begleitet. Die Teilnehmer simulieren die physische Übergabe der Materialien parallel zum digitalen Prozess und erleben dadurch die Prozesslogik

unmittelbar. Diese Verknüpfung von Systemvorgang und realer Handlung stärkt das Prozessverständnis und erhöht die Sorgfalt bei der Durchführung - wodurch die Fehleranfälligkeit reduziert wird.

3.5. Vorbereitung und Logistik - eine unterschätzte Herausforderung

Der Einsatz LEGO-basierter Einheiten im Rahmen der Fallstudie ist mit einem erheblichen organisatorischen und logistischen Vorbereitungsaufwand verbunden. Für jede Lerneinheit werden individuell abgestimmte Bausets zusammengestellt, bei denen benötigte LEGO-Teile pro Teilnehmer exakt vordefiniert und in separaten Tüten verpackt sind. Diese Sets werden in großen Transportboxen an den Ort der Lehrveranstaltung gebracht. Je nach Teilnehmerzahl ergibt sich ein entsprechender Platz- und Transportbedarf, der eine frühzeitige logistische Planung erfordert. Nach Abschluss der Fallstudie müssen die gebauten Modelle für künftige Kurse vollständig auseinandergenommen und neu sortiert werden. Fehlende oder beschädigte Bauteile müssen regelmäßig nachbestellt werden, um die Funktionsfähigkeit der Sets zu gewährleisten. Ergänzend stehen den Teilnehmern während der Einheiten unsortierte LEGO-Steinkisten zur Verfügung, die für kreative Aufgaben jenseits der standardisierten Bauanleitungen genutzt werden können. Trotz des hohen Aufwands bietet dieses Vorgehen einen klaren didaktischen Mehrwert (s. Abb. 4).



Abb. 4: Kreative LEGO-Fabrik der Teilnehmer

Grenzen ergeben sich vor allem durch logistische Rahmenbedingungen, etwa die Anzahl der verfügbaren Bausets, den vorhandenen Stauraum für den Transport sowie durch organisatorische Faktoren wie die Betreuungsrelation, die Größe der Veranstaltungsräume und die technische Infrastruktur im SAP-System. In der Praxis sind es häufig diese äußeren Bedingungen - nicht der LEGO-Anteil selbst -, die die maximale Teilnehmerzahl bestimmen.

3.6. Betreuungsaufwand und Skalierbarkeit

Nach einer kurzen Einführung zu den LEGO-Impuls-Pausen als methodische Ergänzung zur Komplexfallstudie verlaufen die Bauphasen meist betreuungsfrei. LEGO ist den meisten Teilnehmern aus der eigenen Kindheit oder dem Familienalltag vertraut, wodurch die Handhabung intuitiv erfolgt. Die Modelle lassen sich mit geringem Erklärungsbedarf umsetzen, was den Ablauf der Lehrveranstaltung entlastet und gleichzeitig Raum für eigenständiges Arbeiten schafft.

Außerdem bieten die gebauten Modelle eine niedrigschwellige visuelle Hilfe, um bei Rückfragen im Verlauf der Fallstudie komplexe Zusammenhänge anschaulich zu erklären. Sie

dienen somit nicht nur der Aktivierung, sondern auch als didaktische Anker zur Vertiefung einzelner Inhalte.

Gleichzeitig ist zu beachten, dass der LEGO-Einsatz nicht bei allen Teilnehmern auf gleiche Resonanz stößt. Einzelne äußern Vorbehalte gegenüber der spielerischen Form - was im Sinne eines freiwilligen Formats akzeptiert wird und nicht zu einem Teilnahmezwang führt. Hinsichtlich der Skalierbarkeit lässt sich das Konzept systematisch erweitern: Da die simulierte Lieferkette stets aus drei Rollen besteht, kann in festen Dreiergruppen nach oben skaliert werden.

4. Feedback aus Kursen und Wirtschaftspartnern

Erfahrungen aus bisherigen Kursdurchläufen sowie Rückmeldungen von Kooperationspartnern aus der Wirtschaft zeigen, dass der innovative Einsatz neuer Lernformate, insbesondere mit Umgang der Generation Z auf große Resonanz stößt und praxisrelevante Kompetenzen fördert. Nach jedem Kurs führen die Trainer eine Auswertung mit positiven und kritischen Rückmeldungen durch. In diesem Rahmen wird auch gezielt die Meinung zum Einsatz von LEGO im Kursgeschehen erhoben. So äußern Teilnehmer beispielsweise Rückmeldungen, dass der Einsatz von LEGO ihnen dabei geholfen habe, komplexe Inhalte - wie etwa das Szenario der Fallstudie - besser zu erfassen und zu durchdringen. Positiv hervorgehoben wurde zudem die Kombination aus digitalem Arbeiten im SAP-System und der haptischen Abwechslung durch den Einsatz von LEGO. Diese Verbindung aus praktischer Systembedienung und physischem Modellieren wurde von den Teilnehmern als abwechslungsreich und lernförderlich wahrgenommen.

5. Zusammenfassung und Ausblick

Die Untersuchung zeigt, dass kurze, spielerische Impulse - wie sie durch den Einsatz von LEGO-Methoden realisiert werden - eine wirksame Möglichkeit darstellen, die Aufmerksamkeit der Lernenden während einer Lehrveranstaltung zu erhöhen und den Lernerfolg nachhaltig zu fördern. Durch Methoden wie Serious Play werden komplexe ERP-Prozesse greifbar und verständlich gemacht, während Erzählen, Bauen und Zusammenarbeit das Lernen aktiv unterstützen.

Ein weiterer denkbarer und vielversprechender Ansatz ist die Entwicklung individueller Lernpfade durch den Einsatz von Learning Analytics zur Verbesserung der SAP-Ausbildung. Ziel dabei ist es, den individuellen Lernfortschritt der Teilnehmer zu erfassen und die Kursinhalte, beispielsweise auch LEGO entsprechend anzupassen. Learning Analytics umfasst dabei die Messung, Sammlung, Analyse und Berichterstattung von Daten über Lernende, um Lernprozesse und Lernumgebungen besser zu verstehen und gezielt weiterzuentwickeln (Trommer 2023).

Darüber hinaus bietet sich die Übertragung der LEGO-Methoden auf neue SAP-Module an, wobei zu berücksichtigen ist, dass nicht alle Module dafür geeignet sind. So lassen sich beispielsweise Finanzströme deutlich schwerer mit LEGO darstellen als Warenströme. Eine weitere Perspektive ist die Generierung neuer Modelle, möglicherweise unterstützt durch KI, um noch vielfältigere Szenarien abzubilden. Zudem sollte die Wirksamkeit durch den Vergleich mit Kontrollgruppen stärker evaluiert werden.

Fazit: LEGO-gestützte Formate eignen sich besonders für skalierbare Impuls-Pausen innerhalb längerer Schulungseinheiten, da sie Aufmerksamkeit und den nachhaltigen Wissenstransfer fördern - und damit einen wertvollen Beitrag zur *Beantwortung der Forschungsfrage* leisten.

6. Literaturverzeichnis

- Gabryelski, N. et. al. (2022): Digitale Hochschulbildung und Erwachsenenbildung im Vergleich am Beispiel einer SAP S/4HANA-Schulung. In: Mathis, U.; Ondrusch, N.; Kilian, D.; Krcmar, H.; Turowski, K.; Weidner, S. (Hrsg.): Proceedings of the SAP Academic Community Conference 2022 DACH. München: SAP Community Conference. 10.14459/2022md1685828.
- Gallese, V.; Lakoff, G. (2005): The Brain's Concepts: The Role of the Sensory-Motor System in Conceptual Knowledge. In: Cognitive Neuropsychology, 22(3-4), S. 455-479. <https://doi.org/10.1080/02643290442000310>
- Kolb, D. (2015): Experiential Learning: Experience as the Source of Learning and Development. Upper Saddle River, NJ: Pearson FT Press., 7.
- Kristiansen, P.; Rasmussen, R. (2014): Building a Better Business Using the LEGO Serious Play Method. Hoboken: Wiley.
- Mayer, R. E. (2021): Multimedia Learning. 3. Aufl. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108859025>
- Michael, D.; Chen, S. (2006): Serious Games: Games That Educate, Train, and Inform. Boston: Thomson Course Technology.
- Müller, M. et. Al. (2022): Lehrbuch Haptik: Grundlagen in Therapie, Pflege und Medizin. Berlin: Springer., 163 - 183.
- Nieżurawska-Zajac, J et. Al. (2023). Managing Generation Z: Motivation, engagement and loyalty. Abingdon, UK: Routledge. 20-21.
- Open Colleges. (n.d.). The science of attention: How to capture and hold the attention of easily distracted students. Open Colleges. Abgerufen am 10. Juli 2025, von <https://www.opencolleges.edu.au/blogs/articles/the-science-of-attention-how-to-capture-and-hold-the-attention-of-easily-distracted-students>
- Parment, A. (2024). Die Generation Z: Die Hoffnungsträgergeneration in der neuen Arbeitswelt (1. Aufl.). Wiesbaden: Springer Gabler, 35, 125-130.
- Shams, L.; Seitz, A. R. (2008): Benefits of Multisensory Learning. In: Trends in Cognitive Sciences, 12(11), S. 411- 417. <https://doi.org/10.1016/j.tics.2008.07.006>
- Teich, T. et. al. (2020): Umstellung einer integrierten betriebswirtschaftlichen Lehr-Fallstudie auf SAP S/4HANA. In: Gräslund, K.; Kilian, D.; Krcmar, H.; Turowski, K.; Wittges, H. (Hrsg.): Proceedings of the SAP Academic Community Conference D-A-CH 2020. München: SAP Community Conference. 10.14459/2020md1555096.
- Trommer, M. et. al. (2023): WHZ LearnXP: Angewandte Learning Analytics in der SAP-Ausbildung. In: Mathis, U.; Ondrusch, N.; Kilian, D.; Krcmar, H.; Turowski, K.; Weidner, S.; Wittges, H. (Hrsg.): SAP Academic Community Conference 2023 (D-A-CH). München: SAP Community Conference. 10.14459/2023md1719876

Opening the Blackbox: A Socio-Technical Simulation Game for Teaching Resource-Efficient Information Logistics in End-to-End Value Streams

M. Geisthardt, L. Engel & J. Marx Gómez

Carl von Ossietzky University of Oldenburg, Oldenburg, Germany

Jade University of Applied Sciences, Wilhelmshaven, Germany

Abstract: As digital transformation accelerates and industrial production systems grow increasingly complex, optimizing information logistics has emerged as a critical — yet underexplored — capability for achieving resource efficiency in end-to-end value streams. However, due to its abstract and often opaque nature, it is difficult to teach and operationalize effectively. This paper presents a socio-technical simulation game developed to “open the blackbox” of information logistics by making invisible flows, decisions, and dependencies transparent and actionable for learners. Developed using a Cross Industry Standard Process for Data Mining (CRISP-DM)-based gamification framework, the game enables to engage with realistic, interactive value stream scenarios in which learners can explore, redesign, and optimize information flows. The game integrates a Large Multimodal Model (LMM) as a collaborative assistant to guide players through various learning stages, from basic conceptual understanding to training proficiency. To evaluate the game’s effectiveness, a randomized controlled field experiment with a pre-post design was conducted. Learners were randomly assigned to an intervention group (players) or a control group (observers). Statistically significant improvements were observed in knowledge acquisition ($p = 0.005$), practical skills development ($p = 0.183$), and self-efficacy ($p = 0.002$). These findings demonstrate the potential of simulation game-based learning with LMMs to foster Business Process Management (BPM) competencies and promote resource-efficient thinking in academic and professional education.

Keywords: Business Process Management (BPM), Value Stream Analysis and Design, Information Logistics, Socio-Technical Simulation Games, Large Multimodal Models (LMMs)

1. Introduction

In the production of industrial goods—ranging from capital-intensive machinery to consumer-oriented products—human improvement teams are challenged to drive digitization while optimizing resource efficiency (Rother & Shook 2011). While value stream analysis and design has traditionally focused on physical material flows, growing data volumes and complex interdependencies now require a systematic approach to structuring and synchronizing information flows across employees, systems, and processes (Meudt et al. 2017). However, due to its abstract and often opaque nature, information logistics remains difficult to teach and operationalize effectively.

Gamification approaches and simulation games have shown great potential in addressing this educational challenge by engaging learners in interactive, problem-oriented scenarios (Huang et al. 2020, Sailer & Homner 2020). Current research has applied gamification to cultivate understanding of several facets of lean management (Bloechl et al. 2017). Yet, no existing

approaches succeed in making the invisible dynamics of information logistics both visible and actionable for learners.

To address this gap, this paper introduces the extended edition of “North Star® – The Value Stream Game”, a socio-technical simulation game designed to “open the blackbox” of information logistics (Geisthardt et al. 2024). Developed based on the Cross Industry Standard Process for Data Mining (CRISP-DM) and gamification, the game enables learners to explore, redesign, and optimize information flows in interactive value stream scenarios (Wirth & Hipp 2000, Werbach & Hunter 2020). A Large Multimodal Model (LMM) called “NorthStarGPT” supports and guides participants through various learning stages by providing adaptive feedback and intelligent assistance (Yu & Glazieva 2022). The game is designed to educate within five levels of mastery, ranging from level 1 (awareness), level 2 (understanding), level 3 (application), level 4 (effective use), to level 5 (teach others) (Liker 2020).

To evaluate the game’s effectiveness in context of Business Process Management (BPM) education, a randomized controlled field experiment with a pre-post design was conducted, measuring changes in knowledge acquisition, practical skills development, and self-efficacy. Both descriptive and inferential statistical methods were applied to analyze the data and assess the significance and effect size of the game intervention (Rasch et al. 2021).

This paper initially introduces the theoretical background of the socio-technical simulation game and the methodology based on the CRISP-DM and gamification. Subsequently, the game prototype, evaluation setup, and results of the field experiment are presented briefly. Lastly, a conclusion with an outlook on future research is offered.

2. Theoretical Background

2.1. Value Stream Analysis and Design

An end-to-end value stream can be defined as a process chain that encompasses all value-adding and non-value-adding activities required for the main product flow of a company—from sourcing to shipping to the customer (Rother & Shook 2011). Within application of value stream analysis and design, an improvement team maps the current value stream of a relevant product family and illustrates it using standardized value stream elements (Rother & Shook 2011). Wastes can be identified using the concepts of “muda”, “mura”, and “muri” (Ohno 2013). By eliminating the identified wastes and applying the value stream design guidelines, the improvement team designs a resource-efficient target value stream that is optimized in terms of lead time (Erlach 2010). The ubiquitous guideline for all activities is the ideal state known from the Toyota Production System (TPS). Also referred to as the “north star” it incorporates the aspects zero defects, 100 % value creation, one-piece flow on customer demand, and safety for the people (Rother 2009). Inspired by the philosophical ideal state of the TPS, this paper’s socio-technical simulation game is referred to as “North Star® – The Value Stream Game”. Its integrated LMM is referred to as “NorthStarGPT”, emphasizing its role in guiding learners toward the ideal state in the game.

2.2. Information Logistics

To address the increasing complexity of industrial production systems, value stream analysis and design has been extended with additional perspectives (Engel & Kranhold 2021). These enable to create a more holistic understanding of real-world dynamics. One particularly

important extension is information logistics (Meudt 2020). Information logistics refers to the planning, control, and execution of information flows across the value stream. It includes all activities related to how data is collected, stored, transferred, and utilized—using analog and digital memory media. For example, in industrial automotive production, inventory and quality data in the body shop are typically tracked digitally using Manufacturing Execution System (MES), while in final assembly, work instructions or inspection checklists may still be distributed in printed form. These information flows differ in their level of digitization, availability, and utilization, revealing inefficiencies and media gaps across process stages (Meudt et al. 2017). Because information is mostly invisible, cross-functional, and difficult to observe, it is often overlooked in traditional value stream analysis and design. Teaching this concept requires methods that make the dynamics of information logistics actionable. Game-based learning with LMMs can address this challenge by embedding learners in interactive and realistic educational scenarios.

2.3. Game-Based Learning with Large Multimodal Models

Games are structured, goal-oriented activities governed by rules, used for entertainment, competition, or education. Simulation games, a specific game category, replicate real-world systems to allow experimentation and learning in a risk-free environment (Wixon 2015). Game-based learning with Artificial Intelligence (AI) extends this approach by combining interactive game environments with intelligent systems. It offers highly engaging and adaptive learning experiences, supporting motivation, critical thinking, and long-term knowledge retention (Yu & Glazieva 2022). Key elements include personalized feedback, adaptive challenge levels, and real-time learning support (Ma et al. 2022).

A recent and promising advancement in this field is the integration of Large Multimodal Models (LMMs), which extend the capabilities of Large Language Models (LLMs) beyond text by enabling the simultaneous processing of various data types—such as text, images, diagrams, and structured inputs (Baltrušaitis et al. 2018). This multimodality allows for more holistic, context-aware interactions, especially in learning environments that combine visual representations, written inputs, and human decision-making. Within game-based learning, LMMs function as intelligent assistant that provide real-time feedback, answer domain-specific questions, and adaptively support learners across different levels of mastery (Koehn 2020). This technological foundation enables a new instructional format: the socio-technical simulation game. Based on socio-technical systems theory, it combines human-centered learning elements—such as decision-making, collaboration, and systems thinking—with simulation-based experimentation and AI-supported guidance (Brocke et al. 2024). In the context of this paper, the LMM “NorthStarGPT” enhances learner engagement by dynamically interacting with players and facilitating their progression through realistic value stream scenarios.

3. Methodology

To develop and evaluate the extended edition of the socio-technical simulation game “North Star® – The Value Stream Game”, a combined methodological framework was applied, drawing from the six steps to gamification by Werbach and Hunter and the CRISP-DM by Wirth and Hipp (Wirth & Hipp 2000, Werbach & Hunter 2020). While the former provides a structured approach to meaningful game design, the latter introduces an iterative, data-oriented development cycle widely established in big data and data science contexts. This joint application of gamification and the CRISP-DM offer a cyclical information technology (IT)

oriented approach that provides a seamless development flow suited for IT-supported learning systems. The key technological component is the integrated LMM called “NorthStarGPT” based on GPT-4o from OpenAI, which was customized via Retrieval-Augmented Generation (RAG), knowledge modeling, and Chain-of-Thought (CoT) prompt engineering (Lewis et al. 2020). The LMM functions within the simulation game as an adaptive assistant—providing evaluations for the game rounds and context-sensitive guidance. To evaluate the effectiveness of the game in the context of BPM education, a randomized controlled field experiment with a pre-post design was conducted. The study was implemented within the university course “Industrial Engineering” at the Laboratory for Production Management and Technology at Jade University of Applied Sciences, involving 14 master’s students in their second semester ($N = 14$). Among the participants were 4 female and 10 male students. Most participants had prior industrial experience but very few existing knowledge in value stream analysis and design and information logistics. The students were randomly assigned to one of two groups: an intervention group (players, $n = 6$) that engaged with the game, and a control group (observers, $n = 8$) that did not participate in gameplay but completed the same assessment instruments. Both groups took part in identical pre- and post-assessments, allowing for structured comparison of learning effects. The assessment included:

- five knowledge-based questions (theoretical understanding of information logistics),
- one practical task (information logistics design application), and
- three self-assessment questions based on the five levels of mastery (Liker 2020).

To analyze the collected data, both descriptive statistics and inferential statistical methods were applied. Paired-sample t -tests were used to examine significant changes and effects on learning outcomes—specifically in knowledge acquisition, practical skills development, and self-efficacy—within and between groups

4. Implementation and Results

4.1. Implementation

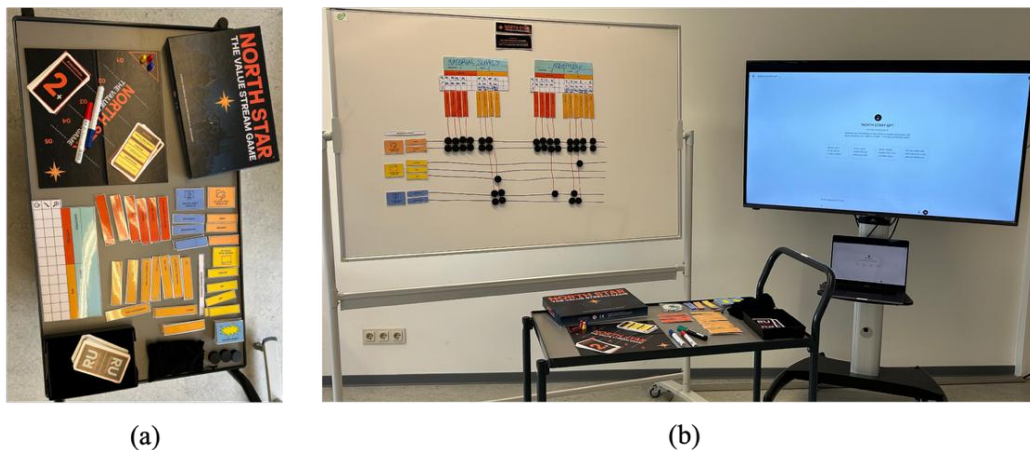
The extended edition of the socio-technical simulation game “North Star® – The Value Stream Game” was developed to teach designing resource efficient information logistics in value streams in a realistic, engaging, and methodologically grounded way. Players—typically learners in higher education or professional development—interactively experiment with value stream scenarios to build methodological competence while fostering resource-efficient thinking. The central learning objective of the simulation game is to enable human improvement teams to apply information logistics design principles in value streams to construct optimized information logistics aligned with the “north star” ideal state. To achieve this, the game presents a realistic scenario with predefined inefficiencies and resource constraints. Teams are challenged to navigate this environment through strategic decision-making, collaborative problem-solving, and critical trade-off analysis between cost, time, and performance.

Gameplay unfolds in five structured steps: First, the facilitator introduces the scenario as a story, the visualized information logistics and the LMM “NorthStarGPT” (*S0*: Initiate). Players are then grouped into three teams of two and get randomly assigned five action cards based on information logistics design guidelines, along with a limited pool of 10 resource units per team for the five game rounds (*S1*: Randomize). Each round, players analyze the current situation,

select a card, invest resource units, and propose a design intervention (*S2*: Analyze and Design). To simulate realistic feedback, the LMM “NorthStarGPT” reviews and evaluates the proposed change and assigns a feedback score from -5 (very bad) to $+5$ (very good) (*S3*: Feedback). This feedback includes both qualitative justification and numerical scoring. The team with the highest score advances one step on the “north star” progression stairs and regains resource units as return on investment. The winning action is then implemented into the shared value stream scenario (*S4*: Implement), and a new round begins (*S5*: Restart). Over five rounds, players iteratively improve the scenario and reflect on the impact of their decisions. The team that wins the most rounds, collects the highest scores, manages resources most effectively, and contributes most to the design of resource-efficient information logistics ultimately wins the game.

4.2. Experimental Setup and Evaluation Results

The evaluation of the extended edition of “North Star® – The Value Stream Game” was conducted as part of the master’s-level course “Industrial Engineering” at Jade University of Applied Sciences. Figure 1 provides insights into the physical game setup, including game materials and the integrated LMM assistant “NorthStarGPT”, as well as the experimental setup.



Setup of “North Star® – The Value Stream Game”. (a) Overview of game components including game box, action cards, resource units, progression stairs, and evaluation cards. (b) In-game setup with game board and integrated LMM displayed on screen for player interaction.

A total of 14 master’s students ($N = 14$) participated in the field experiment. Participants were randomly assigned to either an intervention group (players, $n = 6$), who actively engaged with the simulation game, or a control group (observers, $n = 8$), who did not participate in gameplay but completed the same assessment instruments. All participants completed a pretest and posttest, which included five knowledge-based questions, one paper-based practical task focused on the design of information logistics, and three self-assessment questions (see Section 3).

The game was conducted in a 90-minute session. At the beginning, students were introduced to conventional value stream analysis and design, as well as the theoretical foundations of resource-efficient information logistics. This was followed by the administration of the pretest. After completing the game over five rounds (approximately 60 minutes), the posttest was administered using the same test instruments.

The results of the applied descriptive and inferential statistical methods—conducted using IBM SPSS—are presented in Table 1. Paired-sample *t*-tests were used to analyze changes in learning outcomes across three evaluation dimensions: knowledge acquisition, practical skills development, and self-efficacy. The first two columns of the table display the rounded mean values and standard deviations (SD) for both the intervention and control groups in the pre- and posttest. The third column indicates the relative percentage change from pre- to posttest. The final three columns report the statistical significance (*p*-value) and effect sizes (Cohen's *d*) of the observed changes.

In interpreting these values, *p*-values below 0.05 are considered statistically significant, indicating a meaningful effect of the game. Regarding effect size, values of Cohen's *d* above 0.8 are interpreted as large effects, suggesting strong impact of the game.

Tab. 1: Pre- and posttest results for intervention (*n* = 6) and control groups (*n* = 8) across three evaluation dimensions: knowledge acquisition, practical skills development, and self-efficacy Including Descriptive and Inferential Statistical Key Performance Indicators (*N* = 14).

Group	Pre-Test Mean (SD)	Post-Test Mean (SD)	Mean Δ	<i>p</i> -value	Cohen's <i>d</i>	Significance
<i>Knowledge Acquisition</i>						
Intervention (<i>n</i> = 6)	0.7 (0.5)	2.8 (1.4)	+325.0%	0.005	1.3	Yes
Control (<i>n</i> = 8)	1.4 (1.4)	2.4 (1.5)	+72.7%	0.017	1.1	Yes
<i>Practical Skills Development</i>						
Intervention (<i>n</i> = 6)	1.2 (2.9)	2.2 (3.3)	+85.7%	0.183	2.5	No
Control (<i>n</i> = 8)	0 (0)	0 (0)	±0%	0	0	No
<i>Self-Efficacy</i>						
Intervention (<i>n</i> = 6)	2.2 (0.4)	3.0 (0)	+38.5%	0.002	0.4	Yes
Control (<i>n</i> = 8)	2.0 (0.8)	2.1 (0.5)	5.0%	0.025	0.7	Yes

Concerning knowledge acquisition, the intervention group showed a strong improvement, increasing their mean score from 0.7 to 2.8—an increase of 325%. This change was statistically significant (*p* = 0.005) and associated with a large effect size (*d* = 1.3), indicating a strong learning effect. The control group also improved from 1.4 to 2.4 (72.7% increase), which was statistically significant (*p* = 0.017) with a large effect size (*d* = 1.1), likely reflecting exposure to game observation. However, the effect in the intervention group was more pronounced, suggesting that gameplay contributed significantly to theoretical and conceptual understanding. In the area of practical skills development, only the intervention group showed improvement, increasing from a mean of 1.2 to 2.2. This change did not reach statistical significance (*p* = 0.183), despite a large effect size (*d* = 2.5). The control group showed no measurable change, reinforcing the added value of the game in developing applied skills. The lack of significance may be attributed to the small sample size and task-specific variability. Regarding self-efficacy, the intervention group reported a notable increase in their self-assessed competence level—from 2.2 to 3.0 (+38.5%)—which was statistically significant (*p* = 0.002) with a moderate effect size (*d* = 0.4). The control group showed only a marginal improvement (from 2.0 to 2.1), though this change was statistically significant (*p* = 0.025) with a small-to-moderate effect size (*d* =

0.7), likely due to social desirability or general exposure to the topic. The greater change in the intervention group indicates stronger perceived learning confidence.

5. Conclusion and Outlook

This paper presented the development and evaluation of the extended edition of “North Star® – The Value Stream Game”, a socio-technical simulation game that combines game-based learning with the LMM “NorthStarGPT”. Developed through the integration of the CRISP-DM methodology and gamification, the game offers a structured, iterative, and learner-centered approach to teaching to design resource-efficient information logistics in value streams. One core objective of the game is to “open the blackbox” of information logistics by making invisible flows, decisions, and dependencies within value streams visible, interactive, and actionable. The simulation game was evaluated in a randomized controlled field experiment with 14 master’s students at Jade University of Applied Sciences. Results showed statistically significant improvements in knowledge acquisition and self-efficacy, and a positive—though not statistically significant—trend in practical skills development. While this evaluation provides initial promising insights, it represents a single experimental context. Future research will focus on expanding the study across larger, more diverse learner groups and multiple application settings.

6. References

- Baltrušaitis, T., Ahuja, C., & Morency, L. (2018): Multimodal machine learning: A survey and taxonomy. *IEEE transactions on pattern analysis and machine intelligence* 41 (2): 423-443.
- Bloechl, S. J., Michalicki, M., & Schneider, M. (2017): Simulation game for lean leadership – shopfloor management combined with accounting for lean. *Procedia Manufacturing* 9.
- Brocke, J., van der Aalst, W.M.P., Berente, N., van Dongen, B., Grisold, T., Kremser, W., Mendling, J., Pentland, B., Roeglinger, M., & Rosemann, M. (2024): Process science: the interdisciplinary study of socio-technical change. *Process Science* 2 (1): 1.
- Engel, L., & Kranhold, M. (2021): Lean Management wird zum Green Management, Um Materialflusskosten erweiterte Wertstromanalyse. *Deutsches Ingenieurblatt* 2021 (6): 1-6.
- Erlach, K. (2010): Wertstromdesign, Der Weg zur schlanken Fabrik. 2 ed. Berlin: Springer-Verlag.
- Geisthardt, M, Engel, L, & Marx Gómez, J. (2024): Simulation Game Based Teaching of Value Stream Analysis and Design with Multimodal Large Language Models (MLLMs) Assistance. *SAP Academic Community Conference 2024 (D-A-CH)*, TU München.
- Huang, R., Ritzhaupt, A.D., & Sommer, M. (2020): The impact of gamification in educational settings on student learning outcomes: a meta-analysis. *Educational Technology Research Developmet* 68: 1875-1901.
- Koehn, P. (2020): Neural machine translation. Cambridge University Press.
- Lewis, P., Perez, Ethan, P., Aleksandra, P., Fabio, K., Vladimir, Goyal, N., Küttler, H., Lewis, M., Yih, W. & Rocktäschel, T. (2020): Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in Neural Information Processing Systems* 33: 9459-9474.

- Liker, J. (2020): The toyota way. 2 ed.: Esensi.
- Ma, J., Zhang, Y., Bin, H. & Wang, K. (2022): The Development of Students' Computational Thinking Practices in AI Course Using the Game-Based Learning. ISET 2022, Hong Kong.
- Meudt, T. (2020): Wertstromanalyse 4.0 – Eine Methode zur integrierten Erfassung und Analyse von Material- und Informationsflüssen in Wertströmen, Dissertation Technische Universität Darmstadt. Düren: Shaker Verlag.
- Meudt, T., Metternich, J., & Abele, E. (2017): Value stream mapping 4.0, Holistic examination of value stream and information logistics in production. CIRP Annals.
- Ohno, T. (2013): Das Toyota-Produktionssystem. 3 ed. Frankfurt/Main: Campus Verlag GmbH.
- Rasch, B., Frieze, M., Hofmann, W., & Naumann, E. (2021): Inferenzstatistik. In Quantitative Methoden 1: Einführung in die Statistik für Psychologie, Sozial- & Erziehungswissenschaften, 23-34. Springer.
- Rother, M. (2009): Toyota Kata: Managing People for Improvement, Adaptiveness and Superior Results. New York: McGraw-Hill Global Education Holdings LLC.
- Rother, M., & Shook, J. (2011): Sehen lernen, Mit Wertstromdesign die Wertschöpfung erhöhen und Verschwendung beseitigen. Mülheim: Lean Management Institute GmbH.
- Sailer, M., & Homner, L. (2020): The Gamification of Learning: A Meta-analysis. Educational Psychology Review 32: 77-112.
- Werbach, K., & Hunter, D. (2020): For the win, The power of gamification and game thinking in business, education, government, and social impact. University of Pennsylvania Press.
- Wirth, R., & Hipp, J. (2000): CRISP-DM: Towards a standard process model for data mining. Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining.
- Wixon, D. (2015): What is a game? Interactions 13 (2): 37.
- Yu, D., & Glazieva, A. (2022): Game based learning with artificial intelligence and immersive technologies: an overview. CEUR workshop proceedings.

Competence-Based Assessment for SAP4school: Adaption and Pilot Testing

C. W. Mayer¹, S.-L. Mauch¹, S. Leible² & J. Seifried¹

¹University of Mannheim, Business School, Germany

²University of Hamburg, Germany

Abstract: This paper outlines the development and initial pilot testing of a competence-based assessment tool adapted for the SAP4school teaching and learning material. The aim was twofold: first, to determine whether an existing ERP competence test could be aligned with the learning objectives of SAP4school; and second, to evaluate the tool's applicability and effectiveness in real classroom settings. The adapted instrument covers key competence areas such as legal and commercial principles, process structure, financial accounting and SAP-specific knowledge. The pilot study was conducted with 42 students from a vocational school in Germany. The results indicated that the test performed well in a classroom environment, showing no floor or ceiling effects and a normal distribution of total scores. Performance varied across competence domains, with the highest scores achieved in the 'Process Structure' domain and the lowest in 'Financial Accounting'. Closed-question formats were answered more successfully than open-ended items. These findings support the test's suitability for assessing diverse learner groups and provide a basis for further development. Implications for practice and research include the ongoing psychometric refinement of the test, applying classical test theory and item response theory, and the implementation of a longitudinal study to track vocational students' competence development over time.

Keywords: Competency Assessment, Vocational Schools, Apprentices, Vocational Education and Training, SAP4school.

1. Competency Assessment for ERP in Vocational Education and Training

The digital transformation of business processes has significantly altered the skill sets required in modern workplaces (Billett, 2021). Educational systems, especially in vocational education, must adapt by incorporating digital tools into their curricula (Leppert, 2023; Wilbers, 2021). Enterprise Resource Planning (ERP) systems are becoming increasingly prevalent in vocational education and training (VET), in Germany (Arbeitsgruppe Integrierte Unternehmenssoftware [Integrated Business Software Working Group], 2025) and Austria (HAK.CC, 2025; SAP ACC Wien, 2025). The SAP4school initiative provides vocational schools with a tailored SAP S/4HANA environment and a corresponding curriculum (Leppert, 2023), aiming to enable students a hands-on experience with integrated business processes in SAP. Launched around 2014, this educational program has grown significantly, celebrating its 10th anniversary in 2024 with numerous member schools (SAP4school IUS, 2025). This broad acceptance reflects the recognised importance of ERP skills for vocational students' future careers, i.e., for promoting their employability (Seifried et al., 2019). Despite the growing implementation of the SAP4School curriculum in VET, there is still a lack of research into respective learning outcomes, and considerable discrepancies remain in how vocational students' learning outcomes are assessed. Traditional assessment methods (e.g., assignments, tests, exams) fall short of capturing the competence development fostered through such curricula. Furthermore,

current evaluation practices tend to prioritise procedural knowledge or technical performance over business competencies such as strategic thinking, process integration and decision-making. Currently, there is a lack of a reliable, standardised measure of the learning objectives of the SAP4school curricula. Thus, this study aims to adapt a competence test (Spener, 2023) to align it with the SAP4school teaching material and to pilot the resulting version of this competence-based assessment. To do so, we address the following research questions:

(1) Can a competence test be adapted to align with the SAP4school teaching and learning material?

(2) What do initial pilot results indicate about the adapted test in classroom settings?

This paper outlines the design and pilot testing of a competence-based assessment tool for SAP4school. Next, we provide a brief review of the literature on ERP systems in VET. We then present a definition of competence in VET and the theoretical basis of competency assessment underlying our adaptation. Further, we describe the development of the test instrument and the pilot study. Finally, we analyse the initial pilot data collected from a participating school.

2. Integration of ERP Systems in Vocational Education and Training: Relevance, Implementation and Research Gaps

As digital transformation progresses, the integration of digital tools in education is gaining significance. Consequently, the systematic development of related teaching competencies has become essential (Depré, 2025). In addition, learning situations are becoming increasingly digitized (Leppert, 2023). For educators, this highlights the relevance of understanding digital tools, recognising learning potential, and applying such tools purposefully in their profession (Koehler & Mishra, 2009; Mayer & Seifried, 2022). As noted by Wilbers (2009), ERP systems are considered key tools for representing business processes and allow a meaningful application in educational settings. Classroom integration is therefore particularly well-suited to support competencies related to the use of ERP systems and foster the understanding of basic principles in business processes. Since learners are highly likely to encounter ERP systems in their future workplaces (Spener et al., 2019), it makes pedagogical sense to introduce foundational knowledge of ERP software in VET. This early exposure may help them to develop the essential competencies they need for their professional careers. To this end, there has been an increasing implementation of ERP systems and specific training approaches into vocational teacher education (Gentner et al., 2024; Klatt et al., 2023; Mayer, 2022; Thiel De Gafenco & Klusmeyer, 2023). However, there is still very little research on learning business processes with ERP support in vocational education with regard to the development of the desirable competencies among vocational school students. A study by Spener and Schumann (2021) shows a medium-sized effect for learning business process knowledge with an ERP system, as long as the learning environment was embedded in a more didactically elaborate way.

The competency to understand and apply ERP systems within a profession can be defined as a combination of “[...] intellectual abilities, content-specific knowledge, cognitive skills, domain-specific strategies, routines and subroutines, motivational tendencies, volitional control systems, personal value orientations, and social behaviours” (Weinert, 2001, p. 51). Psychological tests can be used to record such individual abilities, characteristics, skills or states. The classical approach in test theory assumes that every test result is composed of a true value and a random measurement error (Moosbrugger & Kelava, 2020). It provides various models and procedures to assess and optimize psychological measurements. The quality of a

test instrument is evaluated using established criteria that allow for a systematic assessment of their appropriateness. Central to this evaluation are the quality standards: objectivity, reliability, and validity, which are considered essential for ensuring the scientific rigor (Döring & Bortz, 2016). Objectivity refers to the degree to which test results remain unaffected by the individual administering, scoring or interpreting the test. Reliability reflects the consistency and precision with which a test measures a given characteristic. Validity is defined as the extent to which a test accurately measures the construct it purports to assess.

3. Adaption and Pilot Testing of a SAP4school Competence Test

The underlying competency test developed by Spener (2023) was based on the ERP software Navision in cooperation with the working group Integrus NAV4school. A total of 29 items were originally assigned to three competence areas: 17 items addressed legal and commercial principles of business processes, seven items focused on process structure, and five items targeted financial accounting. The items were constructed using both dichotomous formats (correct/incorrect) and multiple answer formats. In terms of reliability, the original items (see Spener & Schumann, 2021) demonstrated the construct of business process knowledge with an acceptable internal consistency. The instrument is anticipated to demonstrate adequate validity, given its theory-driven development and strong conceptual foundation.

The adapted test instrument consists of 30 tasks⁸, which include a total of 42 scorable items (accounting for subtasks) used to evaluate performance. Fundamental changes were made to adapt the original instrument to the SAP4school modules (see Table 1). Competence is assessed based on the total number of points earned, including partial credits, with a maximum possible score of 82 points. Of the 57.5 maximum points attainable in the subdomain of order processing, 21.5 points (37.4%) are allocated to the competence area of financial accounting (eight items), 22 points to legal and commercial fundamentals (eight items), and 10.5 points to process structure (six items). Only one item, worth 4 points, pertains specifically to SAP-related organizational terminology. The Fundamentals of Business Processes section comprises a total of 10 items, including seven multiple-choice questions and two open-ended questions drawn from the original test instrument. One additional item was added, bringing the maximum achievable score for this section to 15.5 points. The second section was entirely redesigned and consisted of SAP-specific multiple-choice questions, with a maximum of nine points. These items primarily assess knowledge of organizational terminology within the SAP system. The third section, Order Processing, is the most content-rich part of the assessment and is largely based on the structure and content of the original test instrument and adapted to the respective module in SAP4school. All associated documents, such as order confirmations, delivery notes and invoices, were adapted to align with the SAP4school IUS content.

⁸ The adaptation and piloting of the instrument was part of the master's thesis of Sarah-Lea Mauch.

Tab. 1: Overview of the adapted instrument.

Domain	Number of Items	Max Score	Example Item
Fundamentals of Business Processes	10	15.5	<i>“How is a purchase contract concluded?” (Single Choice)</i>
SAP S/4HANA specific knowledge	8	9	<i>“What is the highest hierarchical level of an SAP system?” (Single Choice)</i>
SAP4school Business Processes ¹	12	57.5	<i>“There is not enough stock of the air pump to fulfil the order from Cruiser Bikes KG. Therefore, please place the order shown. a) To which process is this operation assigned (Single Choice) b) Is the order posted? If yes, please enter the accounting record with amounts. If no, please justify your answer. (Text answer)”</i>
Total	30	82	

¹ based on SAP4school IUS moduls A1 (Data research and maintenance), A3 (Simple sales process) and A4 (Simple procurement process).

The adapted competence test was piloted with three classes at a commercial vocational school. Students were informed in advance about the data collection. Before administration, all participants received standardised instructions and signed consent forms. For minors, parental consent was obtained beforehand. The paper-based test was conducted in a regular classroom without the need for technical equipment. Although 60 minutes were allocated, most students completed the test within 45 minutes.

A total of 42 vocational students from a Commercial Vocational College participated in piloting the competence test. Their weekly timetable included seven hours of business studies, three hours of management and control, two hours of general economics, and four hours of business processes. All participating vocational students had completed the SAP4school IUS modules A1 (Data research and maintenance), A3 (Simple sales process) and A4 (Simple procurement process), with module A4 concluded two weeks before data collection. For ethical compliance, signed data protection forms from the parents of underage participants were requested. The students' average age was $M = 17.58$ years ($SD = 1.50$), ranging from 16 to 23 years. The sample included 25 female (59.5%) and 17 male (40.5%) participants.

4. Results of Piloting the SAP4school Competence Test

Regarding the ‘Fundamentals of Business Processes’, participants achieved an average score of 7.58 points ($SD = 2.47$), representing 49.97% of the maximum 15.5 points, with individual scores ranging from 2.0 to 12.5. In the ‘SAP-specific section’, the mean score was 4.0 points ($SD = 1.56$), corresponding to 43.45% of the possible nine points. Scores ranged from 1 to 8 points. For the ‘Order Processing’ section, the most extensive part of the test, the average score was 24.31 points ($SD = 6.63$), equating to 42.81% of the maximum 57.5 points. Individual results ranged from 10.5 to 38.5 points. The distribution of the overall test scores ($M = 35.89$, $SD = 8.97$) did not significantly deviate from normality (see Fig. 1), as indicated by the Shapiro-Wilk test ($W = 0.977$, $p = 0.548$). Therefore, the assumption of a normal distribution can be considered met.

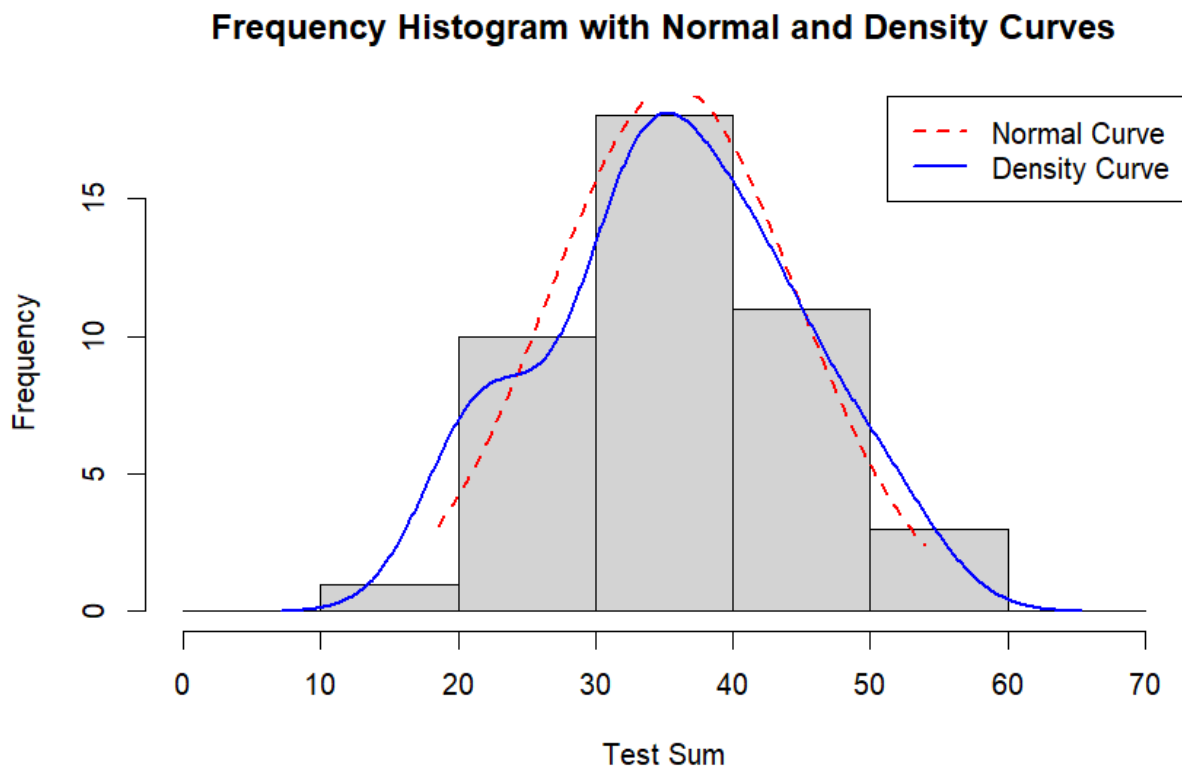


Fig. 1: Frequency Histogram.

Participants scored an average of 49.70 % in the area of legal and commercial principles. The highest performance was observed in the process structure domain, with a mean score of 59.22 %. In contrast, results in financial accounting were notably lower, with participants achieving only 20.29 % of the possible points. For SAP-specific knowledge, the average score was 46.13 %. These findings reflect varying levels of performance across the assessed competence areas. Additionally, average solution rates were analysed by task format. Open-ended items yielded a mean success rate of 30.96 %, while closed-format questions showed a substantially higher average of 49.70 % correct responses.

5. Discussion and Conclusion

This study addressed the research question of how a competence test for NAV4school can be adapted to align with the SAP4school teaching and learning material. The results suggest that such an adaptation is feasible, although it requires substantial modifications, particularly concerning SAP S/4HANA-specific content and the integration of contextually appropriate documents such as receipts and related business documents. The pilot results suggest that the adapted test works effectively in classroom settings. No floor or ceiling effects were detected, and the overall score distribution closely followed a normal pattern. These first findings support the test's general appropriateness for measuring competence levels among a heterogeneous group of learners.

The findings from the pilot study have important implications for research and educational practice. In terms of its practical application, the tool could be used over the long term to evaluate vocational students. Therefore, extensive research is required to analyse the SAP4school adapted competence test. From a psychometric perspective, a descriptive analysis based on classical test theory, examining item difficulty, variance and item-total correlations, are necessary to analyse the quality and diagnostic accuracy of the test items. Items with particularly low/high solution rates will be examined in more detail. Additionally, increasing the sample size allows for more robust analyses applying item response theory, offering a deeper insight into item characteristics and measurement accuracy at different levels of learner competence. Therefore, a longitudinal study is being prepared to examine how students' competencies develop over a school year. Such a study would comprise at least three measurement points: the first at the start of the school year (t_0), the second after completion of the relevant SAP4school module (t_1) and the third to evaluate the application of learning in new or practical contexts (t_2). This design will provide valuable insights into learning progress and the long-term effectiveness of competence-based ERP education.

6. References

- Arbeitsgruppe Integrierte Unternehmenssoftware [Integrated Business Software Working Group]. (2025). Integrus.de. <https://integrus.de/>
- Billett, S. (2021). Mediating worklife learning and the digitalisation of work. *British Journal of Educational Technology*, 52(4), 1580–1593. <https://doi.org/10.1111/bjet.13115>
- Depré, K. (2025). Universitäre Förderung medienfachdidaktischer Wissensfacetten und medienbezogener Überzeugungen angehender Lehrkräfte: Eine Betrachtung der Nutzbarkeit multimedialer Lehr-Lerntools [University promotion of media-related teaching knowledge and media-related beliefs among prospective teachers: A consideration of the usefulness of multimedia teaching and learning tools]. Springer Fachmedien. <https://doi.org/10.1007/978-3-658-48341-8>
- Döring, N. (2023). Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften [Research methods and evaluation in the social and human sciences]. Springer. <https://doi.org/10.1007/978-3-662-64762-2>
- Gentner, S., Ludwig, S., Braunstein, A., Mayer, C., Deutscher, V., Rausch, A., & Seifried, J. (2024). The digital office simulation LUCA from the perspective of teachers and learners: First findings of usability analyses. In Cedefop & OECD (Eds), *Apprenticeships and the digital transition: Modernising apprenticeships to meet digital skill needs*. (pp. 102–117). Publications Office. <https://data.europa.eu/doi/10.2801/074640>
- HAK.CC. (2025). HAK.CC. <https://www.hak.cc/unterricht/hard-und-software/sap>
- Klatt, C., Härtel, S., & Markert, J. (2023). Digitalisierung im Berufsfeld Ernährung und Hauswirtschaft – Einsatz eines Property-Management-Systems aus der Hotellerie in der Lehramtsbildung für berufsbildende Schulen [Digitization in the Vocational Field of Nutrition and Home Economics—Implementation of a property management system adapted from the hotel industry in teacher education for vocational schools]. *bwp@ Berufs- und Wirtschaftspädagogik – online*, 45, 1–20.
- Koehler, M., & Mishra, P. (2009). What is technological pedagogical content knowledge (TPACK)? *Contemporary Issues in Technology and Teacher Education*, 9(1), 60–70.

- Leppert, S. (2023). Vertieft curricular analysieren Eine designbasierte Untersuchung der prozessorientierten Sachanalyse zur digitalen Transformation von Lernsituationen [In-depth curricular analysis A design-based investigation of process-oriented subject analysis for the digital transformation of learning situations] (Texte zur Wirtschaftspädagogik und Personalentwicklung, Vol. 33). epubli.
- Mayer, C. W. (2022). Lehren und Lernen mit Enterprise Resource Planning (ERP) Systemen—Typische Bearbeitungsprobleme als Grundlage der Lernprozessgestaltung [Teaching and learning with Enterprise Resource Planning (ERP) systems Typical processing problems as a basis for designing the learning process]. *Bwp@ Berufs- Und Wirtschaftspädagogik – Online*, 43, 1–17.
- Mayer, C. W., & Seifried, J. (2022). Lehren und Lernen mit ERP-Systemen – Befunde der Evaluation einer Lehrveranstaltung zur Förderung der Technologiekompetenz von Studierenden der Wirtschaftspädagogik [Teaching and learning with ERP systems—Findings from the evaluation of a course to promote the technological competence of business education students]. In U. Mathis, N. Ondrusch, D. Kilian, H. Krcmar, K. Turowski, & S. Weidner (Eds), *Proceedings of the SAP Academic Community Conference 2022 DACH* (pp. 10–24). Technische Universität München. <https://mediatum.ub.tum.de/doc/1685828/document.pdf#page=10>
- Moosbrugger, H., & Kelava, A. (Eds). (2020). *Testtheorie und Fragebogenkonstruktion* [Test theory and questionnaire construction]. Springer. <https://doi.org/10.1007/978-3-662-61532-4>
- SAP ACC Wien. (2025). ACC Portal. <https://portal.acc-wien.at/#/dashboard/kundenPartner>
- SAP4school IUS. (2025). SAP4school IUS - SAP an beruflichen Schulen – Startseite [SAP4school IUS - SAP at vocational schools—Homepage]. <https://sap4school-ius.integrus.de/>
- Seifried, J., Beck, K., Ertelt, B.-J., & Frey, A. (Eds). (2019). *Beruf, Beruflichkeit, Employability* [occupation, professionalism, employability]. wbv.
- Spener, C. (2023). Effekte des Einsatzes integrierter Unternehmenssoftware im kaufmännischen Unterricht an Berufsschulen [Effects of using integrated business software in commercial lessons at vocational schools]. Springer Fachmedien. <https://doi.org/10.1007/978-3-658-42181-6>
- Spener, C., Häuber, G., Horlacher, T., & Schumann, S. (2019). Der berufsschulische Einsatz von ERP-Systemen in der Ausbildung von Industriekaufleuten [The vocational use of ERP systems in the training of industrial clerks]. In K. Wilbers (Ed.), *Digitale Transformation kaufmännischer Bildung: Ausbildung in Industrie und Handel hinterfragt* (pp. 195–216). epubli. https://www.pedocs.de/frontdoor.php?source_opus=17968
- Spener, C., & Schumann, S. (2021). Wissenseffekte des ERP-Einsatzes in der kaufmännischen Berufsschule [Knowledge effects of ERP use in commercial vocational schools]. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 117(3), 395–430. <https://doi.org/10.25162/zbw-2021-0018>
- Thiel De Gafenco, M., & Klusmeyer, J. (2023). Fach- und mediendidaktische Verschränkung digitaler Lern- und Arbeitsmittel: Entwicklungs- und Evaluationsaspekte eines Seminarkonzeptes der Wirtschaftspädagogik [Intertwining digital learning and working tools with subject-specific and media-specific teaching methods: Development and

evaluation aspects of a seminar concept in business education]. In M. Meier, G. Greefrath, M. Hammann, R. Wodzinski, & K. Ziepprecht (Eds), *Lehr-Lern-Labore und Digitalisierung* (pp. 255–268). Springer Fachmedien. <https://doi.org/10.1007/978-3-658-40109-2>

Weinert, F. E. (2001). Concept of competence: A conceptual clarification. In D. S. Rychen & L. H. Salganik (Eds), *Defining and selecting key competencies* (pp. 45–65). Hogrefe.

Wilbers, K. (2009). Integrierte Unternehmenssoftware (ERP-Systeme) im kaufmännischen Unterricht [Integrated business software (ERP systems) in commercial teaching]. In H. Pongratz, P. T. Tramm, & K. Wilbers (Eds), *Prozessorientierte Wirtschaftsdidaktik und Einsatz von ERP-Systemen im kaufmännischen Unterricht* (Vol. 4, pp. 61–76). Shaker.

Wilbers, K. (2021). Kaufmännische Aus- und Weiterbildung in der Industrie im Umbruch Digitale Transformation im Zuge von Industrie 4.0 und künstlicher Intelligenz [Commercial training and further education in industry undergoing radical change Digital transformation in the wake of Industry 4.0 and artificial intelligence]. In S. Seufert, J. Guggemos, D. Ifenthaler, H. Ertl, & J. Seifried (Eds), *Künstliche Intelligenz in der beruflichen Bildung: Zukunft der Arbeit und Bildung mit intelligenten Maschinen?!* (pp. 55–76). Franz Steiner.

PDF oder Video? – Ein empirischer Vergleich zweier Lernformate für E-Learning-basierte SAP-Endanwenderschulungen

M. Wienand

Ruhr Campus Academy gGmbH – We Learn in Bits, Essen, Deutschland

Abstract: Die erfolgreiche Vermittlung von Kompetenzen im Umgang mit SAP-Systemen stellt insbesondere in der Hochschullehre eine didaktische Herausforderung dar. Studierende sehen sich dabei häufig mit komplexen Benutzeroberflächen und abstrakten Systemlogiken konfrontiert – klassische Lernmaterialien wie PDF-Skripte stoßen hier schnell an ihre Grenzen. Der vorliegende Beitrag präsentiert die Ergebnisse einer empirischen Studie, in der zwei asynchrone Lernformate – ein textbasierter PDF-Kurs und ein multimedialer Videokurs – im Rahmen eines SAP-Trainings für Studierende systematisch verglichen wurden.

Das kontrollierte Experiment mit über 400 Teilnehmenden untersuchte sowohl den Lernerfolg (Pre-/Post-Test) als auch die wahrgenommene kognitive Belastung (NASA-TLX, SEQ) bei der Bearbeitung typischer SAP-Aufgaben. Die Ergebnisse zeigen deutlich: Multimediale Lernformate werden von Studierenden nicht nur bevorzugt, sondern führen auch zu höherer Motivation und geringerer kognitiver Belastung – ohne dabei den Lernerfolg zu mindern. Der Beitrag leitet daraus konkrete Empfehlungen für die didaktische Gestaltung künftiger SAP-Schulungsangebote in der Hochschullehre ab, die auf multimediale, modular strukturierte und zielgruppengerechte E-Learning-Formate setzen sollten.

Keywords: Multimediales Lernen, Cognitive Load Theory, Digitale Hochschullehre, ERP-Training, Instructional Design

Nachhaltiger Lernerfolg durch Simulation? Ein Feldversuch zum Vergleich von ERPsim Manufacturing und PP- Fallstudie

Gesine Hilf, Maren Schaal, Wolf Wenger
DHBW Stuttgart, Deutschland

Abstract Enterprise Resource Planning (ERP)-Systeme stellen eine zentrale Komponente moderner Unternehmensanwendungen dar und sind somit ein integraler Bestandteil vieler technischer und betriebswirtschaftlicher Studiengänge. Zur Vermittlung praxisnaher und anwendungsorientierter Kompetenzen im Umgang mit ERP-Systemen setzt die Duale Hochschule Baden-Württemberg (DHBW) Stuttgart auf den Einsatz des ERP-Systems SAP S/4HANA in der Lehre.

Aktuell wird in der Lehrveranstaltung „Produktion und Logistik“ überwiegend die PP-Fallstudie des SAP University Competence Center (UCC) Magdeburg verwendet. Diese Fallstudie ermöglicht eine schrittweise Bearbeitung des gesamten Produktionsplanungs- und -steuerungsprozesses, von der Prognose über die Materialbedarfsplanung bis hin zur Erstellung von Fertigungsaufträgen. Im Rahmen eines Feldversuchs wird untersucht, ob das Planspiel ERPsim Manufacturing der HEC Montréal einen nachhaltigeren Lernerfolg erzielen kann als die Bearbeitung der PP-Fallstudie, die auf einer Schritt-für-Schritt-Anleitung basiert.

ERPsim Manufacturing, welches ebenfalls auf SAP S/4HANA basiert, bietet eine interaktive und wettbewerbsorientierte Lernumgebung. Dabei treffen die Studierenden unter realitätsnahen Bedingungen betriebswirtschaftliche Entscheidungen und erfahren deren Auswirkungen unmittelbar in einem kompetitiven Umfeld. Ähnlich wie in der Fallstudie liegt der Schwerpunkt auf Entscheidungen im Bereich der Produktionsplanung und -steuerung.

Im Rahmen einer begleitenden Untersuchung im Mixed Model Design mit qualitativen und quantitativen Fragen in Form eines schriftlichen Evaluationsbogens sowie Leitfadeninterviews wurde der Einsatz von ERPsim Manufacturing mit der klassischen PP-Fallstudie aus dem UCC-Curriculum verglichen. Erste Ergebnisse deuten darauf hin, dass die Simulation den wahrgenommene Lerneffekt der Studierenden positiv beeinflusst. Auch wenn ERPsim als anspruchsvoller eingestuft wurde, berichteten die Teilnehmenden von einem nachhaltigeren Lerneffekt im Vergleich zur stark geführten Fallstudie. Insbesondere wurde hervorgehoben, dass ERPsim Manufacturing eine präzise Planung unter Berücksichtigung aller relevanten Faktoren und eine zeitpunktgenaue Umsetzung erforderte. Dies führte dazu, dass die Studierenden den Umgang mit SAP und die Auswirkungen ihrer Entscheidungen auf das Ergebnis schneller und effektiver erlernten.

Im Gegensatz dazu wurde bei der Fallstudie die Komplexität der Zusammenhänge als schwieriger verständlich empfunden. Zusätzlich betonten die Studierenden bei ERPsim die höhere Interaktivität, die unmittelbare Rückmeldung auf Entscheidungen sowie eine stärkere Motivation während des Lernprozesses. Zur Validierung dieser ersten Ergebnisse und zur fundierten didaktischen Evaluation sind weiterführende empirische Untersuchungen geplant.

Key Words: SAP S/4HANA, ERPsim, PP Fallstudie, Unternehmenssimulation, Evaluierung,

Vom Lernhindernis zum Lernpfad – Microlearning als Schlüssel für effektive SAP-Endanwenderschulungen

Mareen Wienand

We Learn in Bits (Ruhr Campus Academy gGmbH/ Universität Duisburg-Essen), Deutschland

Abstract: Die Einführung und Nutzung von Enterprise Systems (ES) wie SAP S/4HANA stellt sowohl Unternehmen als auch Bildungseinrichtungen vor didaktische Herausforderungen. Endanwenderschulungen gelten als kritischer Erfolgsfaktor, sind jedoch häufig durch geringe Akzeptanz, hohe Komplexität und ineffiziente Vermittlung geprägt. Der Beitrag zeigt, wie didaktisch fundierte Microlearning-Formate dazu beitragen können, diese Herausforderungen zu adressieren. Auf Basis von Erkenntnissen aus mehreren Studien und Experimenten werden Designprinzipien vorgestellt, die auf Theorien der kognitiven Belastung (Cognitive Load Theory) sowie auf lernpsychologischen Modellen wie dem Dual-Channel-Ansatz basieren.

Im Zentrum steht die Fragestellung, wie multimediale und modular aufgebaute Lernpfade gestaltet sein müssen, um Wissen über SAP-Systeme nachhaltig zu vermitteln und zugleich eine positive Haltung gegenüber der Software zu fördern. Die präsentierten Ergebnisse basieren auf empirischen Untersuchungen zur Medienwirkung sowie auf der praktischen Erprobung in Hochschul- und Weiterbildungskontexten. Der Beitrag bietet praxisrelevante Erkenntnisse zur Gestaltung interaktiver, skalierbarer Lernformate, die sowohl in akademischen Curricula als auch in betrieblicher Weiterbildung Anwendung finden können.

Keywords: Instructional Design, Lerntechnologie, Micro Learning, Multimediales Lernen, Digitale Hochschullehre

SAP-Lehrkonzept „Personalwesen mit ERP-Systemen im Kontext der öffentlichen Verwaltung in Deutschland“

Carolin Hilgers, Dipl.-Volksw.

Hochschule des Bundes für öffentliche Verwaltung, Fachbereich Finanzen, Mecklenbecker Straße
235, 48163 Münster, carolin.hilgers@vit-bund.de

Abstrakt Auch die öffentliche Verwaltung setzt für ein effektives und effizientes Personalmanagement auf den Einsatz von ERP-Systemen. Es sind jedoch keine Lehrmaterialien zum Thema Personalwesen mit ERP-Systemen für den öffentlichen Dienst in der Literatur oder am Markt verfügbar. Ein Forschungsvorhaben der HS Bund soll aktuelle Lehr- und Lerninhalte zum Thema Personalwesen mit ERP-Systemen im öD in einem digitalisierbarem Lernmodul auf Bachelor-Niveau zusammenfassen.

Im Mittelpunkt steht die Entwicklung einer SAP H4S4-Lernumgebung für die Umsetzung von ERP-Fallstudien in öffentlichen Verwaltungen, die die Strukturen und Prozesse der Personalverwaltung im öD abbildet. Die fachlichen Inhalte für eine modulare Fallstudie auf Basis personalrechtlicher Grundlagen des öD (Tarif-/Beamtenrecht, Datenschutz) werden in Kooperation mit den Praxispartnern ITZBund und Dataport erarbeitet. Sie soll sich sowohl an Anwendende als auch an Systembetreuende (Customizing, Berechtigungen, Betrieb) richten, denn die Studierenden der IT-Studiengänge an Hochschulen der öffentlichen Verwaltung sind die zukünftigen Beschäftigten der IT-Dienstleister.

Folgende Forschungsfragen stehen im Mittelpunkt:

- Prozess- vs. Rollenorientierung von Fallstudien: Möglichkeiten entwickeln und untersuchen, die Rollenorientierung zu verbessern und damit den Lerneffekt der Fallstudien zu vergrößern.
- Learning Analytics: Das zu entwickelnde Lernmodul soll dazu beitragen, die Wirksamkeit anhand der konkreten Systeminteraktion der Lernenden zu messen.
- Binnendifferenzierung der Teilnehmenden: Bei größeren Lerngruppen gibt es keine für alle richtige Lehr-/Lern-Geschwindigkeit, darunter leidet die Motivation. Unterschiedliche Lerngeschwindigkeiten sollen aktiv innerhalb eines Lerngruppenverbundes eingesetzt und bzgl. ihrer Wirksamkeit untersucht werden.

NextGen Modelers – Praxisorientierte Lehre von Prozessmodellierung

Prof. Dr. Susanne Glissmann
Berliner Hochschule für Technik (BHT)

Abstract: Wie können Studierende optimal auf die Anforderungen dynamischer Unternehmensumgebungen und zukünftiger BPMN-Systeme vorbereitet werden? Der Beitrag „NextGen Modelers – Praxisorientierte Lehre von Prozessmodellierung“ stellt ein Seminarkonzept vor, mit dem Studierende Teamfähigkeit, kritisches Denken sowie Rollenverständnis als Process Model Reviewer und Transformation Manager erwerben und dazu eine hohe Kontextsensitivität für die Praxis entwickeln. Durch ein modular aufgebautes Format – von BPMN-Basiseinheiten über diskussions- und textbasierte Modellierungsaufgaben bis hin zu praxisnahen Übungen in Unternehmensszenarien – sollen Studierende lernen, Modelle korrekt zu erstellen, kritisch zu hinterfragen und flexibel an unterschiedliche Organisationskontexte anzupassen. Zur semantischen Regelvalidierung werden KIPrompts eingesetzt

Improving Learning Outcomes: A Multi-Stimuli Approach on Teaching ERP Systems in Higher Education

Boris A. Feige, Birgit Lankes

University of Applied Science Niederrhein, Mönchengladbach, Germany

Abstract: Enterprise Resource Planning (ERP) systems are essential for managing business processes across corporate value chains. Given their importance in professional practice, many higher education institutions have included ERP-related courses into their curricula. However, teaching ERP systems presents several challenges such as a high system complexity and a required profound understanding of business processes. To address these challenges and improve learning outcomes of students, this paper presents a conceptual framework based on the Stimulus-Organism-Response (SOR) paradigm. The adapted SOR paradigm for tertiary education is used to structure a multi-stimuli learning environment for ERP competence development. The concept is finally applied to an ERP teaching scenario involving SAP S/4HANA in undergraduate programs offering practical guidance by defining a so-called multi-stimuli sedcard.

Keywords: ERP Systems, Education, SOR paradigm, Learning Stimuli, SAP S/4HANA

1. Introduction

1.1. Problem Statement

Enterprise Resource Planning (ERP) systems are by far the most relevant core systems when it comes to supporting business processes in the value chain of corporates (Gronwald 2025, Sarferaz 2023). The possibility that students in higher education institutions will get in touch with an ERP system during their professional career is most likely (Pfarr, Klima & Winkelmann 2014, Venkatesh 2018). For this reason, many tertiary institutions have included courses on ERP systems in their curricula in various degree programs such as business administration, business informatics or information technology (Schwade & Schubert 2016). Typically, the usage, the technical foundations and the administration of an ERP system are taught. However, this is accompanied by challenges in terms of complexity of an ERP system, the variety of modules and the granularity of individual modules. In addition to learning how to use the system, the course content usually includes links to the numerous business processes and the stakeholders involved from different business functions. These challenges require lecturers to have an appropriate, well-balanced didactic concept to teach and develop the desired competences within a limited scope and time (Schwade & Schubert 2016).

1.2. Research Framework

The underlying research framework is based on the stimulus-organism-response (SOR)-paradigm (Mehrabian & Russell 1974). Our aim is to develop a framework that structures the understanding of how multiple learning stimuli can be set in an ERP system-related learning environment and how such stimuli positively influence the associated learning outcomes of students. A theoretical-conceptual approach is used to adapt the SOR paradigm to teaching in

tertiary education and in particular to address given challenges in student competence development on ERP systems. In conclusion, recommendations for practical teaching activities will be given based on a multi-stimuli learning scenario.

The following research question is guiding the further activities.

Which multiple stimuli can be set to teach ERP systems and how can these stimuli be systematically structured with regard to an adapted SOR paradigm for tertiary education?

2. Theoretical Foundations

2.1. Challenges in Teaching on ERP Systems

Teaching ERP systems results in potential challenges. Administrative challenges arise when the focus is on teaching practical usage and handling of an ERP system. It requires a dedicated ERP system that needs to be financed and, in many cases requires institutional support or third-party cooperation (Becerra-Fernandez et al. 2000). Moreover, system administration involves considerable complexity and lecturers must continuously stay up to date with software developments, system maintenance and evolving system interfaces (Mayer 2022). Service providers such as the SAP University Competence Center Magdeburg have targeted this white spot as a business opportunity and are providing a variety of services for higher education institutions (SAP University Competence Center Magdeburg n.d.). Within the teaching process itself, challenges are mostly related to the complexity of an ERP system serving many business functions and processes. Students often have difficulties recognizing and implementing the correct sequence of steps within the ERP system (Schumann & Schattke 2015). A lack of business fundamentals and a limited understanding of how an ERP system operates means that students do not fully grasp the impact and outcomes of their actions. This impairs the ability to adequately execute complex business processes inside the ERP system. These challenges result in tertiary educational institutions offering courses that focus more often on theoretical aspects of ERP systems rather than on their practical operation (Mayer 2022). However, as the market seeks for graduates with the appropriate theoretical knowledge and the competence to operate relevant modules of an ERP system, lecturers face the challenge of developing a didactic concept that reflects the complexity of the system and addresses students' challenges to achieve the intended learning outcomes.

2.2. The SOR Paradigm

The origins of the Stimulus-Organism-Response paradigm trace back to the early 20th century. At that time, classical behaviorists advocated a reductionist Stimulus-Response (S-R) model, which limited human behavior to observable stimulus-response patterns. Pioneers such as Watson and Pavlov shaped this model through experiments interpreting behavior as a direct reaction to external stimuli (Watson 1913, Pavlov 1927). This model was soon subject to criticism, most notably by Woodworth. Woodworth argued that the S-R model neglected crucial internal processes of the individual, namely cognitive and emotional states. He introduced the organismic variable (O) to the framework, thereby transforming it into the S-O-R model: a stimulus (S) must first be processed by the organism (O) before leading to a behavioral response (R) (Woodworth 1929). A more systematic development of the S-O-R framework was carried out by Mehrabian & Russell, who extended the model to the domain of environmental psychology. They proposed a causal sequence connecting environmental

stimuli, internal affective states and observable behavior (Mehrabian & Russell 1974). In contrast to the traditional S-R model, the S-O-R approach considers the cognitive and affective processes within the organism that mediate between stimulus and response (Mehrabian & Russell 1974, Zhang et al. 2022).

Stimulus (S): The stimulus refers to external or environmental factors that affect an individual. These stimuli are highly context-dependent: In retail environments, stimuli can include atmospheric cues such as interior design, lighting, music, scent, and crowding (Vieira 2013). In educational contexts, stimuli include aspects of classroom layout, instructional methods, learning materials and social interaction between students and teachers (Zhang et al. 2021).

Organism (O): The organism represents internal cognitive and emotional processes that respond to a stimulus. These include perceptions, emotional predispositions and evaluative judgments (Zhang et al. 2021) but also individual perceived uncertainty (Vieira 2013).

Response (R): The response consists of observable behaviors or actions resulting from internal processing of the stimulus. These include positive behaviors such as desire to engage, explore, or remain in a setting but also negative behaviors such as desire to withdraw, disengage or even escape (Vieira 2013). In educational contexts, responses commonly include student satisfaction with learning outcomes and engagement (Zhang et al. 2021).

3. Multi-Stimuli Learning Scenarios

3.1. Tailoring the SOR Paradigm to ERP Competence Development

The SOR paradigm establishes a causal link between measurable external stimuli, its non-observable processing inside an organism and a measurable ultimate reaction. To transfer the concept to the underlying research topic of tertiary education, the following adaptations are made (Figure 1).

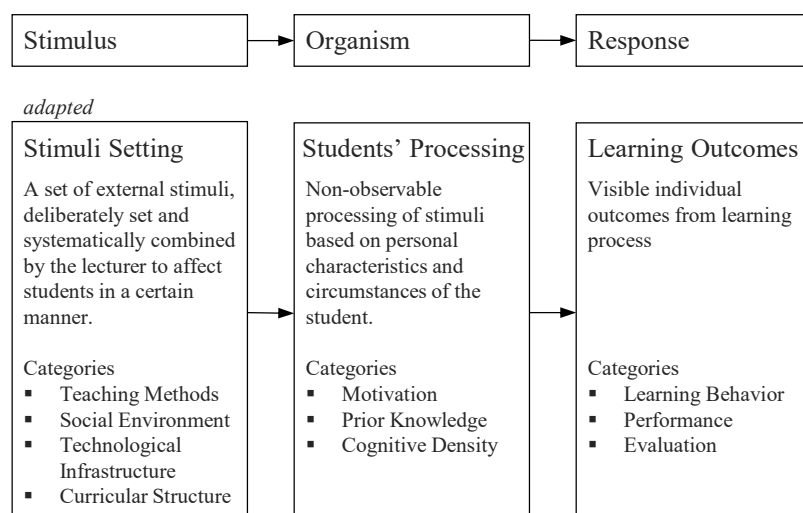


Figure 1: Adapted SOR paradigm for teaching in tertiary education

Stimuli Setting

Multiple stimuli are defined as stimuli that are external, intentionally set and combined in a systematic way by a lecturer, aiming to guide, influence or activate specific cognitive,

emotional or behavioral responses of students. These stimuli can be structured in four superordinate categories: Teaching methods, social environment, technological infrastructure and curricular structure. These categories and its related stimuli are combined in a purposeful manner to create the intended learning conditions.

Different **teaching methods** such as lectures, individual or group work, discussions, guided research activities or case studies promote the learning processes and appeal to students' different learning styles. An appropriate mixture of teaching methods promotes activity and encourages students to gain a lasting understanding of the learning content. The **technical infrastructure** refers to all technical aids that support the learning process. This includes e.g., appropriate classroom media technology, access to required media such as books and papers, shared digital learning spaces or access to business application software for teaching purposes. The **social environment** is characterized by the type of communication and interaction between students and lecturers. Trust and belonging motivates active participation and strengthen the motivation to persevere in a complex enduring learning process.

The **curricular structure** is defined as the didactic structure of a course or a course unit, for example through content that builds on each other, graduated levels of difficulty or the alternation between theory and practice. A well-structured sequence of learning steps facilitates orientation, prevents overload and supports continuous improvement of competences and skills.

Students' Processing

The students' processing is defined as the individual processing of perceived stimuli. This internal, non-observable process is influenced by the students' individual character and circumstances, resulting in three categories: Motivation, prior knowledge and cognitive density.

Motivation acts as a driving factor and increases the willingness to engage with the course and its content. It determines how attentive students are in responding to external stimuli. A high intrinsic motivation strengthens the willingness to actively engage with teaching content based on a self-determined willingness to learn. The extrinsic motivation is triggered by external incentives such as grades or future career benefits and influences engagement primarily in terms of a predefined goal achievement.

Prior Knowledge refers to the existing knowledge and experiences that students bring before starting the learning process. This prior knowledge forms the basis on how new information can be absorbed, categorized and linked to existing concepts. An appropriate prior knowledge base facilitates the understanding of complex content, promotes efficient learning and enables recognizing novel contextual connections more quickly. Given the diverse student body, prior knowledge can have a significant impact on the learning success of an entire group or individual student.

Cognitive density can be divided into two elements. First, it describes the perceived complexity and information density of the learning content to be processed within the course itself, i.e. how many new concepts, details and contexts a student has to process at the same time and across the entire course. On the other hand, it includes the cumulative load on students e.g., by taking multiple courses in the same semester or through individual factors such as a parallel part-time or full-time employment or care work. All of these elements compete for students' attention.

Learning Outcomes

The learning outcomes are defined as achieved knowledge and competences that are measurable by analyzing the students' learning behavior, performance and course evaluation.

Learning behavior describes how actively and purposefully students engage with the learning content, both in classroom situations and in self-study. It is reflected by the level of attention, the participation and quality of classroom discussions, the systematic use of learning resources and the engagement in individual exercises or group work.

The **performance characteristics** include measurable achievements of students. It can be measured based on the grades achieved, the dropout rate of the course or the exam no-show rate.

Evaluation includes the assessment of the entire course offering by students and contains recommendation rate, the course evaluation based on applied feedback methods such as the teaching analysis poll (Schmidt 2024) or standardized course evaluation surveys.

3.2. Creating a Multi-Stimuli Learning Scenario for ERP competence development

In the following section, the adapted SOR paradigm will now be applied to a specific learning scenario. The learning scenario utilizes the structure of the adapted SOR paradigm for teaching in tertiary education, introduced in chapter 3.1. In particular, we develop a set of stimuli to enable the best possible learning outcomes for students.

To this end, the course “Financial Accounting with SAP” and “Introduction to ERP software using SAP” are examined as practical examples. The first course is part of the 3rd semester curriculum of the Bachelor program in Business Informatics. The second course is part of 3rd semester curriculum of the Bachelor program in Business Administration. Both courses are taught at the University of Applied Sciences Niederrhein in Germany by the article’s authors. The set of stimuli is described by using a so-called “Stimuli Sedcard” that summarizes the applied stimuli in a structured way (Figure 2).

Stimuli Sedcard

Scenario Description A creditor payment business process is applied. All necessary steps, from adding the vendor master data and required General Ledger (G/L) accounts to processing the vendor invoice and its payment are performed. The effects on the balance sheet and income statement are analyzed after each step. Reference is made to Global Bike case study FI-AP, Ver. 4.2.	Competences to be developed. Students can understand the business processes related to accounts payable and identify involved stakeholders and their specific roles and tasks ... name the relevant SAP modules, navigate them and enter necessary data correctly ... adapt the applied case study to other accounts payable-related scenarios
Stimuli Setting Curricular Structure - Contextualization of the case study through a specific tangible learning scenario: Bike manufacturer Global Bike Inc. has no warehouse in the north-west of the USA leading to delayed delivery times from production sites to customers. A need for a new decentralize warehouse that is rented from a local real estate company is given - Students listen to a set of fictitious conversations (short audio podcasts) between involved employees from the bike manufacture and the real estate company to gather information about the required data and necessary steps to be executed in SAP S/4HANA - The related documents and related audio files are made available as an Open Educational Resource (OER) Technological Infrastructure - Full ERP environment is accessible via "SAP S/4HANA Fiori" web-based user interface - All required data of fictitious bicycle manufacturer "Global Bike" are available - Exercise from everywhere-opportunity based on VPN access - Web-based live feedback system "Tweedback" - Continuous learning assessment via tool "QuizAcademy" - Teaching Material and asynchronous communication and collaboration via teaching platform "Moodle" - Access to digital library via university infrastructure for ERP-related textbooks or papers	Teaching Methods - Meaningful didactic alternation of frontal teaching, individual work and group work to teach required theory foundations and practical usage of SAP S/4HANA - Multiple learning accounts for each student for repetitive training of this specific case study and related exercises - Bonus points to be achieved by successfully completing additional challenges are credited towards the exam grade as an incentive to work with the system - Learning assessments based on a quiz-format are carried out in student groups in a competitive setting in class (synchronous) or individually outside the class room (asynchronous) Social Environment - Self-perception of lecturer as learning facilitator that guides students stepwise through the case study - Using of web-based live feedback system to integrate introverted students by reducing threshold for participation and improve interaction between students and lecturer, where appropriate

Figure 2: Exemplary stimuli sedcard for creditor payment process in SAP S/4HANA

Scenario Description

The learning scenario focuses on executing a complete creditor payment business process within the SAP S/4HANA environment. The focus is on learning how to navigate and make use of the SAP financial accounting module (SAP-FI) that covers accounts payable, accounts receivable and asset accounting. Students start their learning journey by creating the necessary vendor master data and setting up relevant general ledger accounts required for the transaction. Afterwards, students process the incoming vendor invoice followed by executing the outgoing payment. Each step that impacts the income statement and balance sheet is analyzed allowing to explain implications on financial accounting. Each step is performed manually inside the system to ensure hands-on experience and reinforce conceptual understanding. For further details regarding the applied case study, reference is made to Global Bike case study FI-AP, vers. 4.2 (SAP University Competence Center Magdeburg n.d.).

Competences to be developed

Students should acquire a deeper understanding of the procure-to-pay business process. After completing the case study, they should be able to name the involved stakeholders, roles and organizational units and execute the payment process. At the same time, they should be able to explain related effects of the procure-to-pay process on the balance sheet and profit and loss statement. They should also understand which company-relevant decisions were made and why. Finally, students should be capable of transferring the acquired knowledge to other scenarios related to the process of accounts payable.

Curricular Structure

In order to create a tangible learning scenario, the underlying case study is subjected to intensive contextualization. In our case, the bicycle manufacturer Global Bike Inc. has no warehouse in the northwestern USA, which leads to delayed delivery times from production sites to the customers. There is a need for a new decentralized warehouse, which is rented from a local real estate company. A visual representation with various audio segments depicting fictitious conversations is used to present the scenario e.g., students listen to conversations between involved employees and functional units from the bike manufacturer itself and between the real estate company to gather information about the required data and necessary steps to be executed in SAP S/4HANA. The results from group discussions between students are presented via the massive live feedback system “Tweedback” (Tweedback n.d.), allowing to include a broader student body and a variety of student perspectives. The lecturers initially act more as moderators and encourage the students to work on questions about the scenario in groups. The scenario complexity is increased stepwise. This curricular structure aims to improve the understanding of the underlying business processes and associated execution steps within the ERP system.

Technological Infrastructure

Teaching on ERP systems that aims to build up competences in practical handling of such systems requires the availability of a fully configured hands-on training environment. Therefore, SAP S/4HANA is used, based on the Fiori web-interface. To ensure a close real-world learning scenario, a fictitious corporate, the bike manufacturing company “Global Bike” is made available for teaching and training purposes with a broad variety of preloaded data such as different customers, vendors, employees, manufactured products and the underlying organization with subsidiaries in the United States and Germany as well as the underlying SAP-related organizational structure. In order to ensure the accessibility from everywhere, in particular from outside the university network, an individual VPN access is provided to each student. It allows independent practice beyond scheduled in-class sessions. In addition, web-based tools such as “Tweedback” (Tweedback n.d.) or “QuizAcademy” (QuizAcademy n.d.) are provided to improve collaboration between lecturer and students and to provide group or individual learning assessments. All teaching materials are provided via “Moodle” learning platform (Moodle.org n.d.), which also serves as a space for asynchronous communication and collaboration i.e., to provide videos, slides, a course-related web forum or to organize group work. In addition, students also have access to the university’s digital library offering a wide range of ERP-related textbooks and further academic resources. This broad technological infrastructure ensures a reliable and supportive learning environment, encouraging self-paced learning, repetition and hands-on ERP system engagement.

Social Environment

We recommend creating a social environment where lecturers perceive themselves as facilitators and guide students step-by-step through the underlying scenario. Lecturers should enrich the underlying business process by adding a diverse set of real-world examples and engaging in active discussion with students.

Teaching Methods

Both the theoretical foundations and the practical application of SAP S/4HANA are conveyed through a mixed approach combining frontal teaching, individual tasks and group activities. In particular, each student is given multiple learning accounts to practice the case study and related exercises repeatedly. To encourage active engagement with the ERP system, students can earn

bonus points which are added to their final exam grade by successfully completing optional challenges with the system. In addition, learning assessments take the form of quizzes, either conducted synchronously in-class as group competitions or asynchronously as individual tasks outside the classroom based on the tool “QuizAcademy” (QuizAcademy n.d.).

4. Summary

The combination of a deliberately selected set of stimuli in relation to the categories presented within the adapted SOR paradigm can positively influence and support the development of competences in ERP systems. The combination of a set of stimuli in relation to the presented case study allows a first approach for the didactic structure of a general course program or a specific course unit. In the next steps, the presented concept should be adapted to other scenarios and case studies in an ERP-related context. However, this paper only operates within a theoretical-conceptual framework. Empirical evidence and causal links between applied learning stimuli and the learning outcomes of students are not yet investigated and can be part of further research.

This paper and related research is supported by “Le/Ni: Lehrprojektförderung 2024” of the University of Applied Science Niederrhein in Germany.

5. References

- Becerra-Fernandez, I., Murphy, K. E., & Simon, S. J. (2000): Enterprise Resource Planning: Integrating ERP in the Business School Curriculum. In: Communications of the ACM 43, 39-41.
- Gronwald, KD. (2025): ERP – Enterprise Resource Planning. In: Integrierte Business-Informationssysteme. Springer Vieweg, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-71017-3_3
- Mayer, C. W. F. (2022): Lehren und Lernen mit Enterprise Resource Planning (ERP) Systemen – Typische Bearbeitungsprobleme als Grundlage der Lernprozessgestaltung. In: bwp@ Berufs- und Wirtschaftspädagogik – online 43, 1-17.
- Mehrabian, A., & Russell, J. A. (1974): An Approach to Environmental Psychology. Cambridge, MA, USA: MIT Press.
- Moodle.org. (n.d.): Moodle – Open-source learning platform. Retrieved June 26, 2025, from <https://moodle.org/>
- Pavlov, I.P. (1927): Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex. London, UK: Oxford University Press.
- Pfarr, F., Klima, C. & Winkelmann, A. (2014): ERP-system-environments in IS education: Design and evaluation of a new course concept. In: Interactive Technology and Smart Education 11(2), 136–152. <https://doi.org/10.1108/ITSE-04-2014-0005>
- QuizAcademy. (n.d.): QuizAcademy – Lernsoftware für Quiz, Karteikarten & Umfragen. Retrieved June 26, 2025, from <https://quizacademy.de/>

- SAP University Competence Center Magdeburg. (n.d.): SAP Model Company Global Bike. Magdeburg, Germany: Otto-von-Guericke-Universität Magdeburg. Retrieved June 26, 2025, from <https://portal.ucc.ovgu.de/welcome/sap-model-company-global-bike/>
- SAP University Competence Center Magdeburg. (n.d.): The SAP University Competence Center Magdeburg. Magdeburg, Germany: Otto-von-Guericke-Universität Magdeburg. Retrieved June 26, 2025, from <https://portal.ucc.ovgu.de/>
- Sarferaz, S. (2023): ERP-Software: Funktionalität und Konzepte. Basierend auf SAP S/4HANA. Wiesbaden: Springer Vieweg. <https://doi.org/10.1007/978-3-658-40499-4>
- Schmidt, S. (2024): Teaching Analysis Polls. Berlin, Germany: wbv Publikation.
- Schumann, M., & Schattke, M. (2015): ERP-Systeme in der Hochschullehre: Ein integriertes E-Learning-Konzept zur Vermittlung praktischer ERP-Kompetenzen. In: HMD Praxis der Wirtschaftsinformatik 52, 425-438.
- Schwade, F. & Schubert, P. (2016): The ERP Challenge: An Integrated E-learning Platform for the Teaching of Practical ERP Skills in Universities. In: Procedia Computer Science 100, 147–155. <https://doi.org/10.1016/j.procs.2016.09.134>
- Tweedback. (n.d.). Tweedback. Retrieved June 26, 2025, from <https://tweedback.de/>
- Venkatesh, V. (2018): One-Size-Does-Not-Fit-All: Teaching MBA Students Different ERP Implementation Strategies. In: Journal of Information Systems Education 19(2), 141–146.
- Vieira, V. A. (2013): Stimuli–organism–response framework: A meta-analytic review in the store environment. In: Journal of Business Research 66(9), S. 1420-1426. <https://doi.org/10.1016/j.jbusres.2012.05.009>
- Watson, J. B. (1913): Psychology as the Behaviorist Views It. In: Psychological Review 20, S. 158–177.
- Woodworth, R. S. (1929): Dynamic Psychology. New York, NY: Columbia University Press. <https://ttu-ir.tdl.org/server/api/core/bitstreams/afd9fd96-e3f2-47a3-842f-b6c138b50c62/content>
- Zhang, G., Zhang, J., Zhang, X., Wang, M., & Yu, C. (2021): Understanding the impact of the psychological cognitive process on student learning satisfaction: Combination of the social cognitive career theory and SOR model. In: Frontiers in Psychology, 12, 712323. <https://doi.org/10.3389/fpsyg.2021.712323>
- Zhang, X., Wang, M., & Yu, C. (2022): Perceived environmental corporate social responsibility and employees' innovative behavior: A stimulus–organism–response perspective. In: Frontiers in Psychology, 12, 777657. <https://doi.org/10.3389/fpsyg.2021.777657>

Bridging Theory and Practice in ERP Education: A Didactic Toolbox for Sustainable SAP S/4HANA Teaching

N. Wolf, J. Stifel, & H. Wittges
Technical University of Munich, Munich, Germany

Abstract: Considering increasing demands for practice-oriented digital competencies, this study addresses the ongoing challenge of effectively teaching ERP systems in higher education. This article introduces a modular teaching concept centered around a didactic toolbox designed to foster interactivity and connect theoretical knowledge with practical applications using SAP S/4HANA. Grounded in design science research and enriched by qualitative input, the relevance and feasibility are confirmed by a focus group. Future research will evaluate the proposed concepts and explore gamification elements to further enhance learner engagement.

Keywords: Didactic Toolbox, Education, ERP System, Learning Concept, SAP

1. Introduction

Considering technological advancements, higher education is increasingly challenged to equip students with the competencies required by today's job market. Proficiency in enterprise resource planning (ERP) systems, particularly SAP S/4HANA, which integrates and manages core business processes, is essential (Soellner, 2021). Early, practice-oriented exposure to such technologies during education is therefore vital for a successful transition into professional life (Lea et al., 2022).

Nevertheless, the didactic communication of ERP system functionalities is challenged by their multidimensional complexity (Soellner, 2021). Students often lack prior knowledge of system operations and integrated business processes. Concepts such as horizontal integration frequently stay abstract without hands-on experience, which can hinder effective learning (Sweller, 2016).

To address these challenges, existing teaching approaches employ practice-based methods to reduce system complexity. For instance, Léger (2006) introduces a simulation game where students run a virtual enterprise using a real ERP system. Moreover, institutions such as SAP University Competence Center (UCC) support ERP education by providing a wide range of didactically structured learning resources, including material for SAP S/4HANA.

However, there remains a lack of a flexible, modular teaching concept that presents ERP content in an innovative, accessible, and engaging manner, while being adaptable to diverse educational contexts (Wienand et al., 2024). This article aims to address this gap by developing a modular didactic toolbox for lecturers teaching ERP systems, with the example of the SAP UCC S/4HANA learning material. Accordingly, the study addresses the following research question:

How can an innovative teaching concept for ERP systems be designed to enhance learning outcomes and student motivation?

The subsequent section 2 outlines the methodological framework. The preliminary findings are presented in section 3, including the didactic toolbox. An illustrative application scenario is

shown in section 4. The final section 5 offers an initial interpretation and outlines future research.

2. Methodology

The development of the didactic toolbox for innovative teaching of ERP systems follows a multi-stage process based on the design science research (DSR) methodology by Hevner et al. (2004). To identify relevant challenges, semi-structured interviews are conducted with three students and seven lecturers who utilize SAP UCC teaching materials. These interviews assess current materials, capture user experiences, and gather expectations for a new didactic concept. The insights inform the definition of solution objectives and the development of the didactic toolbox. Moreover, a role-based scenario is developed to illustrate the practical application of one selected method from the toolbox. An initial evaluation is conducted through a focus group involving six experts. A practical application in real teaching contexts is pending.

3. Didactic Toolbox for ERP Education

The didactic toolbox is a structured catalog of teaching methods to foster interactive, collaborative, and realistic ERP education. It is informed by user-centered research, combining qualitative interviews with theoretical input, particularly the ten design principles by Wienand et al. (2024). Interview findings highlight a strong need for modernized delivery formats that incorporate various multimedia tools, such as video content and audio materials. Greater interactivity is desired through gamification and engaging tasks that actively involve learners. Additionally, there is demand for content that brings theory closer to practice, such as real-world examples and simulations.

Didactic Toolbox				
Multimedia-based Learning	Gamified and Motivational Learning	Collaborative Learning	Industry-integrated Learning	Technology-enhanced Learning
(1) Podcast	(5) Synchronous Quizzes	(11) Peer Explanation Task	(13) Conversation with Industry Experts	(15) Practice App
(2) Case Study Walkthrough Videos	(6) Easter Egg Discovery Tasks	(12) Role-based Scenario	(14) Real-world Company Challenges	(16) Embedded Instruction Steps in UI
(3) Learning Snacks (Microlearning Tutorials)	(7) Gamified Tutorial Tasks			(17) Chatbot
(4) Visual Enhancements	(8) Case Study Progress Leaderboard			
	(9) Business Process Card Game			
	(10) ERP Business Simulation Game			

Fig. 1: Didactic Toolbox Methods for ERP Education

Figure 1 presents the developed method catalog, which includes 17 distinct methods categorized into five instructional design dimensions. The toolbox offers lecturers a set of practical and adaptable approaches that can be integrated into existing course formats with ease. It is suitable for

diverse teaching scenarios, ranging from small group settings to large lecture environments. Depending on the pedagogical goals and available resources, methods can be flexibly selected and

Attribute	Attribute Value	Example
Method Category	General classification of the method	Collaborative learning
ID	Sequential identifier	12
Method Name	Title of the method	Role-based scenario
Method Description	Brief explanation of the method's core idea	Students perform tasks in SAP S/4HANA from assigned business roles, simulating real organizational scenarios
Learning Objectives (Bloom)	Cognitive level targeted by the method, based on Bloom's taxonomy (remember, understand, apply, analyze, evaluate, or create)	Apply + Analyze
Learning Modalities (VARK)	Main sensory channel used in the learning process, based on the VARK learning model (auditory, visual, read-write, or kinesthetic)	Visual + Auditory + Read-write + Kinesthetic
Cognitive Engagement (ICAP)	Expected level of learner involvement according to the ICAP framework (passive, active, constructive, or interactive)	Interactive + Constructive
Participant Format	Individual work or group collaboration	Group collaboration
Delivery Mode	Setting in which the method is applied (in-class, out-of-class, or mixed)	In-class
Student Effort	Estimated completion time for students	Medium
Lecturer Effort (Initial Setup)	Lecturer preparation time	Medium to high (scenario design)
Lecturer Effort (Reuse)	Estimated time required for subsequent uses of the method after initial setup	Medium
Setup Requirements	Required or existing resources	Setup effort requires designing role-play scenarios based on SAP case studies (e.g., the sales and distribution module example presented in section 4)
Applied Example in SAP S/4HANA	Example in SAP S/4HANA teaching	Students take on roles (e.g., sales) and execute full process cycles, like creating and processing a sales order

Tab. 1: Didactic Toolbox Method Attributes

combined. The catalog spans from basic tools such as quizzes to more advanced formats like chatbot-based projects and simulation-based exercises. Designed to complement the SAP UCC learning materials for SAP S/4HANA, the toolbox enables straightforward integration with minimal modifications. To support effective implementation, supplementary sheets are provided to assist lecturers in understanding and applying the methods appropriately.

Each method is systematically described using a predefined set of attributes, including classification according to three learning frameworks: Bloom's Taxonomy (Bloom, 1956), VARK (Fleming & Mills, 1992), and ICAP (Chi & Wylie, 2014). These frameworks offer complementary perspectives on cognitive processes, sensory modalities, and levels of learner engagement. Together, they serve as a foundation for selecting or combining methods aligned with specific course objectives and learning outcomes. Table 1 presents the predefined attributes for each method, their description, and an example of one specific method.

4. Application Example of the Didactic Toolbox: Role-based Scenario

A role-based scenario based on the SAP UCC S/4HANA sales and distribution (SD) case study demonstrates the application of one toolbox method. Students take on roles within a fictional model company to simulate a real business process, fostering

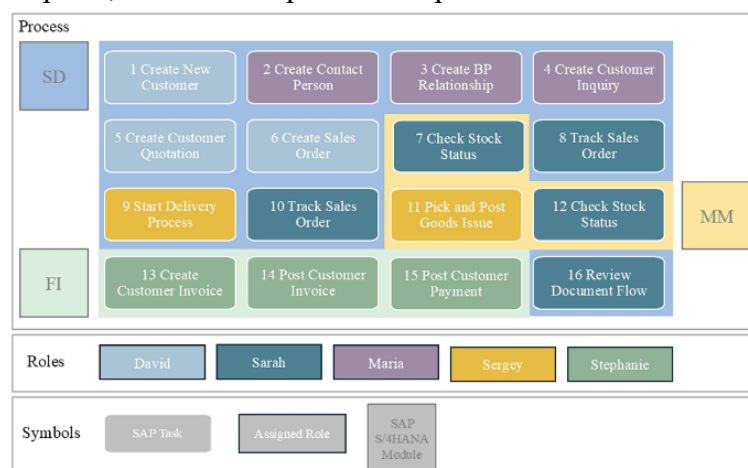


Fig. 2: Process Graphic for Role-based Scenario Example

engagement and understanding of the SD module and its integration. The scenario includes six structured sheets for lecturers and students, guiding preparation, role assignment, execution, and reflection. A central process graphic (Figure 2) supports progress tracking. The method promotes higher-order thinking (Bloom's Taxonomy) and addresses all VARK learning styles. It integrates easily into existing SAP UCC courses and is adaptable to other modules with minimal adjustments.

5. Discussion and Plan for Future Research

This article introduces a didactic toolbox to support sustainable ERP education, using the SAP UCC as a practical example. Grounded in educational theory, the toolbox supports instructors in selecting and combining effective teaching methods. To demonstrate its practical application, a role-based scenario is developed, which enriches SAP UCC materials by introducing interactivity and real-world relevance. A focus group validated the importance and feasibility of the proposed concepts, providing an initial positive response to the research question. Combining qualitative insights with the DSR approach, the article provides actionable strategies for active ERP learning and expands prior work (e.g., Léger, (2006)) through a broader methodological framework.

Future research will focus on further developing the teaching approaches contained in the toolbox and transferring them into concrete units. Additionally, the concepts will be empirically tested in real-world teaching contexts and assessed for effectiveness. The developed role-based scenario will also be validated through pilot studies and refined iteratively. Furthermore, the use of gamification elements, such as narrative structures or randomization, will be explored to enhance learner motivation and immersion. The toolbox's modular structure enables continuous expansion to include new didactic methods and technologies, which subsequent research will also address.

The communication of interim results aims to involve the target audience at an early stage to foster scientific discourse on ERP teaching and identify potential partners for pilot implementations.

Associated content: The developed teaching material is available from the authors upon request.

6. References

- Bloom, B. S. (1956). Taxonomy of educational objectives: The classification of educational goals. Handbook 1: Cognitive domain. New York: Logman.
- Chi, M. T. & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, 49(4).
- Fleming, N. D. & Mills, C. (1992). Not another inventory, rather a catalyst for reflection. *To Improve the Academy: A Journal of Educational Development*, 11(1).
- Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *Management Information Systems Quarterly*, 28(1).

- Lea, B.-R., Mirchandani, D., Sumner, M., & Yu, K. (2022). Personality types in learning enterprise resource planning (ERP) systems. *Journal of Computer Information Systems*, 62(2).
- Léger, P.-M. (2006). Using a simulation game approach to teach enterprise resource planning concepts. *Journal of Information Systems Education*, 17(4).
- Soellner, S. (2021). Digital elements for SAP ERP education and training: Results from a systematic literature review. *International Journal of Engineering Pedagogy*, 11(4).
- Sweller, J. (2016). Working memory, long-term memory, and instructional design. *Journal of Applied Research in Memory and Cognition*, 5(4).
- Wienand, M., Wulfert, T., & Hoang, H. (2024). Design principles for e-learning platforms featuring higher-education students' enterprise systems end-user training. *Discover Education*, 3(82).

Towards a Next-Generation Data Center Curriculum: Teaching AI Hosting Skills on IBM Power

Noah Kim¹, Maximilian Niedermeier¹, Dilser Ünsal¹, Simon Fuchs¹, Holger Wittges¹

¹ TUM School of Computation, Information and Technology, Garching, Germany

Abstract: Artificial Intelligence (AI) has rapidly advanced from theory to practical applications, driven by generative techniques and large language models (LLMs). As AI becomes more prevalent across industries, there is an urgent need to teach AI skills using state-of-the-art solutions. While existing educational research focuses on using and engaging with AI, it overlooks the critical perspective of hosting AI models. This gap is significant, as hosting AI requires new operational skills for computing centers. Against this context, this paper proposes an AI hosting curriculum to teach the specifications and associated skills for AI computing center operations. We describe the implementation, concepts, and contents of said curriculum. Further, we present the evaluations of the curriculum obtained from two teaching sessions with students of diverse backgrounds. Lastly, we conclude that the developed curriculum is suitable for teaching AI operation and hosting skills to students of various backgrounds.

Keywords: Curriculum Design, Higher Education, Hybrid Cloud, LLM, AI

1. Introduction

Artificial Intelligence (AI) solutions have evolved from a theoretical concept into real-world technology applicable to various business and industry use cases (Bharadiya et al. 2023). Drivers of this evolution of AI models are generative techniques, training large deep neural networks to generate new outputs instead of only predicting or classifying results (Bandi et al. 2023; Bansal 2024). The most prominent generative AI models, made acquainted by public chatbots like ChatGPT (OpenAI 2023) or Perplexity.AI (PerplexityAI 2023), are Large Language Models (LLMs) – AI models for Natural Language Processing (NLP) generating Natural Language texts as a response to textual input (Radford et al. 2019).

With the exponential growth of interest in AI applications and their growing saturation across industry use cases, it becomes increasingly urgent to teach AI-related skills and practical application knowledge with state-of-the-art AI solutions (Slade et al. 2025; Ng et al. 2021b; AlKanaan 2022). However, most AI-related education research either focuses on developing skills to engage AI-generated content (AlAfnan et al. 2024; Na-songkhla et al. 2024; Ng et al. 2021a), how to improve education using AI (Neha 2020; Al-Tkhayneh et al. 2023), or how to optimally teach learners to use AI solutions (Barrett/Pack 2023; Kim et al. 2021). However, contemporary AI-related educational research omits the hosting perspective on AI, creating a vital gap in skills and perspective taught.

Hosting AI models and related services significantly changes the needs and requirements in computing centers and deeply affects hosting providers, both hyperscalers and on-premise hosting facilities (Mavani et al. 2024; Vankayalapati 2022). With this, hosting center operations require new skills from their technical staff, while simultaneously these skills will determine promotions and employment (George et al. 2023; Tschang/Almirall 2021). However, the

broader IS-teaching literature currently neglects teaching these AI-related hosting center operational skills.

Against this context, we present an AI hosting curriculum to teach the specifications and associated skills, further adding to a broader computing center operations curriculum (Niedermeier et al. 2024). First, we expound on the course requisites that students should bring to attend. Afterward, we explain the relevant concepts taught in the curriculum and present the course's content. Then, we describe the technical implementation of the landscape used for hosting the course. After that, we present the evaluation of two sessions in which we taught the course to diverse students. Lastly, we reflect on the implications and contribution of the course presented.

2. Course Prerequisites

Our new module about teaching AI hosting expands an existing module catalogue, with all modules collectively focusing on how to teach system administration on IBM Power servers. To be more precise, the existing modules aim to teach students practically how to deploy SAP HANA database systems using IBM hardware, Red Hat Linux, and both IBM and Red Hat virtualization concepts. In a previous publication (Niedermeier et al. 2024), we (1) summarized the topics of the existing curriculum modules in more detail, (2) explained how to implement the container-based teaching environment, and (3) presented the experiences of a very first test run in April 2024.

When starting to engage with our new AI hosting contents, it is important that the students have already learned about containers, both from a theoretical and a practical point of view. Thus, we teach containerization topics in one of the modules preceding the new module, which we present in this publication. We place such high value on learners already having container knowledge because our teaching infrastructure is built on a single logical partition (LPAR). However, the students build containers within this LPAR, each container being an isolated environment for hosting an LLM. This approach is further described in Section 5.

3. Relevant Concepts

3.1. LLMOps

This module displays how Power Servers can effectively host Large Language Models (LLMs) for various applications. This process is known as LLMOps, short for Large Language Model Operations. LLMOps encompasses the specialized practices and workflows that streamline the development, deployment, and management of LLMs across their entire lifecycle (Diaz-De-Arcaya et al. 2024). This course module focuses primarily on selecting and deploying LLM on Power Servers. Fine-tuning is not yet part of this module.

3.2. Granite LLM

For the LLM itself, we use the IBM Granite model⁹. Granite is a family of open-source AI models designed explicitly for business-related tasks (Granite Team 2024). The latest version,

⁹ <https://www.ibm.com/de-de/granite>

Granite 3.3, supports multiple languages, can write code, and integrates Retrieval-Augmented Generation (RAG). One of Granite's key advantages is its open-source availability on platforms like Hugging Face. This fosters a growing community that shares their fine-tuned versions of Granite, resulting in a diverse ecosystem of specialized models tailored for various tasks (Granite Team 2024).

3.3. InstructLab

InstructLab¹⁰ is an open-source framework developed by IBM in collaboration with Red Hat. It is designed to help developers and organizations efficiently deploy and fine-tune LLMs (RedHat 2024). InstructLab simplifies the customization of open LLMs—such as Granite—through easy-to-use ILab command-line tools, eliminating the need for deep machine learning expertise. It supports direct interaction with models via chat and the setup of a web server to serve the LLM through an API (RedHat 2024).

3.4. GGUF

GGUF (Generic GPT Unified Format) is a specialized model storage format designed to efficiently manage quantized LLMs (Kansal 2024). It allows GPT-based models to be compressed and deployed on CPUs, making them more accessible for a broader range of hardware (Kansal 2024). Quantization refers to the process of reducing model size by converting weights to lower-precision formats, such as 8-bit (Q8), 5-bit (Q5), or even 3-bit (Q3). This significantly reduces memory usage and speeds up inference, while maintaining minimal loss in accuracy. In this course, the LLM models are retrieved as GGUF files from Hugging Face and run using llama.cpp through InstructLab (RedHat 2024).

A typical GGUF model file name looks like this:

granite-3.1-8b-instruct-Q5_K_L.gguf

Here is how to interpret it:

- granite-3.1: Indicates the model family (Granite) and version (3.1).
- 8b: Refers to the model size—8 billion parameters.
- instruct: This signifies that the model has been fine-tuned for instruction-following tasks, resulting in more precise and concise responses.
- Q5_K_L: Specifies that the model has been quantized to 5-bit precision per weight using a particular quantization scheme.

4. Course Content

The overall structure of the **LLMOps module** is outlined in **Table 1**. The first Section, Basics, introduces all relevant background knowledge. This includes the LLMOps lifecycle, the InstructLab framework, and the IBM Granite model. It also covers the GGUF model format and the benefits of open-source and locally hosted LLMs. Participants gain foundational theoretical knowledge about LLM hosting.

¹⁰ <https://www.ibm.com/think/topics/instructlab>

Part	Type	Content
Basics	Background Knowledge	What is LLMOps? What is InstructLab and Granite? What is GGUF? What are the prerequisites for a local LLM?
Deploying LLM	Live coding, Hands-On	How to install InstructLab on Power? How to implement open-source LLMs into InstructLab? How to test LLM with different settings? What are zero-shot experiments and temperature?
LLM Use Cases	Group Work, Hands-On	How to connect via the OpenAI API? How to use RAG with InstructLab? How to use AnythingLLM?

Table 1: Outline of the LLMOps Module

Following the theory, the instructor conducts a live coding session demonstrating the deployment of a large language model using InstructLab. The session begins with installing InstructLab on IBM Power Servers and retrieving and running an LLM from the open-source platform Hugging Face (Jain 2022). Participants follow along step-by-step on their machines. Once the model is deployed and running, participants interact with it directly via the terminal. They experiment with zero-shot prompting and adjust temperature settings to explore how different configurations influence the model's responses.

In the final stage, participants connect their deployed LLM to AnythingLLM (Mintplex-Labs 2024), enabling more user-friendly interaction. They begin experimenting with basic Retrieval-Augmented Generation (RAG) techniques using simple CSV files. For example, a CSV containing product sales data allows the LLM to answer questions like “What is the best-selling product?”

5. Implementation

The SAP UCC Munich utilizes IBM Power10 Servers in a server room at the Technical University of Munich for this module. Using the virtualization tool PowerVC¹¹, we created a dedicated workshop LPAR running on Red Hat Enterprise Linux 9.5. We gave special attention to allocating sufficient RAM and virtual CPU cores to ensure optimal performance throughout the course. Within this environment, we use *Podman* (Gantikow et al. 2020) to create individual containers—one for each student group. This containerized setup allows for isolated, parallel model deployment and experimentation.

Each participant or participant group is assigned a unique user account on the LPAR. After logging in, they create their Podman container, each assigned a unique port number. The container is launched using the `-ti` flag to enable an interactive terminal and the `-d` flag to run it in detached mode, which ensures that the container continues to run in the background even after the user logs out.

We based the container on a publicly available image from quay.io¹², including InstructLab and all necessary dependencies. Two folders from the participant's home directory are mounted into

¹¹ <https://www.ibm.com/de-de/products/powervc>

¹² <https://quay.io/repository/mgiessing/ilab?tab=tags&tag=v0.23.1>

the container: one called "models", which contains the downloaded LLMs in GGUF format retrieved from Hugging Face, and another called "taxonomy", which stores knowledge files used for fine-tuning models in future modules.

Once the container runs, participants access its terminal and use `ilab` commands to bind the appropriate GGUF file and launch the LLM. After the configuration, the students can open a chat terminal within the container for initial prompt-based experiments. InstructLab uses `llama.cpp` as the backend for running GGUF models, enabling CPU-only inference without needing GPUs.

To make the LLM accessible to external applications like AnythingLLM, the participant uses the `ilab serve` command, which starts a local HTTP server. This server is compatible with the OpenAI API (Lim et al. 2021) and is accessible through the container's assigned port. In AnythingLLM, the "Generic OpenAI API" option establishes the connection. No API key is required, provided AnythingLLM's computer is on the same network as the LPAR. The Base URL follows this format: `http://<Host-URL>:<Container-Port>/v1`

Once connected to AnythingLLM, participants can create personal workspaces and begin using Retrieval-Augmented Generation (RAG) (Arefeen et al. 2024). Documents and links are stored locally on the participant's computer and integrated into the workspace, allowing the LLM to answer questions using the additional contextual information provided.

6. Technical Evaluation

Until now, we have held the course described in Section 4 two times. The first session (session A) took place on 6 May 2025. It was attended by two master's students in Management and Technology, 22 master's students in Information Systems, and one bachelor's student, also studying Information Systems. In this session, two students had to leave the course early. Also, the students worked in groups, using 10 containers. Additionally, the lecturer used one container for presentation purposes.

The second session (session B) was held on 30 June 2025. It was attended by a Physics master's graduate, a Computer Science master's graduate, a Computer Science master's student, three Computer Science bachelor's students, and two Information Systems bachelor's students. This time, participants deployed their container. Again, the lecturer used one additional container for presentation purposes.

We used an LPAR deployed on an IBM Power E1050 server in both sessions. This LPAR was provided 39 dedicated virtual cores, 800 Gibibyte (~859 gigabytes) of random-access memory (RAM), and 700 Gibibyte (~752 gigabytes) of disk space. During both sessions, we monitored the percentage RAM usage, the percentage disk usage, and the percentage CPU usage of this LPAR. While Figure 3 shows the technical evaluation of session A, Figure 4 shows the technical evaluation of session B.

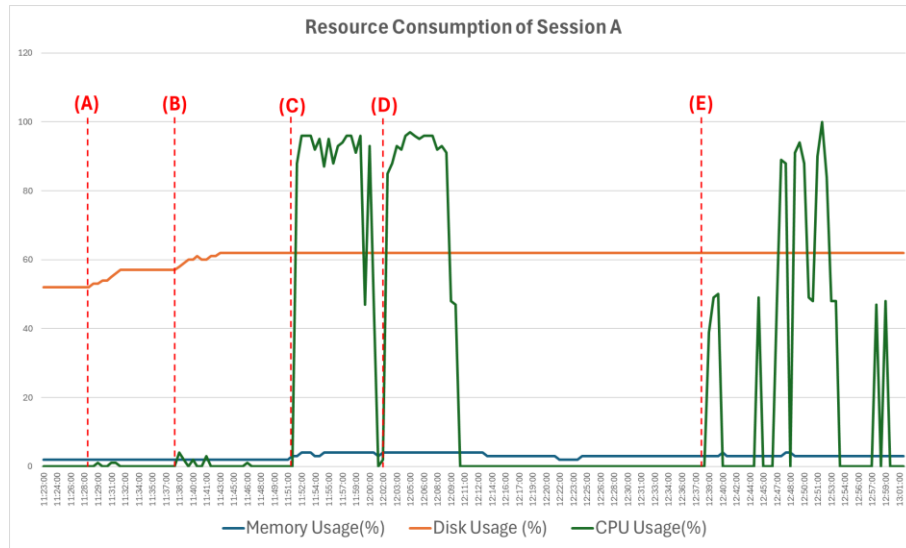


Figure 3: The memory, disk and CPU usage during session A

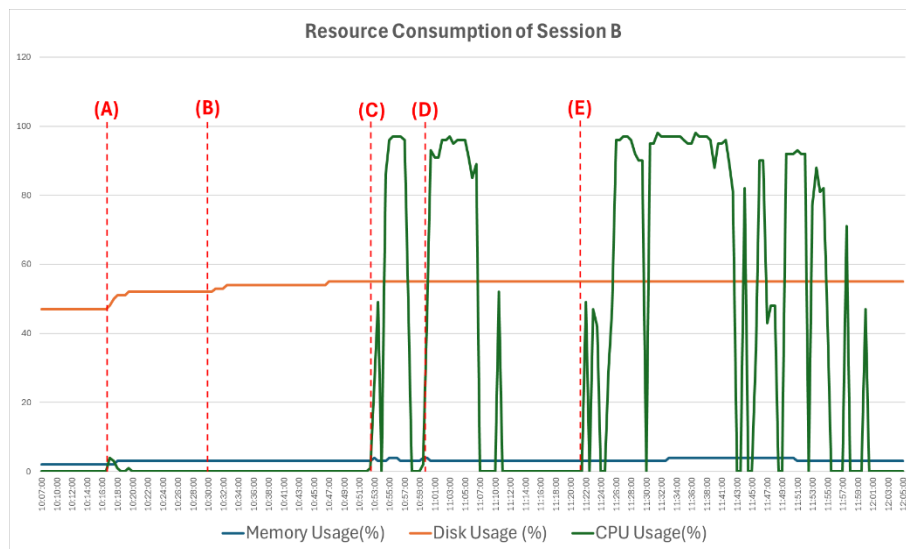


Figure 4: The memory, disk and CPU usage during session B

Below, we explain which events in the curriculum module we can link to the timestamps marked:

- (A): Students create the container.
- (B): Students download the GGUF file.
- (C): Students start chatting with their self-hosted LLMs.
- (D): Students continue chatting with adjusted temperature settings.
- (E): Students serve the model and continue chatting via the OpenAI API and AnythingLLM.

The monitored values show three important findings: First, during each session, the used memory increased by 2% (~17,18 gigabytes). Compared to other modules, where memory-intensive operations are executed, **the amount of memory used in this curriculum module is negligible**. Secondly, the disk space used in session A increased by 10% (~75,2 gigabytes), and the disk space used in session B increased by 8% (~60,16 gigabytes). In session A, 11 containers

were deployed, and in session B, nine containers were deployed. On average, **each container used approximately 6.76 gigabytes of disk space**. Lastly, even if almost 40 virtual cores were used, **the CPU is fully utilized during inference tasks, which results in very high response times when chatting with the self-deployed LLMs**.

7. Qualitative Evaluation

Following session B, we conducted a focus group that included all but one of the students and three instructors. The discussion was recorded and systematically analyzed by identifying key statements and summarizing them into core concepts (Bogner 2009).

A central theme that emerged was the students' concern regarding performance issues with their self-hosted LLMs. Such performance issues encompassed both the quality of answers and the response times. As the response time is directly linked to the CPU resource utilization of the used LPAR, one participant suggested showing a live visualization of the CPU resource consumption – maybe even the electricity consumption. This could help future students to better understand the cause of delays and foster greater tolerance for slower response times.

Also, the performance was compared to ChatGPT, and the question arose why one would not simply use the cloud-based model. Another participant answered that it is important to stress the advantages of hosting an LLM locally. Then, another participant stated that incorporating a use case in which the LLM processes sensitive data would be a good idea. Regardless of the use case topic, it was important to create practice exercises that would not confuse the small model and could be completed by future students.

Next, the use of AnythingLLM, especially the opportunity to use an API that connects the self-hosted LLM with a front-end, was generally approved. It was considered a good showcase that an LLM could read from multiple provided CSV files simultaneously.

Participants also liked the simplicity of creating their own LLM and said understanding where to find the GGUF files was essential to the module. One participant stressed the importance of explaining different infrastructure setups and mentioned that they might still struggle deploying an LLM on another machine, such as their computer. Another participant answered that this was precisely why we used containerization and that the container would run on any infrastructure that can deploy Podman. In reality, containers depend on the server's host architecture, and we realized we need to introduce further explanations of container compatibility.

8. Discussion

First and foremost, we conclude that our developed curriculum is suitable for teaching the concepts described in Section 3 to diverse students. With that, we regard it as a valuable contribution to the larger discussion of how to teach AI-relevant skills, especially in AI operations.

Although technically feasible on IBM Power servers and found interesting by the focus group participants, the inference performance still presents a challenge when aiming for a proper deployment of the new curriculum module. It is an open question whether a different infrastructure set-up – e.g., using many LPARs instead of one large machine – would improve

the response delays. Thus, we aim to repeat the course with different infrastructure configurations and then compare the CPU utilization with the results shown in Figures 3 and 4.

As mentioned in Section 2, the other curriculum modules teach students how to deploy an SAP HANA database system on IBM Power hardware. We believe the newly developed module should not be treated as a standalone topic but rather be integrated into the existing curriculum. For example, a potential integration could involve building a pipeline that enables students to retrieve data using natural language queries. In the future, we thus plan to explore the integration of SAP HANA with IBM Granite, allowing students to directly benefit from a self-hosted LLM.

9. Conclusion

Artificial Intelligence (AI) has rapidly evolved from theoretical to practical technology with widespread business and industry applications. As AI increasingly integrates into various sectors, there is a growing need to teach AI-related skills and practical knowledge using the latest AI solutions. Current educational research in AI typically focuses on developing skills to interact with AI-generated content, leveraging AI to improve education, and teaching learners how to use AI tools effectively. However, these approaches overlook the critical perspective of hosting AI models and services.

Against this context, we developed an AI hosting curriculum, adding to a previous SAP hosting curriculum, teaching critical skills of hosting AI in an IBM Power-based server architecture. This course starts with an LLM Ops module, providing foundational theory on LLM hosting, covering lifecycles, frameworks, and model formats. The course afterward proceeds with a live coding session to deploy and interact with an LLM, experimenting with prompts and settings. Closing, the students connect the installed Granite model to AnythingLLM and practice basic Retrieval-Augmented Generation (RAG) using CSV data.

We evaluated our course in two sessions with students from various fields, including Physics, Computer Science, and Information Systems. Both sessions involved deploying LLMs in containers on an IBM Power E1050 server, utilizing significant resources but negligible memory per container. The evaluation showed the teaching success of our course while highlighting concerns about performance, especially compared to cloud models like ChatGPT; however, it recognized the value of local hosting of LLMs in case of sensitive data. The students appreciated the module's practical aspects, such as using AnythingLLM and containerization, but suggested more precise explanations of infrastructure and container compatibility.

10. Acknowledgments

We want to thank the IBM managers Alexander Loibl, Martin Mähler, and Thomas Heinlein, who made this curriculum possible. Further, we want to thank the IBM employees, Wolfgang Rother, Michael Perzl, and Stephen Lutz, who helped set up the server landscape and contributed variously with their noticeable expertise.

11.Disclaimer of AI

We used Perplexity.AI (PerplexityAI 2023) to help summarize, reformulate, and condense the Abstract and Conclusion sections. Further, we used Grammarly to correct our Spelling, Style, and Grammar.

12.References

- Al-Tkhayneh, K.M.; Alghazo, E.M.; Tahat, D. (2023): The advantages and disadvantages of using artificial intelligence in education. In, (2023).
- AlAfnan, M.A.; Dishari, S.; MohdZuki, S.F. (2024): Developing soft skills in the artificial intelligence era: Communication, business writing, and composition skills. In: Journal of Artificial Intelligence and Technology, Vol. 4 (2024) No. 4, pp. 305-317.
- AlKanaan, H.M.N. (2022): Awareness Regarding the Implication of Artificial Intelligence in Science Education among Pre-Service Science Teachers. In: International Journal of Instruction, Vol. 15 (2022) No. 3, pp. 895-912.
- Arefeen, M.A.; Debnath, B.; Chakradhar, S. (2024): Leancontext: Cost-efficient domain-specific question answering using llms. In: Natural Language Processing Journal, Vol. 7 (2024), pp. 100065.
- Bandi, A.; Adapa, P.V.S.R.; Kuchi, Y.E.V.P.K. (2023): The power of generative ai: A review of requirements, models, input–output formats, evaluation metrics, and challenges. In: Future Internet, Vol. 15 (2023) No. 8, pp. 260.
- Bansal, P. (2024): Prompt Engineering Importance and Applicability with Generative AI. In: Journal of Computer and Communications, Vol. 12 (2024) No. 10, pp. 14-23.
- Barrett, A.; Pack, A. (2023): Not quite eye to AI: Student and teacher perspectives on the use of generative artificial intelligence in the writing process. In: International Journal of Educational Technology in Higher Education, Vol. 20 (2023) No. 1, pp. 59.
- Bharadiya, J.P.; Thomas, R.K.; Ahmed, F. (2023): Rise of artificial intelligence in business and industry. In: Journal of Engineering Research and Reports, Vol. 25 (2023) No. 3, pp. 85-103.
- Bogner, A. (2009): Interviewing experts, Palgrave Macmillan 2009.
- Diaz-De-Arcaya, J.; López-De-Armentia, J.; Miñón, R.; Ojanguren, I.L.; Torre-Bastida, A.I. (2024): Large Language Model Operations (LLMOps): Definition, Challenges, and Lifecycle Management. Paper presented at the 2024 9th International Conference on Smart and Sustainable Technologies (SpliTech), pp. 1-4.
- Gantikow, H.; Walter, S.; Reich, C. (2020): Rootless containers with Podman for HPC. Paper presented at the International Conference on High Performance Computing, pp. 343-354.
- George, A.S.; Sagayarajan, S.; AlMatroudi, Y.; George, A.H. (2023): The Impact of Cloud Hosting Solutions on IT Jobs: Winners and Losers in the Cloud Era. In: Partners Universal International Research Journal, Vol. 2 (2023) No. 3, pp. 1-19.
- Granite Team, I. (2024): Granite 3.0 Language Models.

- Jain, S.M. (2022): Hugging Face. In: Introduction to Transformers for NLP: With the Hugging Face Library and Models to Solve Problems. Eds. Apress, Berkeley, CA 2022, pp. 51-67.
- Kansal, V. (2024): Understanding the GGUF Format: A Comprehensive Guide. <https://medium.com/@vimalkansal/understanding-the-gguf-format-a-comprehensive-guide-67de48848256>, accessed at 03.07.2025.
- Kim, S.; Jang, Y.; Kim, W.; Choi, S.; Jung, H.; Kim, S.; Kim, H. (2021): Why and what to teach: AI curriculum for elementary school. Paper presented at the proceedings of the AAAI Conference on Artificial Intelligence, pp. 15569-15576.
- Lim, R.; Wu, M.; Miller, L. (2021): Customizing GPT-3 for your application. <https://openai.com/index/customizing-gpt-3/>, accessed at 20.05.2024.
- Mavani, C.; Mistry, H.K.; Patel, R.; Goswami, A. (2024): Artificial Intelligence (AI) Based Data Center Networking. In: International Journal on Recent and Innovation Trends in Computing and Communication, Vol. 12 (2024) No. 2, pp. 508-18.
- Mintplex-Labs (2024): Anything-LLM <https://github.com/Mintplex-Labs/>, accessed at 03.07.2025.
- Na-songkhla, J.; Mahakaew, V.; Peytcheva-Forsyth, R. (2024): The Emergence of Generative Artificial Intelligence: Enhancing Critical Thinking Skills in ChatGPT Integrated Cognitive Flexibility Approach. In: Generative Artificial Intelligence in Higher Education: A Handbook for Educational Leaders, Vol. 52 (2024).
- Neha, K. (2020): Role of artificial intelligence in education. In: Alochana Chakra Journal, Vol. 9 (2020) No. 9, pp. 305-309.
- Ng, D.T.K.; Leung, J.K.L.; Chu, K.W.S.; Qiao, M.S. (2021a): AI literacy: Definition, teaching, evaluation and ethical issues. In: Proceedings of the association for information science and technology, Vol. 58 (2021a) No. 1, pp. 504-509.
- Ng, D.T.K.; Leung, J.K.L.; Chu, S.K.W.; Qiao, M.S. (2021b): Conceptualizing AI literacy: An exploratory review. In: Computers and Education: Artificial Intelligence, Vol. 2 (2021b), pp. 100041.
- Niedermeier, M.; Rother, W.; Ünsal, D.; Wittges, H. (2024): Towards a Next-Generation Data Center Curriculum: Utilizing Containerization for Teaching SAP HANA on IBM Power. Paper presented at the SAP Academic Community Conference 2024 (DA-CH).
- OpenAI (2023): ChatGPT. Mar 14 version, <https://chat.openai.com/chat>, accessed at 20.08.2025.
- PerplexityAI (2023): Perplexity [Large language model]. <https://www.perplexity.ai>, accessed at 20.08.2025.
- Radford, A.; Wu, J.; Child, R.; Luan, D.; Amodei, D.; Sutskever, I. (2019): Language models are unsupervised multitask learners. In: OpenAI blog, Vol. 1 (2019) No. 8, pp. 9.
- RedHat (2024): What is InstructLab? <https://www.redhat.com/en/topics/ai/what-is-instructlab>, accessed at 03.07.2025.
- Slade, J.J.; Byers, S.M.; Becker-Blease, K.A.; Gurung, R.A. (2025): Navigating the new frontier: Recommendations to address the crisis and potential of AI in the classroom. In: Teaching of Psychology, Vol. 52 (2025) No. 3, pp. 254-261.

Tschang, F.T.; Almirall, E. (2021): Artificial intelligence as augmenting automation: Implications for employment. In: Academy of Management Perspectives, Vol. 35 (2021) No. 4, pp. 642-659.

Vankayalapati, R.K. (2022): AI Clusters and Elastic Capacity Management: Designing Systems for Diverse Computational Demands. In: Available at SSRN 5115889, (2022).

Effective teaching with integrated analytics case studies & Tools

Nuran Fazlioglu-Adal

Hamburg University of Applied Sciences, Germany

Abstract This contribution introduces a practice-oriented teaching approach that integrates SAP Analytics Cloud (SAC), the Business Builders gamification platform, and case-based learning to develop students' data literacy and business intelligence (BI) competencies. Delivered within the "Management Reporting and Business Intelligence" module of an International Business Master's program at Hamburg University of Applied Sciences, the course blends theoretical foundations with applied learning through interactive elements such as dashboard design, the extended "Global Bike" case study, and the Business Builders: International Expansion game. Students engage with real-world analytics scenarios, make strategic design decisions, and critically reflect on their outcomes. The learning experience is further enriched by self-directed study using the SAP Learning Student Zone and lecture content/videos generated using AI tools. This didactic model illustrates how digital tools can effectively bridge the gap between theory and practice, supporting a transformative, competency-based approach to SAP and BI education.

Teaching Enterprise Software Engineering in the Cloud – An Exemplary Case

Philipp Landler, Alina Hafner, Holger Wittges
Technical University of Munich, Germany

Abstract: Teaching software engineering changes when relying on cloud infrastructure. Updates to services beyond the control of lecturers impact planned sessions and hinder students from executing their tasks. In addition, the shift from self-hosted models to subscription-based services challenges established paradigms of system resource access. Especially in the complex area of enterprise software, lecturers can lose track of which capabilities are covered in which service.

This presentation will illustrate how a university lab course uses an existing partnership to face these challenges. First, we describe the general setup of the course. Second, we discuss our challenges when teaching enterprise software engineering in the cloud and present our solutions. Third, we give an outlook on how the course and the curricula will change in the future.

Keywords: enterprise software engineering, cloud teaching, partnerships.

Experiences with WatsonX Integration in Higher Education: Lessons from a TUM Course

F. Eckhard, N. Wolf, H. Wittges
Technical University of Munich, Munich, Germany

Abstract The adoption of enterprise artificial intelligence platforms in higher education requires careful examination of the tools' scope, pedagogical considerations, and practical outcomes for students. This presentation outlines a case study of IBM WatsonX integration within a master's course at the Technical University of Munich (TUM), documenting the experiences, challenges, and insights gained from incorporating enterprise AI technology into academic teaching. The course covers various WatsonX functionalities across data preparation, classical machine learning, and generative AI. This enables students to gain practical experience in this topic with a leading enterprise offering, especially the integration with further tools like SAP S/4HANA and ERPsim show real-world use cases. Based on the integration, the course shows the opportunities of analyzing enterprise data from a multidimensional perspective. Challenges like cost management, cloud administration, and bridging the gap between academic and enterprise requirements for such tools are also covered. Finally, an outlook for the integration of further WatsonX services in our novel curriculum and how this curriculum can be adapted in other academic environments is given.

TRACK 5: INTERNATIONALISIERUNG

Track Chair: Prof. Dr. Jorge Marx Gómez

Evaluating Large Language Models for Causal Hypothesis Generation: Detecting Training Data Contamination

Sibin Lazar Maliyammavu, Jorge Marx-Gómez, Michael Mattern, Andreas Solsbach
Carl von Ossietzky Universität Oldenburg, Deutschland

Abstract: The evaluation of Large Language Models (LLMs) poses a distinctive risk in the form of data contamination. In the event of an LLM being evaluated on data that it has encountered during the training process, its apparent performance may be artificially inflated. This erodes confidence in the reliability of the benchmarking results. The present paper explores data contamination detection and mitigation strategies, focusing on a specific task (the generation of hypotheses suggesting a causal relationship between KPIs) and drawing on a structured (though not yet complete) literature review.

Keywords: LLM evaluation, data contamination, causal hypothesis, causality, benchmarking

1. Introduction

At the previous year's SAP ACC conference, Marx Gómez, Mattern and Solsbach presented a paper that employed Pearl's do-calculus to quantify causal effects within a "value driver" network (Marx Gómez et al., 2024). However, Pearl's do-calculus necessitates a directed acyclic graph (DAG) of hypothesized causal relationships between variables that cannot be uniquely inferred from observational data alone. In general, observational data yields a Markov equivalence class of graphs. Consequently, the structure of the DAG must stem from theory, expert knowledge, or common sense.

A study is currently underway that explores the potential of generative artificial intelligence (AI) in the generation of such hypotheses. The evaluation plan involves the provision of an LLM with variables and definitions derived from published papers, intending to ascertain the LLM's capacity to propose causal relationships analogous to those initially theorized by the paper's authors. It is imperative to recognize that addressing the issue of data contamination is of the essence for this evaluation. In cases where the target studies have been observed to be present in the training data of the LLM, particularly where these studies are of a highly cited nature, the model may merely reproduce the contents of the paper. This could result in an exaggeration of the related performance. Consequently, the present paper offers a literature-based investigation of strategies to detect and mitigate data contamination in the context of the respective task.

2. Literature Search Methodology

A comprehensive search was conducted to identify scholarly publications addressing data contamination in LLMs using Scopus, Web of Science (WoS), ArXiv, IEEE Xplore, and JSTOR; the respective search strings are shown in Tab 1.

Database	Search Query	Hits
Scopus	TITLE-ABS-KEY("data contamination" AND (llm or "large language model" or "generative ai" or "generative artificial intelligence"))	66
Web of Science	TS="data contamination" AND TS=(llm or "large language model" or "generative ai" or "generative artificial intelligence")	4
arXiv	Query: order: -announced_date_first; size: 50; include_cross_list: True; terms: AND abstract=data contamination; AND abstract=llm OR "large language model" OR "generative ai" OR "generative artificial intelligence"	207
IEEE	("data contamination") AND ("llm" OR "large language model" OR "generative ai" OR "generative artificial intelligence")	6
JSTOR	("data contamination") AND (llm OR "large language model" OR "generative ai" OR "generative artificial intelligence")	5

Tab. 1: Search strings and results

A total of 288 articles were identified, the majority of which, likely due to the novelty of the topic, were found on the arXiv e-print server. Of these, 28 were duplicates. Of the remaining 260 articles, the full text of all but one was available. From this pool, a random sample of 40 articles was drawn. For each of the 40 articles, a 10–15-minute audio summary was generated using Google NotebookLM Pro. This approach was motivated by a desire to facilitate a more comprehensive evaluation than would have been possible by relying solely on abstracts whilst at the same time reducing the time and effort compared to full-text readings.

These summaries were then utilized to ascertain the relevance (21 articles) or irrelevance (19 articles) to the research question. The primary criteria for relevance were as follows:

- Does the article deal with tasks that may be transferable to the research objective of the present study, which is, inter alia, the generation of causal hypotheses? For example, articles focusing on the linguistic characteristics of specific languages, such as Persian or Korean, have been excluded from this study.
- Does the article provide concrete insights into the identification and mitigation of data contamination?

The 21 articles classified as "relevant" form the foundation for this paper. The labelled corpus will also be used in a future study to fine-tune a version of SciBERT, which will then be employed to classify the remaining 219 articles. In Fig. 1, an overview of the screening process is presented in the form of a PRISMA flow diagram.

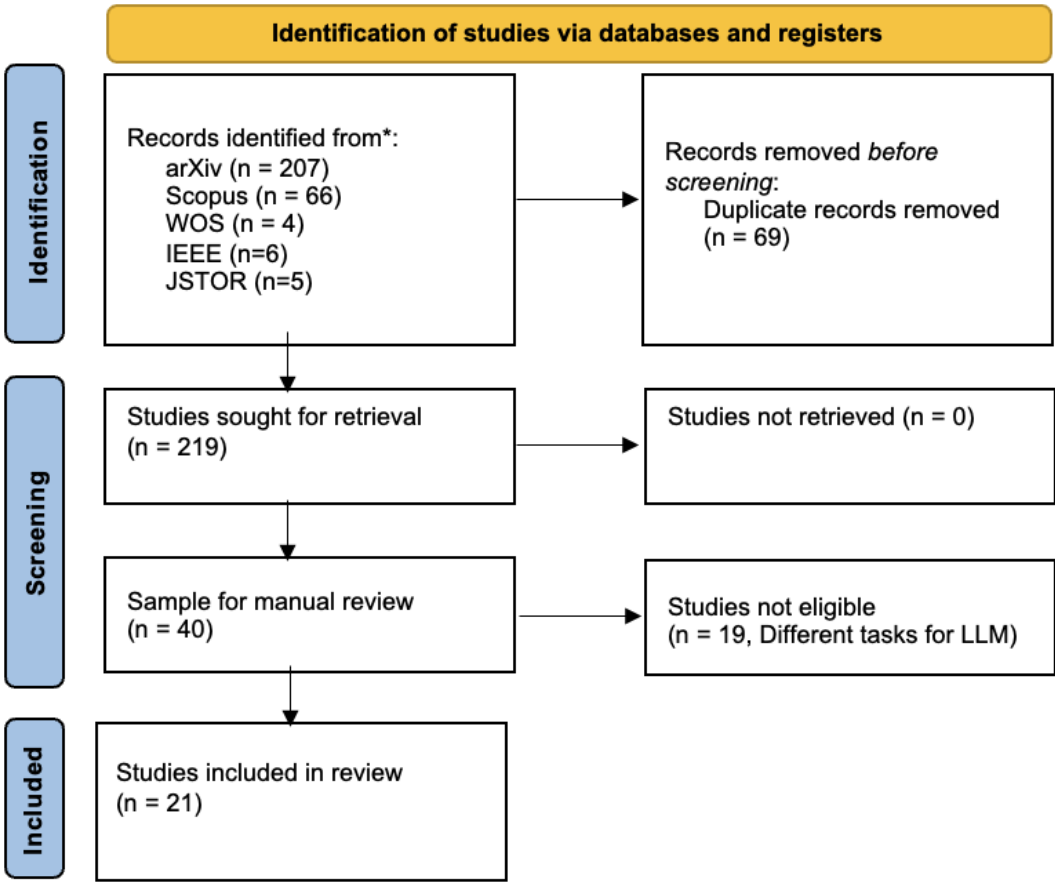


Fig.1 PRISMA Flow Chart

3. Suggested Detection Strategies

In order to assess the relevance of methods for detecting data contamination in relation to our specific research question, we once again utilized Google NotebookLM Pro. The primary objective of this analysis was to determine whether the tasks examined — and consequently, the proposed detection strategies — could help address our research challenge (the use of LLMs to generate hypotheses about causal relationships between variables and the critical challenge of detecting potential contamination in the models' training data that could artificially inflate evaluation performance).

NotebookLM Pro was provided with the following set of instructions and subsequently yielded the results displayed in the initial two columns of Tab. 2. The table was then manually modified to incorporate the studies' designated detection and mitigation strategies. Studies with medium or low relevance were highlighted in grey.

“Consider the following problem context:

- We are using LLMs to hypothesize causal links between variables.
- The LLM will be provided with variables and definitions from academic papers and will be asked to generate hypotheses regarding causal links between these variables.

- We will use the hypotheses found in the respective paper as a benchmark for the LLM’s performance regarding hypothesis generation.
- Our critical issue: Detecting and mitigating contamination when target studies might be present in the LLM’s training data (in which case the LLM has not generated but simply reproduced hypotheses from the respective paper, risking inflated performance metrics.

For each of the 21 scientific studies in your sources, evaluate whether they discuss tasks comparable to our context above (in which case their respective data contamination detection strategies could be useful to help us resolve our critical issue). Provide your assessment as a concise table, with a likelihood (low/medium/high) for each study. Sort the studies alphabetically.”

Tab. 2: Studies and their comparability to our research question

Study	Likelihood of Comparability	Approaches Discussed
Anh et al., 2025	Low	Residual-Noise Fingerprinting (RN-F)
B. Cao et al., 2024	Medium	Structured Evaluation (StructEval)
Dekoninck et al., 2024	High	Performance-Based Contamination Detection (ConStat)
Dong et al., 2024	High	Contamination Detection via Output Distribution (CDD), Trustworthy Evaluation via Output Distribution (TED)
L. Cao et al., 2024	Low	GenCepcion
Fu et al., 2025	High	Instance Similarity, Probability Analysis, Instance Generation and Selection, Answer Memorization
Golchin & Surdeanu, 2025	Medium	Data Contamination Quiz (DCQ)
Hidayat et al., 2025	Medium	Semi-half Method, Permutation Method (R and Q), N-Gram Method
Hou et al., 2025	Medium	Automatic Robustness Checker (AR-CHECKER)
Lopez-Lira et al., 2025	High	Controlled Recall Testing, Anonymization and Masking
Palavalli et al., 2024	High	Controlled Recall Testing, In-domain Data Comparison, Anonymization (Entity Neutering)
Qiu et al., 2025a	Medium	Physics Benchmark (PHYBench)
Ravaut et al., 2025	High	String Matching, Embeddings Similarity, Paraphrase Detection, Dataset-level Performance Analysis, Membership-Inference Attacks (MIA), Model Memorization, Model Confidence Analysis, Hidden Layer Probing
Samuel et al., 2025	High	Min-K% Probability, Canonical Order Statistical Testing, Token Completion Overlap Score, Word Perturbation Quiz (WPQ), Local Order Quiz
Wang et al., 2025	High	Scientific Citation-based Live Assessment of Long-context Academic Reasoning (SCALAR)
Yax et al., 2025	High	Probing and analyzing the log-probabilities of text

		sequences (LogProber)
Ye et al., 2024	High	Polarized Augment Calibration (PAC)
(P. Zhou et al., 2025)	Medium	Multi-disciplinary benchmark assessing reasoning capabilities of Multimodal Large Language Models (MLLMs) at the K-12 level (MDK12-Bench)
J. Zhou et al., 2025	High	Scientific Dynamic Assessor of LLMs (SciDA)
W. Zhu et al., 2024	High	Clean Evaluation on Contaminated Large Language Models (CLEAN-EVAL)
K. Zhu et al., 2024	High	Dynamic evaluation of Large Language Models (DyVal)

Some studies that have received high relevance scores from NotebookLM (e.g., Qiu et al., 2025b) have utilized custom benchmarking datasets. Consequently, the methodologies of these studies do not apply to our research question. Hence, some papers (Dong et al., 2024; Wang et al., 2025; K. Zhu et al., 2024) were excluded from the analysis (and are also highlighted in grey in Tab. 2).

Among these, three statistical approaches—ConStat, PAC, and LogProber—seem to hold particular promise for our purposes. In comparison to other methods, such as Controlled Recall Testing or Membership-Inference Attacks, these approaches were selected on the basis of their wider applicability to the evaluation setting under consideration. ConStat is well-suited for detecting performance-based contamination without requiring extensive auxiliary data, PAC adapts robust membership inference techniques to black-box LLMs, and LogProber provides a scalable way to disentangle memorization from genuine reasoning. While methods such as CLEAN-EVAL or SciDA also demonstrate strong detection capabilities, they either require custom benchmarks or detailed access to the model internals, which limits their practicality in this case. ConStat, PAC, and LogProber collectively address the balance between feasibility, robustness, and compatibility with black-box commercial models.

4. Implications for KPI Hypothesis Benchmarking

PAC (Ye et al., 2024), for example, could be used for our needs.

4.1. Mechanism

Membership inference methods aim to determine if a specific instance, such as a sentence or data point, was included in a model’s training data. PAC adapts classical membership inference attacks from the machine learning domain to large-scale language models, including black-box scenarios.

The core idea is as follows: The model is presented with a suspect input, such as a hypothesis extracted from a paper, and then the input is subjected to small random modifications, or random swap augmentations. PAC then measures the model’s output probabilities, or its “confidence,” for both the original input and its slightly perturbed variants. If the original sentence was part of the training data, the model is likely to assign it a high level of confidence; however, even minor perturbations will be assigned considerably lower probabilities. PAC quantifies this pattern as a polarized distance, which is a measure of how sharply the likelihood of the original deviates from that of its noisy variants. If the deviation exceeds a predefined

threshold, the instance is classified as likely originating from the training corpus, i.e., as contaminated.

4.2. Suitability

In our specific use case, where a hypothesis generated by the model closely matches the published version, PAC is especially valuable. It allows us to directly inspect the hypothesis text and determine if this exact formulation was seen during training. PAC provides an evidence-based answer to this question without requiring access to the model's internal workings. This black-box-compatible method is especially beneficial for closed-source models such as GPT-4o, as it relies solely on querying the model's output. Furthermore, PAC is more robust than simple metrics, such as perplexity thresholds, because it incorporates both near (lightly altered) and distant (randomly distorted) variants of the candidate text. For our application, if GPT-4o has memorized specific KPI-related hypotheses from a paper, it will reproduce those formulations with greater certainty than altered versions. PAC can detect this as a contamination signal. Conversely, suppose the model has not been exposed to the hypothesis during training and must generate it from scratch. In that case, the model's confidence in the original versus perturbed versions should not differ markedly.

For instance, consider a KPI benchmarking task where an LLM is queried about the relationship between increased customer satisfaction and higher renewal rates. Utilizing PAC, the original hypothesis text undergoes a slight perturbation (e.g., substituting "customer" with "satisfaction"). If the model assigns a substantially higher probability to the unaltered version, this would indicate that the specific phrasing in question likely appeared in the training data. This enables researchers to identify and adjust the benchmark accordingly.

4.3. Implementation

Implementing PAC requires access to a model's token-level probability estimates for a given input. Many LLM APIs, including OpenAI's, provide access to logit or log-probability values. In practice, one would start with the hypothesis H as stated in the original paper and generate several slightly modified versions H' —for example, by randomly swapping two words per sentence. Then, the model (e.g., GPT-4o) computes the likelihood (or log-likelihood) for each sequence. PAC recommends assessing the deviation between the original hypothesis (H) and its nearest neighbors (H'), as well as from more distant random variations, to calibrate the detection signal. The result is a score for each hypothesis indicating how familiar it is to the model. A high score suggests that H likely appeared in the model's training data, justifying its exclusion from the benchmark or adjustment of the evaluation results. This method also facilitates the deliberate simulation of clean versus contaminated scenarios. For selected hypotheses that are known to be post-training and therefore guaranteed to be unseen, one can verify whether PAC produces low scores to evaluate its false positive rate.

5. International Considerations

The issue of data contamination in LLMs is not regionally restricted or language-independent, as nowadays LLMs often use all publicly available sources for training and thus also cross these boundaries. The open access culture also contributes to this, as many scientific articles are now available in machine-readable form and can be used as training data. In their article “Data Contamination Can Cross Language Barriers”, (Yao et al., 2024) show that even language

barriers no longer exist today for working with LLMs and thus training data is often used without regional restrictions, whereby due to the open access culture and the large number of English-language publishers, there is a preponderance of English-language data for training sets.

The extensive use of public sources also leads to regional peculiarities being lost in the training sets, for example, susceptibilities in indigenous population groups to diseases in health data that are lost due to the global creation of training data are no longer statistically recognizable as anomalies.

Consequently, the discourse on data contamination must extend beyond the temporal delimitation of training corpora. It is imperative to interrogate the epistemic provenance and geospatial distribution of the underlying datasets. Although end-users and researchers lack direct control over the data curation strategies employed by LLM developers, critical evaluation of training data origins remains essential. This is particularly true in domains involving vulnerable or underrepresented populations, where unbalanced training distributions can propagate systemic bias or erode epistemic pluralism.

Researchers—especially those operating in the Global South—must engage with LLM-generated content through a reflexive lens, recognizing the epistemological asymmetries embedded in model outputs. These asymmetries often reflect the disproportionate representation of knowledge frameworks and discursive norms prevalent in industrialized, English-speaking contexts. Hence, methodological rigor in interpreting LLM behavior must account for the structural biases inherent in the training data, with particular attention to whose knowledge is rendered visible, and whose is marginalized.

6. Conclusion and Outlook

ConStat, PAC, and LogProber offer systematic approaches to identifying potential contamination; however, these methods are not without constraints. ConStat is contingent upon robust statistical baselines; consequently, false signals may be produced if performance variability stems from factors unrelated to contamination. Although the PAC model has proven to be effective in black-box settings, it does incur computational costs due to the repeated perturbation queries that are necessary for its operation. Furthermore, the model assumes that the confidence gaps between contaminated and uncontaminated inputs are consistent. LogProber requires fine-grained probability access and may encounter challenges when confronted with contamination that intersects with authentic reasoning patterns. Consequently, while these methodologies are valuable, their findings should be interpreted with caution and, when feasible, triangulated with complementary approaches.

In recent years, especially within the last twelve months, the body of literature addressing data contamination in LLMs has grown substantially. Among studies focused on detection and mitigation, two broad strategic approaches have emerged. One group relies on yet unpublished and/or custom-built benchmarks. The other group involves interventions at the training data level, presupposing varying degrees of access to or control over the training pipeline. Additionally, the applicability of individual methods depends on the level of transparency of the model in question, i.e., whether it is accessible as a white-box, gray-box, or black-box system. This also includes the extent to which the LLM reveals information about its internals, such as token-level probabilities, through its API.

Based on a labeled sample of peer-reviewed and preprint publications, this paper demonstrates that many detection strategies, particularly those that rely on probability-based inference, are transferable to our specific evaluation task of assessing LLM-generated causal hypotheses regarding dependencies among KPIs.

Before determining whether LLMs can meaningfully contribute to this task beyond merely reproducing memorized causal claims, a thorough and methodical review of the relevant detection techniques is essential. This study provides an initial foundation for such efforts by offering a curated sample and a preliminary mapping of methodological fit.

7. Use of AI Tools

Several AI-based tools were used to help prepare this article. Google NotebookLM Pro was used to generate audio summaries of each paper in the sampled literature. Second, it was used to evaluate the studies' relevance to the research question, particularly regarding data contamination detection in the context of causal hypothesis generation. NotebookLM Pro was also used to answer specific queries about the labeled corpus, such as identifying studies that relied on statistical methods.

To refine the written content, multiple AI writing assistants were consulted, including ChatGPT-4o, ChatGPT-4.5, and DeepL Write. These tools were used to enhance clarity, coherence, and academic tone throughout the manuscript.

8. References

- Anh, L. V., Nguyen, D. D. N., & Nguyen, P. L. (2025). RN-F: A Novel Approach for Mitigating Contaminated Data in Large Language Models (No. arXiv:2505.13249; Issue arXiv:2505.13249). arXiv. <https://doi.org/10.48550/arXiv.2505.13249>
- Cao, B., Ren, M., Lin, H., Han, X., Zhang, F., Zhan, J., & Sun, L. (2024). StructEval: Deepen and Broaden Large Language Model Assessment via Structured Evaluation. In Ku L.-W., Martins A., & Srikumar V. (Eds.), *Proc. Annu. Meet. Assoc. Comput. Linguist.* (pp. 5300–5318). Association for Computational Linguistics (ACL); Scopus. <https://doi.org/10.18653/v1/2024.findings-acl.314>
- Cao, L., Buchner, V., Senane, Z., & Yang, F. (2024). Introducing GenCeption for Multimodal LLM Benchmarking: You May Bypass Annotations. In Ovalle A., Chang K.-W., Cao Y.T., Mehrabi N., Zhao J., Galstyan A., Dhamala J., Kumar A., & Gupta R. (Eds.), *TrustNLP - Workshop Trust. Nat. Lang. Process., Proc. Workshop* (pp. 196–201). Association for Computational Linguistics (ACL); Scopus. <https://doi.org/10.18653/v1/2024.trustnlp-1.16>
- Dekoninck, J., Müller, M. N., & Vechev, M. (2024). ConStat: Performance-Based Contamination Detection in Large Language Models. In Globerson A., Mackey L., Belgrave D., Fan A., Paquet U., Tomczak J., & Zhang C. (Eds.), *Adv. Neural Inf. Proces. Syst. (Vol. 37)*. Neural information processing systems foundation; Scopus. <https://www.scopus.com/inward/record.uri?eid=2-s2.0-105000555768&partnerID=40&md5=7c9e6cec66889f55e6fbae7751505093>

- Dong, Y., Jiang, X., Liu, H., Jin, Z., Gu, B., Yang, M., & Li, G. (2024). Generalization or Memorization: Data Contamination and Trustworthy Evaluation for Large Language Models. In Ku L.-W., Martins A., & Srikumar V. (Eds.), *Proc. Annu. Meet. Assoc. Comput Linguist.* (pp. 12039–12050). Association for Computational Linguistics (ACL); Scopus. <https://doi.org/10.18653/v1/2024.findings-acl.716>
- Fu, Y., Uzuner, O., Yetisgen, M., & Xia, F. (2025). Does Data Contamination Detection Work (Well) for LLMs? A Survey and Evaluation on Detection Assumptions (No. arXiv:2410.18966; Issue arXiv:2410.18966). arXiv. <https://doi.org/10.48550/arXiv.2410.18966>
- Golchin, S., & Surdeanu, M. (2025). Data Contamination Quiz: A Tool to Detect and Estimate Contamination in Large Language Models (No. arXiv:2311.06233; Issue arXiv:2311.06233). arXiv. <https://doi.org/10.48550/arXiv.2311.06233>
- Hidayat, N. S., Kautsar, M. D. A., Wicaksono, A. F., & Koto, F. (2025). Simulating Training Data Leakage in Multiple-Choice Benchmarks for LLM Evaluation (No. arXiv:2505.24263; Issue arXiv:2505.24263). arXiv. <https://doi.org/10.48550/arXiv.2505.24263>
- Hou, Y., Xiao, Z., Yu, F., Jiang, Y., Wei, X., Huang, H., Chen, Y., & Chen, G. (2025). Automatic Robustness Stress Testing of LLMs as Mathematical Problem Solvers (No. arXiv:2506.05038; Issue arXiv:2506.05038). arXiv. <https://doi.org/10.48550/arXiv.2506.05038>
- Lopez-Lira, A., Tang, Y., & Zhu, M. (2025). The Memorization Problem: Can We Trust LLMs' Economic Forecasts? (No. arXiv:2504.14765; Issue arXiv:2504.14765). arXiv. <https://doi.org/10.48550/arXiv.2504.14765>
- Marx Gómez, J., Mattern, M., & Solsbach, A. (2024). Strategic Integration of Value Driver Networks and Causal Inference in Prescriptive Business Analytics. In U. Mathis, N. Ondrusch, D. Kilian, A. Redlein, K. Turowski, S. Weidner, & H. Wittges (Eds.), *Proceedings of the SAP Academic Community Conference 2024 (D-A-CH)* (pp. 176–183). Technische Universität Wien, Institut für Managementwissenschaften. <https://doi.org/10.14459/2024md1752375>
- Palavalli, M., Bertsch, A., & Gormley, M. R. (2024). A Taxonomy for Data Contamination in Large Language Models. In Sainz O., Ferrero I.G., Agirre E., Campos J.A., Jacovi A., Elazar Y., & Goldberg Y. (Eds.), *CONDA - Data Contamination Workshop, Proc. Workshop* (pp. 22–40). Association for Computational Linguistics (ACL); Scopus. <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85204305246&partnerID=40&md5=2a6aee8791f0c701c918581b522d28f7>
- Qiu, S., Guo, S., Song, Z.-Y., Sun, Y., Cai, Z., Wei, J., Luo, T., Yin, Y., Zhang, H., Hu, Y., Wang, C., Tang, C., Chang, H., Liu, Q., Zhou, Z., Zhang, T., Zhang, J., Liu, Z., Li, M., ... Zhu, H. X. (2025a). PHYBench: Holistic Evaluation of Physical Perception and Reasoning in Large Language Models (No. arXiv:2504.16074; Issue arXiv:2504.16074). arXiv. <https://doi.org/10.48550/arXiv.2504.16074>
- Qiu, S., Guo, S., Song, Z.-Y., Sun, Y., Cai, Z., Wei, J., Luo, T., Yin, Y., Zhang, H., Hu, Y., Wang, C., Tang, C., Chang, H., Liu, Q., Zhou, Z., Zhang, T., Zhang, J., Liu, Z., Li, M., ... Zhu, H. X. (2025b). PHYBench: Holistic Evaluation of Physical Perception and Reasoning in Large Language Models (No. arXiv:2504.16074; Issue arXiv:2504.16074). arXiv. <https://doi.org/10.48550/arXiv.2504.16074>

- Ravaut, M., Ding, B., Jiao, F., Chen, H., Li, X., Zhao, R., Qin, C., Xiong, C., & Joty, S. (2025). A Comprehensive Survey of Contamination Detection Methods in Large Language Models (No. arXiv:2404.00699; Issue arXiv:2404.00699). arXiv. <https://doi.org/10.48550/arXiv.2404.00699>
- Samuel, V., Zhou, Y., & Zou, H. P. (2025). Towards Data Contamination Detection for Modern Large Language Models: Limitations, Inconsistencies, and Oracle Challenges. In Rambow O., Wanner L., Apidianaki M., Al-Khalifa H., Di Eugenio B., & Schockaert S. (Eds.), Proc. Main Conf. Int. Conf. Comput. Linguist., COLING: Vol. Part F206484-1 (pp. 5058–5070). Association for Computational Linguistics (ACL); Scopus. <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85218501267&partnerID=40&md5=eff32c9e705593f114968a947c76b1b4>
- Wang, R., Mu, H., Ma, L., Lin, L., Feng, Y., Baldwin, T., Han, X., & Li, H. (2025). SCALAR: Scientific Citation-based Live Assessment of Long-context Academic Reasoning (No. arXiv:2502.13753; Issue arXiv:2502.13753). arXiv. <https://doi.org/10.48550/arXiv.2502.13753>
- Yao, F., Zhuang, Y., Sun, Z., Xu, S., Kumar, A., & Shang, J. (2024). Data Contamination Can Cross Language Barriers. In Al-Onaizan Y., Bansal M., & Chen Y.-N. (Eds.), EMNLP - Conf. Empir. Methods Nat. Lang. Process., Proc. Conf. (pp. 17864–17875). Association for Computational Linguistics (ACL); Scopus. <https://doi.org/10.18653/v1/2024.emnlp-main.990>
- Yax, N., Oudeyer, P.-Y., & Palminteri, S. (2025). LogProber: Disentangling confidence from contamination in LLM responses (No. arXiv:2408.14352; Issue arXiv:2408.14352). arXiv. <https://doi.org/10.48550/arXiv.2408.14352>
- Ye, W., Hu, J., Li, L., Wang, H., Chen, G., & Zhao, J. (2024). Data Contamination Calibration for Black-box LLMs. In Ku L.-W., Martins A., & Srikumar V. (Eds.), Proc. Annu. Meet. Assoc. Comput. Linguist. (pp. 10845–10861). Association for Computational Linguistics (ACL); Scopus. <https://doi.org/10.18653/v1/2024.findings-acl.644>
- Zhou, J., Miao, T., Liao, Y., Wang, Q., Wen, Z., Wang, Y., Huang, Y., Yan, G., Wang, L., Xia, Y., Gao, H., Zeng, Y., Zheng, R., Dun, C., Liang, Y., Yang, T., Huang, W., & Zhang, G. (2025). SciDA: Scientific Dynamic Assessor of LLMs (No. arXiv:2506.12909; Issue arXiv:2506.12909). arXiv. <https://doi.org/10.48550/arXiv.2506.12909>
- Zhou, P., Zhang, F., Peng, X., Xu, Z., Ai, J., Qiu, Y., Li, C., Li, Z., Li, M., Feng, Y., Sun, J., Zhang, H., Li, Z., Mao, X., Zhao, W., Wang, K., Chang, X., Shao, W., You, Y., & Zhang, K. (2025). MDK12-Bench: A Multi-Discipline Benchmark for Evaluating Reasoning in Multimodal Large Language Models (No. arXiv:2504.05782; Issue arXiv:2504.05782). arXiv. <https://doi.org/10.48550/arXiv.2504.05782>
- Zhu, K., Chen, J., Wang, J., Gong, N. Z., Yang, D., & Xie, X. (2024). DyVal: Dynamic Evaluation of Large Language Models for Reasoning Tasks. Int. Conf. Learn. Represent., ICLR. 12th International Conference on Learning Representations, ICLR 2024. Scopus. <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85200544260&partnerID=40&md5=d0d0a3c31e54c1e41bf085b23ff61bc2>
- Zhu, W., Hao, H., He, Z., Song, Y., Zhang, Y., Hu, H., Wei, Y., Wang, R., & Lu, H. (2024). CLEAN-EVAL: Clean Evaluation on Contaminated Large Language Models. In Duh K., Gomez H., & Bethard S. (Eds.), Find. Assoc. Comput. Linguist.: NAACL - Findings (pp.

835–847). Association for Computational Linguistics (ACL); Scopus.
<https://doi.org/10.18653/v1/2024.findings-naacl.53>

Strengthening International Collaboration through Erasmus+ KA171: Insights from East African Partnerships

Sven Lampe, Jorge Marx Gómez
Carl von Ossietzky Universität Oldenburg, Germany

Abstract: This paper reflects on the Erasmus+ Key Action 171 (KA171) partnership between the University of Oldenburg (UOL), the Nelson Mandela African Institution of Science and Technology (NM-AIST, Tanzania), and the University of Mauritius (UOM). With a focus on ICT in disaster management, the project combines student and staff mobility with institutional development and is anchored in UOL's Very Large Business Applications (VLBA) research group. Building on prior collaboration through the CENIT@EA Centre of Excellence, the partnership has evolved from exchange into structural cooperation. Notable outcomes include the International Conference on Resilient and Adaptive Island Systems (ICRAIS), which attracted over 100 submissions, and a Summer School on AI and Sustainability, organised as a hackathon for joint student teams. The partners have also submitted a Horizon Europe proposal on disaster risk governance. These results demonstrate how KA171 can act as a catalyst for long-term academic collaboration, capacity building, and global engagement.

Digital Marketing as a Driver of Business Process Innovation in Higher Education: A Case Study of Tanzanian Institutions

Godfrey Yohana Naman

Nelson Mandela African Institution of Science and Technology, Arusha, Tanzania

Abstract: In the era of digital transformation, technology has brought various challenges and opportunities for the marketers; this study assesses the contribution of digital marketing tools as a driver for accelerating business processes in enrolment at two Tanzania institutions. The study adopts the Technology Acceptance Model to investigate the perceptions and experiences from both institutions. The study was undertaken through a survey research design, and mixed approaches were used. Simple random sampling was used and total of 286 questionnaires were distributed to students and only 190 were returned (66.4%). Primary data comprise closed-ended questionnaires while secondary data are from journals, conference papers. The reliability check (Cronbach $\alpha = 0.74$) was performed. The study reveals both institutions 68.4% of students often encounter the institution on digital platforms, and the same share (68.4%) submit applications online. Usability ratings are positive 28.9% say the process is “almost always easy,” and 22.1% “sometimes easy.” The study shows that over 2500 students submit their application online. The study also reveals that target audience, budget, technology, user support, and priorities are the most important factors to consider when selecting digital marketing platforms. The finding discovers that Business Process Management (BPM) contributes to the education sector by showcasing how digital marketing enhances enrolment efficiency, student engagement, and institutional sustainability by shaping students’ decisions.

Keywords: Digital Marketing, Digital Marketing Platforms, Student Enrolment, Student Choice, Higher Learning Institution.

Smart Cities in Developing Countries: A Context-Adaptive Framework for Sustainable Urban Development

M. Lyimo¹, B. Mgawe¹ & J. Gómez²

¹The Nelson Mandela African Institution of Science and Technology,
Arusha, Tanzania

²The University of Oldenburg, Oldenburg, Germany

Abstract: This paper proposes the Context-Adaptive Smart City Framework (CASCf), an adaptation of a five-layered model adjusted to the context of the developing countries with infrastructure, institutional, and financial limitations. The approach begins with evaluating the infrastructure, then connecting with the community, selecting the right technologies that fit the community's needs, designing systems to match these needs, and then developing methods for achieving sustainability. Unlike traditional top-down smart city models, CASCf integrates community participation, proper technology, service delivery, and sustainable planning. It seeks to overcome significant shortcomings observed in a wide range of global initiatives, including poor scalability, narrow citizen involvement and poor sustainability planning. By integrating the development of smart cities with local needs and conditions, CASCf establishes a pragmatic path to more inclusive, resilient and effective urban change. Future work includes pilot testing to validate real-world impact on service delivery and governance in low-resource environments.

Keywords: Smart cities, developing countries, adaptive framework, digital transformation, urban sustainability

1. Introduction

With the increasing demand for urban solutions, smart cities have attracted much attention worldwide, aiming to improve urban problems through the utilisation of Information and Communication Technologies (ICTs), data analytics, and citizen-centric services (Bozkurt et al., 2025). These are cities seeking more efficient, more responsive, and more inclusive urban governance. Despite the roots in high-income countries, the benefits of smart city investments have not been equitably distributed (Donio & Toch, 2025). Such disparity is mostly related to substantial variance in urban infrastructure requirements, financial resources, institutional capacities and socio-cultural diversity, which are barriers for the transplantation of the smart city paradigm to these settings (OECD, 2024; Moumen et al., 2024; World Bank, 2025).

UN-Habitat (2022) projects that by 2050, around 6.7 billion people will be living in cities, with nearly 90% of this growth expected in developing countries (Neil et al., 2022). This rapid urban growth puts severe pressure on municipalities that already face resource shortages and difficulties in providing basic services. In many areas, people have little or no access to drinking water, electricity, sanitation, or the internet (World Bank, 2023). Local budgets in many developing regions are often too limited to support big digital projects (OECD, 2023).

Many smart city projects also take a top-down approach, focusing on advanced technology while paying little attention to the real challenges faced by local communities (Shao & Min, 2025). Examples from Asia, Africa and Latin America have demonstrated that technology-driven approaches frequently result in situations of vendor lock-in, of white elephants, of lack of stakeholder engagement, and of the inability to move beyond the pilot stage (Kabuya et al., 2024). Singapore and Barcelona are some of the cities that have successfully used sophisticated data systems to enhance public services. Still, their successes depend on robust digital infrastructure and large financial investments, conditions often unavailable in many developing countries (OECD, 2023).

Hence, a rethinking of the smart city paradigm as locally situated, inclusive, sustainable, and adaptable is required. In this paper, the CASCf as a layered model that meets basic infrastructural requirements, involves local communities from the beginning, embeds suitable technologies, provides basic services step by step, and is sustainable over time, is presented. CASCf is not trying to retrofit developing nations into the template of global tech hubs; instead, it is equipping these cities with the right tools and approaches that are tailored to challenges and opportunities found in those areas.

1.1. Key contributions

This paper adds to the existing smart city literature by proposing an innovative five-layer model named the CASCf, which is adapted for the sociotechnical and economic context of developing countries. It attempts to overcome major gaps in traditional smart city models, for example, between technology and local capacity, with respect to engagement of citizens, and in planning for sustainability. CASCf focuses on context-fit, participatory planning and technology that can be scaled up. It is a clear and versatile toolkit for policy-makers, city managers, and implementers of development to realize inclusive and resilient urban transformation. It also provides possibilities to be integrated with digital platforms like SAP and Analytics Cloud, to conduct data-driven governance in resource-constrained settings.

The remainder of the paper addresses the key urban challenges in developing countries in section two, some smart cities initiatives worldwide in section three, analysis of the CASCf framework in section four, along with where and in what form it may be applied in the future in section five, and final considerations for decision makers and researchers in section six.

2. Problem Statement

2.1. Urban challenges in developing countries

Cities in the developing world are expanding at a very fast pace, but many of them lack the infrastructure or systems in place to cope with this expansion. By 2030, more than 1.2 billion additional people are forecast to be living in these cities, straining housing, clean water, electricity and transportation services (Huang, 2024). In Sub-Saharan Africa, for instance, the majority of the urban population resides in informal settlements and has inadequate services (Scheufler et al., 2020).

Financial and institutional weakness worsens the situation, as most local governments have small budgets and little technical know-how, and they struggle to invest in smart technologies. There is also a lack of coordination between various agencies and levels of government.

Moreover, political instability and ambiguous legislation make long-term planning even trickier and discourage private enterprise from investing (OECD, 2023).

Many residents of these cities are also poor and lack digital skills as well as access to the internet and smartphones (Shanahan et al., 2023). This digital divide could result in smart city projects that only benefit a limited number of people, excluding those who need them most.

2.2. Limitations of conventional smart city models

Most smart city blueprints are designed for wealthy countries with well-established digital infrastructure, stable governments and reliable budgets. These models typically rely on emerging technologies, such as artificial intelligence (AI), big data and smart devices (Zaman et al., 2024). While these systems function well in places like Tokyo or Amsterdam, they are often not practical for developing countries, many of which suffer from power outages, weak internet and financial constraints (Shah et al., 2024).

In most developing areas, these models are introduced without consulting the local population or governments. This leads to a lack of citizen involvement, a failure of the fit, and, in many cases, failure of the project (Prakash, 2025). At other times, dependence on international vendors with no focus on capacity building and transfer produces a continuous dance of dependence and unsustainable systems (Andrew, 2024).

Smart city strategies need to be adapted to local circumstances to be successful. They should be basic, inclusive, and developed with available resources in mind. In this paper, the CASCF is suggested as a potential solution to this issue.

3. Related Works

To get an insight into the global existence of smart city initiatives, Table 1 provides a comparative summary of some selected smart city initiatives based on region, strengths and limitations.

Table 1: Overview of global smart city projects and their key strengths and limitations

Reference	Project	Region	Strengths	Limitations/Gaps
(Sipahi & Saayi, 2024)	Smart Nation Singapore	Asia	Well-established infrastructure, strong government leadership, and implementation across platforms	It is costly and challenging to implement in low-resource settings. Limitations of private sector involvement and raising public concerns about data privacy and trust
(Alizadeh et al., 2024)	Barcelona smart city	Europe	Citizen-focused services, open data, and devolved Internet of Things (IoT)	Needs complex infrastructure and mature digital knowledge. Limited citizen participation whereby citizen involvement was highlighted as significant, but during primary phases, citizens were often treated as recipients rather than active partners.
(Santosh et al., 2024)	India's Smart Cities Mission	Asia	General government support, overall national reach, citizen involvement and use of smart technologies	Implementation delays, sustainability challenges

SAP Academic Community Conference 2025 (D-A-CH)

Business Process Management drives Digital Transformation

(UNDP, 2021)	Smart city Rwanda masterplan	Africa	The master plan clearly reflects a strong focus on promoting digital infrastructure, enhancing citizen participation and utilising data-driven urban management to build a smart and inclusive city atmosphere.	The master plan does not adequately address strategies, resources, and institutional arrangements for digital literacy, urban-rural connection, and sustained community.
(John, 2023)	Kenya's Konza Technopolis	Africa	Strong systems, governance, innovation ecosystems at work, and multiple sources of capital.	Inadequate competent human resources, a weak framework to provide motivation, a weak project management mechanism, weak collaborations with industry, and a lack of improved sustainability.
(Mohammad et al., 2022)	Medellín Digital (Colombia)	Latin America	A strong digital innovation environment, strategic geographic location and well-qualified talent pool in Medellín make it a good base to move towards the implementation of a holistic approach to smart city planning.	More coordinated large-scale digital infrastructure upgrades and the implementation of cohesive institutional and regulatory frameworks are needed to scale and sustain smart city initiatives effectively.
(Ogunseye & Bashir, 2022)	Lagos Smart City	Africa	Lagos has shown grassroots citizen participation and knowledge about digital infrastructure projects, as with its fibre optic roll-out plans.	A major challenge in developing a smart city is the combination of low political will, limited stakeholder engagement, a weakly structured implementation approach, and the absence of a clear institutional strategy.
(Kusumastuti & Rouli, 2021)	Indonesia's 100 smart cities	Asia	Indonesia's smart city program uses ICT and social media to resolve urban issues, encourage innovation and engage citizens in urban development.	Constraints on resources exist as well as digital divides, disintegrated governance, inadequate laws and incentives, and the absence of efficient motivational approaches to sustain citizens' intentions to participate.
(AL-Dabbagh, 2022)	Smart Dubai	Middle East	Dubai is advancing its sustainability agenda through green building regulations, renewable energy initiatives, and the development of smart infrastructure.	Obstacles include institutional inertia, human resistance and the emergence of more integrated policies and systems to achieve a smart and sustainable urban development.

These works show the diversity of smart city projects in various areas, but also point out significant differences in their success and conformance. Cities with strong infrastructure and governance have achieved great success with smart city projects. Yet, many emerging regions experience recurring barriers, such as a lack of resources, weak institutional structure, and poor community participation, demonstrating the need for elastic and inclusive smart city models adaptable to context.

4. Proposed Framework

To address the shortcomings found in current smart city frameworks, this paper presents the CASCf, which is a scalable, comprehensive and context-sensitive smart city model developed for the developing countries' environment. In contrast to conventional models that are

characterised by recent technologies and the domination of top-down governance, CASCf presents a bottom-up and modular approach that matches digital strategies with social-economic conditions and existing infrastructure.

CASCf is organised into five layers that are interconnected and address the key aspects of smart city development as illustrated in Figure 1, from infrastructure demand assessment to long-term sustainable development. This multi-layered architecture also enables cities to incrementally adopt and apply the framework, beginning with basic demand and expanding to more sophisticated technologies and governance options.

4.1. CASCf architecture and layers

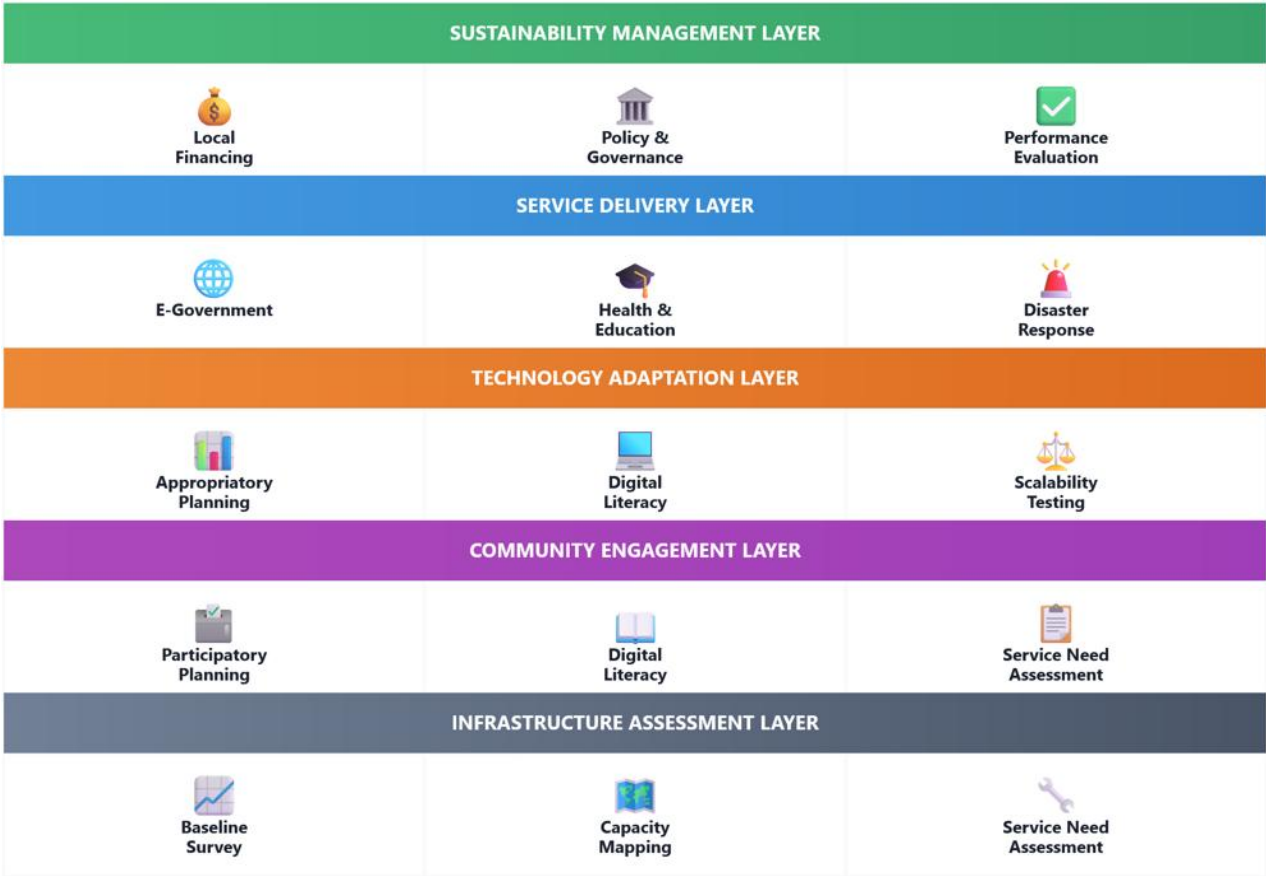


Figure 1: CASCf-a five-layer smart city framework tailored for developing countries.

Infrastructure assessment layer

Infrastructure is a baseline layer that includes mapping of the existing city's physical and digital infrastructure, assessment of service needs, and institutional readiness. It is what makes sure technology does not overstep the ability of the city to keep up. This contrasts with many global frameworks, which expect a level of readiness.

Community engagement layer

In this sense, CASCf emphasises involvement of citizens, local leadership building and cultural integration. With a focus on treating citizens not just as data providers or users of services but rather as co-designers and decision-makers, CASCf enables ownership and long-term engagement in a way that many models do not allow.

Technology adaptation layer

Instead of pushing for expensive, centralised solutions, the CASCf focuses on low-cost, open-source, and modular technologies, affordable IoT devices, edge computing, and mobile-first platforms, for example, solutions that can work with intermittent power or connectivity. This is an advantage in resource-poor settings.

Service delivery layer

This layer involves implementing basic urban services such as e-governance, health, education, water and disaster management. The approach is gradual and demand-driven, which allows early wins to increase trust in the public and confidence in the institutions.

Sustainability management layer

CASCf incorporates local financing modalities, ongoing education, governance feedback, and adaptive policy for long-term sustainability. The majority of current models do not account for this, with many depending on donor funding or on public-private partnerships in the absence of local sustainability plans.

4.2. Uniqueness and value proposition of CASCf

The main feature that distinguishes CASCf is that it is based on the realities of the developing countries. Unlike many smart city models that envision expensive, new infrastructure and significant municipal financial resources, CASCf is modular, scalable and affordable. It prioritizes participation at local levels, leverages open-source and context-appropriate technologies and features an embedded plan for long-term sustainability. This human-centered, flexible methodology provides a more realistic, holistic pathway for economically limited cities to make a start in the smart journey.

4.3. Shifting to CASCf: Linking problems to solutions

Most smart city concepts originate in contexts with strong digital infrastructure, stable funding, and well-established governance. When these same models are transplanted into developing countries, they frequently fall short. Technologies are often mismatched with local conditions, communities are left out of the process, and projects struggle to survive without long-term planning or sufficient resources. That is exactly why we developed the CASCf. Rather than importing complex solutions, CASCf is built around the realities on the ground. Each part of the framework was shaped to solve a problem we have seen in conventional approaches. Table 2 outlines how CASCf bridges these gaps.

Table 2. How CASCf responds to key problems in traditional smart city models.

Key Limitation	CASCf layer	CASCf Solution
Relying too heavily on high-tech systems that do not work well in local conditions	Technology adaptation layer	Suggests low-cost, open-source, modular solutions
Weak community involvement and user resistance	Community engagement layer	Encourages local involvement in planning, co-design, and local ownership
Lack of proper sustainability planning and funding	Sustainability management layer	Recommends diverse financing and long-term maintenance plans
Lack of baseline data and readiness assessment	Infrastructure assessment layer	Ensures local capacity and resource mapping before implementation
Copy-paste the same digital service delivery in every city	Service delivery layer	Allows for flexible, phased services based on what each community actually needs

By moving away from rigid, top-down planning and focusing instead on local needs and capacities, CASCf offers a more practical model for developing countries. Although it hasn't yet been piloted, lessons from similar initiatives in India, Kenya, and Colombia suggest that this kind of approach could reduce implementation costs by 25% to 40%, boost service adoption by up to 50%, and improve sustainability through stronger local capacity and financing. CASCf aims to shift smart city thinking from expensive, tech-heavy projects to realistic, inclusive strategies that work on the ground.

5. Discussion and Future Work

The CASCf model can serve as an organized and flexible tool for smart city implementation in resource-limited areas. By following a structured approach, the framework helps address common challenges such as weak infrastructure, limited local ownership, and sustainability concerns. However, its true value will only be seen if it can successfully move from theory into practice. Cities would need to consider their local context, their institutional preparedness, and community dynamics during the application of CASCf.

Although the model shows promising potential in theory, practical results are often divergent from expected effects in low-resource settings. Consequently, a pilot implementation and evaluation of CASCf in the field is essential before recommending the wider adoption of the application for developing countries. These pilot initiatives should be phased to enhance effectiveness and sustainability. Field validation is an important subsequent step in verifying the practical value of the CASCf model.

Phase 1, Assess: Review infrastructure, map stakeholders. Phase 2, Engage: Capacity building, planning participation. Phase 3, Trial: Roll out specific tech (like digital water meters). Phase 4, Evaluate: Keep track of KPI's such as price, uptake, and satisfaction. Phase 5, Scale: Scale up and replicate successful modules

6. Conclusion

Although smart cities hold much promise for the developing world, current approaches frequently underestimate the specific requirements of this context. To address gaps in conventional practices such as poor contextual fit, limited community involvement, and weak sustainability, this paper introduces CASCf, a framework designed to be both flexible and scalable. CASCf is built around five layers that bring together the right technologies, participatory governance, and capacity building. Although it still needs to be tested in real settings, the framework offers a practical and scalable way to guide smart urban development that is inclusive, sustainable, and shaped by local realities.

7. References

- AL-Dabbagh, R. (2022). Dubai, the sustainable, smart city. *Renewable Energy and Environmental Sustainability*, 7, 3. <https://doi.org/10.1051/rees/2021049>
- Alizadeh, T., Dutia, E., & Clements, R. (2024). Smart Barcelona: the gap between inspiring rhetoric and lackluster implementation in transformative approaches. *Planning Practice and Research*, 39(5), 839–854. <https://doi.org/10.1080/02697459.2024.2369977>

- Andrew, C. (2024). From Smart to Smarter Cities: Enabling Governance and Data-Driven Transformation.
- Bozkurt, Y., Rossmann, A., Pervez, Z., & Ramzan, N. (2025). Assessing data governance models for smart cities: Benchmarking data governance models based on European urban requirements. *Sustainable Cities and Society*, 130. <https://doi.org/10.1016/j.scs.2025.106528>
- Donio, S., & Toch, E. (2025). Neighborhood Disparities in Smart City Service Adoption. <http://arxiv.org/abs/2501.04363>
- Huang, J. (2024). Cities and Climate Action.
- John, O. (2023). Konza Technopolis Development Authority Abridged Version.
- Kabuya, K. I., Ajayi, O. O., & Bagula, A. B. (2024). Smart Cities and Villages: Concept Review and Implementation Perspectives in Developing Cities.
- Kusumastuti, R. D., & Rouli, J. (2021). Smart City Implementation and Citizen Engagement in Indonesia. *IOP Conference Series: Earth and Environmental Science*, 940(1). <https://doi.org/10.1088/1755-1315/940/1/012076>
- UNDP. (2023). Leveraging digitalisation for transformational governance.
- Mohammad, Z., Natalia Currea, D., & Juan Camilo, B. (2022). Smart city technology & the digital opportunity in Medellín.
- Moumen, N., Radoine, H., Nahiduzzaman, K. M., & Jarar Oulidi, H. (2024). Contextualizing the Smart City in Africa: Balancing Human-Centered and Techno-Centric Perspectives for Smart Urban Performance. In *Smart Cities* (Vol. 7, Issue 2, pp. 712–734). Multidisciplinary Digital Publishing Institute (MDPI). <https://doi.org/10.3390/smartcities7020029>
- Neil, K., Ben, A., & Raymond, O. (2022). Envisaging the Future of Cities.
- OECD. (2023). Urban Studies: Smart City Data Governance Challenges and the Way Forward.
- OECD. (2024). Roundtable on Smart Cities and Inclusive-Growth.
- Ogunseye, N. O., & Bashir, O. (2022). Smart Cities Initiatives in Lagos, Nigeria: Are There Lessons to Learn from the Leading Smart Cities? In *Journal of Urban Research and Development* (Vol. 3). <https://www.researchgate.net/publication/371349327>
- Prakash, D. (2025). Why do smart city projects fail to create impact? Understanding decision-making in smart city policy implementation. *Urban Governance*, 5(1), 45–53. <https://doi.org/10.1016/j.ugj.2025.02.004>
- Santosh, K., Ritu, K., Madiha, I., & Aswathy, N. (2024). Vision and Progress of the Smart Cities Mission: Enhancing Urban Life.
- Scheufler, C., Gutierrez Santamarina, L., Plata, A., Saraiva, A., Asad, A., Larrarte, A., Nuñez, A., Elrahzi, A., Sarkar, A., Usoof, A., Paes, B., Kitio, C., Jayasinghe, C., Rollo, C., Ragan, D., Andersdotter Fabre, E., Ali, E., Holz, E., Aziz, F., ... Sithoe, Y. (2020). Stronger women, stronger cities. www.unhabitat.org
- Shah, N. H., Shah, L. H., Ismail, S. A., Azizan, A., & Khan, D. T. (2024). Challenges and Opportunities in Implementing Smart City Initiatives: A Global Perspective. *Renewable*

Energy Technologies and Modern Communications Systems: Future and Challenges Conference, RETMCS 2024. <https://doi.org/10.1109/IEEECONF63577.2024.10881857>

Shanahan, M., Bahia, K., Carboni, I., Sibthorpe, C., Tiel Groenestege, M., Shannon Baxter, A., Fernando Aquije Ballon, H., & Ouattara Boni, A.-N. (2023). The State of Mobile Internet Connectivity 2023. www.gsmainelligence.com

Shao, J., & Min, B. (2025). Sustainable development strategies for Smart Cities: Review and development framework. *Cities*, 158. <https://doi.org/10.1016/j.cities.2024.105663>

Sipahi, E. B., & Saayi, Z. (2024). The world's first "Smart Nation" vision: the case of Singapore. In *Smart Cities and Regional Development Journal*.

World Bank. (2023). Global Smart City Partnership Program phase 2. www.worldbank.org

World Bank. (2025). Overcoming Barriers to Finance for Municipalities in Low-and Middle-Income Countries. www.worldbank.org

Zaman, M., Puryear, N., Abdelwahed, S., & Zohrabi, N. (2024). A Review of IoT-Based Smart City Development and Management. In *Smart Cities* (Vol. 7, Issue 3, pp. 1462–1501). Multidisciplinary Digital Publishing Institute (MDPI). <https://doi.org/10.3390/smartcities7030061>

Toward Sustainable Digital Infrastructure: Decentralized Resource Utilization Using Existing Edge Systems

S. Kokkarachedu, K. Shehzad, M. Farooq & Prof. Dr.-Ing. habil. J. Marx Gómez
Carl von Ossietzky Universität Oldenburg, 26129 Oldenburg, Germany

Abstract: This paper presents a sustainable approach to digital infrastructure by utilizing existing, locally available computational resources within a decentralized architecture. In contrast to the rising energy demands of centralized cloud systems (Verdecchia et al. 2022; Cioca & Schuszter 2022), this model integrates edge-level hardware, decentralized High-Performance Computing (HPC) through HEAppE Middleware, cloud storage, and open-source monitoring tools such as Prometheus and Grafana (Masciari & Napolitano 2022; Aung et al. 2024) to enable efficient task execution and system observability. The configuration, deployed in support of a European initiative developing an open search index, demonstrates the effective reuse of on-premises infrastructure for large-scale digital services. Two servers were used for web crawling, with Prometheus collecting system metrics and Grafana providing performance visualizations (Leppänen 2021; Dewo et al. 2021). Aligned with advancements in hardware efficiency and software optimization, this decentralized setup promotes resource awareness and edge-level computation (Cioca & Schuszter 2022; Saputra et al. 2022), contributing to more sustainable and resilient digital infrastructures.

Keywords: High-Performance Computing (HPC), Sustainable Infrastructure, HEAppE, Cloud Storage, Distributed Systems, Monitoring

1. Acknowledgement

The European Union under the Horizon 2020 program funds the project with funding number the GA 101070014.



Funded by
the European Union

2. References

- Verdecchia, R., Lago, P. & de Vries, C. (2022): The future of sustainable digital infrastructures: A landscape of solutions, adoption factors, impediments, open problems, and scenarios. In: Sustainable Computing: Informatics and Systems.
- Cioca, M. & Schuszter, I.C. (2022): A system for sustainable usage of computing resources leveraging deep learning predictions. In: Applied Sciences 12(8411), August.
- Masciari, E. & Napolitano, E.V. (2022): The Environmental Cost of High Performance Computing System Simulation. In: Proceedings of the IEEE International Conference.

- Aung, T., et al. (2024): Comprehensive Analysis: Monitoring Apache Kafka with Grafana, JMX Exporter, and Prometheus. In: Proceedings of the 5th International Conference on Advanced Information Technologies (ICAIT).
- Leppänen, T. (2021): Data visualization and monitoring with Grafana and Prometheus. Bachelor's Thesis. Finland: [Institution name not specified].
- Dewo, K.T., et al. (2021): IT Infrastructure Dashboard Monitoring Application Development Using Grafana and Prometheus: A Case Study at Astra Polytechnic School. In: Proceedings of the International Conference.
- Saputra, M.Y.E., et al. (2022): Real-Time Server Monitoring and Notification System with Prometheus, Grafana, and Telegram Integration. In: Proceedings of the International Conference, Politeknik Negeri Malang, Indonesia.
- Chan, Y.-W., et al. (2022): Implementation of a Cluster-Based Heterogeneous Edge Computing System for Resource Monitoring and Performance Evaluation. In: IEEE Access 10, 42638–42650.
- Korontanis, I., et al. (2022): Real-time Monitoring and Analysis of Edge and Cloud Resources. In: Proceedings of the ACM Symposium on Applied Computing (SAC).
- Berardi, D., Callegati, F., Melis, A. & Prandini, M. (2020): Sustainable Infrastructure Monitoring for Security-Oriented Purposes. In: Proceedings of the ACM Conference.
- Masciari, E. & Napolitano, E.V. (2024): The Environmental Cost of High Performance Computing System Simulation. In: Proceedings of the 32nd Euromicro International Conference on Parallel, Distributed and Network-Based Processing (PDP). IEEE.
- Park, J., Han, K. & Lee, B. (2023): Green cloud? An empirical analysis of cloud computing and energy efficiency. In: Management Science 69(3), 1639–1664.
- Svaton, V., Martinovic, J., Krenek, J., Esch, T. & Tomanca, P. (2019): HPC-as-a-Service via HEAppE Platform. In: Proceedings of the International Conference on High Performance Computing.
- Zhang, Q., Cheng, L. & Boutaba, R. (2010): Cloud computing: State-of-the-art and research challenges. In: Journal of Internet Services and Applications 1(1), 7–18.
- Shehabi, A., et al. (2016): United States Data Center Energy Usage Report. Berkeley, CA, USA: Lawrence Berkeley National Laboratory.
- Baru, C., et al. (2010): iRODS Primer: Integrated Rule-Oriented Data System. In: Synthesis Lectures on Information Concepts, Retrieval, and Services. Morgan & Claypool Publishers.
- LEXIS Project (2022): LEXIS Platform. Available at: <https://lexis-project.eu/web/lexis-platform/> [Accessed: July 12, 2022].
- HEAppE Platform (2022): LEXIS Project – HEAppE. Available at: <https://heappe.eu/lexis-project/> [Accessed: 2022].
- LEXIS Platform Documentation (2022): Staging, compression and encryption service. Available at: https://docs.lexis.tech/data_system/staging_api.html [Accessed: 2022].

Leveraging German Expertise in Water Utility for Rwandan Education

J. Stüven, D. Schulte

Carl von Ossietzky Universität, Oldenburg, Deutschland

Abstract: Access to clean and reliable water remains a key development challenge in Rwanda. While the government has made considerable progress in infrastructure expansion, the sustainability of water supply systems is hindered by skill gaps, weak operational management, and limited institutional capacity. This paper presents an approach to capacity building in the water sector developed within the NEWW project – Network for Labour-Market Oriented Education in Waste and Water Management for the Sub-Saharan Region. It explores the mechanisms of knowledge transfer, the integration of practice-based insights into university education at UTAB (University of Technology and Arts Byumba), and the emerging impact of this transnational collaboration on institutional and workforce development.

Keywords: Water Supply, Capacity Building, International Collaboration, Vocational Education, Rwanda, Germany