Knowledge-Aware Process Mining: Conceptual Foundations and IT Artifacts for Process Mining in Knowledge-Intensive Processes

Der Fakultät für Wirtschaftswissenschaften der

Universität Paderborn

zur Erlangung des akademischen Grades

Doktor der Wirtschaftswissenschaften

- *Doctor rerum politicarum* -

vorgelegte Dissertation

von

Katharina Brennig (M.Sc.)

geboren am 09.11.1995 in Paderborn

2025

Acknowledgements

It always seems impossible until it's done.

– Nelson Mandela

Before I was asked whether I could imagine pursuing a PhD, I had never even dreamed that this could be an option for me. Now, more than four years later, I can truly say that it was the best decision I could have made. It has been an exciting journey with many highs and lows—a journey I could not have mastered without the people who continuously supported me along the way.

First and foremost, I would like to express my deepest gratitude to Oliver, who gave me the opportunity to start my doctoral journey at his chair. You gave me the chance to grow, to challenge myself, and to learn things I most certainly would not have learned elsewhere. Thank you for your continuous support, guidance, and constructive feedback. These were indispensable in shaping both my research and this thesis. Your trust in me and the freedom you granted to pursue my own ideas opened countless opportunities for my personal and academic growth.

I am also sincerely grateful to Adela, who accompanied me as my mentor throughout this journey. You helped me find my own path and encouraged me to try new things and believe in myself. You once told me that if you want to stay in academia, it is both nice and important when colleagues eventually become friends—something I wholeheartedly agree with and have tried to keep close to heart.

My heartfelt thanks also go to my academic colleagues—and by now, true friends—who made this journey so special: Matthew, Caro, Laura, Miriam, Christian, Carina, David, Sascha, Marcel, Kevin, Haya, Bernd, and Daniel. Thank you for your unwavering support, for the inspiring discussions, the emotional backup, and the genuine friendships we have built. Whether over a glass of wine, a good coffee with cinnamon rolls, or during the weekly running ritual. I am deeply grateful for having shared this chapter of my life with you.

I also want to thank my family—Maximilian, Beate, and Jürgen—whose circle has wonderfully expanded with Tina, Frank, Leah, Kaja, Emma, and Timon. Your constant support, love, and encouragement to follow my goals in life have been the foundation of everything I have achieved. A special thanks goes to my brother Maximilian—thank you for always standing by my side and believing in me. Your love and support have been invaluable every step of the way.

I am also deeply grateful to my friends—some of whom have been by my side since our school days, others I've met along the way. You know who you are. Thank you for standing by me throughout the ups and downs of this dissertation journey, for your encouragement, your understanding when I disappeared for a while, and for reminding me of life beyond academia. Your presence, patience, and constant support have meant more to me than I can express.

A very special thank you also goes to Laura. You have been by my side every single day of this journey—through all its highs and lows. Some crises would hardly have been manageable without your continuous support. Thank you for believing in me and for being such an incredible friend and confidant—always there for me, no matter the time of day.

My greatest thanks go to Tim—my better half. You always manage to ground me, even in the most difficult situations, and never fail to make me smile. Thank you for your boundless patience, your deep understanding—especially when it comes to my moods—your unwavering love, and your constant positivity, no matter the situation. Having you by my side means the world to me, and honestly, I would not have made it without you. Knowing that you are always there for me is something I treasure deeply. Thank you for being part of my life.

Contents

List of Figures

Part A

List of Tables

Part A

Part A

Synopsis

1 Introduction

1.1 Motivation

We are living in an era increasingly defined by processes. Our world is not made up of things; it is made up of processes that continuously transform everything around us. However, adopting a perspective that views the world as flowing—rather than as consisting of stable states—is far from trivial. Viewing the world through the lens of processes is essential not only for understanding societal change, but also for how organizations operate and evolve.

Organizations are increasingly exposed to continuous and rapidly accelerating change (vom Brocke et al., 2021b). The world becomes more interconnected and dynamic, driven by a range of factors, including, for example, exogenous shocks (e.g., pandemics, natural disasters, or geopolitical changes) (Röglinger et al., 2022), technological advancements (Baiyere et al., 2020; Kerpedzhiev et al., 2021), and evolving customer expectations. This has forced organizations to operate across multiple channels, countries, and systems, often involving distributed teams, complex regulatory environments, and heterogeneous IT landscapes. Moreover, advances in digital technologies are fundamentally reshaping how organizations function, collaborate, and deliver value. In such a volatile and fast-paced environment, it becomes increasingly critical for organizations to control, monitor, and adapt their underlying business processes to remain competitive and resilient (Röglinger et al., 2022). Ensuring sustained competitiveness requires organizations to enhance transparency and agility in designing and continuously adapting their processes (Beverungen et al., 2021). These processes—often formalized as standard operating procedures—define how work is structured and executed within organizational contexts (Alter, 2015).

However, not all organizational processes are standardized or digitized. Some processes are characterized by high variability and unpredictable task sequences, primarily because their execution depends on knowledge workers engaging in complex, interconnected decision-making activities (Di Ciccio et al., 2015). Such processes are defined as knowledge-intensive processes (KIPs) (Di Ciccio et al., 2015; Marjanovic and Freeze, 2011). KIPs (e.g., research & development, product innovation, or healthcare treatment processes) enable leveraging specialized knowledge for the competitive advantage of organizations (Nevo and Chan, 2007; Nonaka and Von Krogh, 2009; Rai, 2011; Alavi and Leidner, 2001; Holsapple and Joshi, 2000; Mahapatra and Sarkar, 2000). Unlike operational business processes, KIPs tend to be less structured and more complex due to the need for contextual interpretation, individual judgment, and experiential knowledge (Di Ciccio et al., 2015; Isik et al., 2012; Eppler et al., 2008; Gronau and Weber, 2004). This human-centered nature and the strong reliance on context-specific expertise contribute significantly to the difficulty of standardizing, digitizing, or fully modeling KIPs (Di Ciccio et al., 2015; Isik et al., 2012; Marjanovic and Freeze, 2011).

Organizations have made sustained efforts to improve and manage business processes, grounded in the understanding that process quality is the foundation for operational performance (Dumas et al., 2018) and that well-managed processes represent a key source of competitive advantage (Davenport, 1992). However, this is inherently difficult for KIPs due to their characteristics. Often, they are managed manually by process participants as existing information technology provides only limited functionality to capture the complexity and flexibility of such processes (Di Ciccio et al., 2015). As a result, organizations lack the holistic perspective required to effectively manage KIPs.

To move beyond this, organizations need methods that allow them to observe, reconstruct, and understand the dynamic behavior of KIPs. Methods and tools from the field of Business Process Management (BPM) (Dumas et al., 2018; Weske, 2019) offer valuable foundations for this purpose. The use of BPM enables the targeted control of processes in an organization by providing a structured approach to design, implement, monitor, and improve organizational processes (Dumas et al., 2018). Using different BPM methods, concepts, and techniques that combine knowledge from information technology and management sciences, processes can be improved (van der Aalst, 2016). Among them, process mining has evolved as one state-of-the-art BPM method to manage operational business processes in a data-driven way (van der Aalst et al., 2012; van der Aalst, 2022). Recently, process mining is being adopted across a growing range of business areas such as finance and controlling (44%), customer service (36%), purchasing/procurement (33%), and accounting (31%). However, areas such as production (21%) or logistics (20%) are gaining increasing attention (Deloitte, 2025). Process mining leverages log data from various information systems (e.g., Enterprise Resource Planning (ERP) or BPM systems), which are consolidated into an event log—a structured dataset containing process-related information (van der Aalst, 2016). For process mining techniques to be applicable, this data must satisfy specific formal properties that ensure the traceability and interpretability of process executions (van der Aalst et al., 2012). With this data, analysts are enabled to discover and improve processes (van der Aalst, 2016), helping organizations to eliminate bottlenecks, enable automation, or reduce process costs (van der Aalst, 2022). Consequently, process mining plays a critical role in enabling organizations to manage their processes in a structured and data-driven manner.

1.2 Problem Statement and Research Questions

Nonetheless, applying process mining to KIPs introduces distinct challenges that hinder its straightforward use. This section outlines these core challenges and defines the overarching research aim. To address them, it establishes three focal areas from which the research questions guiding this thesis are derived, presented in Figure 1.1.

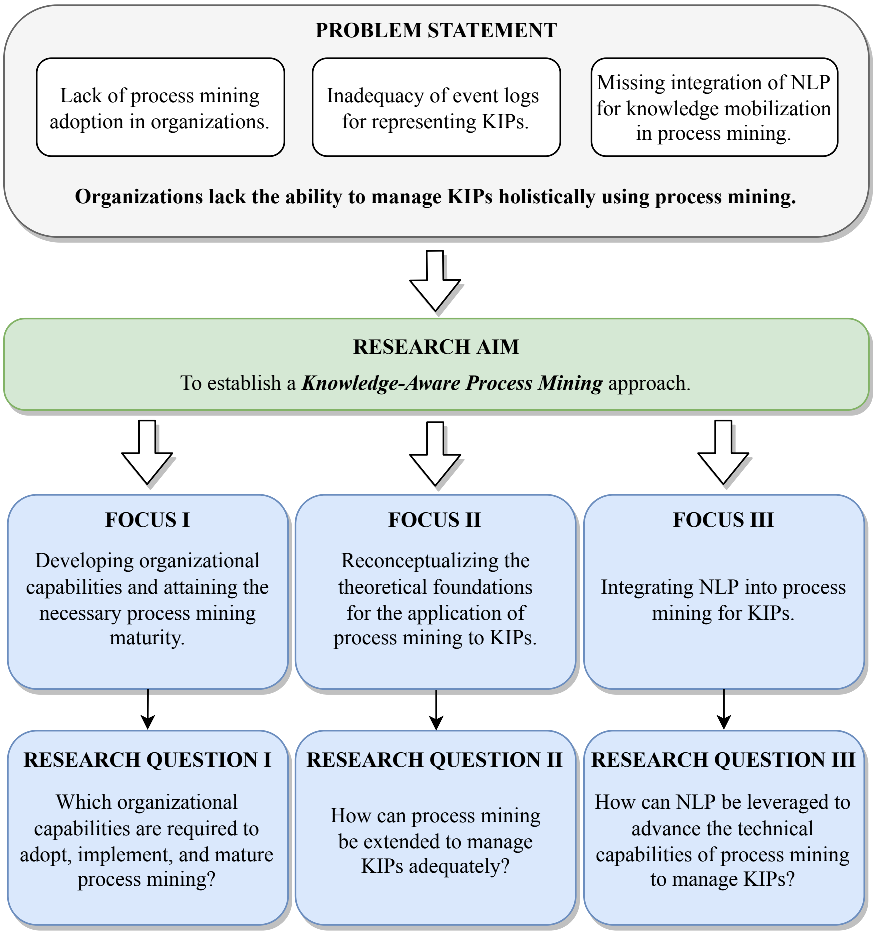

Figure 1.1: Problem Statement and resulting Research Questions

Although 61% of organizations plan to adopt process mining or have initiated pilot projects, widespread and effective implementation remains limited (Daniels, 2022). While organizations aim to leverage process mining for continuous improvement and operational excellence, many face challenges in translating this intention into effective implementation. As a result, many initiatives fail to achieve sustained organizational impact. Even among organizations with established process mining practices, significant obstacles persist, ranging from insufficient managerial support and poor data quality to the complexity of data preparation and integration efforts (Martin et al., 2021).

Further, existing process mining approaches rely on the assumption that all relevant data for analyzing processes is available in event logs. Thus, process mining is particularly applicable to processes with a high degree of standardization and digitalization (e.g., order fulfillment or shipping) (van der Aalst et al., 2012; van der Aalst, 2016, 2022). However, reliance on event logs only containing the activities, control flow, or resource usage associated with a process (van der Aalst et al., 2012; van der Aalst, 2022) presents a significant limitation when applying process mining to KIPs. Event logs cannot adequately represent KIPs in their full breadth and depth as they omit experiences or informal decision-making of individuals that is essential for the execution of KIPs (van der Aalst, 2016; Di Ciccio et al., 2015; van der Aalst et al., 2012). Thus, the more unstructured and tacit parts required for the successful execution, analysis, and understanding of KIPs are missing from event logs.

Additionally, existing process mining approaches are limited in their ability to process and mobilize the process-related knowledge embedded in KIPs. This includes the absence of natural language processing (NLP) approaches in process mining that can extract, represent, and analyze process-related knowledge in a way that reliably supports KIPs. Such NLP-based approaches offer unique potential to process (i.e., generating, storing, retrieving, and transferring (Alavi et al., 2024)) both knowledge embedded in documents, communications, and human reasoning (Feuerriegel et al., 2024; Brown et al., 2020; Alavi et al., 2024).

Taken together, this points to gaps in process mining adoption, the inadequacy of event logs for representing KIPs, and the missing integration of NLP for knowledge mobilization in process mining (see Figure 1.1). Consequently, a holistic management of KIPs through process mining is significantly constrained. To address this, this thesis aims to establish a knowledge-aware process mining approach. Therefore, process mining must be enhanced in three fundamental ways (see Figure 1.1): first, by developing organizational capabilities and attaining the necessary process mining maturity (FOCUS I); second, by reconceptualizing the theoretical foundations for the application of process mining to KIPs (FOCUS II); and third, by integrating NLP into process mining for KIPs (FOCUS III). Addressing these three foci will enable process mining to evolve into a knowledge-aware discipline capable of supporting the dynamic and knowledge-centric nature of KIPs. Building on this rationale, this threefold focus leads to three research questions (see Figure 1.1) that form the foundation of this thesis:

Research Question 1: *Which organizational capabilities are required to adopt, implement, and mature process mining?*

Research Question 2: *How can process mining be extended to manage KIPs adequately?*

Research Question 3: *How can NLP be leveraged to advance the technical capabilities of process mining to manage KIPs?*

1.3 Thesis Structure and Publications

This thesis is structured into two main parts: Part A and Part B. Part A establishes the conceptual and methodological foundations of the thesis. Section 1 introduces and motivates the research endeavor, delineating the relevance and scope of the thesis. Section 2 provides a comprehensive overview of the theoretical and technical underpinnings of process mining, the distinctive characteristics of KIPs, and the principles of NLP-aware approaches. Section 3 presents a summary of the core contributions of each included publication. Section 4 synthesizes these findings in light of the guiding research questions and discusses their theoretical and practical implications. It further concludes the thesis by outlining its limitations and providing directions for future research.

Part B comprises the seven core publications (P1 to P7) that constitute the centerpiece of this cumulative thesis. Table 1.1 provides a detailed overview of each publication. The table includes the publication identifier (#P), the full citation with the author and title details (Author & Title), and the ranking of each publication outlet (JQ4 and CORE 2023). Additionally, it specifies the category of each publication, indicating whether it is a conference or journal paper (Type), and notes its current status as either published or under review (Status). In total, the thesis includes five peer-reviewed conference papers and two journal articles. Among them, six have been published, while one is currently under review.

Table 1.1: Overview of Publications included in this Thesis

#P	Author & Title	Outlet	JQ4	CORE 2023	Type	Status
P1	Brennig, K., Löhr, B., Brock, J., Reineke, M., & Bartelheimer, C. (2024). *Maximizing the Impact of Process Mining Research: Four Strategic Guidelines.*	Americas' Conference on Information Systems	C	/	C	P
P2	Brock, J., Brennig, K., Löhr, B., Bartelheimer, C., von Enzberg, S. & Dumitrescu, R. (2024). *Improving Process Mining Maturity – From Intentions to Actions.*	Business & Information Systems Engineering	B	/	J	P
	Pre-version of the journal publication:[1]					
	Brock, J., Löhr, B., Brennig, K., Seger, T., Bartelheimer, C., von Enzberg, S., Kühn, A., & Dumitrescu, R. (2023). *A Process Mining Maturity Model: Enabling Organizations to assess and improve their Process Mining Activities.*	European Conference on Information Systems	A	/	C	P
P3	Löhr, B., Brennig, K., Bartelheimer, C., Beverungen, D. & Müller, O. (2022). *Process Mining of Knowledge- Intensive Processes: An Action Design Research Study in Manufacturing.*	International Conference on Business Process Management	B	A	C	P
P4	Brennig, K., Bartelheimer, C., Löhr, B., Beverungen, D. & Müller, O. (2025). *Supporting Organizational Knowledge Creation in Knowledge- Intensive Processes through Process Mining.*	Journal of Strategic Information Systems	A	/	J	U
P5	Brennig, K. (2025). *Revealing the Unspoken: Using LLMs to Mobilize and Enrich Tacit Knowledge in Event Logs of Knowledge-Intensive Processes.*	Americas' Conference on Information Systems	C	/	C	P
P6	Brennig, K., Benkert, K., Löhr, B. & Müller, O. (2024). *Text-Aware Predictive Process Monitoring of Knowledge-Intensive Processes: Does Control Flow Matter?.*	Business Process Management Workshops	C	/	C	P
P7	Brennig, K., Kaltenpoth, S. & Müller, O. (2025). *Straight Outta Logs: Can Large Language Models Overcome Preprocessing in Next Event Prediction?.*	Business Process Management Workshops	C	/	C	P

Type: C: Conference Paper, J: Journal Paper

Status: P: Published, U: Under Review

[1] This paper is not included in Part B of the thesis, as its results are already incorporated into the corresponding journal publication. Apart from minor refinements to the definitions of the 23 elements—made in response to reviewer feedback—no substantive changes have been introduced.

2 Research Background

This chapter outlines the relevant research background and theoretical foundations underlying this thesis. Section 2.1 presents a comprehensive overview of process mining, detailing its conceptual and technical foundations. Section 2.2 introduces KIPs and focuses on the specific characteristics and challenges of applying process mining to KIPs. Further, Section 2.3 examines the foundations of NLP and explores its emerging application within the domain of process mining and knowledge management.

2.1 Foundations of Process Mining

Organizations continuously strive to improve and manage their processes, recognizing that process quality is essential for operational performance and a key source of competitive advantage (Dumas et al., 2018; Davenport, 1992). More specifically, a process represents a "*collection of inter-related events, activities and decision points that involve a number of actors and objects, and that collectively lead to an outcome that is of value to at least one customer*" (Dumas et al., 2018, p. 6-7). BPM provides a comprehensive set of methods and tools for the systematic design, execution, monitoring, and optimization of organizational processes (Dumas et al., 2018; Weske, 2019).

Within this context, process mining has evolved as a state-of-the-art, data-driven approach that bridges BPM and data science to uncover and analyze the actual execution of processes based on digital trace data (van der Aalst et al., 2012; van der Aalst, 2022). Existing process mining approaches are particularly applicable to highly digitalized and standardized processes as they work best with high volumes of data (van der Aalst et al., 2012; van der Aalst, 2022). Process mining leverages event data captured in event logs, which are extracted from information systems such as ERP or BPM systems. These logs document structured metadata about process executions (e.g., activity name, timestamp, and resource) and serve as the foundation for various analytical techniques (van der Aalst and Dustdar, 2012). Core process mining techniques include process discovery (i.e., generating as-is process models from event logs, represented for example by Petri nets or directly-follows graphs), conformance checking (i.e., assessing the alignment between observed and prescribed process behavior), performance analysis (i.e., evaluating process execution metrics), and comparative process mining (i.e., benchmarking across multiple event logs) (van der Aalst, 2022, 2016).

Building on these capabilities, more recent advances have introduced predictive and action-oriented forms of process mining, which extend the analytical horizon from descriptive to prescriptive insights (Marquez-Chamorro et al., 2018; Weinzierl et al., 2020; van der Aalst, 2022). Predictive process mining (PPM) aims to anticipate the future behavior of ongoing process instances (Di Francescomarino and Ghidini, 2022; Weske, 2019), such as process outcomes, sequences of future activities, or (remaining) lead time, based on historical patterns (Di Francescomarino and Ghidini, 2022; Marquez-Chamorro et al., 2018). Action-oriented process mining transforms diagnostic insights into concrete actions such as suggesting next steps or reallocating resources, thus closing the loop between analysis and execution (van der Aalst, 2022). These self-learning approaches open the door to an even more dynamic handling of future events (Heinrich et al., 2021) or concept drift (Sato et al., 2022) and enable real-time analysis, outcome forecasting, and recommendations for process optimization (Marquez-Chamorro et al., 2018; Weinzierl et al., 2020; van der Aalst, 2022).

With these methods in hand, process mining enables the design, implementation, control, and analysis of business processes involving people, organizations, documents, and other information sources (Dumas et al., 2018; Weske, 2019; van der Aalst, 2022). In doing so, process mining helps organizations to eliminate bottlenecks, enable automation, or reduce process costs (van der Aalst, 2022). Due to its capability to process and analyze data in real-time (van der Aalst, 2022; Davenport and Spanyi, 2019), live insights on the execution of a process can be gained, and ad hoc evidence-based decisions are possible (Grisold et al., 2020). This enables organizations to cope with the ever-increasing complexity and volatility of today's business processes (Grisold et al., 2020; Pentland et al., 2021; Wurm et al., 2021; Kipping et al., 2022) and generate business value (Badakhshan et al., 2022). Still, many organizations struggle to turn their ambitions into lasting impact due to obstacles such as lacking management support, poor data quality, and complex data preparation (Martin et al., 2021).

2.2 Process Mining in Knowledge-Intensive Processes

In addition to well-structured and standardized processes, organizations also engage in KIPs, such as product development, innovation, or healthcare treatment processes (Di Ciccio et al., 2015; Isik et al., 2012; Eppler et al., 2008; Bahrs and Müller, 2005). KIPs can be defined as "*processes whose conduct and execution are heavily dependent on knowledge workers performing various interconnected, knowledge-intensive decision-making tasks. KIPs are genuinely knowledge, information and data centric and require substantial flexibility at design- and run-time*" (Di Ciccio et al., 2015, p. 5). Thus, they rely heavily on human-centered, knowledge-intensive activities, contextual decision-making, and the individual experience of process participants (Di Ciccio et al., 2015; Little and Deokar, 2016; Marjanovic and Freeze, 2011; Isik et al., 2012; Eppler et al., 2008; Gronau and Weber, 2004). This reliance results in high complexity, ambiguous inputs and outputs, and considerable variability. As such, KIPs require significant flexibility and are inherently difficult to predict, formalize, and manage (Di Ciccio et al., 2015; Isik et al., 2012; Marjanovic and Freeze, 2011; Eppler et al., 2008; Gronau and Weber, 2004). Accordingly, the depth and availability of process-related knowledge in KIPs evolve dynamically through the actions and decisions of the individuals involved (De Almeida Rodrigues Gonçalves et al., 2023).

Process-related knowledge can be categorized into explicit and tacit process-related knowledge (Bahrs and Müller, 2005; Di Ciccio et al., 2015). Explicit process-related knowledge can be codified and stored in a knowledge base, defining relevant knowledge objects, data, information, and artifacts that constitute process context and execution state. In contrast, tacit process-related knowledge resides in the skills and experiences of process participants and is reflected in their everyday practices and decision-making (Di Ciccio et al., 2015). Thus, each process participant possesses tacit process-related knowledge linked to the execution of their specific tasks within a given process instance, underscoring its central importance in KIPs (Marjanovic and Freeze, 2011; Di Ciccio et al., 2015). While explicit process-related knowledge can be codified, expressed in formal language, and documented in text or visuals, enabling its use across contexts (Nonaka and Takeuchi, 1995; Nonaka and Von Krogh, 2009), tacit process-related knowledge is more difficult to formalize and communicate as it is personal and context-specific (Nonaka and Takeuchi, 1995; Polanyi, 1967). Clinical decision-making, as part of KIPs in healthcare, exemplifies this distinction. It draws on explicit process-related knowledge such as medical evidence, clinical guidelines, and patient records, as well as on tacit process-related knowledge embedded in the clinician's expertise and experiential judgment (Di Ciccio et al., 2015). Thus, effective utilization of tacit process-related knowledge is crucial for achieving the overarching objectives of KIPs.

As a consequence, organizations should foster a culture of knowledge exchange in which tacit process-related knowledge is mobilized and transformed into explicit process-related knowledge—and vice versa—that can be shared, analyzed, and reused (Nonaka and Takeuchi, 1995; Gold et al., 2001). Through this, a well-grounded knowledge base, represented in KIP-specific event logs, should be established. To facilitate this, organizations can draw upon established approaches from the field of knowledge management (Di Ciccio et al., 2015). The theory of organizational knowledge creation, particularly the SECI-model proposed by Nonaka and Takeuchi (1995), provides a conceptual foundation for this transformation. The model describes four modes of knowledge conversion. In the mode of socialization, individuals acquire tacit process-related knowledge by learning from the experiences of others. For example, in healthcare processes, a junior doctor learns diagnostic intuition by observing and shadowing senior doctors during patient rounds. Similarly, in product development, tacit insights are transferred during collaborative workshops or informal discussions. Externalization involves articulating tacit process-related knowledge to make it explicit and shareable. This includes activities such as concept creation, quality improvement meetings, or documenting expert insights during prototyping for product development. In healthcare contexts, experienced doctors may formalize treatment heuristics into clinical guidelines. Combination refers to synthesizing explicit process-related knowledge into structured knowledge systems. In product development, this might involve compiling market research, patent databases, customer feedback, and past project reports to create a new product concept or feature roadmap. In healthcare, doctors combine patient records, diagnostic results, and clinical evidence to create personalized treatment plans. Finally, internalization describes how individuals absorb explicit process-related knowledge through experience, such as learning by doing, thereby enriching their own tacit process-related knowledge base. Medical staff, for instance, internalize clinical protocols by applying them in real cases, gradually developing intuitive judgment that informs complex decision-making (Nonaka and Takeuchi, 1995; Nonaka et al., 2000; Nonaka and Von Krogh, 2009). These four modes serve as the driving force behind the entire process of knowledge creation and are not independent of each other as they cause a spiral in their interaction over time. This conversion process begins at the individual level and progressively scales across organizational units to create a shared knowledge base (Nonaka and Takeuchi, 1995; Nonaka et al., 2000).

However, mobilizing the tacit process-related knowledge inherent in KIPs remains a persistent challenge, as this type of process-related knowledge is rarely externalized or systematically shared within organizations. Thereby, its broader dissemination and integration into organizational knowledge structures are impeded (Di Ciccio et al., 2015). The effectiveness of knowledge management systems (KMS) in this regard is limited, as it depends on their integration into broader organizational systems and processes (Alavi and Leidner, 2001; Nevo and Chan, 2007). Ideally, such systems should interact with both technological and social process-related infrastructures to contextualize and operationalize process-related knowledge in KIPs effectively (Bhatt, 2001). Therefore, organizations must go beyond static data storage and retrieval. Instead, they need to establish dynamic, integrated systems that support continuous knowledge exchange and contextual interpretation. Without these capabilities, the knowledge essential for achieving the goals of KIPs remains fragmented or inaccessible, impeding both process improvement and innovation.

These limitations in knowledge management also affect data-driven approaches such as process mining, which rely heavily on structured, explicit event logs. Although first process mining approaches target KIPs (Remus and Lehner, 2000; Bahrs and Müller, 2005; Richetti et al., 2017; Pérez-Castillo et al., 2011; Khanbabaei et al., 2019; Benner-Wickner et al., 2015; Terziev et al., 2015; Dunzer et al., 2021; Berriche et al., 2015; Munoz-Gama et al., 2022), the available solutions only use a subset of the knowledge relevant to KIPs (Isik et al., 2012; Bahrs and Müller, 2005), thereby compromising the quality of process mining results (van der Aalst, 2015). This stems from the assumption that the essential process-related data, information, and knowledge for process analysis are captured in event logs. This typically contains the activities, control flow, and resource usage (van der Aalst and Dustdar, 2012; van der Aalst, 2022; Di Ciccio et al., 2015), while overlooking the experiences or informal decision-making of individuals (i.e., tacit process-related knowledge) critical to KIPs (Di Ciccio et al., 2015). Consequently, process mining struggles to capture the full breadth and depth of KIPs. Subsequently, abstracting data from specific instances and their context, deviations in a KIP cannot be interpreted on a process level anymore as order and time of activities are emphasized over context (van der Aalst, 2016; Rosemann et al., 2008; van der Aalst and Dustdar, 2012). This leads to an ever-increasing spiral in which KIPs can only be analyzed, redesigned, and executed on the basis of their explicit process-related knowledge.

2.3 Enhancing Process Understanding and Knowledge Work with Natural Language Processing

The field of NLP comprises all approaches that enable machines to understand, interpret, or generate human language (Jurafsky and Martin, 2025). Over the last decades, the field of NLP has undergone diverse paradigm shifts. Initially, rule-based approaches (e.g., hand-crafted grammars or pattern-matching systems such as ELIZA) were prevalent (Jurafsky and Martin, 2025). These were later complemented and largely superseded by statistical approaches (e.g., n-gram language models or Naive Bayes classifiers), which leveraged large corpora to model language probabilistically (Manning and Schütze, 2000; Jurafsky and Martin, 2025). Over the years, however, the field has evolved from machine learning (e.g., logistic regression, support vector machines) (Sebastiani, 2002; Jurafsky and Martin, 2025; Young et al., 2018) to deep learning NLP approaches (e.g., recurrent neural networks, long short-term memory) (Jurafsky and Martin, 2025). The advent of deep learning methods in NLP has markedly advanced the field (Brown et al., 2020; Jurafsky and Martin, 2025; Young et al., 2018). Particularly, the Transformer architecture has led to improvements in training efficiency and natural language understanding and generation (Vaswani et al., 2017).

The Transformer's encoder-decoder architecture encodes input sequences into contextual representations, which the decoder then uses to generate coherent outputs (Vaswani et al., 2017). This effectively supports sequence-to-sequence tasks such as translation (Vaswani et al., 2017). In the early developments of transformer-based models, encoder-only models such as Bidirectional Encoder Representations from Transformers (BERT) advanced natural language understanding, including text classification, and sentiment analysis (Devlin et al., 2019). In contrast, decoder-only models, such as Generative Pre-trained Transformers (GPT), were capable of natural language generation, including summarization (Radford et al., 2018), revolutionizing the field of NLP (Feuerriegel et al., 2024). Leveraging the scaling laws (Kaplan et al., 2020) and developing specialized models that comprise only the decoder part of the Transformer architecture has led to the emergence of large language models (LLMs).

Notably, decoder-only LLMs enable the processing of large-scale textual data and the generation of contextually relevant outputs (Vaswani et al., 2017; Feuerriegel et al., 2024). These generative models are designed to produce meaningful and coherent text by predicting the next token (i.e., word) t_{n+1} in a sequence. They achieve this by learning the conditional probability distribution $P(t_{n+1}|t_1...t_n)$, allowing them to generate text autoregressively based on the preceding context (Shanahan et al., 2023). Thus, almost any NLP task, including natural language understanding, can be modeled as word prediction (Jurafsky and Martin, 2025).

Typically, LLMs comprise billions of parameters and are trained on large-scale text corpora, including books, websites, and code (Brown et al., 2020; OpenAI et al., 2023). This enables them to understand language, its context, and its meaning, allowing near-human performance in answering questions, performing tasks, and generating code or markup (e.g., HTML, XML) (Brown et al., 2020; Sui et al., 2024; Chen et al., 2021; Ouyang et al., 2022; Feuerriegel et al., 2024). However, LLMs are prone to generating outputs that, while syntactically and semantically plausible, may be factually incorrect—a phenomenon known as hallucination (Ji et al., 2023; Feuerriegel et al., 2024). Thus, LLMs may produce content that lacks factual grounding (Feuerriegel et al., 2024). These outputs can mislead users by presenting misinformation indistinguishable from accurate content (Spitale et al., 2023). However, to apply LLMs on specific domain data, to specific output representations, or to mitigate hallucinations, they can also be fine-tuned (Feuerriegel et al., 2024; Alavi et al., 2024; Wu et al., 2024).

LLMs have found widespread adoption in process mining (Feuerriegel et al., 2024; Vidgof et al., 2023; Dumas et al., 2023). They can process directly-follows graphs, Petri nets (Berti et al., 2024b), and XES-formatted event logs (Berti et al., 2024a; Berti and Qafari, 2023). Their capabilities span answering process-related questions (Berti et al., 2024b), mitigating bias (Berti et al., 2024a), improving outcomes (Berti et al., 2024b), supporting log abstraction (Brzychczy et al., 2025), and generating artificial event logs (Redis et al., 2024). LLMs also contribute to predictive and prescriptive process monitoring by recommending and explaining process interventions (Berti et al., 2024a; Kubrak et al., 2024; Käppel et al., 2024), thereby improving transparency of the origin of predictions and recommendations (Kubrak et al., 2024) and decision-making (Käppel et al., 2024). To improve the effectiveness of LLMs in process mining applications, techniques such as fine-tuning, prompt engineering, and the incorporation of iterative feedback can be employed (Jessen et al., 2023; Berti et al., 2024a).

Beyond enhancing existing process mining capabilities, LLMs offer potential to bridge the gap between explicit and tacit process-related knowledge—particularly in human-centered processes where conventional data-driven methods fall short. Their growing influence on knowledge management is reflected in their impact on some phases of the SECI model (Alavi et al., 2024; Korzynski et al., 2023; Sumbal and Amber, 2024) and its use in traditional KMS (Alavi et al., 2024). As conversational agents, LLMs support socialization by fostering collaboration and knowledge sharing (Sumbal and Amber, 2024; Zheng et al., 2024; Alavi et al., 2024). For externalization, they enable natural language querying and learn from user feedback (Alavi et al., 2024). They also support knowledge combination by synthesizing insights, updating databases, and identifying knowledge gaps (Korzynski et al., 2023; Sumbal and Amber, 2024; Alavi et al., 2024). In doing so, LLMs enhance organizational learning and decision-making by extracting, structuring, and retrieving knowledge (Alavi et al., 2024; Zhang et al., 2024; Korzynski et al., 2023). However, hallucinations, i.e., the generation of plausible yet factually incorrect information, (Feuerriegel et al., 2024) pose a significant challenge to knowledge creation with LLMs (Alavi et al., 2024). Without critical evaluation, LLM-generated misinformation may enter organizational knowledge bases (Alavi et al., 2024).

3 Research Contributions

The limited ability of organizations to holistically manage KIPs with process mining is the central challenge addressed by the seven research papers (P1-P7) presented in this dissertation. Each paper tackles a distinct aspect of how process mining can be advanced toward a knowledge-aware discipline, ultimately enabling its effective application in KIPs. Figure 3.1 visualizes how the individual papers interrelate, showing their cumulative contribution to developing a knowledge-aware process mining approach.

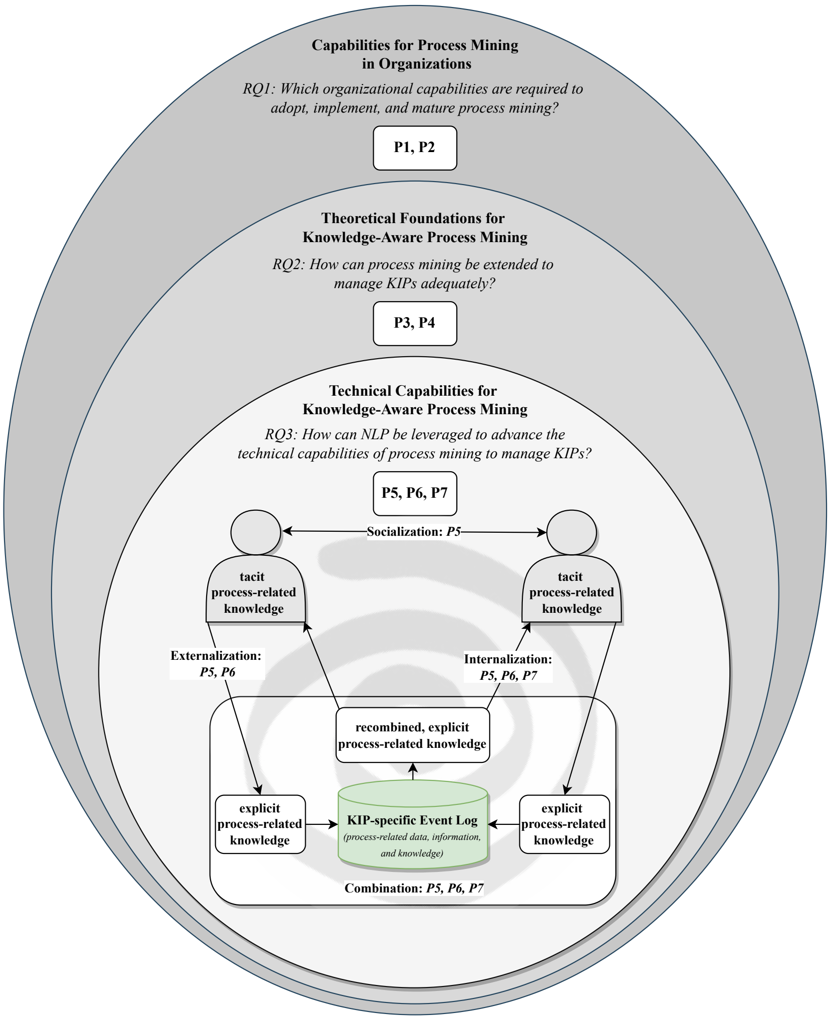

Figure 3.1: Overview of Research Contributions and their Interrelationships

The outer layer focuses on the capabilities for process mining in organizations. The middle layer focuses on establishing theoretical foundations for knowledge-aware process mining, grounded in the SECI model (Nonaka and Takeuchi, 1995). The inner layer presents technical capabilities that integrate NLP to mobilize process-related knowledge in KIPs. This layer represents the circular and iterative nature of knowledge mobilization in KIPs (Nonaka and Takeuchi, 1995). Thus, this layer extends the theoretical foundations by introducing NLP as key enabler, given its ability to process and extract insights from textual data. Tacit process-related knowledge must first be externalized to become explicit. Through combination, this explicit process-related knowledge—along with relevant data and information—can be integrated into a KIP-specific event log. Based on this, a knowledge base containing recombined, explicit process-related knowledge (e.g., predictions or recommended actions) can be constructed. Such a system supports process participants in internalizing new process-related knowledge through practical application.

Each layer aligns with one of the three research questions guiding this thesis. Accordingly, each paper can be mapped to one of the research questions, structured as follows:

- P1 and P2 address RQ1: *Which organizational capabilities are required to adopt, implement, and mature process mining?*
- P3 and P4 address RQ2: *How can process mining be extended to manage KIPs adequately?*
- P5, P6, and P7 address RQ3: *How can NLP be leveraged to advance the technical capabilities of process mining to manage KIPs?*

An overview of the main research contributions addressing the research questions is provided in Table 3.1. The following sections summarize the key contributions of each paper, highlighting how they address the challenges of building organizational capabilities, developing theoretical foundations, and implementing technical approaches for mobilizing tacit process-related knowledge and enriching event logs. Thus, advancing process mining toward a knowledge-aware approach moving from organizational readiness (P1, P2), through conceptual design (P3, P4), to technical realization (P5-P7).

Table 3.1: Main Research Contributions

RQ	#P	Contributions
RQ1	P1	Four strategic guidelines to assess and enhance the relevance of process mining results.
	P2	A process mining maturity model comprising 5 factors and 23 elements and a set of 30 possible actions to support organizations in advancing their process mining capabilities.
RQ2	P3	Five design principles that establish theoretical foundations for applying process mining in KIPs.
	P4	Extends the conceptual groundwork laid in P3 and develops five improved design principles that establish theoretical foundations for applying process mining in KIPs. The design principles are grounded in the theory of organizational knowledge creation.
RQ3	P5	A LLM-based framework to operationalize the mobilization of tacit process-related knowledge, technically realizing the SECI-based conceptual approach.
	P6	A text-aware PPM approach applying encoder-only models for natural language understanding and neglecting the control-flow of KIPs.
	P7	A LLM-based approach which can directly generate next event predictions from XES-formatted, knowledge-enriched event logs.

3.1 Paper 1 – Maximizing the Impact of Process Mining Research: Four Strategic Guidelines

The IS literature offers limited insights into how organizations can effectively adopt, integrate, and assess process mining to generate business value (vom Brocke et al., 2021a; Badakhshan et al., 2022), partly due to the scarcity of studies examining implementation within existing organizational structures (van Eck et al., 2015; Aguirre et al., 2017). To guide organizations and clarify the value of process mining adoption, P1 formulates the following research question: *"To what extent does the process mining literature in IS reflect on the utility of its results for organizations and what are the implications for future research?"*

To address this question, P1 conducts a systematic literature review (SLR) (Simons et al., 2009; Liberati et al., 2009). The SLR focuses on literature addressing the integration of process mining within real-world organizational contexts and its practical impacts. The search has been limited to contributions in the IS knowledge base and concentrated on papers published since 2011, when the process mining manifesto was published by van der Aalst et al. (2012), as this was the cornerstone of the common understanding of process mining in the community. 28 final papers have been identified and organized in a concept matrix. The concept matrix revealed that the vast majority of process mining contributions in the IS knowledge base focus on applying process mining techniques and artificially developed artifacts. They often neglect to identify requirements from an application domain and to evaluate the usefulness of an artifact.

P1 develops four strategic guidelines to assess and enhance the relevance of process mining results. These guidelines aim to guide researchers better in carving out the practical implications of their contributions:

- **Guideline 1:** Start with identifying a use case in a real-world organization.
- **Guideline 2:** Actively report on design objectives and decisions.
- **Guideline 3:** Evaluate artifacts with field data.
- **Guideline 4:** Discuss implications and quantify the business value.

These guidelines emphasize the need for research that yields practically relevant and managerially impactful insights, as the utility and effectiveness of process mining often remain invisible in organizational practice. This aligns with the dual mission of design-oriented IS research: developing theory for design and action while addressing business problems (Hevner et al., 2004; Gregor, 2006). The lack of clarity regarding the practical utility of many process mining artifacts exposes a gap in the IS knowledge base, despite the prevalence of the Design Science Research (DSR) paradigm, which emphasizes practical relevance (Hevner, 2007). Integrating the relevance cycle in research endeavors is fundamental, especially in application-oriented research such as process mining.

3.2 Paper 2 – Improving Process Mining Maturity: From Intentions to Actions

Although organizations aim to use process mining to enable agile and adaptive BPM, widespread implementation is often hindered by deficiencies in process mining readiness across multiple dimensions (Daniels, 2022; Martin et al., 2021; Badakhshan et al., 2022; Reinkemeyer, 2020; van der Linden, 2021; Reinkemeyer et al., 2022). Even among organizations with established process mining practices, significant obstacles persist, ranging from insufficient managerial support and poor data quality to the complexity of data preparation and integration efforts (Martin et al., 2021). Addressing this intention-action gap requires clear guidance for organizations on enhancing their process mining readiness (Dunzer et al., 2021; Martin et al., 2021; Beverungen et al., 2021).

Therefore, P2 develops a Process Mining Maturity Model (P3M), designed to support organizations in evaluating and systematically advancing their process mining capabilities. The model was developed using an IT-related maturity model methodology (Becker et al., 2009) grounded in the DSR paradigm (Hevner et al., 2004). The method is based on the properties and development history of previous maturity models and covers the steps from ideation to publishing and usage, emphasizing understandability, reproducibility, and iterative refinement based on literature and implications from the specific context. As maturity models are context-sensitive and can become outdated, regular review ensures continued relevance. Design decisions were further informed by Kühn et al. (2013) to enhance accessibility. A 30-month development project was initiated, structured into five phases. Phase one included problem scoping, requirements elicitation, and the selection of a development strategy. Phase two developed the initial P3M in collaboration with a manufacturing company to ensure both rigor and relevance (Hevner, 2007). Phase three focused on refining the model's factors and elements in close cooperation with the organizational partner. Phase four applied and evaluated the P3M with a second organization, while phase five focused on identifying actions to enhance organizational readiness for process mining. The actions are derived from eleven qualitative interviews (Myers and Newman, 2007) conducted with two practitioner groups: "internal" users, driving process mining within their organizations, and "external" users, whose organizations support others in adopting process mining.

The P3M[1] comprises five factors (i.e., *Organization, Data Foundation, Peoples' Knowledge, Scope of the PM Activity*, and *Governance*) with 23 elements (see Figure 3.2).

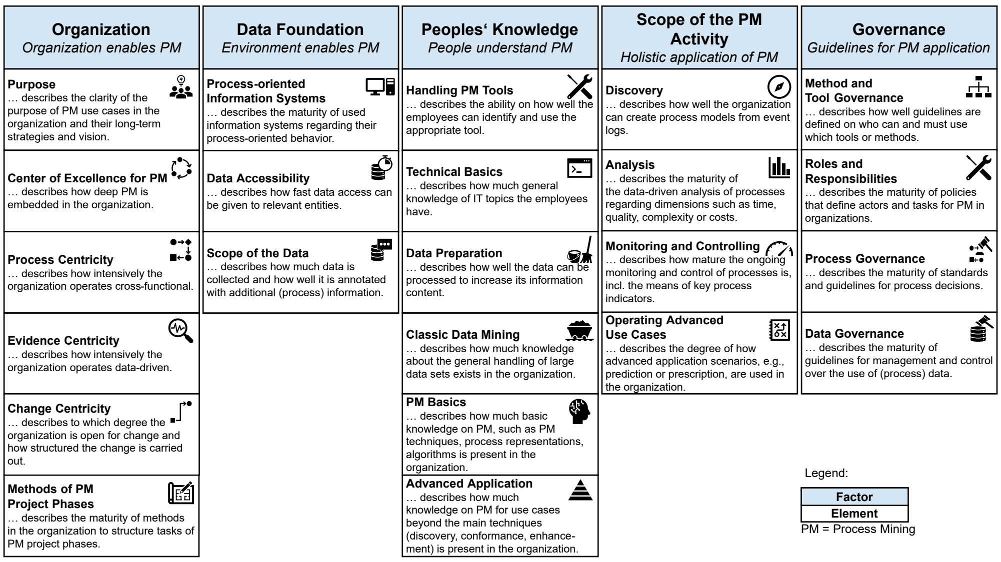

Figure 3.2: P2 - Overview of the Process Mining Maturity Model (P3M) (Brock et al., 2024)

Each element has five maturity stages, ranging from *Initial* to *Optimizing* (e.g., adapted to the maturity stages of Paulk et al. (1991) and Rosemann and de Bruin (2005)). At

[1] Detailed description: `https://www.its-owl.de/process-mining-maturity-model/`.

the highest maturity level, organizations demonstrate awareness of improvement opportunities, adaptability to evolving business needs and market dynamics, and maintain structures to advance process mining. Progressing to the next maturity stage requires fulfilling the previous one (Rosemann and de Bruin, 2004). As the maturity stages are ordinal, the effort to advance varies by each element and level. The maturity model supports assessing as-is organizational maturity and prioritizing improvement areas. Additionally, to guide organizations in advancing their maturity, 30 possible actions have been defined. Organizations can use these actions to establish a more responsive and dynamic BPM environment utilizing process mining. A comprehensive overview of these actions is depicted in Table 3.2.

Table 3.2: P2 - Identified Actions to Improve the Process Mining Maturity of Organizations (Brock et al., 2024)

Factor		Action	Explanation
Organization	A-O1	Raise awareness in management workshops	Let the top management experience process mining, e.g., by conducting a workshop where they mine an artificial process.
	A-O2	Raise awareness in management circles	Present findings from process mining projects to the management.
	A-O3	Anchor initiative centrally	One centralized initiative bundles all process mining activities within an organization.
	A-O4	Anchor initiative in a hybrid setup	Connect a centralized initiative to multiple decentralized initiatives, e.g., in the corporate head quarters and branches.
	A-O5	Human-centered communication campaign about processes and data	Highlight the importance of data and processes, e.g., by posting short videos on the intranet.
	A-O6	Involve the IT department early	Communicate with the IT department early, and identify benefits for the them.
	A-O7	Combine the process mining initiative with larger digital transformation projects	Use larger digital transformation initiatives to validate and utilize process mining in an innovative environment.
Data Foundation	A-D1	Implement a central data repository	Initialize a central data infrastructure, e.g., a data lake, to gather event data from different sources.
	A-D2	Iteratively include new information systems	Do not try to include all information systems at once, but step-by-step.
	A-D3	Implement the connectors to data sources	Utilize the connectors offered by process mining vendors to connect with the leading information system.
	A-D4	Increase data quality with automation	Implement automation (e.g., RPA) to automatically maintain master data and increase data quality.
	A-D5	Manually export first data for validation	Especially for on premise systems, using connectors is difficult. Manually export data for a low-effort start.
	A-D6	Strive for perfection, deploy pragmatism	Most data can be used as is. Focus on that data to begin with.
Peoples' Knowledge	A-P1	Determine an internal multiplier	Identify a person or group of persons to collect and share knowledge about process mining within the organization.
	A-P2	Store knowledge in a wiki	Externalize knowledge by documenting it in a knowledge base.
	A-P3	Train domain experts by involving them in the analysis	Train domain experts by including them in the data pre-processing and mining phases.
	A-P4	Conduct trainings with vendors	Conduct the trainings of the respective process mining tool vendor.
	A-P5	Specify trainings for the respective departments	Customize training to the specific needs of different departments (e.g., IT vs. business department).
	A-P6	Utilize online-classes for self-study	Various online classes on process mining exist, where practitioners can train on particular aspects of process mining.
	A-P7	Create a technical and a functional documentation	Create documentations for technical aspects and functional aspects, such as methods used or analysis steps taken.
Scope of PM Activity	A-S1	Systematically identify new use cases for your initiative	Determine use cases on the basis of benefits, interests, and data availability.
	A-S2	Showcase previous use cases to gain attention of business units	Utilize regular meetings to demonstrate the possibilities of process mining to draw business attention.
	A-S3	Start with process discovery	Start with process mining by applying process discovery, because it is the foundation for other techniques.
	A-S4	Utilize classical data analytics techniques to gain general insights	Generate ordinary descriptive data analysis plots for domain experts.
	A-S5	Gradually add new use cases to the initiative	Do not overload the organization with too many use cases, but work on them in a step-by-step fashion.
Governance	A-G1	Create short-term data usage agreement	Create a written document in collaboration, e.g., with the works council, for conducting first PoC projects.
	A-G2	Create long-term data usage agreement	Create a written document concerning multiple process mining projects, addressing data and privacy concerns.
	A-G3	Involve the works council early	Involve the works council in the project to show that no individual performance is measured and jobs will not necessarily be rationalized.
	A-G4	Develop a clear set of roles	Aside from classical roles such as process mining or domain expert, also consider roles such as an (analysis dashboard) user.
	A-G5	Select the right process mining tool	Consider various aspects when selecting a vendor, and do not hesitate to test multiple vendors.

3.3 Paper 3 – Process Mining of Knowledge-Intensive Processes: An Action Design Research Study in Manufacturing

Existing process mining methods are designed for highly digitalized and standardized processes (van der Aalst, 2016). However, organizations also conduct processes that are more knowledge-intensive and involve creative activities, require flexibility and decision autonomy, and target external goals like customer satisfaction (Di Ciccio et al., 2015). As a result, KIPs rely heavily on the tacit process-related knowledge of process participants, which often resists codification and remains unavailable for process mining—unlike explicit process-related knowledge (Di Ciccio et al., 2015; Nonaka and Takeuchi, 1995). Thus, KIPs often lack extensive digital event logs, limiting the direct application of common process mining techniques without adaptation.

Building on the organizational foundations established in P1 and P2, P3 designs and evaluates a process mining approach for KIPs. Therefore, P3 conducts Action Design Research (ADR), which emphasizes the organizational context in shaping research and artifact development (Sein et al., 2011). Rooted in both action and design research, ADR comprises four phases: (1) problem formulation, (2) building, intervention, and evaluation (BIE), (3) reflection and learning, and (4) formalization of learning. Through this, ADR promotes close collaboration between researchers and practitioners (Sein et al., 2011). Applying ADR, P3 presents two manufacturing cases (i.e., product innovation and engineer-to-order) aiming to analyze and improve their KIPs through process mining. Insights from the literature and the use cases (i.e., data from 27 interviews and 49 workshops) informed the development of initial propositions and a prototypical IT artifact. The prototype represents a set of mock-ups of a process analytics tool that is focused on the needs of KIPs. Based on the preliminary results, an evaluation was conducted with practitioners of the core ADR team, which enabled the development of the resulting design principles (see Table 3.3) and the IT artifact.

The design principles are structured in line with Gregor et al. (2020), presenting the aim, implementer and user, the mechanism, and the rationale. The user or implementer can be the organization, a process analyst (e.g., process manager), and process participants (e.g., process engineer, process executor) (Dumas et al., 2018). Table 3.3 shows the developed design principles (DP). DP1-DP3 build the foundation for DP4 and DP5, addressing KIP challenges such as heterogeneous, incomplete, and variable event logs (Nguyen, 2017). To address this, DP1 suggests decomposing processes into stages and gates for analysis while maintaining flexibility (Seidel et al., 2010). DP2 and DP3 pro-

Table 3.3: P3 - Design Principles (Löhr et al., 2022)

Design Principle 1: Decision- vs. Knowledge-intensive Activities	
Implementer	Organizations using process mining for KIPs
User	Process participants & process analyst
Aim	To balance between quality insurance and flexibility in performing KIPs.
Mechanism	Distinguish decision- from knowledge-intensive activities.
Rationale	We draw on Seidel et al. (2010) framework of *pockets of creativity* that aims at conceptualizing creativity within business processes.
Design Principle 2: Capture Domain-Specific Knowledge	
Implementer	Organizations using process mining for KIPs
User	Process participants
Aim	To capture process-relevant domain knowledge.
Mechanism	Enhance the event log with relevant unstructured data (e.g., codifiable knowledge, including business documents, drawings, notes).
Rationale	In *adaptive case management*, actionable knowledge is collected from process participants and required information for processing the case is stored (Osuszek and Stanek, 2015). The idea has already been transferred to KIPs by Herrmann and Kurz (2011).
Design Principle 3: Define Process-External Goals	
Implementer	Organizations using process mining for KIPs
User	Algorithms
Aim	To learn relationships between process execution and higher-order business goals.
Mechanism	Annotate event logs with process-external goals.
Rationale	Goals of process instances can vary and decisions on process goals are often based on incomplete knowledge. Considering *adaptive case management* (Osuszek and Stanek, 2015), we extend this idea to also take the integration of external factors into account.
Design Principle 4: Retrieve Process Knowledge	
Implementer	Organizations using process mining for KIPs.
User	Process participants
Aim	To make experiences from past process instances accessible.
Mechanism	Retrieve and analyze similar past instances.
Rationale	The principle is grounded in the first phases of the *case-based reasoning* lifecycle (Kolodner, 1992; Osuszek and Stanek, 2015). The idea has been used in process mining by Berriche et al. (2015).
Design Principle 5: Derive Actionable Interventions	
Implementer	Organizations using process mining for KIPs
User	Process participants & process analyst
Aim	To use experiences from past process instances to perform new instances.
Mechanism	Consider and implement prescriptive actions derived from similar instances, and provide reasoning on decisions to evaluate the effects that past decisions had on process goals.
Rationale	In *case-based reasoning*, reusing information from similar past cases, revising proposed solutions, and retaining experiences can support solving new problems (Kolodner, 1992; Osuszek and Stanek, 2015). We extend this idea by taking the integration of recommended actions into consideration.

pose enriching event logs with unstructured domain knowledge and process-external goals. KIPs require knowledge transfer (Gronau and Weber, 2004), but relevant knowledge is often uncodified and siloed (Pentland et al., 2020). The interviews revealed that process participants repurposed their product lifecycle management system to capture contextual knowledge, highlighting the need to codify and integrate such knowledge into event logs (DP2) (Osuszek and Stanek, 2015). Additionally, since process goals can vary and are often based on incomplete information, process-external goals should be defined and encoded as dependent variables in the log (DP3). Though deriving such labels remains challenging due to limited evaluation data. Given the unpredictability of KIPs and their high structural variance (Isik et al., 2012; Marjanovic and Freeze, 2011; Di Ciccio et al., 2015), DP4 suggests leveraging knowledge from similar instances to support analysis and learning. Finally, DP5 acknowledges the role of human judgment in complex decisions (Marjanovic and Freeze, 2011), proposing that learning from prior cases can support decision-making by offering prescriptive guidance and feedback loops (Kolodner, 1992; Osuszek and Stanek, 2015).

3.4 Paper 4 – Supporting Organizational Knowledge Creation in Knowledge-Intensive Processes through Process Mining

P4 represents an advancement of the conceptual groundwork laid in P3. This study focuses more on the transformation of tacit into explicit process-related knowledge—and vice versa—to make it usable for process mining of KIPs. Its objective is to develop abstract design knowledge (i.e., design principles) for the design of a new class of IT artifacts coined process mining for KIPs. In doing so, P4 builds on the theory of organizational knowledge creation as a kernel theory (Nonaka and Takeuchi, 1995) that guides the design of this solution class. The developed design principles provide reusable knowledge, aligning with Type V theory of design and action (Hevner et al., 2004; Gregor, 2006). This forms the basis for the technical capabilities developed in P5–P7 to operationalize knowledge-aware process mining for KIPs.

P4 extends the ADR study from P3 with three iterations, each involving a BIE phase followed by a reflection and learning phase (Sein et al., 2011). In addition to the two manufacturing organizations from P3, a software company specialized in product lifecycle and workflow management systems was included. Across all iterations, 27 interviews, 72 workshops, seven focus groups, and three simulation games were conducted—27 interviews and 49 workshops also informed P3. The first iteration developed a conceptual IT artifact with an initial graphical user interface, evaluated through focus groups in terms of feasibility (also reported in P3). In the second iteration, the enhanced artifact was developed into a working prototype evaluated with end users in a simulation game in terms of utility. In the third iteration, five initial propositions were derived and refined through an internal and external evaluation for better reusability and generalizability according to Iivari et al. (2021).

From these iterations, five theory-ingrained design principles were synthesized, drawing on organizational knowledge creation theory, process mining, and the insights gained from the research project. The developed design principles are also structured in line with Gregor et al. (2020), similar to P3. Table 3.4 presents the developed design principles.

Table 3.4: P4 - Design Principles (Brennig et al., 2025a)

Design Principle 1: Externalize Tacit Process-Related Knowledge to Information	
Implementer	Process Participant
User	IT Artifact & Process Analyst
Aim	To externalize a process participants' tacit knowledge related to the KIP and add it as unstructured information to the artifact.
Mechanism	Write down knowledge as text, make a drawing from scratch, or add other documents that might exist but have not been made available to co-workers or the organization yet. Filter the "right" knowledge and relate it to the specific process instances so that it is visible which knowledge belongs to which process and to which instance to use it for process mining.
Rationale	The externalization of tacit knowledge, as described in the SECI model, is crucial for sharing and building new knowledge (Nonaka and Takeuchi, 1995), especially in knowledge-intensive processes reliant on skilled participants (Di Ciccio et al., 2015; Isik et al., 2012; Eppler et al., 2008; Gronau and Weber, 2004). However, many organizations lack effective mechanisms to convert tacit process-related knowledge into explicit, shareable information.
Design Principle 2: Enhance Event Log with Categorized Information	
Implementer	IT Artifact & Process Analyst
User	IT Artifact & Process Analyst
Aim	To categorize and classify information and store them as data in the event log related to a specific process instance.
Mechanism	Transform the information into data by classifying and assigning it as attributes to the event log, ensuring both machine-readability and human interpretability.
Rationale	Unstructured data are challenging to analyze efficiently and are often excluded from event logs, limiting their potential for contextual insights (Pentland et al., 2020). Additionally, while algorithms can find correlations in unstructured data, their lack of explainability hinders their use in making decisions that require interpretation and rationale.
Design Principle 3: Retrieve Information from the Event Log through Process Mining	
Implementer	IT Artifact & Process Analyst
User	Process Participant
Aim	To (re-) combine data to information by retrieving information of past process instances from the event log.
Mechanism	Identify similar past process instances (e.g., through clustering). Extend and implement process mining techniques (especially predictive and action-oriented techniques), to support process participants. State the effect of the actions taken based on the provided information and integrate a feedback loop so that the system can learn further.
Rationale	Decision-making and performance in knowledge-intensive activities are influenced by availability heuristics, where the subjective recall of past events biases judgment (Tversky and Kahneman, 1973). Combining explicit knowledge from diverse sources is essential (Nonaka and Takeuchi, 1995; Nonaka et al., 2000), as even experienced participants struggle to overcome these biases when recalling similar past events.
Design Principle 4: Internalize Information to Create new Process-Related Knowledge	
Implementer	IT Artifact & Process Analyst
User	Process Participant
Aim	To enable process participants to internalize information and create new process-related knowledge.
Mechanism	Present process participants with relevant process-related information that is useful to their specific context and situation.
Rationale	Knowledge is created through internalization, where individuals relate information to their experiences and beliefs (Nonaka and Takeuchi, 1995; Nonaka et al., 2000). Our study found that the information presented to process participants during tasks significantly influences their performance in business processes.
Design Principle 5: Identify and Network Knowledge Carriers through Process Mining	
Implementer	IT Artifact & Process Analyst
User	Process Participant
Aim	To identify knowledge carriers and and connect them.
Mechanism	Use previously added and recombined information to reverse engineer the knowledge transformation process and trace back the knowledge to the original contributor.
Rationale	Individual knowledge contributes to the organizational knowledge base only when shared (Bhatt, 2000; Nonaka and Takeuchi, 1995; Nonaka et al., 2000), yet knowledge carriers often remain hidden and difficult to identify. Our study found that experienced process participants with valuable knowledge from similar past instances were frequently unknown and inaccessible to others.

DP1 proposes externalizing tacit process-related knowledge into shareable information (e.g., words, images, text) to enable others to build on it and create new knowledge. Although many organizations recognize this need, mechanisms to convert tacit process-related knowledge into explicit process-related knowledge are often lacking. Building upon this, DP2 suggests enhancing the event log with categorized information. Storing information in event logs requires structuring and categorizing it into machine-readable, human-interpretable attributes, extending event logs beyond their typical attributes. Further, retrieving information from the event log through process mining (cf. DP3) is essential to reduce bias in human recall and improve performance. To enable this, process mining should be extended to handle unstructured data using NLP and computer vision, with a focus on predictive and action-oriented methods to support decision-making and execution in processes. In addition, DP4 proposes to internalize information to create new process-related knowledge by presenting process participants with context-relevant information, enabling them to improve performance and take informed action within KIPs. Lastly, DP5 suggests identifying and networking knowledge carriers through process mining. A drill-down mechanism allows knowledge to be traced back to the original contributor, as prior knowledge entries are stored, transformed, and recombined in the artifact for later retrieval by others.

3.5 Paper 5 – Revealing the Unspoken: Using LLMs to Mobilize and Enrich Tacit Knowledge in Event Logs of Knowledge-Intensive Processes

P3 and P4 tackled the issue of developing theoretical foundations for a knowledge-aware process mining approach. However, this issue requires not only a theoretical basis but also a concrete technical implementation for mobilizing tacit process-related knowledge in KIPs. LLMs are well-suited for this purpose as they can extract process knowledge from event logs and answer process questions (Feuerriegel et al., 2024; Vidgof et al., 2023; Berti et al., 2024a) due to their capabilities of understanding and generating meaningful and coherent text (Feuerriegel et al., 2024). Thus, LLMs have the potential to surface tacit process-related knowledge and enrich event logs in this context (Aureli et al., 2019; Seidler-de Alwis and Hartmann, 2008).

Building on the identified problems and design principles developed in P3 and P4, P5 introduces a LLM-based framework to operationalize the mobilization of tacit process-related knowledge, technically realizing the SECI-based conceptual approach. Therefore, P5 adopts a design-oriented approach (Hevner et al., 2004) to iteratively develop and refine the framework. The LLM-based framework is informed by theoretical and practical considerations from knowledge management and process mining. As a design artifact, the framework structures the problem space, supports adaptation across use cases, and facilitates both theoretical reflection and practical application. Figure 3.3 illustrates the *Bidirectional T2XES-Framework* comprising two key components that work in tandem to mobilize tacit process-related knowledge in KIPs: (1) knowledge externalization and event log enrichment and (2) knowledge retrieval and internalization.

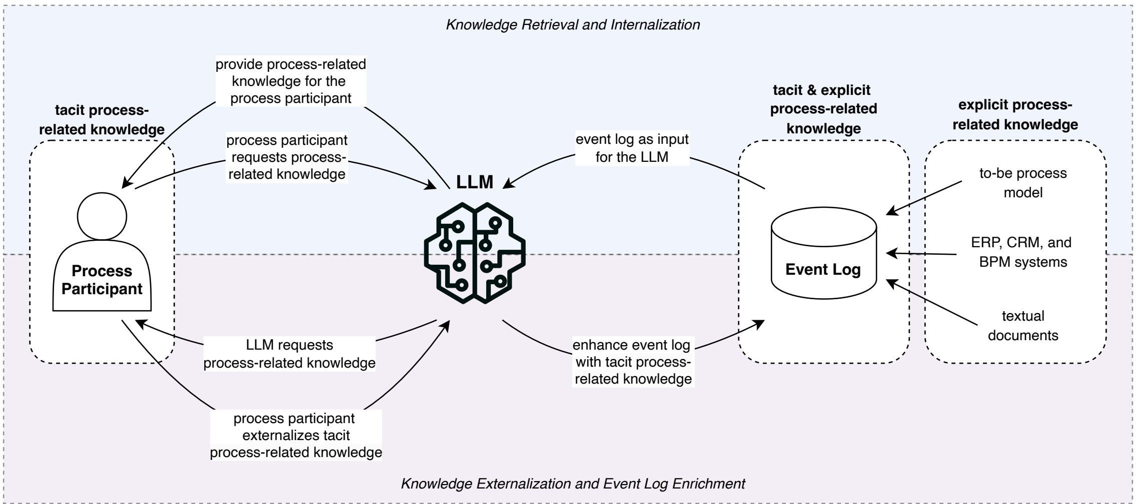

Figure 3.3: P5 - Bidirectional T2XES-Framework for Mobilizing Tacit process-related Knowledge in KIPs (Brennig, 2025)

The first path of the framework enables process participants to externalize their tacit process-related knowledge, enriching the underlying event log of the respective KIP. The LLM analyzes the event log, identifying patterns and gaps, and suggests areas where tacit process-related knowledge is missing. Through targeted adaptation (Jessen et al., 2023; Berti et al., 2024a), the LLM engages participants with targeted prompts to externalize relevant tacit process-related knowledge. Neglecting this aspect may result in the loss of knowledge and the event log being augmented with non-essential data. As the responses are aligned with explicit data, experiential knowledge is converted into structured entries linked to process instances, creating a comprehensive event log combining explicit and tacit process-related knowledge.

The second path enables process participants to retrieve and internalize both tacit (i.e., expertise and insights from process participants) and explicit (i.e., different information from textual documents, process information systems (e.g., ERP, CRM, or BPM systems), or to-be process models) knowledge from the enriched event log. By interacting with the LLM, participants gain insights from past instances to guide future process executions (Berti et al., 2024b). Expert intuition, decision rationales, and experiential knowledge are transformed into actionable insights (Feuerriegel et al., 2024; Kubrak et al., 2024; Käppel et al., 2024; Alavi et al., 2024; Zhang et al., 2024; Korzynski et al., 2023). Therefore, LLMs compress the information layer, generating knowledge directly by processing the underlying data (Alavi et al., 2024).

For evaluation, a proof-of-concept demonstrates the framework through a KIP-specific LLM-based conversational agent (Feuerriegel et al., 2024) using GPT-4o[2], allowing early assessment in a realistic yet controlled setting. A system prompt was designed to guide the conversational agent's behavior and contextualize its interactions with process participants. It first defines the agent's role as a process specialist engaging with participants (e.g., doctors, nurses) within the KIP. Second, it specifies interaction protocols, enabling the agent to analyze the event log, identify deviations and knowledge gaps, and formulate targeted questions to elicit tacit process-related knowledge. Participants can contribute and acquire knowledge, while the agent generates recommendations and seeks approval before updating the event log. Third, with participant consent, the event log is enriched with new activities or attributes, ensuring continuous refinement.

3.6 Paper 6 – Text-Aware Predictive Process Monitoring of Knowledge-Intensive Processes: Does Control Flow Matter?

Primarily, information generated through PPM techniques can support process participants in making decisions and in executing new process activities and process instances. PPM enables the forecasting of process behavior such as the remaining lead time (Di Francescomarino et al., 2018; Marquez-Chamorro et al., 2018) which is particularly relevant for KIPs that often span extended durations. However, the dynamic and flexible nature of KIPs often excludes a fixed process structure (Di Ciccio et al., 2015), making control-flow-based prediction less effective. As such, PPM in KIPs benefits from leveraging structured and unstructured data beyond control flow, as shown in P3. Further, as knowledge plays an important role in KIPs, PPM needs to be combined with NLP techniques so that the embedded knowledge can be used as input to generate predictions, as proposed in P4. Therefore, P6 extends this line of research by developing a text-aware PPM approach applying encoder-only models for natural language understanding while neglecting the control-flow of KIPs.

[2] https://openai.com/index/introducing-gpts/

Figure 3.4 illustrates the five-step framework tailored to KIPs, especially those with a stage-gate structure (Cooper, 1990). Formally, a case with n gates is represented along a timeline from $t = 0$ (start) to $t = n$ (completion). The steps include: (1) defining the prediction target at each stage; (2) applying a time-based train/test split to avoid temporal leakage (Kapoor and Narayanan, 2023); (3) incremental feature engineering for each $t < n$, using structured and textual data encoded into one-dimensional vectors; (4) training models per stage using algorithms such as Random Forest, Ridge Regression, Gradient Boosted Trees, and Multi-Layer Perceptron (MLP); and (5) evaluating model performance on unseen test data using feature transformations learned solely on the training set. For text encoding, three methods were applied: latent dirichlet allocation (LDA) for topic modeling, term frequency-inverse document frequency (TF-IDF) for Bag-of-Words (BoW) features, and contextual embeddings using the SentenceTransformers model distiluse-base-multilingual-cased-v1 (Reimers and Gurevych, 2019).

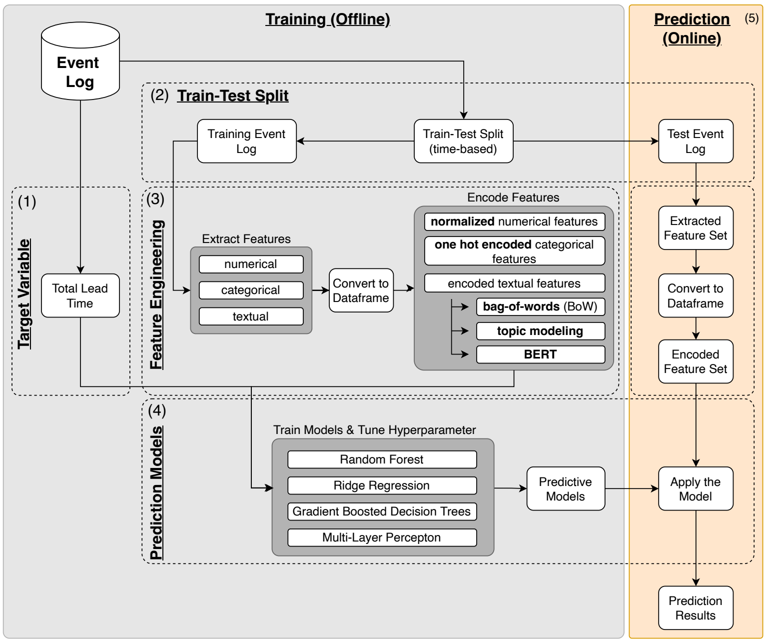

Figure 3.4: P6 - Machine Learning Pipeline (Brennig et al., 2024)

The framework is evaluated on two real-world event logs from a manufacturing organization. The mean absolute error (MAE) is used as the evaluation metric across all $t < n$. Hyperparameter tuning is performed via Optuna Search with Bayesian optimization and 5-fold cross-validation (Akiba et al., 2019). To enhance model performance and interpretability, recursive feature elimination (RFE) is applied to the Random Forest, Ridge Regression, and Gradient Boosted models. Compared to the benchmark (Cabrera et al., 2023), the proposed approach, based on stage-wise aggregation and exclusion of control-flow information, shows clear conceptual and performance advantages. The results (see Table 3.5) show that the proposed method consistently outperforms the benchmark across stages. The benchmark cannot provide predictions at $t = 0$ due to its reliance on event prefixes. In contrast, the proposed approach improves prediction accuracy incrementally over time. BERT achieves the best results at early stages, with BoW and topic modeling occasionally outperforming it later. However, differences in MAE between text representations are generally minor. Among learning algorithms, Random Forest and Ridge Regression perform best regardless of the encoding method. However, despite improvements, the MAE values remain high in absolute terms, reflecting the inherent complexity of KIPs.

Table 3.5: P6 - MAE (in days) of Predicting the Total Lead Time of a Running Case at Different Gates. For each Gate, the lowest MAE Value is underlined (Brennig et al., 2024).

Language Model	Product Development				Product Modification		
	t=0	t=1	t=2	t=3	t=0	t=1	t=2
Benchmark (based on (Cabrera et al., 2023))							
LSTM	N/A	169.247	163.229	163.036	N/A	66.803	65.266
Models with BoW							
Random Forest	107.776	126.274	68.036	45.041	45.431	42.889	**27.308**
Ridge	113.446	106.832	97.356	94.032	45.245	43.571	36.593
Gradient Boosted Decision Trees	131.427	209.513	80.327	49.725	50.021	45.140	27.833
Multi-Layer Percepton	125.962	108.655	112.309	110.526	46.607	46.327	46.429
Models with Topic Modeling							
Random Forest	112.454	127.281	66.676	**44.138**	45.183	**42.724**	27.467
Ridge	117.872	110.751	97.654	89.963	45.360	43.784	37.168
Gradient Boosted Decision Trees	126.960	171.267	80.340	54.013	50.249	44.660	28.173
Multi-Layer Percepton	186.053	111.725	110.727	123.186	46.544	46.566	46.372
Models with BERT							
Random Forest	**101.729**	111.828	**64.009**	48.745	44.790	43.029	28.771
Ridge	103.539	**106.342**	99.403	86.279	**44.734**	43.419	37.430
Gradient Boosted Decision Trees	145.253	146.917	94.964	55.671	45.770	44.200	30.357
Multi-Layer Percepton	117.895	107.424	102.197	133.601	46.446	46.520	49.662

3.7 Paper 7 – Straight Outta Logs: Can Large Language Models Overcome Preprocessing in Next Event Prediction?

Additionally to predicting the remaining lead time of KIPs, getting information about the next possible event in a KIP also supports process participants in making decisions and executing the process while mitigating biases in judgment and performance (cf. P4). To provide meaningful information based on the knowledge contained in the process, semantic relationships and nuanced dependencies in historical cases need to be captured, providing contextually relevant next event predictions (NEPs) beyond mere control-flow patterns. However, this requires methods that can directly process and understand XES-formatted event logs, and also generate natural language. This is important as existing PPM methods for NEP require extensive data preparation and encoding (van der Aalst, 2022; Di Francescomarino and Ghidini, 2022; Marquez-Chamorro et al., 2018). To overcome this challenge, LLMs with their capabilities in understanding and generating natural language, programming code, and markup languages (Chen et al., 2021; Feuerriegel et al., 2024; Sui et al., 2024) present a suitable approach.

Therefore, P7 develops a LLM-based approach to directly generate NEPs from XES-formatted event logs. This eliminates the need for preprocessing, advancing existing PPM approaches. Using the open-source Llama 3 (8 billion parameters) model (Meta, 2024; AI@Meta, 2024), fine-tuned on five benchmark event logs, the NEP pipeline iteratively predicts the next event in XES format, maintaining all relevant attributes (see Figure 3.5). While P7 focuses on predicting and evaluating the attribute *concept:name*, the pipeline generates additional attributes (e.g., timestamp, org:resource), enabling broader tasks. The LLM has been fine-tuned on the next token prediction (Gururangan et al., 2020). The prompt contains the maximum of past events that fit in the context length, separated by end-of-sequence (EOS) tokens (see Figure 3.5).

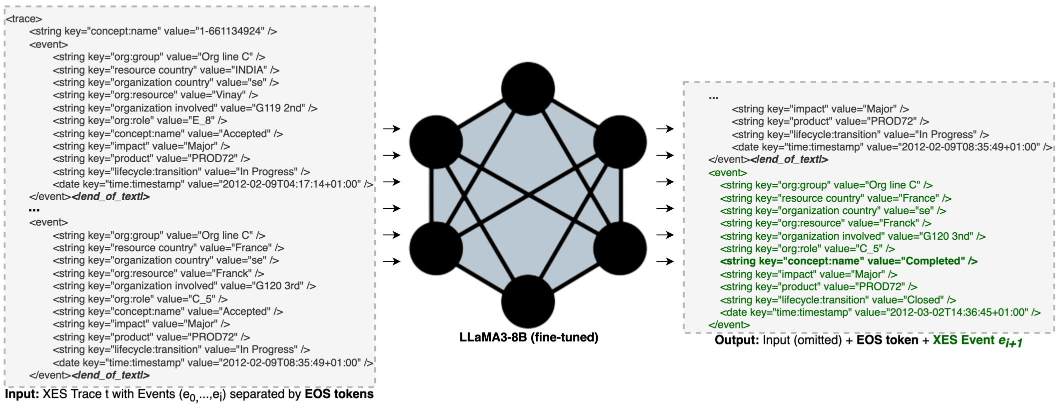

Figure 3.5: P7 - NEP Inference Pipeline (Brennig et al., 2025b)

Evaluating the approach comprised a 5-fold cross-validation with consistent splits across datasets, using accuracy as the primary metric for comparability (Oved et al., 2025; Rama-Maneiro et al., 2022). Table 3.6 summarizes the accuracies of the related benchmark studies compared to the approach developed in P7. The results show that while not outperforming the state-of-the-art, the approach achieves above-median accuracy with only a 3% average gap, while nearly eliminating preprocessing requirements. To generate a complete event, the developed pipeline only needs 3 to 11 seconds. The

LLM-generated outputs demonstrated near-perfect syntactic correctness, with an error rate below 0.005% across datasets, and minimal hallucinations (0.13%), typically involving semantically similar but non-identical events. The results show that LLMs are capable of identifying syntactical and semantical relationships between events and within event attributes. Based on that, they enable the generation of NEPs of a process instance using an event log with negligible preprocessing of the data.

Table 3.6: P7 - Accuracy Comparison (Brennig et al., 2025b)

	BP113cp		BPI13in		Sepsis		Helpdesk		BPI12		Average		Total
	Accuracy	Rank	Accuracy	Rank	Accuracy	Rank	Accuracy	Rank	Accuracy	Rank	Accuracy	Rank	Rank
Camargo[3]	0.547	8	0.667	7	0.610	5	0.829	5	0.833	6	0.697	6	6
Evermann[3]	0.588	7	0.668	6	0.400	11	0.336	11	0.593	10	0.517	9	10
Hinnka[3]	0.635	5	0.747	3	0.635	3	0.831	4	0.867	2	0.743	3.4	4
Khan[3]	0.436	11	0.519	9	0.210	12	0.800	8	0.429	12	0.479	10.4	12
Mauro[3]	0.249	12	0.367	12	0.615	4	0.318	12	0.847	5	0.479	9	10
Pasquadibisceglie[3]	0.475	10	0.460	11	0.562	9	0.840	3	0.833	6	0.634	7.8	7
Tax[3]	0.640	4	0.701	5	0.642	2	0.842	2	0.855	3	0.736	3.2	2
Theis[3]	0.595	6	0.594	8	0.557	10	0.788	10	0.829	9	0.673	8.6	8
Venugopal[4]	0.484	9	0.496	10	0.610	5	0.797	9	0.547	11	0.587	8.8	9
Rama-Maneiro[4]	0.675	2	0.777	1	0.610	5	0.852	1	0.871	1	0.757	2	1
Oved[4]	0.696	1	0.774	2	0.655	1	0.816	6	0.833	6	0.755	3.2	2
Ours	**0.645**	3	**0.724**	4	**0.610**	5	**0.816**	6	**0.848**	4	0.729	4.4	5
Median	**0.592**		**0.668**		**0.610**		**0.816**		**0.833**		**0.685**		
Deviation from max	0.051		0.053		0.045		0.036		0.023		0.028		
Ours[2] (Weighted F1)	0,637		0,701		0,599		0,803		0,845				
Ours[2] (Macro F1)	0,377		0,449		0,382		0,413		0,675				

[1] We assume the median on missing values to provide a fair comparison.
[2] We report the weighted and macro average F1-Score for future comparability.
[3] Rama-Maneiro et al. (2022)
[4] Oved et al. (2025)

4 Discussion and Conclusion

4.1 Implications for Research and Practice

This thesis contributes to establishing a knowledge-aware process mining approach. The results show how current process mining needs to be extended to develop into a more knowledge-aware discipline. For this purpose, foundational requirements for process mining in organizations, theoretical foundations, and technical implementations are presented. In practice, this thesis offers organizations guidance to build a process mining basis and to manage their KIPs holistically using process mining.

A key focus of this thesis is strengthening organizational capabilities and achieving the necessary process mining maturity within organizations (FOCUS I). To leverage process mining in KIPs, organizations must first develop foundational capabilities and reach sufficient maturity before its knowledge-awareness can be meaningfully integrated. Addressing this need, this thesis investigates the organizational capabilities required to adopt, implement, and mature process mining in organizations (RQ1). Therefore, P1 provides strategic guidelines to assess and enhance the relevance of process mining outcomes, while P2 introduces a maturity model with actions for progressing between maturity stages. These findings support researchers in planning, executing, and evaluating process mining projects and allow them to theorize about the role of process mining and data-driven methods in achieving a more flexible, evidence-based BPM approach. For practitioners, the findings offer guidance on managing the development and implementation of process mining research projects to address important business problems and generate business value. It helps organizations allocate resources effectively and focus on the most promising activities, overcoming common challenges in adopting and managing process mining (vom Brocke et al., 2021a; Martin et al., 2021). By enhancing their process mining maturity and capabilities, organizations improve their ability to respond dynamically to changing process dynamics (Pentland et al., 2021; Wurm et al., 2021) and gain valuable insights into process performance (Grisold et al., 2020; Wurm et al., 2021; Kipping et al., 2022).This is particularly critical for KIPs, which are characterized by high variability, unpredictable task sequences, and complex process flows

(Di Ciccio et al., 2015; Isik et al., 2012; Eppler et al., 2008; Gronau and Weber, 2004). P2 further reveals that organizations that advance their process mining maturity have become more flexible in handling their business processes. Overall, improving organizational process mining maturity lays the groundwork for a knowledge-aware process mining approach and represents a first essential step toward effectively managing KIPs.

Further, this thesis also focuses on reconceptualizing the theoretical foundations for applying process mining to KIPs (FOCUS II), as existing process mining approaches primarily target standardized, digitalized processes (van der Aalst et al., 2012; van der Aalst, 2016, 2022). This thesis therefore examines how process mining can be extended to manage KIPs adequately (RQ2). To this end, papers P3 and P4 conceptualize a new class of process mining systems for KIPs, by deriving design principles that enable the integration of tacit process-related knowledge and the mobilization of embedded process-related knowledge for analysis. This establishes a theoretical basis for applying process mining in knowledge-intensive contexts, grounded in the theory of organizational knowledge creation, especially the SECI model (Nonaka and Takeuchi, 1995). From a theoretical perspective, the results demonstrate how knowledge management capabilities—particularly the SECI model—can extend process mining by enabling systematic knowledge sharing and access to KIP-related knowledge. This approach transcends traditional process mining by externalizing tacit process-related knowledge and fostering continuous, spiral cycles of knowledge creation. Knowledge evolves from individual instances to process-type levels, crystallizing at higher ontological levels (Nonaka and Takeuchi, 1995). This reflects the theoretical underpinnings required to conceptualize process mining in environments where tacit process-related knowledge plays a central role. In practice, the developed design principles offer blueprints for applying process mining to KIPs. They guide organizations in mobilizing and transforming tacit process-related knowledge into explicit, analyzable formats, showing that event logs can capture process-related knowledge typically excluded from analysis. This supports the ongoing adaptation and improvement of KIPs, fosters a culture of knowledge exchange, and gives guidance on establishing a fundamental knowledge base. This enables process participants to make more objective and strategic decisions within KIPs.

Lastly, this thesis focuses on integrating NLP into process mining for KIPs (FOCUS III). The previous findings underpin the technical implementations that are important to advance knowledge-aware process mining (see Figure 3.1 in section 3). Therefore, this thesis further investigates how NLP can be leveraged to advance the technical capabilities of process mining to manage KIPs (RQ3). P5 demonstrates the feasibility of enriching event logs with tacit process-related knowledge and mobilizing the embedded knowledge in KIPs using LLMs. P6 builds on this by showing how this enriched knowledge base can improve PPM using encoder-based models. P7 advances this further by testing whether LLMs can directly generate NEPs from XES-formatted, knowledge-enriched event logs. Collectively, these papers operationalize the SECI model within process mining. They transform tacit process-related knowledge into explicit process-related knowledge that is combined in a KIP-specific event log with process-related data and information and reintegrated into process execution—enabling continuous learning and organizational improvement. Table 4.1 outlines how these papers contribute to the modes of knowledge conversion (Nonaka and Takeuchi, 1995) in the context of knowledge-aware process mining and offers guidance on technically advancing the design principles developed in P3 and P4.

Table 4.1: Contribution to the four Modes of Knowledge Conversion

Knowledge Conversion Mode	Paper	Contribution
Socialization	P5	The LLM-based approach enables process participants to communicate in natural language—via text or speech—to efficiently capture process-related knowledge in real-time, reducing the need for peer consultation. This reduces tacit process-related knowledge loss during process execution.
Externalization	P5	The LLM-based approach helps process participants to articulate tacit process-related knowledge by using custom prompts to encourage reflection and capture decision rationales.
	P6	It encourages participants to externalize context-specific knowledge—typically hidden—by highlighting its importance for accurate predictions in KIPs.
Combination	P5	The LLM-based approach synthesizes knowledge from multiple sources through real-time updates and analytics, linking externalized tacit and explicit process-related knowledge directly to process events.
	P6	The text-aware PPM approach aggregates and encodes the event log of a KIP by focusing on textual information provided by participants rather than traditional control-flow structures.
	P7	It processes explicit event log data into structured NEPs, enabling explicit-to-explicit process-related knowledge transformation and simplifying prediction workflows across cases, supporting multi-case learning and reuse.
Internalization	P5	It allows users to query the enriched process knowledge base and interact with the LLM to receive real-time, context-relevant insights, supporting reflection and learning.
	P6, P7	It provides predictive insights that inform decision-making and task execution, fostering reflection, learning, and continuous improvement.

From a theoretical perspective, the results show that text-aware approaches can capture syntactic and semantic relationships between events and attributes. P5 illustrates how LLMs support knowledge mobilization in KIPs, extend process mining with human-AI interaction, and enhance AI's role in decision-making. Leveraging their contextual understanding (Feuerriegel et al., 2024; Grohs et al., 2024), LLMs enable the integration of PPM for KIPs, enhancing support for process participants in executing upcoming process instances. However, existing PPM methods require extensive preprocessing (van der Aalst, 2022; Di Francescomarino and Ghidini, 2022; Marquez-Chamorro et al., 2018) and mainly focus on the control-flow (Marquez-Chamorro et al., 2018; Pegoraro et al., 2021; Verenich et al., 2019). P6 and P7 show how integrating text-aware models with PPM allows for richer, more context-sensitive predictions in KIPs. Notably, P7 demonstrates that LLMs can directly process XES-formatted event logs and generate valid XES outputs. This simplifies data preparation while capturing semantic relationships and nuanced dependencies within historical cases. This approach supports multi-task predictions (e.g., event name, resource, timestamp) and enables direct downstream processing, such as simulation or further analysis. By reducing complexity and demonstrating practical feasibility, these findings strengthen the technical foundation established in P5. From a managerial perspective, the framework introduced in P5 helps organizations build a robust knowledge base for KIPs, facilitating real-time knowledge sharing, reducing knowledge loss, and improving decision-making. P6 and P7 provide process participants with actionable insights drawn from historical data and textual information, helping them mitigate biases in judgment and task execution. These results underscore the importance of contextual knowledge for accurate predictions in KIPs and encourage participants to reflect on and document context-specific insights that otherwise remain hidden. This guides organizations in designing information systems that enhance data quality for future applications. Moreover, incorporating LLMs into process mining supports adaptation to new domains through transfer learning, as demonstrated by Liessmann et al. (2024). By learning the structure and format of process traces, LLMs enable cross-domain application without retraining from scratch. This allows easy integration into existing monitoring systems—requiring only an XES-formatted event log—and facilitates scalable deployment. Altogether, these contributions lower the barriers for organizational adoption and support enhanced decision-making and knowledge mobilization in KIPs.

4.2 Limitations

While this thesis advances process mining toward a more knowledge-aware discipline, it poses several limitations. First, this thesis is limited in its generalizability regarding its context and the observed processes. The developed theoretical foundations, such as the guidelines (P1), maturity model (P2) and the developed design principles (P3, P4), offer valuable guidance. However, their effectiveness in different organizational contexts remains to be further tested and implemented in other real-world environments as most results are derived from studies conducted within specific organizational contexts, industries (e.g., manufacturing, healthcare), and with selected partners. While the ADR approach justifies this close practitioner collaboration, it may limit the broader applicability of the findings to other domains or sectors. This is also essential for the LLM-based framework (P5) and predictive models (P6, P7). They were evaluated through prototypes and specific datasets. However, their performance, scalability, and adoption in large-scale, real-world environments remain underexplored.

Second, certain technical limitations must be considered. Specifically, process mining applications require extensive context windows, as event logs often contain a vast number of cases and events. The limited token capacity of text-aware approaches may pose constraints when analyzing KIP-specific event logs, which typically feature more extensive textual descriptions than standard event logs. Addressing this challenge is crucial for ensuring the scalability and efficacy of text-based process mining applications. Further, despite LLMs' ability to mobilize and enrich process-related knowledge, inherent risks like hallucination and factually incorrect outputs remain a known issue. The impact of such risks on process execution and decision-making in practice was not deeply explored.

4.3 Future Research Directions and Outlook

The limitations discussed in this dissertation provide a foundation for future research avenues. The generalizability of the proposed theoretical foundations requires further empirical validation across a broader spectrum of organizational contexts and industries. While the findings offer valuable insights, their applicability beyond the studied cases remains to be systematically examined. Moreover, the contributions of P3-P5 offer prototypical components for the development of knowledge-aware process mining tools. Future research should extend these initial artifacts into comprehensive systems that incorporate both the conceptual frameworks and the technical approaches presented in this thesis. Such efforts would benefit from deployment in real-world organizational settings to assess both technical feasibility and user acceptance. Given the centrality of human knowledge in KIPs, it is particularly important to investigate whether process participants are willing to adopt such tools and actively contribute their process-related knowledge. Consequently, future research should also address the organizational conditions required to foster a culture of knowledge sharing, thereby supporting the practical adoption of knowledge-aware process mining in knowledge-intensive environments.

Part B

Research Papers

Maximizing the Impact of Process Mining Research: Four Strategic Guidelines

Katharina Brennig
Paderborn University

Bernd Löhr
Paderborn University

Jonathan Brock
Fraunhofer IEM

Malte Reineke
Paderborn University

Christian Bartelheimer
Paderborn University

Abstract—While most organizations recognize the potential of process mining and plan to start process mining initiatives, significant challenges for applying process mining in organizations remain unsolved. In this paper, we investigate the utility of process mining in organizations. Although process mining papers in the current Information Systems (IS) knowledge base deal with developed artifacts based on real-world scenarios, they often do not adequately reflect on their results' utility and effectiveness in the application context, diminishing their contributions' practical and theoretical implications. By discussing the results from the systematic literature review in the backdrop to the existing knowledge base, we develop four strategic guidelines for conducting process mining research with high relevance and managerial impact. Fellow researchers can follow these guidelines to rigorously plan, execute, and evaluate process mining research projects to generate business value and achieve maximum organizational impact.

Full Citation: *Brennig, K., Löhr, B., Brock, J., Reineke, M., & Bartelheimer, C. (2024). Maximizing the Impact of Process Mining Research: Four Strategic Guidelines. In AMCIS 2024 Proceedings.*

Improving Process Mining Maturity: From Intentions to Actions

Jonathan Brock
Fraunhofer IEM

Katharina Brennig
Paderborn University

Bernd Löhr
Paderborn University

Christian Bartelheimer
Paderborn University

Sebastian von Enzberg
Fraunhofer IEM

Roman Dumitrescu
Fraunhofer IEM

Abstract—Process mining is advancing as a powerful tool for revealing valuable insights about process dynamics. Nevertheless, the imperative to employ process mining to enhance process transparency is a prevailing concern for organizations. Despite the widespread desire to integrate process mining as a pivotal catalyst for fostering a more agile and flexible Business Process Management (BPM) environment, many organizations face challenges in achieving widespread implementation and adoption due to deficiencies in various dimensions of process mining readiness. The current Information Systems (IS) knowledge base lacks a comprehensive framework to aid organizations in augmenting their process mining readiness and bridging this intention-action gap. This paper presents a Process Mining Maturity Model (P3M), refined through multiple iterations, which outlines five factors and 23 elements that organizations must address to increase their process mining readiness. The maturity model advances the understanding of how to close the intention-action gap of process mining initiatives in multiple dimensions. Furthermore, insights from a comprehensive analysis of data gathered in eleven qualitative interviews are drawn, elucidating 30 possible actions that organizations can implement to establish a more responsive and dynamic BPM environment by means of process mining.

Full Citation: *Brock, J., Brennig, K., Löhr, B., Bartelheimer, C., von Enzberg, S. & Dumitrescu, R. (2024). Improving Process Mining Maturity – From Intentions to Actions. Business & Information Systems Engineering: Vol. 66: Iss. 5, 585-605.*

Process Mining of Knowledge-Intensive Processes: An Action Design Research Study in Manufacturing

Bernd Löhr
Paderborn University

Katharina Brennig
Paderborn University

Christian Bartelheimer
Paderborn University

Daniel Beverungen
Paderborn University

Oliver Müller
Paderborn University

Abstract—Existing process mining methods are primarily designed for processes that have reached a high degree of digitalization and standardization. In contrast, the literature has only begun to discuss how process mining can be applied to knowledge-intensive processes—such as product innovation processes—that involve creative activities, require organizational flexibility, depend on single actors' decision autonomy, and target process-external goals such as customer satisfaction. Due to these differences, existing Process Mining methods cannot be applied out-of-the-box to analyze knowledge-intensive processes. In this paper, we employ Action Design Research (ADR) to design and evaluate a process mining approach for knowledge-intensive processes. More specifically, we draw on the two processes of product innovation and engineer-to-order in manufacturing contexts. We collected data from 27 interviews and conducted 49 workshops to evaluate our IT artifact at different stages in the ADR process. From a theoretical perspective, we contribute five design principles and a conceptual artifact that prescribe how process mining ought to be designed for knowledge-intensive processes in manufacturing. From a managerial perspective, we demonstrate how enacting these principles enables their application in practice.

Full Citation: *Löhr, B., Brennig, K., Bartelheimer, C., Beverungen, D. & Müller, O. (2022). Process Mining of Knowledge-Intensive Processes: An Action Design Research Study in Manufacturing. In: Di Ciccio, C., Dijkman, R., del Río Ortega, A., Rinderle-Ma, S. (eds) Business Process Management. BPM 2022. Lecture Notes in Computer Science, vol 13420. Springer, Cham.*

Supporting Organizational Knowledge Creation in Knowledge-Intensive Processes through Process Mining

Katharina Brennig
Paderborn University

Christian Bartelheimer
Paderborn University

Bernd Löhr
Paderborn University

Daniel Beverungen
Paderborn University

Oliver Müller
Paderborn University

Abstract—Knowledge-intensive processes (KIPs) are complex, strategic core processes that drive organizational competitive advantage. These processes rely on explicit and tacit knowledge. While explicit knowledge can be codified and leveraged—often through technologies such as process mining—tacit knowledge remains embedded in individual process participants, limiting knowledge transfer and organizational learning. Process mining, a data-driven approach to analyze process data, works best for standard processes that are managed for consistency, costs, and time but is insufficiently equipped to enhance KIPs, which depend on dynamic, experience-based decision-making. We present findings from a 39-month Action Design Research (ADR) project to conceptualize a new class of IT artifacts that enable process mining for KIPs. This class of IT artifacts integrates richer process-related information, facilitating knowledge transfer by allowing participants to learn from similar process instances and engage in socialization. We propose five theory-ingrained design principles that guide the development of such systems and examine their role in fostering knowledge creation within organizations. Our research bridges critical gaps between business process management and knowledge management, offering theoretical and managerial insights. For practitioners, our findings provide a foundation for improving KIPs, ultimately upgrading strategic decision-making and organizational performance.

Full Citation: *Brennig, K., Bartelheimer, C., Löhr, B., Beverungen, D., & Müller, O. (2025). Supporting Organizational Knowledge Creation in Knowledge-Intensive Processes through Process Mining. Working Papers Dissertations, 148, Paderborn University, Faculty of Business Administration and Economics.*

Revealing the Unspoken: Using LLMs to Mobilize and Enrich Tacit Knowledge in Event Logs of Knowledge-Intensive Processes

Katharina Brennig
Paderborn University

Abstract—Large Language Models (LLMs) excel in understanding, generating, and processing human language, with growing adoption in process mining. Process mining relies on event logs that capture explicit process knowledge; however, knowledge-intensive processes (KIPs) in domains such as healthcare and product development depend on tacit knowledge, which is often absent from event logs. To bridge this gap, this study proposes a LLM-based framework for mobilizing tacit process knowledge and enriching event logs. A proof-of-concept is demonstrated using a KIP-specific LLM-driven conversational agent built on GPT-4o. The results indicate that LLMs can capture tacit process knowledge through targeted queries and systematically integrate it into event logs. This study presents a novel approach combining LLMs, knowledge management, and process mining, advancing the understanding and management of KIPs by enhancing knowledge accessibility and documentation.

Full Citation: *Brennig, K. (2025). Revealing the Unspoken: Using LLMs to Mobilize and Enrich Tacit Knowledge in Event Logs of Knowledge-Intensive Processes. AMCIS 2025 Proceedings. 11.*

Text-Aware Predictive Process Monitoring of Knowledge-Intensive Processes: Does Control Flow Matter?

Katharina Brennig
Paderborn University

Kay Benkert
Paderborn University

Bernd Löhr
Paderborn University

Oliver Müller
Paderborn University

Abstract—Predictive process monitoring (PPM) enables organizations to predict the behavior of ongoing processes, e.g., the lead time. This is of great interest for knowledge-intensive processes (KIPs), which often cover long time spans. With such insights, resource allocation or customer relationship management could be improved. While already many PPM methods exist, they have not yet been applied to KIPs. Thus, we extend PPM research by using machine learning and natural language processing (NLP) to develop and evaluate a novel text-aware PPM approach tailored towards monitoring KIPs. By developing suitable features and considering various time intervals, our approach encodes and aggregates the event log. Using two real-world event logs, we assess our methodology. We demonstrate that the MAE improves as compared to state-of-the-art PPM methods. It shows that the control flow perspective of KIPs should primarily be neglected, while considering more structured features and unstructured textual information is essential.

Full Citation: *Brennig, K., Benkert, K., Löhr, B. & Müller, O. (2024). Text-Aware Predictive Process Monitoring of Knowledge-Intensive Processes: Does Control Flow Matter?. In: De Weerdt, J., Pufahl, L. (eds) Business Process Management Workshops. BPM 2023. Lecture Notes in Business Information Processing, vol 492. Springer, Cham.*

Straight Outta Logs: Can Large Language Models Overcome Preprocessing in Next Event Prediction?

Katharina Brennig
Paderborn University

Sascha Kaltenpoth
Paderborn University

Oliver Müller
Paderborn University

Abstract—Predictive process monitoring (PPM) aims to predict the future behavior of process instances to mitigate process violations or take preventive measures. Current PPM methods for next event prediction (NEP) often utilize machine learning techniques, while first approaches also use deep learning techniques, especially natural language processing (NLP). Hence, these approaches often require extensive data preprocessing. To counteract this, we train and evaluate a fine-tuned large language model (LLM) to directly generate NEPs from XES-formatted event logs without any preprocessing. The results suggest that the proposed PPM approach performs comparably to the state-of-the-art in ML-based PPM, while contributing a simplified prediction process for NEP with minimal data preprocessing. Additionally, our LLM-driven approach produces valid XES outputs in nearly all cases, facilitating the direct export of predictions as event logs to be processed downstream (e.g., to employ process mining techniques or simulation). Further, our method offers easy integration into existing organizational infrastructures.

Full Citation: *Brennig, K., Kaltenpoth, S. & Müller, O. (2025). Straight Outta Logs: Can Large Language Models Overcome Preprocessing in Next Event Prediction?. In: Gdowska, K., Gómez-López, M.T., Rehse, JR. (eds) Business Process Management Workshops. BPM 2024. Lecture Notes in Business Information Processing, vol 534. Springer, Cham.*

Bibliography

Aguirre, S., Parra, C., and Sepúlveda, M. (2017). Methodological proposal for process mining projects. *International Journal of Business Process Integration and Management*, 8(2):102.

AI@Meta (2024). Llama 3 Model Card.

Akiba, T., Sano, S., Yanase, T., Ohta, T., and Koyama, M. (2019). Optuna: A Next-generation Hyperparameter Optimization Framework. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 2623–2631, Anchorage AK USA. ACM.

Alavi, M., Georgia Institute of Technology, Leidner, D. E., University of Virginia, Mousavi, R., and University of Virginia (2024). Knowledge Management Perspective of Generative Artificial Intelligence. *Journal of the Association for Information Systems*, 25(1):1–12.

Alavi, M. and Leidner, D. E. (2001). Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues. *MIS Quarterly*, 25(1):107.

Alter, S. (2015). A Workaround Design System for Anticipating, Designing, and/or Preventing Workarounds. In Gaaloul, K., Schmidt, R., Nurcan, S., Guerreiro, S., and Ma, Q., editors, *Enterprise, Business-Process and Information Systems Modeling*, volume 214, pages 489–498. Springer International Publishing, Cham. Series Title: Lecture Notes in Business Information Processing.

Aureli, S., Giampaoli, D., Ciambotti, M., and Bontis, N. (2019). Key factors that improve knowledge-intensive business processes which lead to competitive advantage. *Business Process Management Journal*, 25(1):126–143.

Badakhshan, P., Wurm, B., Grisold, T., Geyer-Klingeberg, J., Mendling, J., and vom Brocke, J. (2022). Creating business value with process mining. *The Journal of Strategic Information Systems*, 31(4):101745.

Bahrs, J. and Müller, C. (2005). Modelling and Analysis of Knowledge Intensive Business Processes. In Hutchison, D., Kanade, T., Kittler, J., Kleinberg, J. M., Mattern, F., Mitchell, J. C., Naor, M., Nierstrasz, O., Pandu Rangan, C., Steffen, B., Sudan, M., Terzopoulos, D., Tygar, D., Vardi, M. Y., Weikum, G., Althoff, K.-D., Dengel, A., Bergmann, R., Nick, M., and Roth-Berghofer, T., editors, *Professional Knowledge Management*, volume 3782, pages 243–247. Springer Berlin Heidelberg, Berlin, Heidelberg. Series Title: Lecture Notes in Computer Science.

Baiyere, A., Salmela, H., and Tapanainen, T. (2020). Digital transformation and the new logics of business process management. *European Journal of Information Systems*, 29(3):238–259.

Becker, J., Knackstedt, R., and Pöppelbuß, J. (2009). Developing Maturity Models for IT Management: A Procedure Model and its Application. *Business & Information Systems Engineering*, 1(3):213–222.

Benner-Wickner, M., Brückmann, T., Gruhn, V., and Book, M. (2015). Process mining for knowledge-intensive business processes. In *Proceedings of the 15th International Conference on Knowledge Technologies and Data-driven Business*, pages 1–8, Graz Austria. ACM.

Berriche, F.-Z., Zeddini, B., Kadima, H., and Riviere, A. (2015). Combining Case-Based Reasoning and Process Mining to improve collaborative decision-making in products design. In *2015 IEEE/ACS 12th International Conference of Computer Systems and Applications (AICCSA)*, pages 1–7, Marrakech, Morocco. IEEE.

Berti, A., Kourani, H., Häfke, H., Li, C.-Y., and Schuster, D. (2024a). Evaluating Large Language Models in Process Mining: Capabilities, Benchmarks, and Evaluation Strategies. In van der Aa, H., Bork, D., Schmidt, R., and Sturm, A., editors, *Enterprise, Business-Process and Information Systems Modeling*, volume 511, pages 13–21. Springer Nature Switzerland, Cham. Series Title: Lecture Notes in Business Information Processing.

Berti, A. and Qafari, M. S. (2023). Leveraging Large Language Models (LLMs) for Process Mining (Technical Report). Version Number: 1.

Berti, A., Schuster, D., and van der Aalst, W. M. P. (2024b). Abstractions, Scenarios, and Prompt Definitions for Process Mining with LLMs: A Case Study. In De Weerdt, J. and Pufahl, L., editors, *Business Process Management Workshops*, volume 492, pages 427–439. Springer Nature Switzerland, Cham. Series Title: Lecture Notes in Business Information Processing.

Beverungen, D., Buijs, J. C. A. M., Becker, J., Di Ciccio, C., van der Aalst, W. M. P., Bartelheimer, C., vom Brocke, J., Comuzzi, M., Kraume, K., Leopold, H., Matzner, M., Mendling, J., Ogonek, N., Post, T., Resinas, M., Revoredo, K., del Río-Ortega, A., La Rosa, M., Santoro, F. M., Solti, A., Song, M., Stein, A., Stierle, M., and Wolf, V. (2021). Seven Paradoxes of Business Process Management in a Hyper-Connected World. *Business & Information Systems Engineering*, 63(2):145–156.

Bhatt, G. D. (2000). Organizing knowledge in the knowledge development cycle. *Journal of Knowledge Management*, 4(1):15–26.

Bhatt, G. D. (2001). Knowledge management in organizations: examining the interaction between technologies, techniques, and people. *Journal of Knowledge Management*, 5(1):68–75.

Brennig, K. (2025). Revealing the Unspoken: Using LLMs to Mobilize and Enrich Tacit Knowledge in Event Logs of Knowledge-Intensive Processes. In *AMCIS 2025 Proceedings*.

Brennig, K., Bartelheimer, C., Löhr, B., Beverungen, D., and Müller, O. (2025a). Supporting Organizational Knowledge Creation in Knowledge-Intensive Processes through Process Mining. *Working Papers Dissertations*, 148, Paderborn University, Faculty of Business Administration and Economics.

Brennig, K., Benkert, K., Löhr, B., and Müller, O. (2024). Text-Aware Predictive Process Monitoring of Knowledge-Intensive Processes: Does Control Flow Matter? In De Weerdt, J. and Pufahl, L., editors, *Business Process Management Workshops*, volume 492, pages 440–452. Springer Nature Switzerland, Cham. Series Title: Lecture Notes in Business Information Processing.

Brennig, K., Kaltenpoth, S., and Müller, O. (2025b). Straight Outta Logs: Can Large Language Models Overcome Preprocessing in Next Event Prediction? In Gdowska, K., Gómez-López, M. T., and Rehse, J.-R., editors, *Business Process Management Workshops*, volume 534, pages 197–208. Springer Nature Switzerland, Cham. Series Title: Lecture Notes in Business Information Processing.

Brock, J., Brennig, K., Löhr, B., Bartelheimer, C., von Enzberg, S., and Dumitrescu, R.

(2024). Improving Process Mining Maturity – From Intentions to Actions. *Business & Information Systems Engineering*, 66(5):585–605.

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I., and Amodei, D. (2020). Language Models are Few-Shot Learners. Version Number: 4.

Brzychczy, E., Kluza, K., and Szała, L. (2025). Enhancement of Low-Level Event Abstraction with Large Language Models (LLMs). In Gdowska, K., Gómez-López, M. T., and Rehse, J.-R., editors, *Business Process Management Workshops*, volume 534, pages 209–220. Springer Nature Switzerland, Cham. Series Title: Lecture Notes in Business Information Processing.

Cabrera, L., Weinzierl, S., Zilker, S., and Matzner, M. (2023). Text-Aware Predictive Process Monitoring with Contextualized Word Embeddings. In Cabanillas, C., Garmann-Johnsen, N. F., and Koschmider, A., editors, *Business Process Management Workshops*, volume 460, pages 303–314. Springer International Publishing, Cham. Series Title: Lecture Notes in Business Information Processing.

Chen, M., Tworek, J., Jun, H., Yuan, Q., Pinto, H. P. d. O., Kaplan, J., Edwards, H., Burda, Y., Joseph, N., Brockman, G., Ray, A., Puri, R., Krueger, G., Petrov, M., Khlaaf, H., Sastry, G., Mishkin, P., Chan, B., Gray, S., Ryder, N., Pavlov, M., Power, A., Kaiser, L., Bavarian, M., Winter, C., Tillet, P., Such, F. P., Cummings, D., Plappert, M., Chantzis, F., Barnes, E., Herbert-Voss, A., Guss, W. H., Nichol, A., Paino, A., Tezak, N., Tang, J., Babuschkin, I., Balaji, S., Jain, S., Saunders, W., Hesse, C., Carr, A. N., Leike, J., Achiam, J., Misra, V., Morikawa, E., Radford, A., Knight, M., Brundage, M., Murati, M., Mayer, K., Welinder, P., McGrew, B., Amodei, D., McCandlish, S., Sutskever, I., and Zaremba, W. (2021). Evaluating Large Language Models Trained on Code. Version Number: 2.

Cooper, R. G. (1990). Stage-gate systems: A new tool for managing new products. *Business Horizons*, 33(3):44–54.

Daniels, J. (2022). Trends In Process Improvement And Data Execution. How Organizations Are Improving Processes And Turning Process Data Into Real-Time Action. Accessed 08 Nov 2023.

Davenport, T. H. (1992). *Process innovation: reengineering work through information technology*. Harvard Business School Press, Boston, Mass, repr. edition.

Davenport, T. H. and Spanyi, A. (2019). What process mining is, and why companies should do it. Accessed 08 Nov 2023.

De Almeida Rodrigues Gonçalves, J. C., Baião, F. A., Santoro, F. M., and Guizzardi, G. (2023). A cognitive BPM theory for knowledge-intensive processes. *Business Process Management Journal*, 29(2):465–488.

Deloitte (2025). Global Deloitte Process Mining Study.

Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Di Ciccio, C., Marrella, A., and Russo, A. (2015). Knowledge-Intensive Processes: Characteristics, Requirements and Analysis of Contemporary Approaches. *Journal on Data Semantics*, 4(1):29–57.

Di Francescomarino, C. and Ghidini, C. (2022). Predictive Process Monitoring. In van der Aalst, W. M. P. and Carmona, J., editors, *Process Mining Handbook*, volume 448, pages 320–346. Springer International Publishing, Cham. Series Title: Lecture Notes in Business Information Processing.

Di Francescomarino, C., Ghidini, C., Maggi, F. M., and Milani, F. (2018). Predictive Process Monitoring Methods: Which One Suits Me Best? In Weske, M., Montali, M., Weber, I., and vom Brocke, J., editors, *Business Process Management*, volume 11080, pages 462–479. Springer International Publishing, Cham. Series Title: Lecture Notes in Computer Science.

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., and Weber, I. (2023). AI-augmented Business Process Management Systems: A Research Manifesto. *ACM Transactions on Management Information Systems*, 14(1):1–19.

Dumas, M., La Rosa, M., Mendling, J., and Reijers, H. A. (2018). *Fundamentals of Business Process Management*. Springer Berlin Heidelberg, Berlin, Heidelberg.

Dunzer, S., Zilker, S., Marx, E., Grundler, V., and Matzner, M. (2021). The Status

Quo of Process Mining in the Industrial Sector. In Ahlemann, F., Schütte, R., and Stieglitz, S., editors, *Innovation Through Information Systems*, volume 48, pages 629–644. Springer International Publishing, Cham. Series Title: Lecture Notes in Information Systems and Organisation.

Eppler, M. J., Seifried, P., and Röpnack, A. (2008). Improving Knowledge Intensive Processes through an Enterprise Knowledge Medium (1999). In Meckel, M. and Schmid, B. F., editors, *Kommunikationsmanagement im Wandel*, pages 371–389. Gabler, Wiesbaden.

Feuerriegel, S., Hartmann, J., Janiesch, C., and Zschech, P. (2024). Generative AI. *Business & Information Systems Engineering*, 66(1):111–126.

Gold, A. H., Malhotra, A., and Segars, A. H. (2001). Knowledge Management: An Organizational Capabilities Perspective. *Journal of Management Information Systems*, 18(1):185–214.

Gregor, S. (2006). The Nature of Theory in Information Systems. *MIS Quarterly*, 30(3):611–642.

Gregor, S., Kruse, L., and Seidel, S. (2020). Research Perspectives: The Anatomy of a Design Principle. *Journal of the Association for Information Systems*, 21:1622–1652.

Grisold, T., Wurm, B., Mendling, J., and vom Brocke, J. (2020). Using Process Mining to Support Theorizing About Change in Organizations.

Grohs, M., Abb, L., Elsayed, N., and Rehse, J.-R. (2024). Large Language Models Can Accomplish Business Process Management Tasks. In De Weerdt, J. and Pufahl, L., editors, *Business Process Management Workshops*, volume 492, pages 453–465. Springer Nature Switzerland, Cham. Series Title: Lecture Notes in Business Information Processing.

Gronau, N. and Weber, E. (2004). Management of Knowledge Intensive Business Processes. In Desel, J., Pernici, B., and Weske, M., editors, *Business Process Management*, volume 3080, pages 163–178. Springer Berlin Heidelberg, Berlin, Heidelberg. Series Title: Lecture Notes in Computer Science.

Gururangan, S., Marasović, A., Swayamdipta, S., Lo, K., Beltagy, I., Downey, D., and Smith, N. A. (2020). Don't Stop Pretraining: Adapt Language Models to Domains and Tasks. In *Proceedings of the 58th Annual Meeting of the Association for*

Computational Linguistics, pages 8342–8360, Online. Association for Computational Linguistics.

Heinrich, K., Zschech, P., Janiesch, C., and Bonin, M. (2021). Process data properties matter: Introducing gated convolutional neural networks (GCNN) and key-value-predict attention networks (KVP) for next event prediction with deep learning. *Decision Support Systems*, 143:113494.

Herrmann, C. and Kurz, M. (2011). Adaptive Case Management: Supporting Knowledge Intensive Processes with IT Systems. In Schmidt, W., editor, *S-BPM ONE - Learning by Doing - Doing by Learning*, volume 213, pages 80–97. Springer Berlin Heidelberg, Berlin, Heidelberg. Series Title: Communications in Computer and Information Science.

Hevner, A. R. (2007). A Three Cycle View of Design Science Research. *Scandinavian Journal of Information Systems*, 19: Iss. 2, Article 4.

Hevner, A. R., March, S. T., Park, J., and Ram, S. (2004). Design science in information systems research. *MIS Quarterly*, 28(1):75–105.

Holsapple, C. and Joshi, K. (2000). An investigation of factors that influence the management of knowledge in organizations. *The Journal of Strategic Information Systems*, 9(2-3):235–261.

Iivari, J., Rotvit Perlt Hansen, M., and Haj-Bolouri, A. (2021). A proposal for minimum reusability evaluation of design principles. *European Journal of Information Systems*, 30(3):286–303.

Isik, O., van den Bergh, J., and Mertens, W. (2012). Knowledge Intensive Business Processes: An Exploratory Study. In *2012 45th Hawaii International Conference on System Sciences*, pages 3817–3826, Maui, HI, USA. IEEE.

Jessen, U., Sroka, M., and Fahland, D. (2023). Chit-Chat or Deep Talk: Prompt Engineering for Process Mining. Version Number: 1.

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., and Fung, P. (2023). Survey of Hallucination in Natural Language Generation. *ACM Computing Surveys*, 55(12):1–38.

Jurafsky, D. and Martin, J. H. (2025). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition* *with Language Models*. 3rd edition edition. Online manuscript released January 12, 2025.

Kaplan, J., McCandlish, S., Henighan, T., Brown, T. B., Chess, B., Child, R., Gray, S., Radford, A., Wu, J., and Amodei, D. (2020). Scaling Laws for Neural Language Models. Version Number: 1.

Kapoor, S. and Narayanan, A. (2023). Leakage and the reproducibility crisis in machine-learning-based science. *Patterns*, 4(9):100804.

Kerpedzhiev, G. D., König, U. M., Röglinger, M., and Rosemann, M. (2021). An Exploration into Future Business Process Management Capabilities in View of Digitalization: Results from a Delphi Study. *Business & Information Systems Engineering*, 63(2):83–96.

Khanbabaei, M., Alborzi, M., Sobhani, F. M., and Radfar, R. (2019). Applying clustering and classification data mining techniques for competitive and knowledge-intensive processes improvement. *Knowledge and Process Management*, 26(2):123–139.

Kipping, G., Djurica, D., Franzoi, S., Grisold, T., Marcus, L., Schmid, S., Brocke, J. V., Mendling, J., and Röglinger, M. (2022). How to Leverage Process Mining in Organizations - Towards Process Mining Capabilities. In Di Ciccio, C., Dijkman, R., Del Río Ortega, A., and Rinderle-Ma, S., editors, *Business Process Management*, volume 13420, pages 40–46. Springer International Publishing, Cham.

Kolodner, J. L. (1992). An introduction to case-based reasoning. *Artificial Intelligence Review*, 6(1):3–34.

Korzynski, P., Mazurek, G., Altmann, A., Ejdys, J., Kazlauskaite, R., Paliszkiewicz, J., Wach, K., and Ziemba, E. (2023). Generative artificial intelligence as a new context for management theories: analysis of ChatGPT. *Central European Management Journal*, 31(1):3–13.

Kubrak, K., Botchorishvili, L., Milani, F., Nolte, A., and Dumas, M. (2024). Explanatory Capabilities of Large Language Models in Prescriptive Process Monitoring. In Marrella, A., Resinas, M., Jans, M., and Rosemann, M., editors, *Business Process Management*, volume 14940, pages 403–420. Springer Nature Switzerland, Cham. Series Title: Lecture Notes in Computer Science.

Käppel, M., Ackermann, L., Jablonski, S., and Härtl, S. (2024). Attention Please: What

Transformer Models Really Learn for Process Prediction. In Marrella, A., Resinas, M., Jans, M., and Rosemann, M., editors, *Business Process Management*, volume 14940, pages 203–220. Springer Nature Switzerland, Cham. Series Title: Lecture Notes in Computer Science.

Kühn, A., Bensiek, T., and Gausemeier, J. (2013). Framework for the development of maturity based self-assessments for process improvement. In *ICED13. Proceedings of the 19th International Conference on Engineering Design. Volume 1: Design processes*, pages 119–128.

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gotzsche, P. C., Ioannidis, J. P. A., Clarke, M., Devereaux, P. J., Kleijnen, J., and Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ*, 339(jul21 1):b2700–b2700.

Liessmann, A., Wang, W., Weinzierl, S., Zilker, S., and Matzner, M. (2024). Transfer Learning for Predictive Process Monitoring. *ECIS 2024 Proceedings. 4.*

Little, T. A. and Deokar, A. V. (2016). Understanding knowledge creation in the context of knowledge-intensive business processes. *Journal of Knowledge Management*, 20(5):858–879.

Löhr, B., Brennig, K., Bartelheimer, C., Beverungen, D., and Müller, O. (2022). Process Mining of Knowledge-Intensive Processes: An Action Design Research Study in Manufacturing. In Di Ciccio, C., Dijkman, R., Del Río Ortega, A., and Rinderle-Ma, S., editors, *Business Process Management*, volume 13420, pages 251–267. Springer International Publishing, Cham. Series Title: Lecture Notes in Computer Science.

Mahapatra, R. K. and Sarkar, S. (2000). The Role of Information Technology in Knowledge Management. In *AMCIS 2000 Proceedings*.

Manning, C. D. and Schütze, H. (2000). *Foundations of statistical natural language processing*. MIT Press, Cambridge, Mass, 2e éd. avec des corrections edition.

Marjanovic, O. and Freeze, R. (2011). Knowledge Intensive Business Processes: Theoretical Foundations and Research Challenges. In *2011 44th Hawaii International Conference on System Sciences*, pages 1–10, Kauai, HI. IEEE.

Marquez-Chamorro, A. E., Resinas, M., and Ruiz-Cortes, A. (2018). Predictive Monitoring of Business Processes: A Survey. *IEEE Transactions on Services Computing*, 11(6):962–977.

Martin, N., Fischer, D. A., Kerpedzhiev, G. D., Goel, K., Leemans, S. J. J., Röglinger, M., van der Aalst, W. M. P., Dumas, M., La Rosa, M., and Wynn, M. T. (2021). Opportunities and Challenges for Process Mining in Organizations: Results of a Delphi Study. *Business & Information Systems Engineering*, 63(5):511–527.

Meta (2024). Introducing Meta Llama 3: The most capable openly available LLM to date.

Munoz-Gama, J., Martin, N., Fernandez-Llatas, C., Johnson, O. A., Sepúlveda, M., Helm, E., Galvez-Yanjari, V., Rojas, E., Martinez-Millana, A., Aloini, D., Amantea, I. A., Andrews, R., Arias, M., Beerepoot, I., Benevento, E., Burattin, A., Capurro, D., Carmona, J., Comuzzi, M., Dalmas, B., De La Fuente, R., Di Francescomarino, C., Di Ciccio, C., Gatta, R., Ghidini, C., Gonzalez-Lopez, F., Ibanez-Sanchez, G., Klasky, H. B., Prima Kurniati, A., Lu, X., Mannhardt, F., Mans, R., Marcos, M., Medeiros De Carvalho, R., Pegoraro, M., Poon, S. K., Pufahl, L., Reijers, H. A., Remy, S., Rinderle-Ma, S., Sacchi, L., Seoane, F., Song, M., Stefanini, A., Sulis, E., Ter Hofstede, A. H., Toussaint, P. J., Traver, V., Valero-Ramon, Z., Weerd, I. V. D., van der Aalst, W. M., Vanwersch, R., Weske, M., Wynn, M. T., and Zerbato, F. (2022). Process mining for healthcare: Characteristics and challenges. *Journal of Biomedical Informatics*, 127:103994.

Myers, M. D. and Newman, M. (2007). The qualitative interview in IS research: Examining the craft. *Information and Organization*, 17(1):2–26.

Nevo, D. and Chan, Y. E. (2007). A Delphi study of knowledge management systems: Scope and requirements. *Information & Management*, 44(6):583–597.

Nguyen, H. (2017). Stage-based Business Process Mining. In *CAiSE-Forum-DC*.

Nonaka, I. and Takeuchi, H. (1995). *The knowledge-creating company: how Japanese companies create the dynamics of innovation.* Oxford University Press, New York.

Nonaka, I., Toyama, R., and Konno, N. (2000). SECI, Ba and Leadership: a Unified Model of Dynamic Knowledge Creation. *Long Range Planning*, 33(1):5–34.

Nonaka, I. and Von Krogh, G. (2009). Tacit Knowledge and Knowledge Conversion:

Controversy and Advancement in Organizational Knowledge Creation Theory. *Organization Science*, 20(3):635–652.

OpenAI, Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., Avila, R., Babuschkin, I., Balaji, S., Balcom, V., Baltescu, P., Bao, H., Bavarian, M., Belgum, J., Bello, I., Berdine, J., Bernadett-Shapiro, G., Berner, C., Bogdonoff, L., Boiko, O., Boyd, M., Brakman, A.-L., Brockman, G., Brooks, T., Brundage, M., Button, K., Cai, T., Campbell, R., Cann, A., Carey, B., Carlson, C., Carmichael, R., Chan, B., Chang, C., Chantzis, F., Chen, D., Chen, S., Chen, R., Chen, J., Chen, M., Chess, B., Cho, C., Chu, C., Chung, H. W., Cummings, D., Currier, J., Dai, Y., Decareaux, C., Degry, T., Deutsch, N., Deville, D., Dhar, A., Dohan, D., Dowling, S., Dunning, S., Ecoffet, A., Eleti, A., Eloundou, T., Farhi, D., Fedus, L., Felix, N., Fishman, S. P., Forte, J., Fulford, I., Gao, L., Georges, E., Gibson, C., Goel, V., Gogineni, T., Goh, G., Gontijo-Lopes, R., Gordon, J., Grafstein, M., Gray, S., Greene, R., Gross, J., Gu, S. S., Guo, Y., Hallacy, C., Han, J., Harris, J., He, Y., Heaton, M., Heidecke, J., Hesse, C., Hickey, A., Hickey, W., Hoeschele, P., Houghton, B., Hsu, K., Hu, S., Hu, X., Huizinga, J., Jain, S., Jain, S., Jang, J., Jiang, A., Jiang, R., Jin, H., Jin, D., Jomoto, S., Jonn, B., Jun, H., Kaftan, T., Kaiser, , Kamali, A., Kanitscheider, I., Keskar, N. S., Khan, T., Kilpatrick, L., Kim, J. W., Kim, C., Kim, Y., Kirchner, J. H., Kiros, J., Knight, M., Kokotajlo, D., Kondraciuk, , Kondrich, A., Konstantinidis, A., Kosic, K., Krueger, G., Kuo, V., Lampe, M., Lan, I., Lee, T., Leike, J., Leung, J., Levy, D., Li, C. M., Lim, R., Lin, M., Lin, S., Litwin, M., Lopez, T., Lowe, R., Lue, P., Makanju, A., Malfacini, K., Manning, S., Markov, T., Markovski, Y., Martin, B., Mayer, K., Mayne, A., McGrew, B., McKinney, S. M., McLeavey, C., McMillan, P., McNeil, J., Medina, D., Mehta, A., Menick, J., Metz, L., Mishchenko, A., Mishkin, P., Monaco, V., Morikawa, E., Mossing, D., Mu, T., Murati, M., Murk, O., Mély, D., Nair, A., Nakano, R., Nayak, R., Neelakantan, A., Ngo, R., Noh, H., Ouyang, L., O'Keefe, C., Pachocki, J., Paino, A., Palermo, J., Pantuliano, A., Parascandolo, G., Parish, J., Parparita, E., Passos, A., Pavlov, M., Peng, A., Perelman, A., Peres, F. d. A. B., Petrov, M., Pinto, H. P. d. O., Michael, Pokorny, Pokrass, M., Pong, V. H., Powell, T., Power, A., Power, B., Proehl, E., Puri, R., Radford, A., Rae, J., Ramesh, A., Raymond, C., Real, F., Rimbach, K., Ross, C., Rotsted, B., Roussez, H., Ryder, N., Saltarelli, M., Sanders, T., Santurkar, S., Sastry, G., Schmidt, H., Schnurr, D., Schulman, J., Selsam, D., Sheppard, K., Sherbakov, T., Shieh, J., Shoker, S., Shyam, P., Sidor, S., Sigler, E., Simens, M., Sitkin, J., Slama, K., Sohl, I., Sokolowsky, B., Song, Y., Staudacher, N., Such, F. P., Summers, N., Sutskever, I., Tang, J., Tezak, N., Thompson, M. B., Tillet, P., Tootoonchian, A., Tseng, E., Tuggle, P., Turley, N., Tworek, J., Uribe, J. F. C., Vallone, A., Vijayvergiya, A., Voss, C., Wainwright, C., Wang, J. J., Wang, A., Wang, B., Ward, J., Wei, J., Weinmann, C., Welihinda, A., Welinder, P., Weng, J., Weng, L., Wiethoff, M., Willner, D., Winter, C., Wolrich, S., Wong, H., Workman, L., Wu, S., Wu, J., Wu, M., Xiao, K., Xu, T., Yoo, S., Yu, K., Yuan, Q., Zaremba, W., Zellers, R., Zhang, C., Zhang, M., Zhao, S., Zheng, T., Zhuang, J., Zhuk, W., and Zoph, B. (2023). GPT-4 Technical Report. Version Number: 6.

Osuszek, and Stanek, S. (2015). Case Based Reasoning as an improvement of decision making and case processing in Adaptive Case Management systems. pages 217–223.

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., Schulman, J., Hilton, J., Kelton, F., Miller, L., Simens, M., Askell, A., Welinder, P., Christiano, P. F., Leike, J., and Lowe, R. (2022). Training language models to follow instructions with human feedback. In Koyejo, S., Mohamed, S., Agarwal, A., Belgrave, D., Cho, K., and Oh, A., editors, *Advances in Neural Information Processing Systems*, volume 35, pages 27730–27744. Curran Associates, Inc.

Oved, A., Shlomov, S., Zeltyn, S., Mashkif, N., and Yaeli, A. (2025). SNAP: Semantic Stories for Next Activity Prediction. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(28):28871–28877.

Paulk, M. C., Curtis, B., Chrissis, M. B., and Weber, C. V. (1991). *Capability maturity model for software*. Citeseer.

Pegoraro, M., Uysal, M. S., Georgi, D. B., and van der Aalst, W. M. (2021). Text-Aware Predictive Monitoring of Business Processes. *Business Information Systems*, pages 221–232.

Pentland, B., Vaast, E., and Wolf, J. R. (2021). Theorizing Process Dynamics with Directed Graphs: A Diachronic Analysis of Digital Trace Data. *MIS Quarterly*, 45(2):967–984.

Pentland, B. T., Recker, J., Wolf, J., and Wyner, G. (2020). Bringing Context Inside Process Research with Digital Trace Data. *Journal of the Association for Information Systems*, 21(5):1214–1236.

Polanyi, M. (1967). *The tacit dimension*. Anchor books philosophy. Doubleday, Garden City, New York.

Pérez-Castillo, R., Weber, B., De Guzmán, I.-R., and Piattini, M. (2011). Process mining through dynamic analysis for modernising legacy systems. *IET Software*, 5(3):304–319.

Radford, A., Narasimhan, K., Salimans, T., and Sutskever, I. (2018). Improving Language Understanding by Generative Pre-Training.

Rai, R. K. (2011). Knowledge management and organizational culture: a theoretical integrative framework. *Journal of Knowledge Management*, 15(5):779–801.

Rama-Maneiro, E., Vidal, J., and Lama, M. (2022). Deep Learning for Predictive Business Process Monitoring: Review and Benchmark. *IEEE Transactions on Services Computing*, pages 1–1.

Redis, A. C., Sani, M. F., Zarrin, B., and Burattin, A. (2024). ProcessTBench: An LLM Plan Generation Dataset for Process Mining. Version Number: 2.

Reimers, N. and Gurevych, I. (2019). Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 3980–3990, Hong Kong, China. Association for Computational Linguistics.

Reinkemeyer, L., editor (2020). *Process Mining in Action: Principles, Use Cases and Outlook*. Springer International Publishing, Cham.

Reinkemeyer, L., Grindemann, P., Egli, V., Röglinger, M., Marcus, L., and Fabri, L. (2022). Accelerating Business Transformation with a Process Mining Center of Excellence (CoE). Technical report, Fraunhofer-Institut für Angewandte Informationstechnik FIT.

Remus, U. and Lehner, F. (2000). The Role of Process-oriented Enterprise Modeling in Designing Process-oriented Knowledge Management Systems. In *Designing Process-oriented Knowledge Management Systems. 2000 AAAI Spring Symposium, Bringing Knowledge to Business Processes*.

Richetti, P. H. P., De A.R. Gonçalves, J. C., Baião, F. A., and Santoro, F. M. (2017). Analysis of Knowledge-Intensive Processes Focused on the Communication Perspective. In Carmona, J., Engels, G., and Kumar, A., editors, *Business Process Management*, volume 10445, pages 269–285. Springer International Publishing, Cham. Series Title: Lecture Notes in Computer Science.

Rosemann, M. and de Bruin, T. (2004). Application of a Holistic Model for Determining BPM Maturity. *Proceedings of the 3rd Pre-ICIS Workshop on Process Management and Information Systems.*

Rosemann, M. and de Bruin, T. (2005). Towards a Business Process Managment Maturity Model. *ECIS 2005 Proceedings. 37.*

Rosemann, M., Recker, J., and Flender, C. (2008). Contextualisation of business processes. *International Journal of Business Process Integration and Management*, 3(1):47.

Röglinger, M., Plattfaut, R., Borghoff, V., Kerpedzhiev, G., Becker, J., Beverungen, D., vom Brocke, J., van Looy, A., del Río-Ortega, A., Rinderle-Ma, S., Rosemann, M., Santoro, F. M., and Trkman, P. (2022). Exogenous Shocks and Business Process Management: A Scholars' Perspective on Challenges and Opportunities. *Business & Information Systems Engineering*, 64(5):669–687.

Sato, D. M. V., De Freitas, S. C., Barddal, J. P., and Scalabrin, E. E. (2022). A Survey on Concept Drift in Process Mining. *ACM Computing Surveys*, 54(9):1–38.

Sebastiani, F. (2002). Machine learning in automated text categorization. *ACM Computing Surveys*, 34(1):1–47.

Seidel, S., Müller-Wienbergen, F., and Rosemann, M. (2010). Pockets of Creativity in Business Processes. *Communications of the Association for Information Systems*, 27.

Seidler-de Alwis, R. and Hartmann, E. (2008). The use of tacit knowledge within innovative companies: knowledge management in innovative enterprises. *Journal of Knowledge Management*, 12(1):133–147.

Sein, M. K., Henfridsson, O., Purao, S., Rossi, M., and Lindgren, R. (2011). Action Design Research. *MIS Quarterly*, 35(1):37.

Shanahan, M., McDonell, K., and Reynolds, L. (2023). Role play with large language models. *Nature*, 623(7987):493–498.

Simons, A., Niehaves, B., Niehaves, B., and Riemer, K. (2009). Reconstructing the

Giant: On the Importance of Rigour in Documenting the Literature Search Process. In *ECIS 2009 Proceedings. 161.*

Spitale, G., Biller-Andorno, N., and Germani, F. (2023). AI model GPT-3 (dis)informs us better than humans. *Science Advances*, 9(26):eadh1850.

Sui, Y., Zhou, M., Zhou, M., Han, S., and Zhang, D. (2024). Table Meets LLM: Can Large Language Models Understand Structured Table Data? A Benchmark and Empirical Study. In *Proceedings of the 17th ACM International Conference on Web Search and Data Mining*, pages 645–654, Merida Mexico. ACM.

Sumbal, M. S. and Amber, Q. (2024). ChatGPT: a game changer for knowledge management in organizations. *Kybernetes*.

Terziev, Y., Benner-Wickner, M., Brückmann, T., and Gruhn, V. (2015). Ontology-based recommender system for information support in knowledge-intensive processes. In *Proceedings of the 15th International Conference on Knowledge Technologies and Data-driven Business*, pages 1–8, Graz Austria. ACM.

Tversky, A. and Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5(2):207–232.

van der Aalst, W. (2016). *Process Mining - Data Science in Action.* Springer Berlin Heidelberg, Berlin, Heidelberg.

van der Aalst, W., Adriansyah, A., De Medeiros, A. K. A., Arcieri, F., Baier, T., Blickle, T., Bose, J. C., Van Den Brand, P., Brandtjen, R., Buijs, J., Burattin, A., Carmona, J., Castellanos, M., Claes, J., Cook, J., Costantini, N., Curbera, F., Damiani, E., De Leoni, M., Delias, P., Van Dongen, B. F., Dumas, M., Dustdar, S., Fahland, D., Ferreira, D. R., Gaaloul, W., Van Geffen, F., Goel, S., Günther, C., Guzzo, A., Harmon, P., Ter Hofstede, A., Hoogland, J., Ingvaldsen, J. E., Kato, K., Kuhn, R., Kumar, A., La Rosa, M., Maggi, F., Malerba, D., Mans, R. S., Manuel, A., McCreesh, M., Mello, P., Mendling, J., Montali, M., Motahari-Nezhad, H. R., Zur Muehlen, M., Munoz-Gama, J., Pontieri, L., Ribeiro, J., Rozinat, A., Seguel Pérez, H., Seguel Pérez, R., Sepúlveda, M., Sinur, J., Soffer, P., Song, M., Sperduti, A., Stilo, G., Stoel, C., Swenson, K., Talamo, M., Tan, W., Turner, C., Vanthienen, J., Varvaressos, G., Verbeek, E., Verdonk, M., Vigo, R., Wang, J., Weber, B., Weidlich, M., Weijters, T., Wen, L., Westergaard, M., and Wynn, M. (2012). Process Mining Manifesto. In Daniel, F., Barkaoui, K., and Dustdar, S., editors, *Business Process Management Workshops*, volume 99, pages 169–194. Springer Berlin Heidelberg, Berlin, Heidelberg.

van der Aalst, W. M. and Dustdar, S. (2012). Process Mining Put into Context. *IEEE Internet Computing*, 16(1):82–86.

van der Aalst, W. M. P. (2015). Extracting Event Data from Databases to Unleash Process Mining. In vom Brocke, J. and Schmiedel, T., editors, *BPM - Driving Innovation in a Digital World*, pages 105–128. Springer International Publishing, Cham.

van der Aalst, W. M. P. (2022). Process Mining: A 360 Degree Overview. In van der Aalst, W. M. P. and Carmona, J., editors, *Process Mining Handbook*, volume 448, pages 3–34. Springer International Publishing, Cham. Series Title: Lecture Notes in Business Information Processing.

van der Linden, E.-J. (2021). *Successful Process Improvement: a Practice-Based Method to Embed Process Mining in Enterprises*. Tilia Cordata IT B.V, Nijmegen.

van Eck, M. L., Lu, X., Leemans, S. J. J., and van der Aalst, W. M. P. (2015). PM2: A Process Mining Project Methodology. In Zdravkovic, J., Kirikova, M., and Johannesson, P., editors, *Advanced Information Systems Engineering*, volume 9097, pages 297–313. Springer International Publishing, Cham. Series Title: Lecture Notes in Computer Science.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I. (2017). Attention Is All You Need. Version Number: 7.

Verenich, I., Dumas, M., Rosa, M. L., Maggi, F. M., and Teinemaa, I. (2019). Survey and Cross-benchmark Comparison of Remaining Time Prediction Methods in Business Process Monitoring. *ACM Transactions on Intelligent Systems and Technology*, 10(4):1–34.

Vidgof, M., Bachhofner, S., and Mendling, J. (2023). Large Language Models for Business Process Management: Opportunities and Challenges. In Di Francescomarino, C., Burattin, A., Janiesch, C., and Sadiq, S., editors, *Business Process Management Forum*, volume 490, pages 107–123. Springer Nature Switzerland, Cham. Series Title: Lecture Notes in Business Information Processing.

vom Brocke, J., Jans, M., Mendling, J., and Reijers, H. A. (2021a). A Five-Level Framework for Research on Process Mining. *Business & Information Systems Engineering*, 63(5):483–490.

vom Brocke, J., van der Aalst, W. M., Grisold, T., Kremser, W., Mendling, J., Pentland, B., Recker, J., Roeglinger, M., Rosemann, M., and Weber, B. (2021b). Process Science: The Interdisciplinary Study of Continuous Change. *SSRN Electronic Journal*.

Weinzierl, S., Dunzer, S., Zilker, S., and Matzner, M. (2020). Prescriptive Business Process Monitoring for Recommending Next Best Actions. In Fahland, D., Ghidini, C., Becker, J., and Dumas, M., editors, *Business Process Management Forum*, volume 392, pages 193–209. Springer International Publishing, Cham. Series Title: Lecture Notes in Business Information Processing.

Weske, M. (2019). *Business Process Management: Concepts, Languages, Architectures*. Springer Berlin Heidelberg, Berlin, Heidelberg.

Wu, Z., Arora, A., Wang, Z., Geiger, A., Jurafsky, D., Manning, C. D., and Potts, C. (2024). Re{FT}: Representation Finetuning for Language Models.

Wurm, B., Grisold, T., Mendling, J., and vom Brocke, J. (2021). Business Process Management and Routine Dynamics. In Feldman, M. S., Pentland, B. T., D'Adderio, L., Dittrich, K., Rerup, C., and Seidl, D., editors, *Cambridge Handbook of Routine Dynamics*, pages 513–524. Cambridge University Press, 1 edition.

Young, T., Hazarika, D., Poria, S., and Cambria, E. (2018). Recent Trends in Deep Learning Based Natural Language Processing [Review Article]. *IEEE Computational Intelligence Magazine*, 13(3):55–75.

Zhang, Y., Wang, X., Liang, J., Xia, S., Chen, L., and Xiao, Y. (2024). Chain-of-Knowledge: Integrating Knowledge Reasoning into Large Language Models by Learning from Knowledge Graphs. Version Number: 1.

Zheng, D., Lapata, M., and Pan, J. Z. (2024). How Reliable are LLMs as Knowledge Bases? Re-thinking Facutality and Consistency. Version Number: 2.