Knowledge-Aware Process Mining: Conceptual Foundations and IT Artifacts for Process Mining in Knowledge-Intensive Processes (Extended Abstract)*

Katharina Brennig[1]

[1]Paderborn University, Faculty of Business Administration and Economics, Department of Information Systems, Warburger Str. 100, 33098 Paderborn, Germany

Abstract

Business processes are the backbone of organizational value creation, yet not all processes lend themselves to standardization and full digitalization. Knowledge-intensive processes (KIPs)–such as product innovation or healthcare treatment–depend on the tacit knowledge, judgment, and experience of knowledge workers and are therefore poorly represented in the structured event logs on which conventional process mining is based. This cumulative dissertation develops a *Knowledge-Aware Process Mining* approach that extends process mining along three complementary layers: organizational capabilities, theoretical foundations, and technical realization. The seven included papers (i) provide strategic guidelines and a maturity model that help organizations build process mining capabilities, (ii) derive, design principles for process mining of KIPs, grounded in the theory of organizational knowledge creation, and (iii) demonstrate how natural language processing (NLP), especially large language models (LLMs), can mobilize tacit process-related knowledge, enrich event logs, and generate predictions with minimal preprocessing. Together, the contributions show how process mining can evolve from a technique for standardized, digitalized processes into a knowledge-aware discipline that supports the dynamic, human-centered nature of knowledge-intensive work.

Keywords

Process Mining, Knowledge-Intensive Processes, Business Process Management, Natural Language Processing, Organizational Knowledge Creation, Design Science Research, Action Design Research

1. Introduction

Not all organizational processes are standardized or digitized: knowledge-intensive processes (KIPs) (e.g., product innovation or healthcare treatment) [1] represent strategic assets of modern organizations [2], enabling them to leverage specialized knowledge for competitive advantage [3]. However, KIPs are often executed manually, as existing information technology provides only limited functionality to capture their knowledge-intensity, complexity, and flexibility [1]. To move beyond this, organizations need methods to observe, reconstruct, and understand the dynamic behavior of KIPs–for which process mining, a data-driven method to manage business processes, offers valuable foundations [4].

Nonetheless, applying process mining to KIPs introduces distinct challenges that hinder its straightforward use. First, many organizations plan to adopt process mining or have piloted it, yet widespread, effective implementation remains limited, as ambitions for continuous improvement often fail to translate into sustained impact. This is due to obstacles ranging from insufficient managerial support and poor data quality to complex data preparation [5]. Second, existing process mining approaches assume all relevant data is available in event logs, making them suited to standardized, digitalized processes but limiting their applicability to KIPs, as event logs omit the experiences and informal decision-making essential to executing and understanding them [4, 1]. Third, existing process mining approaches are limited in their ability to process and mobilize the process-related knowledge embedded in KIPs, largely due to the absence of natural language processing (NLP) approaches that can extract, represent, and analyze such knowledge in a way that reliably supports KIPs [6, 7]. Taken together,

Proceedings of the BPM 2026 Best Dissertation Award, Doctoral Consortium, and Demonstration & Resources Forum (BPM 2026) , Toronto, Canada, September 27th to October 2nd, 2026.
katharina.brennig@uni-paderborn.de (K. Brennig)
0000-0003-0007-2824 (K. Brennig)

a holistic management of KIPs through process mining is constrained. To address this, this thesis aims to establish a *Knowledge-Aware Process Mining* approach along three dimensions: developing organizational capabilities and process mining maturity, reconceptualizing the theoretical foundations for applying process mining to KIPs, and integrating NLP into process mining for KIPs–leading to three research questions:

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

2. Research Background

Organizations continuously strive to improve their processes, recognizing process quality as a key source of competitive advantage [8]. Process mining has evolved as a data-driven business process management (BPM) approach that uses event logs from information systems (e.g., ERP or BPM systems) to discover, monitor, and improve actual execution of business processes, working best with highly digitalized, standardized processes and large data volumes [4]. By enabling real-time analysis and evidence-based decisions, process mining helps organizations eliminate bottlenecks, enable automation, and reduce costs [4]. However, many still struggle to translate these ambitions into lasting impact due to obstacles such as lack of management support, poor data quality, and complex data preparation [5].

This reliance on highly digitalized, standardized processes, however, sets process mining apart from how organizations manage KIPs (e.g., product innovation or healthcare treatment), which are less structured and more complex, relying on contextual interpretation, individual judgment, and experiential knowledge that makes them difficult to standardize, digitize, or fully model [1]. Process-related knowledge in KIPs is accordingly either explicit, i.e., codified and storable, or tacit, i.e., residing in participants' skills and experience, which is difficult to formalize and communicate as it is personal and context-specific [1, 9]. Organizations should therefore foster a culture of knowledge exchange in which tacit knowledge is mobilized into explicit knowledge–and vice versa–for which the SECI model offers a conceptual foundation. It describes four interdependent modes of knowledge conversion (socialization, externalization, combination, internalization) that form a spiral scaling from individuals to a shared organizational knowledge base [9]. This dependence on tacit knowledge, however, exposes a core limitation of process mining: since event logs typically capture only activities, control flow, and resource usage, they systematically overlook the tacit knowledge that drives KIPs, leaving KIPs to be analyzed, redesigned, and executed only on the basis of their explicit knowledge [1, 4].

Closing this gap requires methods capable of processing the very knowledge event logs leave out–a role increasingly filled by NLP, and especially large language models (LLMs), which are currently applied in process mining, e.g., to process XES-formatted event logs and support predictive and prescriptive monitoring [10]. Beyond this, LLMs show potential to bridge tacit and explicit process-related knowledge as conversational agents that support all four modes of knowledge conversion within knowledge management systems, provided their hallucination risk is critically managed [7].

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 presented in this dissertation. Table 1 provides an overview of each paper's contribution and its mapping to the three research questions guiding this thesis. Figure 1 visualizes how the individual papers interrelate across three nested layers, moving from organizational readiness to technical realization, showing their cumulative contribution toward *Knowledge-Aware Process Mining*.

Table 1
Main Research Contributions

RQ	#P	Contributions	Ref.
RQ1	P1	Four strategic guidelines to assess and enhance the relevance of process mining results.	[11]
	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.	[12]
RQ2	P3	Five design principles for applying process mining in KIPs.	[13]
	P4	Extends P3 and develops five improved design principles, grounded in the theory of organizational knowledge creation, that establish theoretical foundations for applying process mining in KIPs.	[14]
RQ3	P5	A LLM-based framework to operationalize the mobilization of tacit process-related knowledge, technically realizing the SECI-based conceptual approach.	[15]
	P6	A text-aware PPM approach applying encoder-only models for natural language understanding and neglecting the control-flow of KIPs.	[16]
	P7	A LLM-based approach which can directly generate next event predictions from XES-formatted, knowledge-enriched event logs.	[17]

The outer layer addresses RQ1, focusing on organizational capabilities for process mining. P1's guidelines and P2's maturity model help researchers plan and evaluate process mining projects and help practitioners allocate resources and overcome adoption barriers [5]–particularly critical for KIPs, given their variability and complex process flows [1]. P2 further shows that advancing process mining maturity increases organizational flexibility, laying essential groundwork for *Knowledge-Aware Process Mining*.

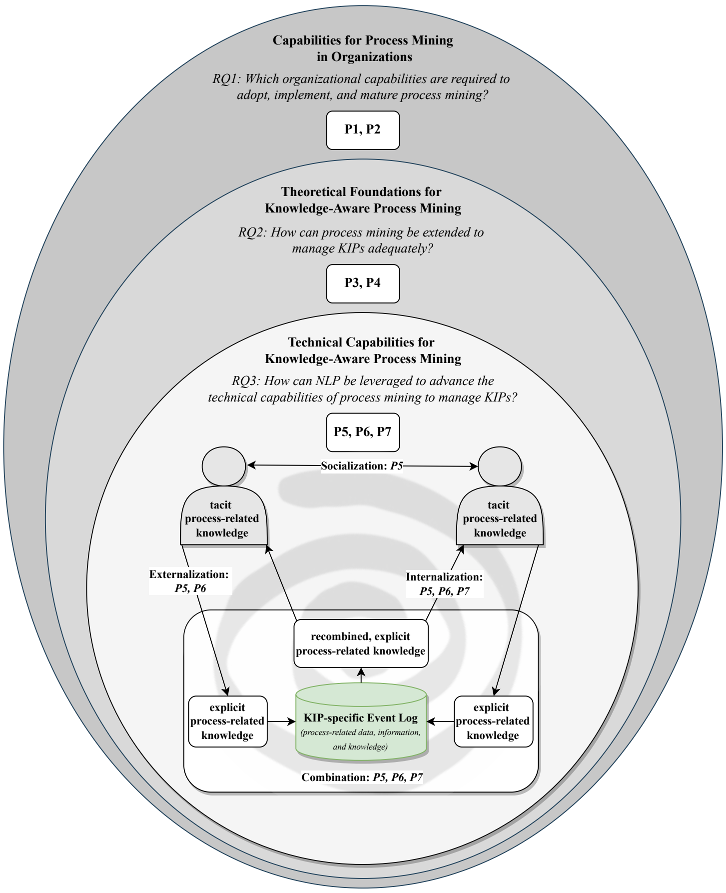

Figure 1: Overview of Research Contributions and their Interrelationships

The middle layer addresses RQ2, establishing theoretical foundations for *Knowledge-Aware Process Mining*. P3 and P4 conceptualize a new class of process mining artifacts for KIPs through design principles grounded in the SECI model [9]. This extends process mining through continuous, spiral cycles of knowledge creation that evolve from individual instances to process-type levels and offer blueprints for transforming tacit knowledge into explicit, analyzable formats to support more objective, strategic decision-making in KIPs. The inner layer addresses RQ3, presenting the technical capabilities that integrate NLP as a key enabler to mobilize process-related knowledge in KIPs, capturing the circular, iterative knowledge mobilization cycle: tacit knowledge is externalized, combined with relevant data into a KIP-specific event log, and internalized by participants as recombined, explicit knowledge (e.g., predictions or recommended actions). P5 to P7 operationalize this cycle technically: P5 shows the feasibility of externalizing tacit knowledge into event logs via LLMs; P6 demonstrates that this enriched knowledge base improves predictive process monitoring using encoder-based models; and P7 shows that LLMs can generate next event predictions directly from XES-formatted, knowledge-enriched logs. These text-aware approaches capture syntactic and semantic relationships beyond control-flow-only methods. P7 further shows that LLMs can generate valid XES output directly, simplifying data preparation. Managerially, these frameworks help organizations build robust knowledge bases, reduce knowledge loss and judgment bias, lowering barriers to adoption. Together, these contributions advance process mining toward a knowledge-aware approach, moving from organizational readiness (P1, P2), through conceptual design (P3, P4), to technical realization (P5–P7).

4. Limitations and Future Outlook

While this thesis advances process mining toward a more knowledge-aware discipline, its generalizability is limited, as most findings stem from specific organizational contexts (manufacturing, healthcare) and close ADR partnerships. Further, the LLM-based framework and predictive models were evaluated only through prototypes and specific datasets, leaving scalability and real-world adoption underexplored. Technical limitations include constrained context windows for the extensive textual descriptions typical of KIP event logs and the persistent risk of LLM hallucination, whose practical impact on process execution was not deeply examined. Future research should therefore empirically validate the theoretical foundations across broader organizational contexts, extend the prototypical artifacts into deployable systems, and assess their feasibility and user acceptance in real-world settings. Given the centrality of human knowledge in KIPs, it is particularly important to investigate participants' willingness to contribute their knowledge, and to identify the organizational conditions that foster sustained knowledge sharing.

Declaration on Generative AI

The author used a generative AI tool to assist with grammar, spelling, and rephrasing. The author has reviewed and edited all AI-generated content and take full responsibility for the publication's content.

References

[1] C. Di Ciccio, A. Marrella, A. Russo, Knowledge-Intensive Processes: Characteristics, Requirements and Analysis of Contemporary Approaches, Journal on Data Semantics 4 (2015) 29–57. doi:10.1007/s13740-014-0038-4.

[2] J. C. De Almeida Rodrigues Gonçalves, F. A. Baião, F. M. Santoro, G. Guizzardi, A cognitive BPM theory for knowledge-intensive processes, Business Process Management Journal 29 (2023) 465–488. doi:10.1108/BPMJ-11-2021-0746.

[3] M. Alavi, D. E. Leidner, Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues, MIS Quarterly 25 (2001) 107. doi:10.2307/3250961.

[4] W. M. P. van der Aalst, Process Mining: A 360 Degree Overview, in: W. M. P. van der Aalst, J. Carmona (Eds.), Process Mining Handbook, volume 448, Springer International Publishing, Cham, 2022, pp. 3–34. doi:10.1007/978-3-031-08848-3_1, series Title: Lecture Notes in Business Information Processing.

[5] N. Martin, D. A. Fischer, G. D. Kerpedzhiev, K. Goel, S. J. J. Leemans, M. Röglinger, W. M. P. van der Aalst, M. Dumas, M. La Rosa, M. T. Wynn, Opportunities and Challenges for Process Mining in Organizations: Results of a Delphi Study, Business & Information Systems Engineering 63 (2021) 511–527. doi:10.1007/s12599-021-00720-0.

[6] S. Feuerriegel, J. Hartmann, C. Janiesch, P. Zschech, Generative AI, Business & Information Systems Engineering 66 (2024) 111–126. doi:10.1007/s12599-023-00834-7.

[7] M. Alavi, D. E. Leidner, R. Mousavi, Knowledge Management Perspective of Generative Artificial Intelligence, Journal of the Association for Information Systems 25 (2024) 1–12. doi:10.17705/1jais.00859.

[8] M. Dumas, M. La Rosa, J. Mendling, H. A. Reijers, Fundamentals of Business Process Management, Springer Berlin Heidelberg, Berlin, Heidelberg, 2018. doi:10.1007/978-3-662-56509-4.

[9] I. Nonaka, H. Takeuchi, The knowledge-creating company: how Japanese companies create the dynamics of innovation, Oxford University Press, New York, 1995.

[10] A. Berti, H. Kourani, H. Häfke, C.-Y. Li, D. Schuster, Evaluating Large Language Models in Process Mining: Capabilities, Benchmarks, and Evaluation Strategies, in: H. van der Aa, D. Bork, R. Schmidt, A. Sturm (Eds.), Enterprise, Business-Process and Information Systems Modeling, volume 511, Springer Nature Switzerland, Cham, 2024, pp. 13–21. doi:10.1007/978-3-031-61007-3_2, series Title: Lecture Notes in Business Information Processing.

[11] K. Brennig, B. Löhr, J. Brock, M. Reineke, C. Bartelheimer, Maximizing the Impact of Process Mining Research: Four Strategic Guidelines, in: AMCIS 2024 Proceedings, 2024.

[12] J. Brock, K. Brennig, B. Löhr, C. Bartelheimer, S. von Enzberg, R. Dumitrescu, Improving Process Mining Maturity – From Intentions to Actions, Business & Information Systems Engineering 66 (2024) 585–605. doi:10.1007/s12599-024-00882-7.

[13] B. Löhr, K. Brennig, C. Bartelheimer, D. Beverungen, O. Müller, Process Mining of Knowledge-Intensive Processes: An Action Design Research Study in Manufacturing, in: C. Di Ciccio, R. Dijkman, A. Del Río Ortega, S. Rinderle-Ma (Eds.), Business Process Management, volume 13420, Springer International Publishing, Cham, 2022, pp. 251–267. doi:10.1007/978-3-031-16103-2_18, series Title: Lecture Notes in Computer Science.

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

[15] K. Brennig, Revealing the Unspoken: Using LLMs to Mobilize and Enrich Tacit Knowledge in Event Logs of Knowledge-Intensive Processes, in: AMCIS 2025 Proceedings, 2025.

[16] K. Brennig, K. Benkert, B. Löhr, O. Müller, Text-Aware Predictive Process Monitoring of Knowledge-Intensive Processes: Does Control Flow Matter?, in: J. De Weerdt, L. Pufahl (Eds.), Business Process Management Workshops, volume 492, Springer Nature Switzerland, Cham, 2024, pp. 440–452. doi:10.1007/978-3-031-50974-2_33, series Title: Lecture Notes in Business Information Processing.

[17] K. Brennig, S. Kaltenpoth, O. Müller, Straight Outta Logs: Can Large Language Models Overcome Preprocessing in Next Event Prediction?, in: K. Gdowska, M. T. Gómez-López, J.-R. Rehse (Eds.), Business Process Management Workshops, volume 534, Springer Nature Switzerland, Cham, 2025, pp. 197–208. doi:10.1007/978-3-031-78666-2_15, series Title: Lecture Notes in Business Information Processing.