UNIVERSITÄT BAYREUTH

Advancing Process Mining from the Core: Managing Process Mining Project Portfolios from Data Processing to Process Improvement

Dissertation

zur Erlangung des Grades eines Doktors der Wirtschaftswissenschaft
der Rechts- und Wirtschaftswissenschaftlichen Fakultät
der Universität Bayreuth

vorgelegt
von
Dominik Andreas Fischer
aus
Wörth an der Donau

Dekan:	Prof. Dr. Michael Grünberger
Erstberichterstatter:	Prof. Dr. Maximilian Röglinger
Zweitberichterstatterin:	Prof. Dr. Moe Thandar Wynn
Tag der mündlichen Prüfung:	10.05.2023

Not using process mining is a sign of self-neglect, showing an inability or unwillingness to manage processes properly.

Wil van der Aalst - "Godfather" of Process Mining

2014 startete ich hochmotiviert in mein Bachelorstudium. Wer hätte gedacht, dass meine Reise an der Universität Bayreuth fast neun Jahre dauern und mit Abschluss einer Promotion enden würde - ich jedenfalls nicht. Diese Zeit hat mich nachhaltig geprägt, ich durfte unglaublich viel lernen und wunderbare Menschen und Persönlichkeiten kennenlernen. So blicke ich auf eine außergewöhnliche Zeit zurück und bin froh und stolz, diesen Weg gegangen zu sein. Es sind aber insbesondere die Menschen, denen man auf diesem Weg begegnet, die diese Zeit unvergesslich und einzigartig machen. Daher gilt mein besonderer Dank all jenen, die mich in den vergangenen Jahren stets unterstützt haben und die ich auch in Zukunft nicht missen möchte.

An meinen Doktorvater Max Röglinger: Ich danke dir für dein unendliches Vertrauen, dein Fordern und dein Fördern. Ich schätze dich nicht nur als Wissenschaftler, sondern vor allem als Mensch. Ich bewundere deinen Ehrgeiz, deine Zielstrebigkeit und dein Einfühlungsvermögen. Ich durfte in der engen Zusammenarbeit mit dir unglaublich viel lernen, konnte, wann immer nötig, auf deinen wertvollen Rat zählen, und den Freiraum genießen, auch meinen eigenen Interessen zu folgen. Ganz besonders danke ich dir auch für die Vermittlung nach Brisbane. Ich blicke hier auf eine einzigartige Zeit zurück, in der ich eine hervorragende Grundlage für meine Promotion legen konnte.

An meine Freund:innen, Co-Autor:innen und Kolleg:innen am Institutsteil Wirtschaftsinformatik des Fraunhofer FIT, am FIM Forschungsinstitut, an der Universität Bayreuth und an der Queensland University of Technology: Ich danke euch für die tägliche Zusammenarbeit, die mir auch in intensiven Zeiten immer großen Spaß bereitet hat, die vielen abendlichen Forschungs-Sessions, die tollen Konferenzreisen, sowie die unvergesslichen Events. Allen aktuellen sowie zukünftigen Doktorand:innen wünsche ich viel Durchhaltevermögen und die selbe Unterstützung, wie ich sie erfahren durfte.

An meine Partnerin Franziska: Danke für für deine bedingungslose Unterstützung und deine unendliche Liebe. Danke, dass du auch in turbulenten Zeiten immer Verständnis hattest, wenn mal wieder etwas weniger Zeit für uns übrig blieb. Du gibst mir den Ausgleich, den ich brauche. Du gibst mir die Zuneigung, die ich mir wünsche. Auch wenn wir individuell anstregende Zeiten durchleben, hat mich die Zeit gelehrt, dass wir gemeinsam stärker sind. Ich freue mich auf die gemeinsame Zukunft mit dir!

Zuletzt an meine Familie - Alois, Lydia, Alexander und Hannah: Ihr habt mir immer den Rücken frei gehalten, mich bei all meinen Vorhaben unterstützt und dafür, wenn nötig, auch eigene Lasten auf euch genommen. Ich konnte mich nach meinen eigenen Interessen frei entfalten und meinen eigenen Weg gehen. Dafür bin ich euch unendlich dankbar. Ihr akzeptiert mich so, wie ich bin. Man kann sich seine Familie nicht aussuchen. Selbst wenn ich könnte, würde ich mir nichts anderes wünschen.

Ihr alle habt mich zu dem gemacht, was ich heute bin.
Euch allen widme ich diese Dissertation!

Copyright Statement

The following sections are partly comprised of content from the research papers included in this thesis. To improve the readability of the text, I omit the standard labeling of these citations.

Abstract

Process mining is a specialized form of data-driven process analysis that organizations use to understand and improve their business processes. Applying process mining techniques such as process discovery, conformance checking, or enhancement using event logs as the central data source generates insights into process behavior, performance, and compliance. Turning these insights into action supports evidence-based process improvement and strategic decision-making. Therefore, process mining supports multiple phases of the business process management lifecycle (i.e., process discovery, process analysis, process improvement and implementation, and process monitoring and controlling) using data about the execution of a process. The groundwork for these outstanding developments has been laid in academia, where a huge research stream focuses on developing and improving new process mining algorithms for various use cases, resulting in a strong technology core for process mining and analysis techniques.

The overall purpose of this dissertation is to advance process mining by building on its solid technological core around the numerous process mining analysis algorithms by adding the missing pieces in preceding and subsequent steps of end-to-end process mining projects. Furthermore, this dissertation also abstracts from a single-project perspective and contributes on the managerial side to the broad applicability of process mining in organizations. Applying design science research principles, the research objectives of this thesis are primarily addressed through design-oriented research by creating and evaluating multiple artifacts in the form of reference architectures, methods, and instantiations. Ultimately, researchers in the process mining field as well as practitioners on the vendor and adopter side should benefit equally from the contributions of this thesis. Therefore, this cumulative dissertation comprising five research papers addresses three challenges that slow down the widespread adoption of process mining in organizations.

First, research on adopting process mining at the enterprise level is somewhat fragmented, leading to a call for better guidance on managing process mining project portfolios, complemented by a holistic understanding of the opportunities and challenges of using PM in organizational settings. Therefore, this dissertation provides two deliverables to address this research need: Research Paper P1 provides a holistic overview of the opportunities and challenges of using process mining in organizations. Further, Research Paper P2 developed a method to manage portfolios of process mining projects in a value-oriented manner.

Second, for process data quality management, there is a need for a dedicated environment focused on detecting, measuring, and repairing data quality problems. Research Paper

P3 proposes a reference architecture for process data quality management to address this research need. The reference architecture is designed to be comprehensive and flexible enough to incorporate current and future contributions to the multifaceted problem of process data quality. Before applying process mining techniques, the high quality of the underlying process data should be validated. However, to my knowledge, research on event log quality assessment of event logs remains scarce. Therefore, Research Paper P4 proposes a second contribution to process data quality management by proposing a user-guided and semi-automated approach for detecting and quantifying process data quality issues in event logs.

Third, this dissertation aims to contribute to the data-driven process improvement enabled by process mining. There is still a lack of research on balancing tool-based automation and guidance of business process improvement tasks by incorporating process data and domain expertise. Research Paper P5 addresses this gap by proposing a reference architecture that guides users in improving business processes by leveraging existing process data.

The dissertation concludes with a reflection on some limitations that stimulate future research. Overall, this dissertation and the embedded research papers contribute to the identified gaps and aim to enable organizations to realize the full potential of process mining. First, this thesis provides a foundation that enables organizations to tailor their process mining project roadmaps to take advantage of the relevant opportunities and be fully aware of the prevailing challenges. Second, this thesis provides two artifacts for process data quality management to support the preceding data processing step before applying process mining. Third, this thesis presents an approach for assisted business process redesign to support the subsequent process improvement step that can leverage insights from applying process mining.

Table of Contents

List of Figures

List of Tables

Acronyms

aBPR assisted business process redesign.

AI artificial intelligence.

AL automation level.

BPI business process improvement.

BPM business process management.

BPMN business process model and notation.

BPR business process redesign.

CAGR compound annual growth rate.

CRediT Contributor Roles Taxonomy.

DO design objective.

DSR design science research.

FEDS framework for evaluation in design science.

GUI graphical user interface.

IS information systems.

PM process mining.

PPM project portfolio management.

RA reference architecture.

SME situational method engineering.

TAM technology acceptance model.

I Introduction

I.1 Motivation

Process mining (PM) is a specialized form of data-driven process analysis that organizations use to understand and improve their business processes (Martin, D. A. Fischer, et al., 2021; vom Brocke et al., 2021). Applying PM techniques such as process discovery, conformance checking, or enhancement using event logs as the central data source generates insights into process behavior, performance, and compliance (van der Aalst, 2016). Turning these insights into action supports evidence-based process improvement and strategic decision-making (Martin, D. A. Fischer, et al., 2021). Therefore, PM supports multiple phases of the business process management (BPM) lifecycle (i.e., process discovery, process analysis, process improvement and implementation, and process monitoring and controlling) using data about the execution of a process (van der Aalst, Adriansyah, et al., 2012; van der Aalst, 2016; Dumas et al., 2018).

In recent years, PM has gained tremendous industry attention and has become a highly sought-after technology. Although the first PM startups entered the market only a decade ago, Gartner (2023) already monitors 35 PM vendors, and the PM ecosystem continues to increase. Founded in 2011, Celonis, the undisputed PM market leader, is already valuated at $13 billion (Metinko, 2022). This rapid growth is also reflected in the global PM market, which has grown at a compound annual growth rate (CAGR) of approximately 70% in recent years and is projected to exceed $15 trillion (Kerremans et al., 2021; Fortune Business Insights, 2022). Meanwhile, several tech giants such as SAP, Microsoft, and IBM also recognized the huge market potential as evidenced by their recent acquisitions of PM vendors (see Signavio (2021), Graham (2022), and Lunden (2022)).

The groundwork for these outstanding developments has been laid in academia. A vast research stream focuses on developing and improving ever-new PM algorithms for various use cases, resulting in a strong technology core for process mining and analysis techniques. Initially, PM research focused on the control-flow discovery, i.e., retrieving a process flow model from an event log. While the control-flow discovery remains an important use case (Augusto et al., 2018), PM research has broadened its scope over time to include techniques for checking conformance between a control-flow model and an event log (Carmona et al., 2018), gaining insights in the involvement of resources in a process (Song and van der Aalst, 2008), or connecting PM to other techniques such as simulation and predictive process monitoring (Kratsch et al., 2020; Martin, Depaire, et al., 2016; Teinemaa et al., 2019). While many of the state-of-the-art PM algorithms have been integrated into the open-source platform ProM, the use of PM in organizations has been stimulated by the development of commercial tools (van der Aalst, 2016).

Since these PM algorithms in isolation do not provide value to organizations, van Eck et al. (2015) present a PM project methodology that guides the execution of individual PM projects end-to-end and is divided into three phases with corresponding sub-activities (see Figure 1): (a) The *initialization* phase covers the planning of a project (i.e., identifying research questions, selecting business processes, and composing the project team) and the data extraction (i.e., determining the scope, extracting event data, and transferring process knowledge). (b) The *analysis* phase includes the data processing (i.e., creating views, aggregating events, enriching logs, and filtering logs), mining & analysis (i.e., process discovery, conformance checking, enhancement, and process analytics), and evaluation (i.e., diagnose, and verify & validate). (c) The *process improvement & support* phase includes implementing improvements and supporting operations.

While the technological maturity of PM algorithms, i.e., the basis for the "mining & analysis" activity of van Eck et al. (2015) (Figure 1), is considered sufficiently high due to the significant investments in the development of new or better algorithms over the past decades, support for the preceding and subsequent steps is still fragmented. To achieve reliable PM results, event logs with sufficient data quality generated by the "data processing" activity (Figure 1) are a crucial requirement (van der Aalst, 2016; Andrews, C. G. J. van Dun, et al., 2020; van der Aalst, Adriansyah, et al., 2012). In practice, event logs are often

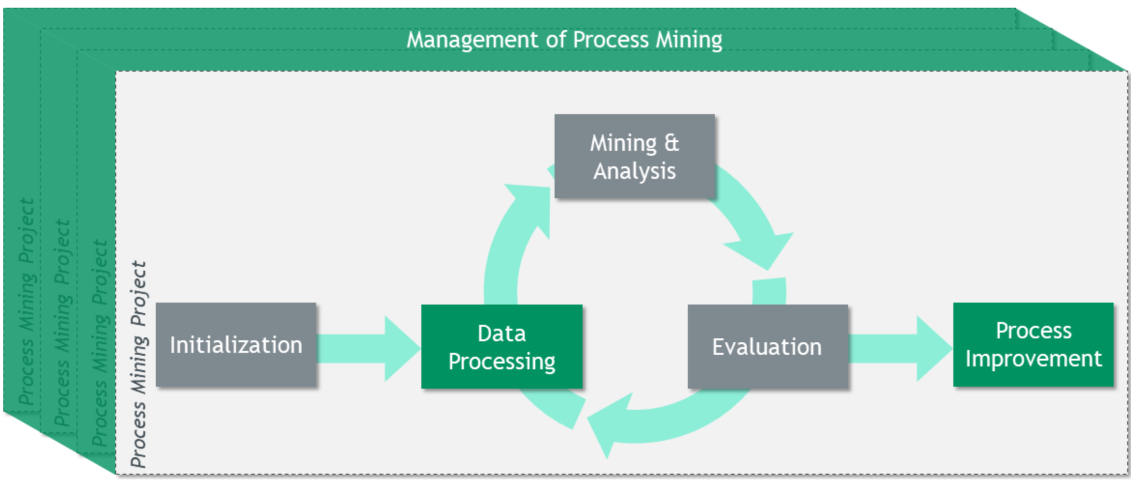

Figure 1: Focus areas of the dissertation adapted from van Eck et al. (2015)

far from the desired quality (Bose et al., 2013; Suriadi et al., 2017). Therefore, event logs should not be used naively for PM without ensuring adequate event data quality (van der Aalst, 2016). Data scientists spend up to eighty percent of their work identifying, assessing, and remediating data quality issues (Wynn and Sadiq, 2019). Thus, there is a growing interest in the BPM community to explore the roots of data quality issues and the associated assurance of high-quality log data (Wynn and Sadiq, 2019; van der Aalst, Bichler, et al., 2017).

The problem of data quality is multifaceted. Research has emphasized quality problems as they manifest in event logs (e.g., Bose et al. (2013) and Suriadi et al. (2017)). Managing process data has its own requirements, some of which are different from other types of data, which means that process data governance needs to consider the nature of PM (Goel et al., 2021). Thus, managing process data quality requires a deep understanding of the business environment, culture, processes, and goals. The complex and rich nature of process data quality management means that technological solutions should be able to evolve as our understanding of the field increases. New prevention and detection strategies should not require a complete overhaul of the architecture of such solutions. While existing PM environments, such as ProM and Disco, provide support for managing specific data quality issues typically found in event logs, such as filtering and abstraction, a dedicated environment focused on detecting, measuring, and repairing data quality issues is a sine qua non for the next generation of process data quality management.

Selected real-world case reports show that organizations can realize significant value through PM by continuously identifying opportunities for "process improvement" (Figure 1) (Grisold et al., 2020; Reinkemeyer, 2020). Transforming business processes at an accelerated pace is essential for organizations to meet increasing competition and customer demands (Beverungen et al., 2021; Huang et al., 2015). In BPM, business process improvement (BPI) (often referred to as "process redesign" or "process reengineering") is concerned with improving business processes to address previously identified process-related issues, for example, by applying PM techniques (Dumas et al., 2018). BPI projects involve significant human and technical investments but also yield promising returns (Huang et al., 2015). Therefore, BPI is generally considered the most value-creating phase in the BPM lifecycle (Dumas et al., 2018; Gross et al., 2021; Reijers and Limam Mansar, 2005; Zellner, 2011). Despite the importance of BPI projects and the abundant availability of BPI methods, 60-80% are reported to fail (vom Brocke et al., 2021; Gross et al., 2021;

Zuhaira and Ahmad, 2021). The failure of BPI projects is rooted in the fact that BPI itself still “happens in a black box” (Zellner, 2011, p. 217). Therefore, the quality and effectiveness of BPI depend on the creativity and expertise of the project team to find valuable solutions (Essam and Limam Mansar, 2012).

Alongside methods, tools are essential for managing the complexity of business processes and supporting their improvement and deployment (Zuhaira and Ahmad, 2021). While most of the literature presenting BPI methods does not provide tool support, some approaches build on redesign patterns to generate tool-based suggestions for their application to business processes (Fellmann et al., 2019; Netjes et al., 2010; Zuhaira and Ahmad, 2021). However, they have limitations, such as (1) they rely on data that is difficult to retrieve, (2) they are inflexible due to hard-coded assumptions, and (3) only a few approaches provide the ability to incorporate a variety of redesign patterns (Essam and Limam Mansar, 2012). Therefore, it is questionable to what extent such tool-based approaches can handle the complexity and the rich information about business processes and provide actionable suggestions for BPI (Essam and Limam Mansar, 2012). While this research gap has been recognized in the literature, no assistive approach combines both worlds in a guided process by leveraging process data, e.g., aggregated by applying PM techniques (Röglinger, C. van Dun, et al., 2021; Essam and Limam Mansar, 2012): tool-based automation and guidance of BPI tasks on the one hand, and the incorporation of process data and domain expertise on the other hand.

To understand the organizational impact of PM and its solid technical core, it is essential to zoom out from a single-project perspective to the enterprise level and focus on questions such as how organizations adopt PM, how they integrate PM to support BPM, and how organizations generate value from PM (vom Brocke et al., 2021; Martin, D. A. Fischer, et al., 2021; Grisold et al., 2020). Recently, case studies have also been published that report on the application of PM (e.g., Andrews, Wynn, Vallmuur, ter Hofstede, and Bosley (2020)). These case studies typically provide rich insights into the use of one or more PM techniques in specific organizational contexts and draw lessons regarding points of attention for the use of PM in the organization (e.g., Reinkemeyer (2020)).

Although such insights are precious, they are often limited to the boundaries of a single organization. Furthermore, when implementing PM, identifying and selecting valuable business processes and use cases for applying PM is a critical challenge that remains largely unsolved and continues to be a struggle for those responsible (Thiede et al., 2018;

Grisold et al., 2020). From a project selection perspective, Rott and Böhm (2022) provide a first approach to identify suitable processes for PM pilot projects. From a project execution perspective, van Eck et al. (2015) provide a methodology for the end-to-end execution of an individual PM project. However, there is a lack of support beyond piloting for scaling and managing PM project portfolios (Reinkemeyer et al., 2022). Better guidance on managing PM project portfolios, complemented by a holistic understanding of the opportunities and challenges of using PM in organizational settings, would improve state of the art in research and practice. Contributions to this gap would help organizations to tailor their PM project roadmaps while taking advantage of the relevant opportunities and being fully aware of the prevailing challenges. Furthermore, such an understanding would guide research by revealing barriers to PM adoption and highlighting avenues that the PM research community should explore to contribute to successful PM projects.

I.2 Research Objectives

Based on the identified research needs, this dissertation aims to contribute in three areas: First, research on the adoption of PM at the enterprise level (i.e., the management layer in Figure 1) is somewhat fragmented, leading to a call for better guidance on managing PM project portfolios, complemented by a holistic understanding of the opportunities and challenges of using PM in organizational settings. Therefore, this dissertation provides two deliverables to address this research need: On the one hand, a holistic overview of the opportunities and challenges of using PM in organizations is created. On the other hand, a method is developed to manage portfolios of PM projects in a value-oriented manner.

Second, for process data quality management, there is a need for a dedicated environment focused on detecting, measuring, and repairing data quality problems. This dissertation proposes a reference architecture for process data quality management to address this research need. The reference architecture is designed to be comprehensive and flexible enough to incorporate current and future contributions to the multifaceted problem of process data quality. Before applying PM techniques, the high quality of the underlying process data should be validated. However, to my knowledge, research on event log quality assessment of event logs remains scarce. Therefore, this dissertation proposes a second contribution to process data quality management by proposing a user-guided and semi-automated approach for detecting and quantifying process data quality issues in event logs based on multiple data quality dimensions and metrics. These research outputs contribute to the "data processing" activity (Figure 1).

Third, this dissertation aims to contribute to the data-driven process improvement (last activity in Figure 1) enabled by PM. It was found that there is still a lack on balancing tool-based automation and guidance of BPI tasks by incorporating process data and domain expertise. This dissertation addresses this gap by proposing a reference architecture that guides users in improving business processes by leveraging existing process data.

The overall purpose of this dissertation is to advance PM by building on its solid technological core around the numerous PM analysis algorithms by adding the missing pieces in preceding and subsequent steps of end-to-end PM projects. Furthermore, this dissertation also abstracts from a single-project perspective and contributes on the managerial side to the broad applicability of PM in organizations. Applying design science research (DSR) principles, the research objectives of this thesis are addressed through design-oriented research by creating and evaluating multiple artifacts in the form of reference architectures, methods, and instantiations. Ultimately, researchers in the PM field as well as practitioners on the vendor and adopter side should benefit equally from the contributions of this thesis.

I.3 Structure of the Thesis and Embedding of the Research Papers

This dissertation comprises five research papers contributing to the stated research objectives. Table 1 provides an overview of the structure of this thesis and the embedding of the research papers. These papers contribute to current research on PM at the enterprise level, process data quality management, and data-driven BPI.

After motivating the scope of this dissertation and defining the research objectives (Section I), Section II (including Research Papers P1 and P2) presents research that addresses the topic of PM at the enterprise level. Therefore, the thesis provides guidance on managing PM project portfolios combined with a broad view of the opportunities and challenges of PM. Research Paper P1 provides a holistic view of the opportunities and challenges for PM in organizations identified in a Delphi study with 40 international experts from academia and industry. Apart from proposing a set of 30 opportunities and 32 challenges, the paper provides insights into the comparative relevance of individual items and differences in the perceived relevance between academics and practitioners. Research Paper P2 proposes a management method to assist organizations in determining portfolios of PM projects that generate business value by improving business processes. The developed method consists of five activities that outline techniques, roles, and tools: *strategize*, *identify*, *select*, *implement*, and *monitor*. These two research papers enable organizations to tailor their PM

Table 1: Structure of this thesis and embedding of the research papers

I.	**Introduction**
II.	**Process Mining at the Enterprise Level**
P1	**Opportunities and Challenges for Process Mining in Organisations – Results of a Delphi Study** *Martin N, Fischer DA, Kerpedzhiev GD, Goel K, Leemans SJJ, Röglinger M, van der Aalst WMP, Dumas M, La Rosa M, Wynn MT*
P2	**A Portfolio Management Method for Process Mining-enabled Business Process Improvement Projects** *Fischer DA, Marcus L, Röglinger M*
III.	**Process Data Quality Management**
P3	**PraeclarusPDQ: A Framework for Process Data Quality** *Sadeghianasl S, Fischer DA, ter Hofstede AHM, Adams M, Andrews R, Comuzzi M, Ko Y, Koschmider A, Wynn MT, Ziolkowski T*
P4	**Towards Interactive Event Log Forensics: Detecting and Quantifying Timestamp Imperfections in Event Logs** *Fischer DA, Goel K, Andrews R, van Dun CGJ, Wynn MT, Röglinger M*
IV.	**Data-Driven Business Process Improvement**
P5	**An Assisted Approach to Business Process Redesign** *Fehrer T, Fischer DA, Leemans SJJ, Röglinger M, Wynn MT*
V.	**Summary and Future Research**
VI.	**References**

project roadmaps to take advantage of the relevant opportunities and be fully aware of the prevailing challenges.

Section III (including Research Papers P3 and P4) presents two artifacts for process data quality management. Research Paper P3 proposes PraeclarusPDQ, a reference architecture for process data quality management. The reference architecture is designed to be comprehensive and flexible enough to incorporate current and future contributions to the multifaceted problem of process data quality. Research Paper P4 presents a user-guided and semi-automated approach for detecting and quantifying timestamp-related problems in event logs. The approach includes multiple metrics related to timestamp quality dimensions for a systematic and interactive event log quality assessment during the data processing phase of PM projects. The artifacts of both research papers have been instantiated as open-source software to galvanize the PM community.

Section IV (including Research Paper P5) presents research that addresses the topic of data-driven BPI. The thesis provides an approach for assisted and semi-automated BPI using process data, e.g., generated from PM. Research Paper P5 presents a conceptualization of assisted business process redesign (aBPR) to enhance the quality and effectiveness of BPI. The aBPR concept guides users in improving business processes based on redesign patterns. Depending on the available data, the aBPR concept classifies four types of recommendations that differ in their level of automation. Furthermore, this paper proposes a reference architecture that provides operational support for implementing aBPR tools. Finally, the reference architecture has been instantiated as a prototype.

Section V concludes this thesis with a summary of the work, outlining limitations and highlighting avenues for future research. Section VII provides an index of the research papers, my contributions to the papers, and abstracts of the research papers.

II Process Mining at the Enterprise Level

As motivated above, a holistic understanding of the opportunities and challenges of using PM in organizational settings is needed to guide research by revealing barriers to PM adoption and highlighting avenues that the PM research community should explore to contribute to successful PM initiatives. Furthermore, there is a lack of support beyond piloting PM when scaling and dealing with PM project portfolios. Therefore, this thesis provides research that enables organizations to manage their PM project roadmaps (Section II.2; Research Paper P2) to take advantage of the relevant opportunities and be fully aware of the prevailing challenges (Section II.1; Research Paper P1).

II.1 Opportunities and Challenges of Process Mining

Over the past decade, a significant amount of research has been conducted in the area of PM (Thiede et al., 2018). While previous research has primarily focused on the development and improvement of PM algorithms, case studies reporting on the application of PM have also been published (e.g., Andrews, Wynn, Vallmuur, ter Hofstede, and Bosley (2020)). These case studies typically provide rich insights into the use of one or more PM techniques in specific organizational contexts and draw lessons regarding points of attention for the use of PM in the organization (e.g., Reinkemeyer (2020)). While highly valuable, such insights are often limited to the boundaries of a single organization. A more general and holistic understanding of the opportunities and challenges of using PM in organizational settings would complement existing insights based on case studies. Against this background, Research Paper P1 investigates the following research question: *What are the opportunities and challenges of using PM in organizations?*

To approach the research question, Research Paper P1 performs a Delphi study with PM experts from both academia and industry (Paré et al., 2013; Schmidt, 1997). Delphi studies are a well-established method in information systems (IS) and BPM research that strives for consensus on a specific topic with a panel of experts over multiple rounds utilizing questionnaires interspersed with feedback (Kerpedzhiev et al., 2021; Gupta and Clarke, 1996; Skinner et al., 2015). Tables 2 and 3 summarise the main results of the Delphi study: the shortlisted opportunities and challenges, including the rating distributions regarding the comparative relevance of the items in the academic and industry subpanels. The opportunities and challenges are structured along the BPM core elements (i.e., *strategic alignment, governance, methods/information technology (IT), people,* and *culture* (de

Table 2: Shortlisted opportunities for the use of PM in organisations

ID	Strategic alignment	ER		MR		SR		IR	
O.1	**Enabling inter-organisational value creation**** PM enables value creation by fostering inter-organisational interaction and collaboration.	6.25%	56.25%	62.50%	12.50%	31.25%	25.00%	0.00%	6.25%
O.2	**Facilitating strategic decision-making** PM facilitates strategic decision-making by objectively assessing the congruency of operational practices and corporate strategy.	6.25%	12.50%	75.00%	62.50%	18.75%	18.75%	0.00%	6.25%
O.3	**Supporting digital transformation***** PM supports identifying digital transformation initiatives as well as designing digital transformation strategies.	0.00%	56.25%	68.75%	31.25%	31.25%	12.50%	0.00%	0.00%
	Governance								
O.4	**Maintaining an up-to-date business process repository** PM enables maintaining an up-to-date business process repository.	12.50%	25.00%	62.50%	25.00%	25.00%	50.00%	0.00%	0.00%
O.5	**Supporting data management** PM helps identify relevant data and highlights potential data management issues.	31.25%	18.75%	56.25%	43.75%	12.50%	37.50%	0.00%	0.00%
	Methods/IT - Overall								
O.6	**Complementing management approaches and techniques** PM instils process and data awareness into other management approaches and techniques.	6.25%	37.50%	68.75%	31.25%	25.00%	31.25%	0.00%	0.00%
O.7	**Supporting IT management** PM helps derive insights that are useful for the selection, implementation, and improvement of IT systems, tools, and interfaces.	6.25%	12.50%	75.00%	68.75%	18.75%	12.50%	0.00%	6.25%
	Methods/IT - Process discovery								
O.8	**Accelerating as-is business process modelling*** PM accelerates as-is business process modelling and makes it more objective compared to data-agnostic methods.	37.50%	81.25%	62.50%	18.75%	0.00%	0.00%	0.00%	0.00%
O.9	**Enhancing business process transparency** PM increases the transparency of business processes by visualising the actual business process flows based on real-life data.	87.50%	81.25%	12.50%	18.75%	0.00%	0.00%	0.00%	0.00%
	Methods/IT - Process analysis								
O.10	**Analysing business processes from the resource perspective** PM permits the retrieval of actionable insights into the resource involvement and collaboration patterns in a business process.	75.00%	56.25%	18.75%	18.75%	6.25%	25.00%	0.00%	0.00%
O.11	**Analysing business process variants and exceptions*** PM allows the analysis of business process variants and exceptional business process instances, supporting initiatives such as business process standardisation.	68.75%	100.00%	25.00%	0.00%	6.25%	0.00%	0.00%	0.00%
O.12	**Understanding business process compliance** PM allows efficient and comprehensive compliance checking of business process executions as well as understanding the reasons for deviant behaviour and fraud.	93.75%	75.00%	6.25%	18.75%	0.00%	6.25%	0.00%	0.00%
O.13	**Detecting business process drift*** PM enables the detection of business process changes and getting insights into the evolution of business processes over time.	25.00%	50.00%	68.75%	18.75%	6.25%	31.25%	0.00%	0.00%
O.14	**Enabling business process comparison and benchmarking** PM enables comparative analysis and benchmarking of business processes or business process variants.	62.50%	75.00%	25.00%	25.00%	12.50%	0.00%	0.00%	0.00%
O.15	**Enhancing business process risk management** PM enables assessing business process risks and supports the definition of risk mitigation actions.	6.25%	31.25%	62.50%	43.75%	31.25%	25.00%	0.00%	0.00%
O.16	**Identifying business process waste** PM supports the identification of business process waste such as non-value-added tasks or bottlenecks.	75.00%	75.00%	18.75%	18.75%	6.25%	6.25%	0.00%	0.00%
	Methods/IT - Process redesign and implementation								
O.17	**Enabling business process automation*** PM supports the identification of automation potential in business processes.	43.75%	25.00%	37.50%	75.00%	18.75%	0.00%	0.00%	0.00%
O.18	**Enhancing business process improvement and redesign** PM enhances business process improvement and redesign, ranging from the identification of improvement options to the evaluation of its effects.	93.75%	93.75%	6.25%	6.25%	0.00%	0.00%	0.00%	0.00%
O.19	**Improving resource assignment in business processes**** PM allows organisations to improve resource assignments at the levels of tasks and team composition.	62.50%	18.75%	37.50%	43.75%	0.00%	37.50%	0.00%	0.00%
	Methods/IT - Process monitoring and controlling								
O.20	**Enabling decision-making at run-time** PM enables run-time decision-making as well as resource assignment.	75.00%	43.75%	18.75%	18.75%	6.25%	37.50%	0.00%	0.00%
O.21	**Evaluating business process performance** PM supports organisations in assessing and continuously monitoring the performance of business processes.	100.00%	81.25%	0.00%	18.75%	0.00%	0.00%	0.00%	0.00%
O.22	**Predicting outcomes of running cases**** PM supports prediction at run-time regarding expected business process paths and outcomes.	50.00%	56.25%	50.00%	6.25%	0.00%	37.50%	0.00%	0.00%
	People								
O.23	**Enhancing employee training*** PM supports the assessment and improvement of business process training.	6.25%	12.50%	12.50%	50.00%	81.25%	37.50%	0.00%	0.00%
O.24	**Enriching domain knowledge through data** PM encourages domain experts to actively analyse business process data.	50.00%	43.75%	50.00%	37.50%	0.00%	12.50%	0.00%	6.25%
O.25	**Generating intuitive visualisations for business users** PM generates intuitive business process visualisations for business users without technical expertise.	68.75%	81.25%	31.25%	6.25%	0.00%	12.50%	0.00%	0.00%
O.26	**Supporting knowledge management** PM helps make implicit knowledge explicit by unveiling good and bad practices in business processes.	62.50%	31.25%	25.00%	50.00%	12.50%	12.50%	0.00%	6.25%
	Culture								
O.27	**Fostering a business process- and data-centric mindset** PM fosters a cross-functional process- and data-centric mindset by visualising business processes and providing data-backed insights.	37.50%	62.50%	62.50%	37.50%	0.00%	0.00%	0.00%	0.00%
O.28	**Fostering a continuous improvement mindset***** PM stimulates a continuous improvement mindset by encouraging employees to systematically scrutinise business processes.	12.50%	87.50%	62.50%	12.50%	25.00%	0.00%	0.00%	0.00%
O.29	**Nurturing evidence-based communication and decision-making** PM acts as a catalyst for evidence-based communication and decision-making, encouraging objective conversations related to business processes.	56.25%	87.50%	31.25%	12.50%	12.50%	0.00%	0.00%	0.00%
O.30	**Supporting a culture of customer centricity** PM supports a culture of customer centricity when retrieving insights in business processes with an explicit focus on the customer's perspective.	6.25%	18.75%	18.75%	50.00%	68.75%	25.00%	6.25%	6.25%

*A: academics(); P: practitioners(); significance codes: p0.001: ***, p0.01: **, p0.05: *;*

Table 3: Shortlisted challenges for the use of PM in organisations

ID	Strategic alignment	ER		MR		SR		IR	
C.1	**Elusive business value** The business value of PM is difficult to determine with regard to the alignment of strategic and operational goals as well as the quantification of costs and benefits.	56.25%	25.00%	31.25%	62.50%	12.50%	12.50%	0.00%	0.00%
C.2	**Lack of management support** Initiating, funding, and conducting PM initiatives requires a strong management commitment.	62.50%	87.50%	37.50%	6.25%	0.00%	6.25%	0.00%	0.00%
C.3	**Unclear success factors** It is unknown which organisational setups and properties ensure an efficient and effective use of PM.	50.00%	31.25%	37.50%	62.50%	12.50%	6.25%	0.00%	0.00%
	Governance								
C.4	**Constraining data access barriers** Limited data access across departmental and organisational boundaries restricts PM.	50.00%	12.50%	43.75%	68.75%	6.25%	12.50%	0.00%	6.25%
C.5	**Lack of interdisciplinary and cross-functional teams*** PM suffers from a lack of interdisciplinary and cross-functional teams covering sponsors, IT, and data specialists as well as business users and project managers.	25.00%	68.75%	56.25%	12.50%	18.75%	18.75%	0.00%	0.00%
C.6	**Missing implementation guidance** There is a lack of comprehensive guidance on the implementation of PM for different organisations, domains, contexts, and strategic goals.	18.75%	12.50%	68.75%	56.25%	12.50%	31.25%	0.00%	0.00%
C.7	**Poor data quality*** Source or event data are often inaccurate, noisy, and/or incomplete.	93.75%	56.25%	6.25%	43.75%	0.00%	0.00%	0.00%	0.00%
C.8	**Restricting data privacy regulations** Compliance with data privacy and security regulations limits the detail of what can be discovered and analysed through PM.	37.50%	25.00%	56.25%	37.50%	6.25%	31.25%	0.00%	6.25%
C.9	**Unavailability of data***** The availability of event data needed for PM is limited.	93.75%	31.25%	0.00%	50.00%	6.25%	12.50%	0.00%	6.25%
C.10	**Unclear organisational anchoring***** It is unclear how PM expertise should be anchored within the organisation.	12.50%	37.50%	25.00%	62.50%	56.25%	0.00%	6.25%	0.00%
	Methods/IT								
C.11	**Challenging (real-time) system integration**** Insufficient real-time system connectivity or integration into existing IT infrastructure negatively impacts deriving insights through PM.	37.50%	43.75%	62.50%	25.00%	0.00%	31.25%	0.00%	0.00%
C.12	**Complex data preparation** Substantial effort is required for data extraction and pre-processing.	62.50%	50.00%	37.50%	25.00%	0.00%	25.00%	0.00%	0.00%
C.13	**Difficult analysis of process exceptions** PM lacks support for deriving insights from process exceptions.	25.00%	12.50%	50.00%	43.75%	25.00%	31.25%	0.00%	12.50%
C.14	**Difficult handling of unstructured data** PM provides limited support for exploiting unstructured data that is not available in activity-based semantics or event format.	37.50%	43.75%	56.25%	18.75%	6.25%	31.25%	0.00%	6.25%
C.15	**Fragmented solutions**** There is a lack of comprehensive PM solutions supporting a wide range of conceivable use cases.	43.75%	6.25%	50.00%	31.25%	6.25%	56.25%	0.00%	6.25%
C.16	**Incomprehensible outcomes***** Non-standard visualisation techniques used in PM may lead to overcomplicated and hardly understandable business process models.	62.50%	6.25%	37.50%	37.50%	0.00%	50.00%	0.00%	6.25%
C.17	**Insufficient prescriptive capabilities** PM tools are limited regarding their prescriptive capabilities.	25.00%	25.00%	56.25%	31.25%	18.75%	37.50%	0.00%	6.25%
C.18	**Lack of advanced features** PM lacks advanced features such as automation, simulation, and data anonymisation.	18.75%	18.75%	37.50%	37.50%	43.75%	43.75%	0.00%	0.00%
C.19	**Underrepresentation of declarative models** PM relies disproportionately on imperative business process models and largely disregards declarative/hybrid process models.	0.00%	0.00%	56.25%	18.75%	37.50%	75.00%	6.25%	6.25%
	People								
C.20	**Insufficient domain expertise** The lack of comprehensive domain and business expertise inhibits the ability to customise PM as well as to adequately interpret the results.	50.00%	37.50%	37.50%	50.00%	12.50%	12.50%	0.00%	0.00%
C.21	**Insufficient analytical skills** The lack of fundamental analytical skills, including business process modelling and optimisation, impedes deriving value from PM.	75.00%	37.50%	18.75%	31.25%	6.25%	25.00%	0.00%	6.25%
C.22	**Insufficient technical skills**** The lack of sufficient training in technical skills required to implement PM is detrimental to setting up and conducting PM.	81.25%	25.00%	12.50%	50.00%	6.25%	18.75%	0.00%	6.25%
	Culture								
C.23	**Aversion to transparency** PM leads to an undesired level of transparency, revealing unpleasant results and triggering defensive mechanism in employees.	62.50%	37.50%	37.50%	50.00%	0.00%	12.50%	0.00%	0.00%
C.24	**Insufficient data orientation***** A lack of data orientation causes doubts regarding the validity of PM outcomes.	62.50%	6.25%	37.50%	43.75%	0.00%	50.00%	0.00%	0.00%
C.25	**Insufficient process orientation***** A lack of process orientation causes doubts regarding the value of PM.	75.00%	12.50%	25.00%	81.25%	0.00%	6.25%	0.00%	0.00%
C.26	**Invasive work monitoring** PM is perceived as intrusive and raises concerns about privacy and individual performance controlling.	25.00%	6.25%	56.25%	50.00%	18.75%	31.25%	0.00%	12.50%
C.27	**Lack of continuous incorporation** PM is perceived as a one-off initiative, creating a barrier for scaling up and establishing continuous PM.	12.50%	43.75%	75.00%	50.00%	12.50%	6.25%	0.00%	0.00%
C.28	**Lack of trust in insights** PM results and their potential to generate value are discredited since applied techniques are not understood or perceived as a black box.	56.25%	18.75%	25.00%	62.50%	18.75%	18.75%	0.00%	0.00%
C.29	**Misleading overconfidence** Overconfidence in current business process performance downplays the improvement potential through PM.	31.25%	18.75%	50.00%	43.75%	18.75%	31.25%	0.00%	6.25%
C.30	**Resistance to change** Unwillingness to break down long-established routines negatively affects acting on PM insights.	37.50%	75.00%	56.25%	25.00%	6.25%	0.00%	0.00%	0.00%
C.31	**Unsubstantiated expectations** More is projected into PM than can realistically be achieved leading to false expectations and disappointment with the obtained results.	43.75%	31.25%	37.50%	43.75%	18.75%	25.00%	0.00%	0.00%
C.32	**Unwillingness to share domain knowledge*** PM stakeholders are unwilling to share domain knowledge due to the fear of providing too much business information or becoming obsolete.	37.50%	25.00%	50.00%	18.75%	12.50%	56.25%	0.00%	0.00%

*A: academics( ); P: practitioners( ); significance codes: p0.001: ***, p0.01: **, p0.05: *;*

Bruin and Rosemann, 2007)) and, for the opportunities regarding *methods/IT*, along the phases of the BPM lifecycle (i.e., *process discovery, process analysis, process redesign and implementation,* and *process monitoring and controlling* (Dumas et al., 2018)).

The *primary contribution* of the Delphi study is a list of 30 opportunities and 32 challenges that academics and practitioners consider relevant to the use of PM in organizations. The opportunities and challenges are very diverse, covering each of the BPM core elements and addressing technical, managerial, and cultural aspects. For the opportunities, an additional distinction is made between the phases of the BPM lifecycle, and items are present in each phase. Despite the historically strong focus of PM research on the technical side, this study shows that many opportunities and challenges are non-technical. Only 17 out of 30 opportunities (57%) and 9 of 32 challenges (28%) are related to the *methods/IT* core element. The fact that relatively more opportunities than challenges are related to the *methods/IT* core element is consistent with the observation that PM is still primarily perceived as technical practice but also indicates that PM techniques are maturing. The fact that many challenges are located in core elements other than *methods/IT* opens up perspectives for various strands of non-technical PM research.

The *secondary contribution* of the study is its insights into the comparative relevance of opportunities and challenges, deliberately separating the views of academics and practitioners. This perspective has not yet been considered in previous literature. To compare the subpanels used, the median and mode rating values for both subpanels were calculated, as well as the p-value of Fisher's exact test, which indicates the significance of inhomogeneity between the subpanels for the rating of the respective opportunity or challenge. In Tables 2 and 3, asterisks are used to label the denomination of an item in case of significant inhomogeneity of the subpanel rating distributions based on Fisher's exact test. When discussing similarities and differences in the subpanel rating distributions, the median serves as the primary criterion, as it is more robust to outliers than the mode (von der Gracht, 2012). In the case of different medians, the mode and Fisher's exact test statistics are used to check for significant differences.

Regarding strong agreement between the two subpanels, academics and practitioners rated ten key opportunities as extremely relevant according to the median and mode (Table 4). These tend to be in the area of *methods/IT*, suggesting that there appears to be an agreement between the two subpanels on the potential of PM from a technical perspective. In contrast, there are significant differences between the subpanels regarding

Table 4: Opportunities and challenges rated as extremely relevant by academics and practitioners

Opportunities	• Enhancing business process transparency (O.9) • Analysing business processes from the resource perspective (O.10) • Analysing business process variants and exceptions (O.11) • Understanding business process compliance (O.12) • Enabling business process comparison and benchmarking (O.14) • Identifying business process waste (O.16) • Enhancing business process improvement and redesign (O.18) • Evaluating business process performance (O.21) • Generating intuitive visualisations for business users (O.25) • Nurturing evidence-based communication and decision-making (O.29)
Challenges	• Lack of management support (C.2) • Poor data quality (C.7) • Complex data preparation (C.12)

challenges, as reflected in their rating distributions. Academics and practitioners agree on the high relevance of data quality (C.7) and preparation (C.12) as well as management support (C.2). These are challenges that this thesis actively addresses (Research Papers 2 - 4; Sections II.2, III.1, and III.2). The remaining challenges show varying degrees of difference in relevance between the two subpanels. This assertion is supported by the observation that only 9 out of 32 challenges have the same median and mode, and the p-value does not indicate a significant difference. This means that academics and practitioners differ in their assessment of the relevance of many challenges. Only three key challenges are rated as extremely relevant by both subpanels (Table 4).

The observed differences outlined above indicate a perceived disconnect between academics and practitioners. To push the boundaries of what PM can do in organizational contexts, a strong partnership between academia and industry is needed, and a disconnect should be avoided. Such a disconnect would prevent organizations from reaping the full benefits of academic work because of prevailing challenges that academics are unaware of or consider less relevant given the predominantly technical research interests of many PM researchers. Consequently, engaging in bidirectional knowledge transfer through roundtables or communities of practice could ensure that academia solves scientifically challenging problems with real-world relevance, allowing the industry to benefit from the latest scientific knowledge.

A holistic understanding of the opportunities and challenges of using PM in organizations has been missing. Research Paper P1 provides a structured overview of such opportunities and challenges based on a Delphi study with a panel of academic and industry experts. Finally, it also reveals insights into the comparative relevance of opportunities and challenges, distinguishing between the views of academics and practitioners. Therefore, this thesis provides foundational work that enables organizations to take advantage of the relevant opportunities and be fully aware of the prevailing challenges of PM.

II.2 Management of Process Mining Project Portfolios

The Delphi study presented (Research Paper P1) identified key challenges to using PM in organizations. This study highlights the lack of management support and the elusive business value of PM as key challenges (Martin, D. A. Fischer, et al., 2021). When implementing PM, identifying and selecting valuable business processes and use cases for applying PM is a key challenge that remains largely unresolved and continues to plague responsible managers (Thiede et al., 2018; Grisold et al., 2020). When establishing PM, it is often difficult to know which use cases to start with (Rozinat, 2021). From a project selection perspective, Rott and Böhm (2022) provide a first approach to identifying suitable processes for PM pilot projects. From a project execution perspective, van Eck et al. (2015) provide a methodology for the end-to-end execution of an individual PM project. However, there is a lack of support beyond piloting when scaling and dealing with PM project portfolios (Reinkemeyer et al., 2022). Therefore, Research Paper P2 addresses this challenge with the following research question: *How can organizations manage process mining project portfolios?*

To approach the research question, Research Paper P2 proposes a method for managing portfolios of so-called PM value cases, which are defined as PM-enabled BPI projects. The method is developed based on DSR combined with situational method engineering (SME) (Hevner et al., 2004; Peffers et al., 2007; Ralyté et al., 2003). In line with DSR and SME principles, the method is based on justificatory knowledge about project portfolio management (PPM) and the business value of PM. It is built around method requirements and three design objectives (DOs) (*structured guidance*, *consideration of process and context factors*, *comparability of PM value cases*). The overall goal of the method is to help organizations manage PM project portfolios to generate value for the organization through data-driven BPI. A prototypical instantiation complements the method.

Figure 2 shows the artifact resulting from the DSR project, including adjustments made

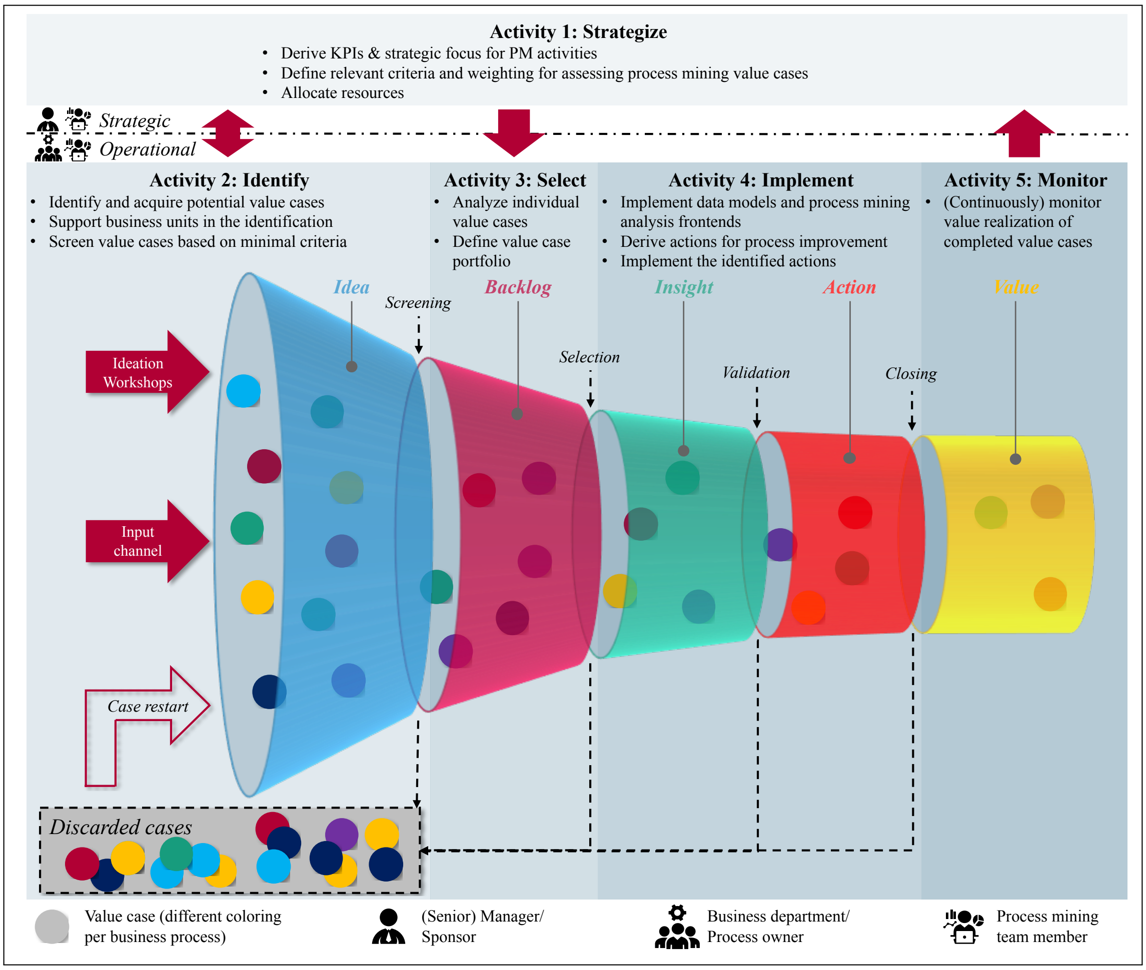

Figure 2: Overview of the method for managing PM project portfolios

following extensive evaluation. The method for managing PM project portfolios consists of five activities (*strategize*, *identify*, *select*, *implement*, *monitor*) that guide potential users through the required actions by suggesting techniques, roles, and tools for each activity. These activities are derived from the literature (see Archer and Ghasemzadeh (1999), Stettina and Hörz (2015), Dumas et al. (2018), and van Eck et al. (2015)) and supported by a panel of twelve experts from research and practice involved throughout the design and evaluation of the method. Activity 1 (*"Strategize"*) is preparatory, with the aim of gathering and processing input for subsequent activities. At the strategic level, relevant strategic goals and criteria are derived for each decision point during the value case journey. IT and human resources are also allocated. Activities 2 to 5 take a value case-centric perspective and define the PM value case journey: Activity 2 (*"Identify"*) involves identifying potential value cases and screening for minimum criteria that must be met for a value case to be considered. Activity 3 ("*Select*") takes place before value case implementation and concerns evaluating candidate value cases and defining a project portfolio for the subsequent activity. Activity 4 ("*Implement*") concerns the implementation of individual value cases and is divided into two phases: the insight phase, which covers the implementation of the analysis tools and the derivation of actions for BPI, and the action phase, which covers the implementation of the derived BPI actions and thus marks the beginning of value realization. Activity 5 ("*Monitor*") takes place after the value case implementation and concerns monitoring the performance improvement and value realization of implemented value cases. Overall, the method is structured in a decision-support and user-guiding manner that requires humans in the loop. For each activity, the method guides the user through the required actions and relevant decision points, thereby providing a framework for making informed decisions about whether proceed with or discard individual value cases. However, the calculated results are intended only as decision support for the user and not as a definite result.

The method is designed to meet the attributes of *goal orientation*, *systematic approach*, *principles orientation*, and *repeatability* (Denner et al., 2018). It assembles five activities that support decision-making for a valuable portfolio of PM projects to reflect goal orientation. The method's structured approach and activities guide execution to account for the systematic approach. The detailed specification of each activity, including techniques, tools, roles, and outputs, ensures the method's repeatability in different contexts. Furthermore, the application of the method in a real-world case study shows that the method is repeatable in a real-world context in addition to the artificial context used during the development. Regarding principles orientation, the method considers three DOs derived from the literature and evaluated with the expert panel.

Overall, DSR projects should aim to design useful artifacts (Gregor and Hevner, 2013). To achieve this goal, the evaluation of the method was designed based on three phases according to Venable et al. (2016) and Sonnenberg and vom Brocke (2012). The ex-ante evaluation concerned the justification of the research problem and the DOs through a survey within the expert panel to underpin the importance and novelty of the research project. The intermediate evaluation was structured around an interview-based validation of the method's design specification with the expert panel to underpin the applicability and real-world fidelity of the method. Finally, a case study as an ex-post evaluation completed the overall evaluation. For this purpose, the method was applied at one of the world's leading semiconductor manufacturers. This case study helped to assess the usefulness of an artifact instantiation of the method in a real-world environment.

The increasing adoption of PM in practice requires prescriptive knowledge on managing PM project portfolios. Research Paper P2 contributes to this need by proposing a method for managing PM project portfolios. A key distinction of our method from existing PPM knowledge is the consideration of the data-driven nature of the projects. By iteratively collecting data along different value case states, the method supports the successful confirmation of the value of each project. This results in a funnel-like structure accompanied by a gradual sorting process. Existing concepts that have been evaluated do not cover the full scope of the presented method, nor do they exploit new potentials from the data. In particular, existing work completely neglects data-driven monitoring, which the expert panel repeatedly emphasized as a critical step in assessing impact and gaining management support for further projects. Finally, it became clear throughout the process that an agile approach to value case management provides the most accurate picture of current approaches in practice.

In summary, this section has presented two significant contributions on the organizational side of PM. First, a structured overview of opportunities and challenges based on a Delphi study with a panel of academic and industry experts was presented. Second, a method for managing PM project portfolios was developed. Together, these two contributions provide solid guidance for managing PM project portfolios, complemented by a holistic understanding of the opportunities and challenges of using PM in organizational settings.

III Process Data Quality Management

Zooming from a portfolio perspective to the PM project level, Section I showed that process data quality management is critical to PM. As the saying goes - *garbage in, garbage out* - reliable PM results depend on high-quality process data (van der Aalst, 2016; Andrews, C. G. J. van Dun, et al., 2020; van der Aalst, Adriansyah, et al., 2012). However, process logs are often far from the desired quality, resulting in data scientists spending up to eighty percent of their work on identifying, assessing, and remediating data quality issues (Wynn and Sadiq, 2019; Bose et al., 2013; Suriadi et al., 2017). Unfortunately, research on process data quality has somewhat lagged behind the development of algorithms. Therefore, this thesis contributes to this gap and provides a reference architecture for process data quality management (Section III.1; Research Paper P3). In addition, it provides a framework for process data quality measurement, a component that has received remarkably little attention to date (Section III.2; Research Paper P4).

III.1 A Framework for Process Data Quality Management

The problem of data quality is multi-faceted. Research has emphasized quality problems as they manifest in event logs (e.g., Bose et al. (2013) and Suriadi et al. (2017)). However, managing process data has its own requirements, some different from other types of data, meaning that process data governance needs to take the nature of PM into consideration (Goel et al., 2021). The complex and rich nature of process data quality management means that technological solutions should be able to evolve as our understanding of the field increases. New prevention and detection strategies should not require a complete overhaul of the architecture of such solutions. While existing PM environments, e.g., ProM and Disco, provide support for managing certain data quality problems typically encountered in event logs, e.g., through filtering and abstraction, a dedicated environment focusing on the detection, measurement, and repair of data quality problems is a sine qua non for the next generation of process data quality management. Therefore, Research Paper P3 aims for the following research question: *How to design a reference architecture for process data quality management?*

This thesis introduces the *PraeclarusPDQ* reference architecture to address the stated research question. Based on DSR and guidelines for empirically based reference architectures, the artifact is designed to satisfy seven design principles that capture desired capabilities and have been defined ex-ante (Peffers et al., 2007; Gregor and Hevner, 2013;

Galster and Avgeriou, 2011). Based on the design principles, a conceptual reference architecture was formulated that is shown in Figure 3. The architecture is designed to accommodate new developments in data quality management strategies, as well as future developments in mitigation and prevention. It is built to anticipate new methods and strategies beyond current algorithms and techniques. Therefore, the key underlying principle is abstraction and extensibility, not only for data quality improvement but also for visualization techniques that allow analysts to interact with data in new ways.

The reference architecture provides two interfaces. The first is a plugin interface, an abstraction that (currently) allows for four types of plugins: readers, writers, patterns, and actions. The second interface is a public-facing interface that facilitates user interaction. It also provides an API for external applications and services. This API is used by the authentication component, which authorizes permitted activities and manages active sessions. The core of the reference architecture consists of three components: (1) **Workflow Management:** This component is responsible for the creation, management, and execution of workflows. A workflow is a 'chain' of one or more plugin objects, where the output of one becomes the input of the next (except for a reader object, which provides output only). (2) **Data Management:** This component is responsible for the maintenance and internal

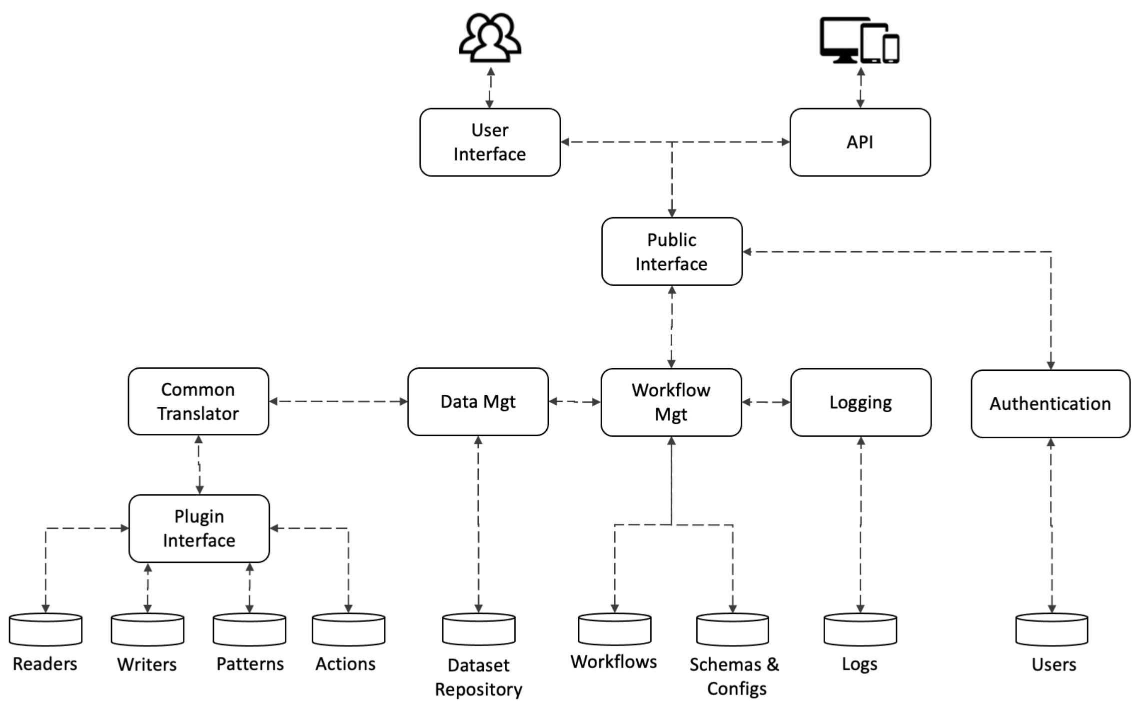

Figure 3: A high-level reference architecture of the PraeclarusPDQ framework

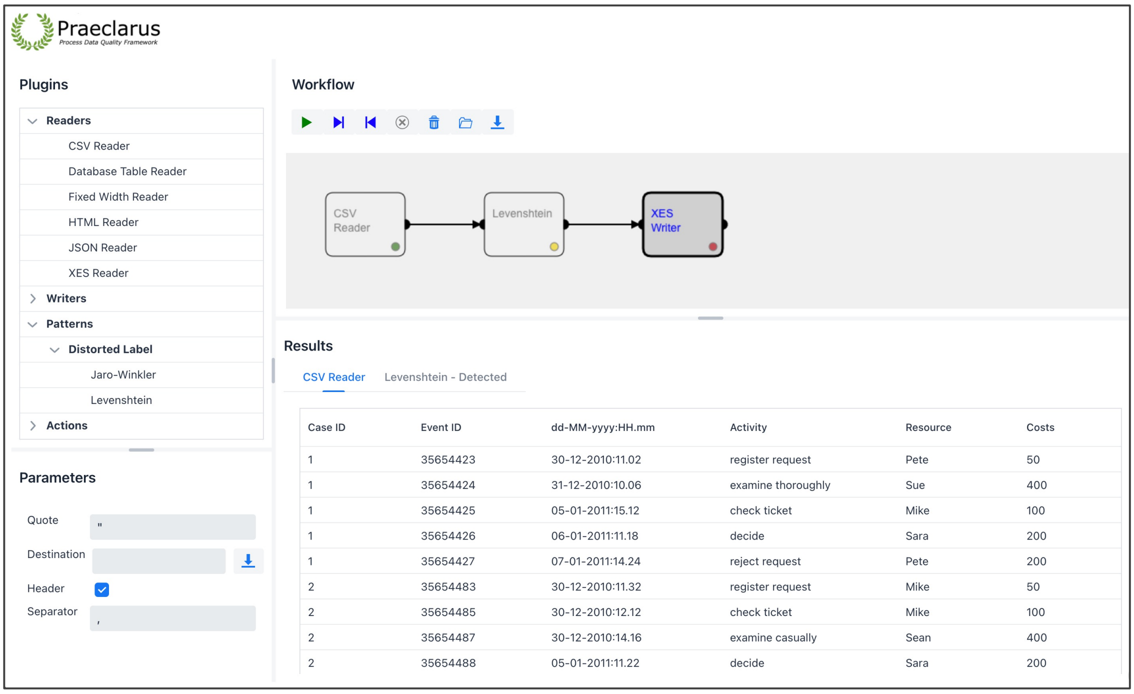

Figure 4: Web UI of the PraeclarusPDQ prototype

transport of all process log records. This includes versioning and storing all records to meet data provenance requirements. (3) **Logging:** In conjunction with the other main components, the logging module ensures that all actions are recorded, including user actions that can be used for future analysis tasks.

The presented reference architecture is the basis for the framework, which has been implemented as a web service written primarily in Java. All working code is available on GitHub[1]. The repository also references a web interface that reflects the latest release and an introductory video that complements the implementation. Figure 4 shows the web-based user interface for the implementation. The top-left panel lists the currently loaded plugins, grouped by type: *Readers, Writers, Patterns*, and *Actions*. Pattern plugins are further subcategorized by their imperfection pattern type. Eventually, the framework will support the input and output of data from any potential data source. New plugins can be added at any time and are automatically listed without requiring any configuration changes. Plugins can be dragged onto the workflow canvas to create new workflow nodes.

In line with DSR and reference architecture development principles, multiple evaluation phases were integrated into the design and development process (Peffers et al., 2007; Galster and Avgeriou, 2011). Therefore, the framework for evaluation in design science

[1] `https://github.com/praeclaruspdq/PraeclarusPDQ/`

(FEDS) by Venable et al. (2016) was adopted and extended with selected components from the DSR evaluation framework by Sonnenberg and vom Brocke (2012). Complementing the overarching goal of developing useful artifacts (Gregor and Hevner, 2013), the evaluation strategy was divided into three phases: For the *ex-ante evaluation*, a literature scan was conducted to provide ex-ante justification of the research problem, the research gap, and the derivation of design objectives. For the *intermediate evaluation*, the applicability of the PraeclarusPDQ framework was tested by implementing two process data quality management plugins from Sadeghianasl et al. (2019) and Ko and Comuzzi (2021) from their previous research dealing with data quality management activities. For the *ex-post evaluation*, the PraeclarusPDQ was presented to a group of researchers actively engaged in process data quality management to evaluate the artifact in terms of perceived "usefulness" and "ease of use" (Davis, 1989).

Research Paper P3 presented the PraeclarusPDQ reference architecture and an associated open-source software environment designed to incorporate future solutions for various aspects of process data quality management. The reference architecture and associated software framework have been tested and evaluated for their applicability, usefulness, and ease of use. The PraeclarusPDQ framework is publicly available, which means that it can serve as a rallying point for process data quality researchers. The framework will help them develop their research contributions in the area of process data quality management and make them available to the public for feedback. Based on the open-source software approach, it is the ambition that the PraeclarusPDQ framework will evolve into an artifact that facilitates research collaboration in the area of process data quality management and will become *the* hub for software contributions in this area.

III.2 Measurement of Process Data Quality

Although adequate data quality is often assumed, in practice, process logs are often far away from the desired quality (Wynn and Sadiq, 2019; Bose et al., 2013; Suriadi et al., 2017). Therefore, event logs should not be used naively for PM without ensuring adequate event data quality (van der Aalst, 2016). Hence, it is essential to have the means to detect and quantify the quality of event logs (Wynn and Sadiq, 2019). However, research dealing with (semi-automated) quality assessment of event logs remains scarce (Andrews, Wynn, Vallmuur, ter Hofstede, Bosley, et al., 2019). This work aims to fill this research gap specifically for timestamp-related data quality issues, as timestamps are at the core of many PM use cases (van der Aalst, Adriansyah, et al., 2012; Dixit et al., 2018; Gschwandtner et al., 2012). Accurate timestamps are essential for reproducing the correct sequence of activities, obtaining accurate process models (*discovery*), measuring the alignment between the process model and the actual process flow (*conformance*), and determining the effectiveness and efficiency in the execution of activities (*performance*) (Dixit et al., 2018; Gschwandtner et al., 2012). In contrast, inaccurate and coarse timestamps often result in convoluted process models that can lead to incorrect analyses (Dixit et al., 2018). Research Paper P4, therefore, focuses on the following research question: *How can timestamp-related data quality issues in event logs be detected and quantified?*

Following the DSR paradigm (Peffers et al., 2007), the paper builds on four essential DOs for a timestamp-related quality quantification approach from mature knowledge about data and event log quality. Table 5 shows the final framework for detecting and quantifying timestamp imperfections in an event log, resulting from the design and development phase and the subsequent evaluation. To address the stated DOs, the framework consists of two axes: On the one hand, standard abstraction levels of event logs (event, activity, trace, log) (van der Aalst, 2016) are used to evaluate the timestamp quality. On the other hand, four data quality dimensions are used: *accuracy*, *completeness*, *consistency*, and *uniqueness*. The derived framework forms the basis for positioning metrics for timestamp quality assessment. However, since multiple units are required to evaluate equality, measuring consistency or uniqueness for a single event at the event level is

Table 5: Timestamp quality assessment framework

	TIMESTAMP QUALITY			
	QD_1: Accuracy	QD_2: Completeness	QD_3: Consistency	QD_4: Uniqueness
Log Level		M_5: Missing Trace[c]	M_9: Mixed Granularity of the Log[c] M_{10}: Format[b]	M_{13}: Duplicates within Log[c]
Trace Level	M_1: Infrequent Activity Ordering[a] M_2: Overlapping Activities per Resource[a]	M_6: Missing Activity[b]	M_{11}: Mixed Granularity of Traces[a]	M_{14}: Duplicates within Trace[b]
Activity Level		M_7: Missing Event[c]	M_{12}: Mixed Granularity of Activities[c]	M_{15}: Duplicates within Activity[c]
Event Level	M_3: Future Entry[c] M_4: Granularity[c]	M_8: Missing Timestamp[c]		

□: metric can be allocated; ■: **no** metric can be allocated

a: pre-existing detection approach used; b: modification of pre-existing detection approach; c: new development impossible. Therefore, the background of the corresponding cells is gray.

Within the designed framework, 15 quality metrics are positioned. These metrics are either derived from existing detection approaches (marked with [a]), modifications of existing detection approaches ([b]), or detection approaches that have been designed by the authors based on insights from the literature ([c]). However, no evidence of timestamp quality issues affecting accuracy at the log or activity level was found. Nonetheless, the framework should be seen as an extensible basis for quantifying event log quality, and thus additional metrics or dimensions can be integrated. Four components are used to describe each metric: (1) A *description* of the metric and the quality issue under investigation is presented. Additionally, possible reasons for the presence of such quality issues are mentioned. (2) An approach for the *detection* of the investigated quality issue is presented. (3) An approach is presented to aggregate the detected quality issues per metric into *patterns*. This helps the user to keep the overview and to systematically clean the detected quality issues (Suriadi et al., 2017). (4) Finally, an approach for the *quantification* of a score between 0 and 1 per metric is shown to support the user in decision-making.

Another goal of the approach is to incorporate domain and use case knowledge from domain experts. The approach for detecting and quantifying timestamp imperfections has been implemented as open-source software[2]. The prototype displays the metrics and calculated scores, computed fully automatically after importing an event log in XES format. It also visualizes aggregated scores for each quality dimension and event log level. In addition, several additional components are added to allow for the integration of domain and user input. The components allow for domain and use case independence of the solution and a human in the loop to optimize the detection of timestamp imperfections and to customize the timestamp quality quantification for specific use cases and domains: (1) *Quantification configuration.* The approach provides a configuration option to suppress unimportant metrics or dimensions and adjust the metrics' weight. (2) *Errorlists.* To increase transparency and traceability, the user can investigate the detected quality issues for each metric. (3) *Whitelisting.* The presented approach for detecting and quantifying timestamp imperfections is deliberately designed to minimize the risk that existing issues remain undetected (false positives). To reduce the number of such false positives, the user can "allowlist" incorrectly detected patterns and quality issues. (4) *Quality information.* In

[2]available in the ProM nightly build which can be downloaded here: `http://bit.ly/38KVKvJ` (Verbeek et al., 2011). The source code is available in the ProM package "LogQualityQuantification" (`http://bit.ly/39OAgj0`) or on Github (`https://bit.ly/3iiEnub`)

the approach, the user can add the current quality information to the metadata of the event log under consideration.

The approach has been evaluated according to the framework of Sonnenberg and vom Brocke (2012), which consists of four evaluation activities (EVAL1 to EVAL4) structured along two dimensions, i.e., ex-ante/ex-post and artificial/naturalistic evaluation (Pries-Heje et al., 2008; Venable et al., 2012). For EVAL1, the research gap and the derived DOs were justified based on a literature scan. For EVAL2, the design specification of the approach was compared with competing artifacts to support the significant added value of the approach to the existing literature. For EVAL3, the approach was implemented as a software prototype. In addition, experiments were conducted with experts from research and practice using real-world event logs to refine the metrics and demonstrate the real-world fidelity and consistency of the approach. As for EVAL4, a survey study with PM experts from academia and industry validated the perceived ease-of-use and usefulness of the approach based on the technology acceptance model (TAM) (Davis, 1989).

Following DSR principles, this thesis presents an approach for detecting and quantifying timestamp imperfections in event logs based on 15 novel data quality metrics structured along four data quality dimensions and log levels. The framework focuses on timestamp quality issues and provides a first step toward quantifying event log quality. The approach is domain-agnostic and assists process stakeholders in determining the suitability of an event log for PM analysis. In summary, the approach and its implementation adequately support users in detecting and quantifying timestamp quality issues in event logs.

This section has presented two major contributions to process data quality management. First, a reference architecture and software environment for process data quality management. Second, a framework for process data quality assessment. Therefore, this thesis, including these two contributions, provides a solid base to pave the way for future research on process data quality management and assessment.

IV Data-Driven Business Process Improvement

Transforming business processes at an accelerated pace is essential for organizations to meet increasing competition and customer needs (Beverungen et al., 2021; Huang et al., 2015). In BPM, business process redesign (BPR) (or BPI) is concerned with improving business processes to address previously identified process-related issues, for example, through PM (Dumas et al., 2018). BPR consumes significant resources but also yields promising returns (Huang et al., 2015). Therefore, BPR is generally considered to be the most value-adding phase in the BPM lifecycle (Dumas et al., 2018; Gross et al., 2021; Reijers and Limam Mansar, 2005; Zellner, 2011). As a subsequent step to the PM analysis, it was found that there is still a lack of research on how to use the process data generated by PM to improve business processes. Therefore, this thesis aims to contribute to this gap to enable the end-to-end execution of PM projects. This section presents an approach to guide BPI tasks by incorporating process data and domain expertise.

Organizations often conduct workshops with consultants and various process stakeholders to analyze challenges and opportunities and manually generate BPR options (Zellner, 2011). Therefore, the quality and effectiveness of BPR depend on the creativity and expertise of the project team to find valuable solutions (Essam and Limam Mansar, 2012). Tools are an essential means to manage the complexity of business processes and assist in their improvement and implementation (Zuhaira and Ahmad, 2021). While most of the literature presenting BPR methods does not include tool support, some approaches build on redesign patterns to generate tool-based suggestions for their application on business processes (Fellmann et al., 2019; Netjes et al., 2010; Zuhaira and Ahmad, 2021). However, they have limitations, such as (1) they rely on data that is difficult to obtain, (2) they are inflexible due to hard-coded assumptions, and (3) only a few approaches provide the ability to incorporate a variety of redesign patterns (Essam and Limam Mansar, 2012). Therefore, it is questionable to what extent such tool-based approaches can handle the complexity and richness of business process information and provide actionable suggestions for BPR (Essam and Limam Mansar, 2012). While this research gap has been recognized in the literature, no interactive and assistive approach combines both worlds in a guided process (Röglinger, C. van Dun, et al., 2021; Essam and Limam Mansar, 2012): tool-based automation and guidance of BPR tasks on the one hand and the incorporation of domain expertise on the other hand. Thus, Research Paper P5 addresses the following research question: *How can assistive tools improve BPR?*

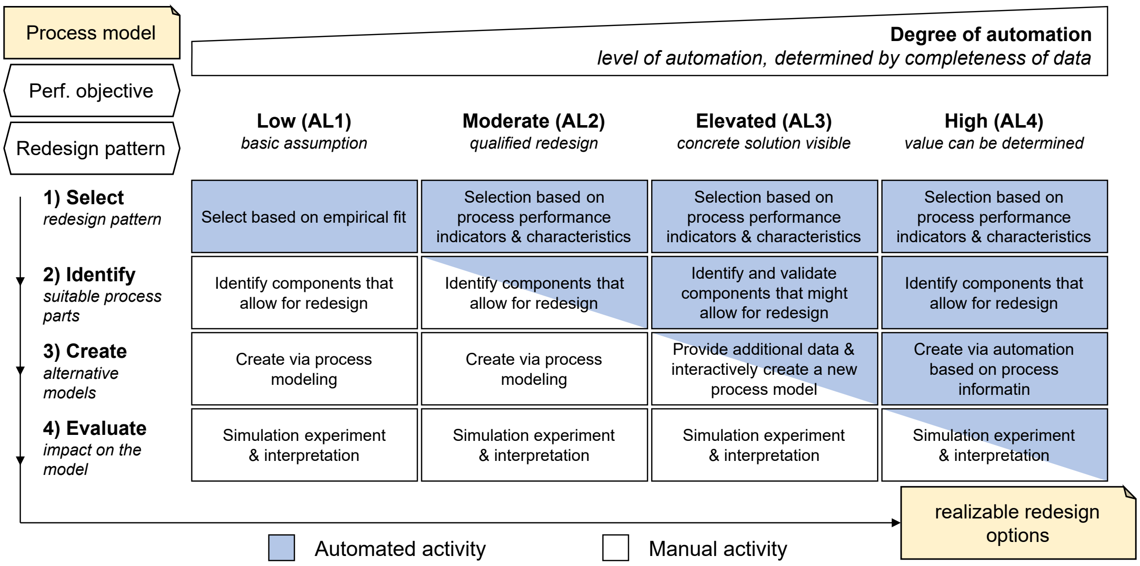

Figure 5: Conceptualization of aBPR

Adopting the DSR paradigm (Hevner et al., 2004; Peffers et al., 2007) in conjunction with reference architecture (RA) development (Galster and Avgeriou, 2011), Research Paper P5 presents the aBPR concept, an RA design specification, and a prototypical instantiation. Four activities derived from related work guide the development of redesign options using patterns in a step-by-step manner. as shown in Figure 5: Step 1) *select* suitable redesign patterns, Step 2) *identify* suitable process parts, Step 3) *create* alternative models, and Step 4) *evaluate* the performance of these alternative models. The execution of these four steps results in redesign options that can improve the process under study, depending on the evaluation outcome. aBPR tools, as envisioned, deeply integrate these steps and guide users through their structured application. Using automation potential, tools that implement the aBPR concept perform these steps in the background and present their results as redesign recommendations. Users complete the remaining steps manually using their expertise to transform the recommendations into redesign options. Combinations of (semi-)automated and manual steps lead to different recommendations that automate more and more individual steps. In Figure 5, four types of recommendation are defined in increasing automation level (AL) (Parasuraman et al., 2000).

In addition, the approach presents a diversified and ranked selection of top recommendations to the user while initially retaining less valuable or too similar recommendations. To evaluate their potential and similarity, a scoring function implemented in aBPR tools estimates the impact of each recommendation according to the selected performance objec-

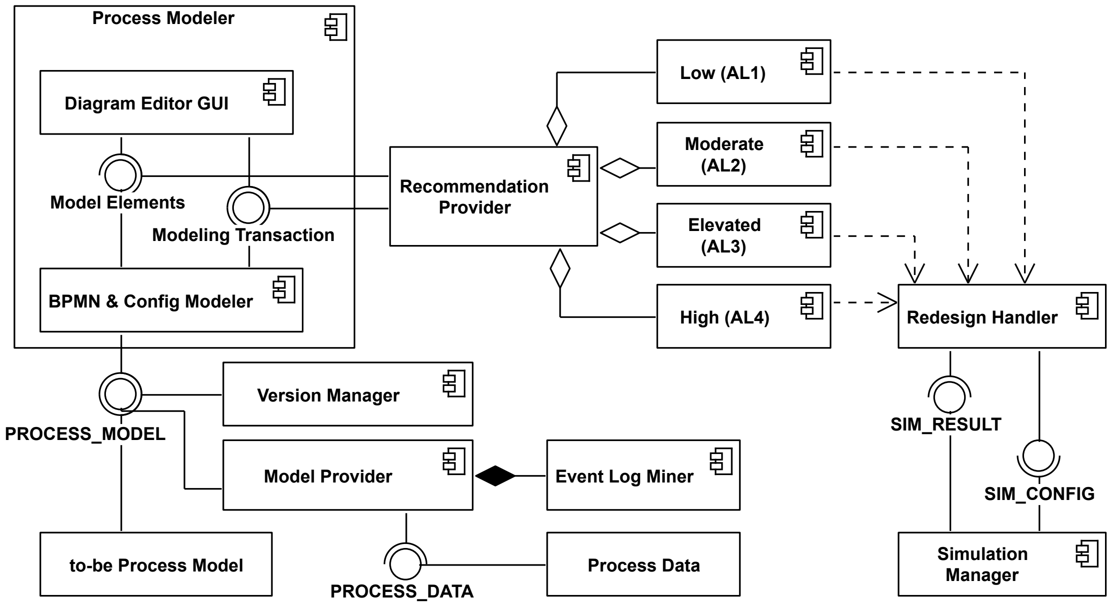

Figure 6: aBPR reference architecture

tive. The impact of recommendations is not directly comparable across recommendation types because recommendations at higher ALs are more specific than recommendations at lower ALs. The scoring function uses empirical information to estimate the potential impact when no specific impact is measurable. The similarity is calculated as a measure integrating information about the redesign pattern (e.g., process aspect, pattern identifier) and the specific recommendation (e.g., the overlap of affected elements).

To implement aBPR tools, Figure 6 shows the aBPR RA as a component diagram. The RA consists of several components: (1) The **model provider** component serves as an external interface for data. (2) The **process modeler** component provides user interaction and modeling capabilities. (3) The **simulation manager** component provides an interface for process models and executes simulation experiments according to the simulation configuration. (4) The **redesign handler** component ensures that the four steps shown in Figure 5 are followed in sequence for each redesign option. (5) Triggered by changes in a process model, the **recommendation provider** repeatedly checks the potential of redesign handlers and diversifies them to create a list that encourages user creativity. (6) The **version manager** tracks the evolution of all redesign options.

aBPR has also been implemented as a prototypical instantiation of the RA (Galster and Avgeriou, 2011). Except for the simulation manager, which is outsourced as a cloud service for load balancing, and the event log miner, which is not part of this implementation, all

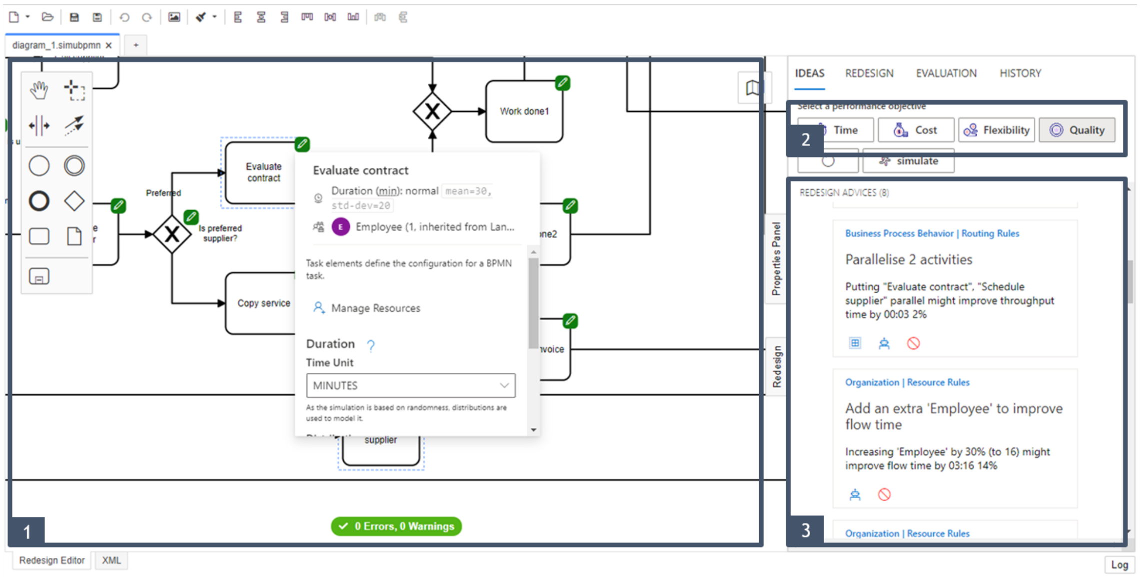

Figure 7: Software prototype - general overview with graphical user interface (GUI) elements (1) diagram editor, (2) performance objective selection, and (3) list of recommendations

components are implemented as modules of a desktop application, as shown in Figure 7. The source code and technical documentation are available online[3]. The prototype supports all patterns from Reijers and Limam Mansar (2005). The application starts with an empty canvas or an existing business process model and notation (BPMN) diagram. It allows the user to edit the process model and provides recommendations for its redesign after selecting a unique performance objective, such as time, cost, flexibility, or quality. The top recommendations are displayed in a list. Each recommendation details the process aspect, the heuristic category, its name, a description, and optionally the expected impact and affected process elements. The user can accept or reject recommendations and evaluate their impact through simulation experiments and expert judgment. This process is repeated until satisfaction with the process is achieved, and the improved process model is exported.

In line with the design principles of both DSR and RA development, several evaluation activities were integrated into the design process (Peffers et al., 2007; Galster and Avgeriou, 2011). Again, the DSR evaluation framework (EVAL1 to EVAL4) by Sonnenberg and vom Brocke (2012), was applied. For EVAL1, a literature scan was performed to justify the research problem, the research gap, and the derivation of DOs. For EVAL2 and EVAL3, the design specification of the aBPR concept was validated through expert interviews. Furthermore, the prototypical instantiation of the aBPR RA was provided, and BPM experts from academia and industry were involved in evaluating the applicability of the

[3]https://github.com/dtdi/assisted-bpr-modeler

artifact. Finally, for EVAL4, the usefulness of the artifact in naturalistic settings was underpinned via a case study at KUKA, a global automation company.

BPR is a key to long-term business success for many organizations. Therefore, Research Paper P5 addresses how assistive tools can improve BPR and proposes a conceptualization for aBPR. The approach takes process data as input and interactively assists users in iteratively improving the business process to achieve a specified performance objective. This research adds to the prescriptive knowledge on BPR by building on and extending existing approaches (Gregor and Hevner, 2013). aBPR provides a novel approach for applying redesign recommendations with varying levels of automation and interactivity. aBPR provides a way to categorize and embed existing pattern application approaches into a structured process and provides a framework to guide their implementation. In summary, the contribution of this thesis to BPI marks a cornerstone for actively using process data as it can come from PM analysis to improve business processes.

V Conclusion

V.1 Summary

The evolution of PM over the past decades can be considered a true BPM success story. Over the past years, the massive investment in developing and improving ever-new PM algorithms for different use cases has created a strong technology core for process mining and analysis techniques. Although this technological strength provides a solid foundation for industrial applications, organizations still face several hurdles due to the limited support for the preceding and subsequent steps surrounding PM analysis techniques. In addition, the management issues that arise when scaling PM have received little research attention. This dissertation and the embedded research articles contribute to the identified gaps and aim to enable organizations to realize the full potential of PM. First, this thesis provides a foundation that enables organizations to tailor their PM project roadmaps to take advantage of the relevant opportunities and be fully aware of the prevailing challenges. Second, this thesis provides two artifacts for process data quality management to support the preceding data processing step before applying PM. Third, this thesis presents an approach for assisted business process redesign (aBPR) to support the subsequent process improvement step that can leverage insights from applying PM.

The PM research field primarily focuses on technical topics such as algorithm development. However, to support the adoption of these algorithms, it is also essential to gain insight into the use of PM in organizations and to guide its application at scale. Section II presents two research papers that address this need. Research Paper P1 provides a structured overview of opportunities and challenges based on a Delphi study with a panel of academic and industry experts. It also provides insights regarding the comparative relevance of opportunities and challenges, distinguishing between the views of academics and practitioners. Research Paper P2 investigated how to design a PM PPM method. In line with DSR and SME as the primary research methods, the artifact design builds upon justificatory knowledge on PPM and the value of PM. The method assists users in evaluating, comparing, and selecting appropriate PM project portfolios. It also guides managing PM value cases throughout their lifecycle, i.e., from initiation to completion and monitoring, based on five activities: strategize, identify, select, implement, and monitor. The method was developed and evaluated with an expert panel of academics and practitioners. Based on a conclusive evaluation in a real-world application, the applicability and usefulness of the artifact in naturalistic settings are underpinned. Together, these two contributions provide solid guidance for managing PM project portfolios, complemented by a holistic understanding of the opportunities and challenges of using PM in organizational settings.

At the project level, there are gaps in process data quality management, even though it is well known that the reliability of PM analyses is highly dependent on the quality of the imported data. Thus, process data quality can be a key obstacle for PM to gain further traction in practice. Section III presents two DSR-based artifacts that contribute to this research gap. Research Paper P3 presents the PraeclarusPDQ reference architecture and an associated open-source software environment to incorporate future solutions for various process data quality management aspects. The reference architecture and its associated software framework have been tested and evaluated for applicability, usefulness, and ease of use. Based on the open-source software approach, the PraeclarusPDQ framework is intended to evolve into an artifact that facilitates research collaboration in process data quality management and becomes *the* hub for software contributions in this area. Before applying PM techniques, the high quality of the underlying process data should be validated. Therefore, Research Paper P4 presents an approach to detect and quantify timestamp imperfections in event logs based on 15 novel data quality metrics structured along four data quality dimensions and log levels. The framework focuses on timestamp quality issues and provides a first step toward quantifying event log quality. The approach can identify common timestamp-related problems and measure the quality of timestamp information in event logs. In addition, the approach is domain-agnostic (e.g., by suppressing irrelevant metrics or adjusting the weight of metrics). As a result, it helps process stakeholders interactively determine the suitability of an event log for PM analysis.

Finally, it was found that there is still a lack of research on balancing tool-based automation and guidance of BPI tasks with the incorporation of process data and domain expertise. Therefore, Section IV proposes an approach to leverage process data that can be generated by PM for semi-automating BPI. Research Paper P5 presents a conceptualization for aBPR that takes process data as input and interactively assists the user in improving the business process to achieve a specified performance objective. Four types of recommendations assist the user by leveraging increasing domain and use case knowledge. The proposed aBPR reference architecture is a template for new instantiations to address the lack of tools. The paper provides evidence of the novelty of the approach and demonstrates functionalities on artificial process data using the prototypical instantiation. Furthermore, the design specification and the prototype were discussed with experts from academia and industry to demonstrate their applicability. Finally, a case study was conducted in a naturalistic setting to demonstrate the utility of aBPR. The research adds to the prescriptive knowledge on BPI by building on and extending existing approaches. aBPR provides a novel approach for applying redesign recommendations with varying levels of automation and interactivity.

V.2 Limitations and Future Research

The results of this thesis need to be reflected against some limitations that also stimulate future research. This section zooms out from the individual limitations of each research paper that can be found in the articles (see Sections VII.3 to VII.7). The following presents limitations and avenues for future research in the preceding and subsequent steps of the core PM analysis and on the managerial side of PM adoption.

Methodologically, the DSR and Delphi methods were central to the development of the results. Accordingly, some limitations are inherent in the nature of these methodological frameworks. First, as with any Delphi study, the results are based on the perceptions of a limited number of experts. Since the size of the panel consulted aligns with the exploratory nature of Delphi studies but is relatively small for statistical purposes, panel bias cannot be formally excluded. Therefore formal claims about the representativeness of the results are not possible. Nevertheless, the structured approach used to assemble the panel (with explicit selection criteria), as well as the positive feedback and high level of satisfaction throughout the study, support confidence in the validity of the results. Second, in this thesis, various artifacts resulted from DSR. A known challenge in the IS literature is the limited advice on measuring the level of contribution of DSR artifacts (Gregor and Hevner, 2013). In general, DSR strives to design artifacts that address fundamental unsolved problems innovatively and are useful to a specific user group (Hevner et al., 2004; Gregor and Hevner, 2013). To address this shortcoming, the design of the artifacts has consistently been complemented by extensive evaluations to provide some evidence of the applicability and usefulness of the artifacts.

Given that PM is a process-oriented form of data science, there have been occasional questions about the exclusivity of the management contributions to PM. While it can be confirmed that the results overlap with findings in related disciplines, it was not the scope of this thesis to systematically identify exclusive components that only apply to PM. Instead, the goal was also to learn from related disciplines and to obtain holistic results with specific relevance to PM or a specific interpretation in the context of PM, e.g., a holistic overview of challenges and opportunities relevant to PM. Therefore, the thesis aimed at validating the specific relevance of the results, which can be seen in the expert ratings in the Delphi study or the evaluation results of the managerial artifacts reflecting the particular interest of the PM field in the results. Accordingly, future research should be encouraged to learn from related fields to generate results relevant to the PM field.

For the design of the process data quality framework, it has to be considered that it is at best possible to anticipate medium-term developments (e.g., object-centric event logs) in research and practice. Although the PraeclarusPDQ framework was designed to be as flexible as possible to accommodate innovative approaches to process data quality management, the framework will likely require repeated minor slight updates to reflect the latest developments in the field. Furthermore, process data quality management has proven to be very domain dependent, which significantly hampers the ability to design generalizable and automated artifacts. Accordingly, a strong focus on inputs that systematically incorporate expert knowledge to account for domain dependencies is required. While data quality management approaches in this thesis build on the involvement of a human in the loop as a major strength, they also have limitations in that they require additional effort from the users to improve the detection, assessment, and resolution of process data quality issues.

For data-driven BPI, user input and a human in the loop are also mandatory for the aBPR approach. The evaluation showed that the process data needed to create accurate simulation models are not readily available yet and must be generated and imported mostly manually. Although BPI is a multi-dimensional construct, the current approach is designed to optimize the process towards a single performance objective, such as time or cost. In addition, aBPR can only improve business processes incrementally. Therefore, process owners must complement this process exploration tool with more radical process innovation concepts to operate business processes successfully over the long term.

Beyond the limitations, this thesis opens additional avenues for future research. Regarding management topics, the work presented provides a first step in adding content and structure to the field. Nevertheless, many organizational issues that can benefit from this thesis remain unsolved. Therefore, the identified organizational challenges for PM and the future directions by vom Brocke et al. (2021) provide a holistic foundation on major challenges to stimulate future research in the field. For example, the results of the Delphi study can be extended to assess the maturity of organizations, e.g., by determining the extent to which organizations seize opportunities and face challenges. Furthermore, there is still ambiguity about the (business) value of PM. While this thesis provides some directions on the value of PM, an in-depth scientific evaluation of the PM value concept and subsequent validation of value assessment criteria would help research and practice to understand how PM can create business value and enhance the current body of research. Finally, effective and efficient implementation and adoption of PM requires establishing governance structures to understand responsibilities, roles, and communication channels clearly.

Regarding process data quality management, the PraeclarusPDQ framework invites researchers for future contributions by design. Therefore, this thesis encourages future research to use the provided framework to add new approaches supporting process data quality management. The approach of detecting timestamp issues and measuring the event log quality is only a first step, but there are still many blind spots for process data quality management. The vision is to provide an integrated approach to detect, quantify, repair, and track process data quality issues. However, current research is heavily focused on identifying and resolving quality issues in process data. Future research could also focus more on guidelines and approaches for preventing and mitigating process data quality issues to reduce the effort required for process data preprocessing in the long run.

While the aBPR approach provides some foundation for prescriptive BPI, there is still potential to increase the level of automation of BPI and to reduce manual effort. Prescriptive BPI concerns (semi-)automated guidance on beneficial process changes to be implemented. However, a human is still required to implement the process changes. In addition, the aBPR approach requires the manual collection and import of process data for process simulations. A comprehensive end-to-end approach should integrate components to derive the initial process models from existing process data, such as event logs or other documents. The ultimate level of BPI is introduced as *"Augmented BPM"* by Dumas (2021) and concerns adaptive BPI at runtime. Therefore, data analytics and artificial intelligence (AI) techniques can drive the future of BPM and BPI to continuously monitor and improve business processes based on automated decision-making and adaptation (Dumas, 2021).

Overall, I am confident that my dissertation will contribute to the current body of knowledge and pave the way for future research on the organizational side of PM adoption and the end-to-end execution of PM projects.

VI References

Allen, L., A. O'Connell, and V. Kiermer (2019). "How Can We Ensure Visibility and Diversity in Research Contributions? How the Contributor Role Taxonomy (CRediT) Is Helping the Shift from Authorship to Contributorship". In: *Learned Publishing* 32.1, pp. 71–74. DOI: 10.1002/leap.1210.

Andrews, R., C. G. J. van Dun, M. T. Wynn, W. Kratsch, M. K. E. Röglinger, and A. H. M. ter Hofstede (2020). "Quality-Informed Semi-Automated Event Log Generation for Process Mining". In: *Decision Support Systems* 132.3. DOI: 10.1016/j.dss.2020.113265.

Andrews, R., M. T. Wynn, K. Vallmuur, A. H. M. ter Hofstede, and E. Bosley (2020). "A Comparative Process Mining Analysis of Road Trauma Patient Pathways". In: *International Journal of Environmental Research and Public Health* 17.10, p. 3426. DOI: 10.3390/ijerph17103426.

Andrews, R., M. T. Wynn, K. Vallmuur, A. H. M. ter Hofstede, E. Bosley, M. Elcock, and S. Rashford (2019). "Leveraging Data Quality to Better Prepare for Process Mining: An Approach Illustrated Through Analysing Road Trauma Pre-Hospital Retrieval and Transport Processes in Queensland". In: *International Journal of Environmental Research and Public Health* 16.7, p. 1138. DOI: 10.3390/ijerph16071138.

Archer, N. P. and F. Ghasemzadeh (1999). "An Integrated Framework for Project Portfolio Selection". In: *International Journal of Project Management* 17.4, pp. 207–216. DOI: 10.1016/S0263-7863(98)00032-5.

Augusto, A., R. Conforti, M. Dumas, M. La Rosa, F. M. Maggi, A. Marrella, M. Mecella, and A. Soo (2018). "Automated Discovery of Process Models from Event Logs: Review and Benchmark". In: *IEEE Transactions on Knowledge and Data Engineering* 31.4, pp. 686–705. DOI: 10.1109/TKDE.2018.2841877.

Beerepoot, I. et al. (2023). "The Biggest Business Process Management Problems to Solve before We Die". In: *Computers in Industry* 146, p. 103837. DOI: 10.1016/j.compind.2022.103837.

Beverungen, D. et al. (2021). "Seven Paradoxes of Business Process Management in a Hyper-Connected World". In: *Business & Information Systems Engineering* 63.2, pp. 145–156. DOI: 10.1007/s12599-020-00646-z.

Bose, R. P. J. C., R. S. Mans, and W. M. P. van der Aalst (2013). "Wanna Improve Process Mining Results?" In: *2013 IEEE Symposium on Computational Intelligence and Data Mining (CIDM)*. IEEE, pp. 127–134. DOI: 10.1109/CIDM.2013.6597227.

Carmona, J., B. van Dongen, A. Solti, and M. Weidlich (2018). *Conformance Checking: Relating Processes and Models*. Cham: Springer. DOI: 10.1007/978-3-319-99414-7.

Davis, F. D. (1989). "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology". In: *MIS Quarterly* 13.3, pp. 319–340. DOI: 10.2307/249008.

de Bruin, T. and M. Rosemann (2007). "Using the Delphi Technique to Identify BPM Capability Areas". In: *ACIS 2007 Proceedings*. URL: https://aisel.aisnet.org/acis2007/42.

Denner, M.-S., L. C. Püschel, and M. Röglinger (2018). "How to Exploit the Digitalization Potential of Business Processes". In: *Business & Information Systems Engineering* 60.4, pp. 331–349. DOI: 10.1007/s12599-017-0509-x.

Dixit, P. M., S. Suriadi, R. Andrews, M. T. Wynn, A. H. M. ter Hofstede, J. C. A. M. Buijs, and W. M. P. van der Aalst (2018). "Detection and Interactive Repair of Event Ordering Imperfection in Process Logs". In: *Advanced Information Systems Engineering*. Ed. by J. Krogstie and H. A. Reijers. Vol. 10816. Lecture Notes in Computer Science. Cham: Springer, pp. 274–290. DOI: 10.1007/978-3-319-91563-0_17.

Dumas, M. (2021). *From Process Mining to Augmented Business Process Management*. URL: https://www.linkedin.com/pulse/from-process-mining-augmented-business-management-marlon-dumas/.

Dumas, M., M. La Rosa, J. Mendling, and H. A. Reijers (2018). *Fundamentals of Business Process Management*. Second. Berlin, Heidelberg: Springer. DOI: 10.1007/978-3-662-56509-4.

Essam, M. M. and S. Limam Mansar (2012). "Towards a Software Framework for Automatic Business Process Redesign". In: *ACEEE International Journal on Communication*, pp. 23–28.

Fehrer, T., D. A. Fischer, S. J. J. Leemans, M. Röglinger, and M. T. Wynn (2022). "A Tool for Assisted Business Process Redesign". In: *Proceedings of the Best Dissertation Award, Doctoral Consortium, and Demonstration & Resources Track at BPM 2022*. Ed. by C. Janiesch, C. Di Francescomarino, T. Grisold, A. Kumar, J. Mendling, B.

Pentland, H. Reijers, and M. Weske. Vol. 3216. CEUR Workshop Proceedings. CEUR, pp. 97–101. URL: https://ceur-ws.org/Vol-3216/paper_243.pdf.

Fellmann, M., A. Koschmider, R. Laue, A. Schoknecht, and A. Vetter (2019). "Business Process Model Patterns: State-of-the-Art, Research Classification and Taxonomy". In: *Business Process Management Journal* 25.5, pp. 972–994. DOI: 10.1108/BPMJ-01-2018-0021.

Fortune Business Insights (2022). *Process Mining Software Market Size, Growth & Report*. URL: https://www.fortunebusinessinsights.com/process-mining-software-market-104792.

Galster, M. and P. Avgeriou (2011). "Empirically-Grounded Reference Architectures: A Proposal". In: *Proceedings of the Joint ACM SIGSOFT Conference – QoSA and ACM SIGSOFT Symposium – ISARCS on Quality of Software Architectures – QoSA and Architecting Critical Systems – ISARCS*. New York: Association for Computing Machinery, pp. 153–158. DOI: 10.1145/2000259.2000285.

Gartner (2023). *Process Mining Reviews and Ratings*. URL: https://www.gartner.com/reviews/market/process-mining.

Goel, K., F. Emamjome, and A. H. M. ter Hofstede (2021). "Data Governance for Managing Data Quality in Process Mining". In: *ICIS 2021 Proceedings*. URL: https://aisel.aisnet.org/icis2021/governance/governance/9.

Graham, J. (2022). *Microsoft Acquires Minit to Strengthen Process Mining Capabilities*. URL: https://blogs.microsoft.com/blog/2022/03/31/microsoft-acquires-minit-to-strengthen-process-mining-capabilities/.

Gregor, S. and A. R. Hevner (2013). "Positioning and Presenting Design Science Research for Maximum Impact". In: *MIS Quarterly* 37.2, pp. 337–355. DOI: 10.25300/MISQ/2013/37.2.01.

Grisold, T., J. Mendling, M. Otto, and J. vom Brocke (2020). "Adoption, Use and Management of Process Mining in Practice". In: *Business Process Management Journal* 27.2, pp. 369–387. DOI: 10.1108/BPMJ-03-2020-0112.

Gross, S., K. Stelzl, T. Grisold, J. Mendling, M. Röglinger, and J. vom Brocke (2021). "The Business Process Design Space for Exploring Process Redesign Alternatives". In: *Business Process Management Journal* 27.8, pp. 25–56. DOI: 10.1108/BPMJ-03-2020-0116.

Gschwandtner, T., J. Gärtner, W. Aigner, and S. Miksch (2012). "A Taxonomy of Dirty Time-Oriented Data". In: *Multidisciplinary Research and Practice for Information Systems*. Ed. by G. Quirchmayr, J. Basl, I. You, L. Xu, and E. Weippl. Vol. 7465. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, pp. 58–72. DOI: 10.1007/978-3-642-32498-7_5.

Gupta, U. G. and R. E. Clarke (1996). "Theory and Applications of the Delphi Technique: A Bibliography (1975–1994)". In: *Technological Forecasting and Social Change* 53.2, pp. 185–211. DOI: 10.1016/S0040-1625(96)00094-7.

Hevner, A. R., S. T. March, J. Park, and S. Ram (2004). "Design Science in Information Systems Research". In: *MIS Quarterly* 28.1, pp. 75–105. DOI: 10.2307/25148625.

Huang, S. Y., C.-H. Lee, A.-A. Chiu, and D. C. Yen (2015). "How Business Process Reengineering Affects Information Technology Investment and Employee Performance under Different Performance Measurement". In: *Information Systems Frontiers* 17.5, pp. 1133–1144. DOI: 10.1007/s10796-014-9487-4.

Kerpedzhiev, G. D., U. M. König, M. Röglinger, and M. Rosemann (2021). "An Exploration into Future Business Process Management Capabilities in View of Digitalization: Results from a Delphi Study". In: *Business & Information Systems Engineering* 63, pp. 83–96. DOI: 10.1007/s12599-020-00637-0.

Kerremans, M., T. Srivastava, and F. Choudhary (2021). *Market Guide for Process Mining*. URL: https://www.signavio.com/downloads/analyst-reports/gartner-market-guide-process-mining/.

Ko, J. and M. Comuzzi (2021). "Detecting Anomalies in Business Process Event Logs Using Statistical Leverage". In: *Information Sciences* 549, pp. 53–67. DOI: 10.1016/j.ins.2020.11.017.

Kratsch, W., J. Manderscheid, and M. Röglinger (2020). "Machine Learning in Business Process Monitoring: A Comparison of Deep Learning and Classical Approaches Used for Outcome Prediction". In: *Business & Information Systems Engineering* 63.3, pp. 261–276. DOI: 10.1007/s12599-020-00645-0.

Lunden, I. (2022). "IBM Acquires Italy's myInvenio to Integrate Process Mining Directly into Its Suite of Automation Tools". In: *TechCrunch*. URL: https://techcrunch.com/2021/04/15/ibm-acquires-italys-myinvenio-to-integrate-process-mining-directly-into-its-suite-of-automation-tools/.

Martin, N., B. Depaire, and A. Caris (2016). “The Use of Process Mining in Business Process Simulation Model Construction: Structuring the Field”. In: *Business & Information Systems Engineering* 58.1, pp. 73–87. DOI: 10.1007/s12599-015-0410-4.

Martin, N., 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, and M. T. Wynn (2021). “Opportunities and Challenges for Process Mining in Organizations: Results of a Delphi Study”. In: *Business & Information Systems Engineering* 63.5, pp. 511–527. DOI: 10.1007/s12599-021-00720-0.

Metinko, C. (2022). *Celonis Raises $1B In Funding At $13B Valuation*. URL: https://news.crunchbase.com/web3/celonis-raises-1billion/.

Netjes, M., H. A. Reijers, and W. M. P. van der Aalst (2010). “The PrICE Tool Kit: Tool Support for Process Improvement”. In: *Proceedings of the Business Process Management 2010 Demonstration Track*. Ed. by M. La Rosa. Vol. 615. CEUR Workshop Proceedings. CEUR, pp. 58–63. URL: http://ceur-ws.org/Vol-615/paper18.pdf.

Parasuraman, R., T. B. Sheridan, and C. D. Wickens (2000). “A Model for Types and Levels of Human Interaction with Automation”. In: *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* 30.3, pp. 286–297. DOI: 10.1109/3468.844354.

Paré, G., A. Cameron, P. Poba-Nzaou, and M. Templier (2013). “A Systematic Assessment of Rigor in Information Systems Ranking-Type Delphi Studies”. In: *Information & Management* 50.5, pp. 207–217. DOI: 10.1016/j.im.2013.03.003.

Peffers, K., T. Tuunanen, M. A. Rothenberger, and S. Chatterjee (2007). “A Design Science Research Methodology for Information Systems Research”. In: *Journal of Management Information Systems* 24.3, pp. 45–77. DOI: 10.2753/MIS0742-1222240302.

Pries-Heje, J., R. Baskerville, and J. Venable (2008). “Strategies for Design Science Research Evaluation”. In: *ECIS 2008 Proceedings*. 87, pp. 1–13. URL: http://aisel.aisnet.org/ecis2008/87.

Ralyté, J., R. Deneckère, and C. Rolland (2003). “Towards a Generic Model for Situational Method Engineering”. In: *Advanced Information Systems Engineering*. Ed. by J. Eder and M. Missikoff. Vol. 2681. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, pp. 95–110. DOI: 10.1007/3-540-45017-3_9.

Reijers, H. A. and S. Limam Mansar (2005). "Best Practices in Business Process Redesign: An Overview and Qualitative Evaluation of Successful Redesign Heuristics". In: *Omega* 33.4, pp. 283–306. DOI: 10.1016/j.omega.2004.04.012.

Reinkemeyer, L., ed. (2020). *Process Mining in Action: Principles, Use Cases and Outlook*. First. Cham: Springer. DOI: 10.1007/978-3-030-40172-6.

Reinkemeyer, L., P. Grindemann, V. Egli, M. Röglinger, L. Marcus, and L. Fabri (2022). *Accelerating Business Transformation with Process Mining Centers of Excellence (CoEs)*. URL: https://www.celonis.com/report/fraunhofer-study/ty/.

Röglinger, M., D. Fischer, and F. Baumgarte (2021). *Prozessdigitalisierung für das „New Normal" - Branchenübergreifende Studie zu Herausforderungen und Chancen der Prozessoptimierung*. URL: https://digital.abbyy.com/fraunhofer-studie.html.

Röglinger, M., C. van Dun, T. Fehrer, D. A. Fischer, L. Moder, and W. Kratsch (2021). "Automated Process (Re-)Design". In: *Proceedings of the International Workshop on BPM Problems to Solve Before We Die (PROBLEMS 2021)*. Ed. by I. Beerepoot, C. Di Ciccio, A. Marrella, H. A. Reijers, S. Rinderle-Ma, and B. Weber. Vol. 2938. CEUR Workshop Proceedings. CEUR, pp. 28–33. URL: https://ceur-ws.org/Vol-2938/paper-PROBLEMS-28.pdf.

Rott, J. and M. Böhm (2022). "Mine the Right Process – towards a Method for Selecting a Suitable Use Case for Process Mining Adoption". In: *ECIS 2022 Research Papers*. URL: https://aisel.aisnet.org/ecis2022_rp/45.

Rozinat, A. (2021). *Project Guide – Part 1: Process Selection*. URL: https://fluxicon.com/blog/2021/04/project-guide/.

Sadeghianasl, S., A. H. M. ter Hofstede, M. T. Wynn, and S. Suriadi (2019). "A Contextual Approach to Detecting Synonymous and Polluted Activity Labels in Process Event Logs". In: *On the Move to Meaningful Internet Systems: OTM 2019 Conferences*. Ed. by H. Panetto, C. Debruyne, M. Hepp, D. Lewis, C. A. Ardagna, and R. Meersman. Vol. 11877. Lecture Notes in Computer Science. Cham: Springer, pp. 76–94. DOI: 10.1007/978-3-030-33246-4_5.

Schmidt, R. C. (1997). "Managing Delphi Surveys Using Nonparametric Statistical Techniques". In: *Decision Sciences* 28.3, pp. 763–774. DOI: 10.1111/j.1540-5915.1997.tb01330.x.

Signavio (2021). *SAP to Acquire Business Process Intelligence Company Signavio - SAP Signavio*. URL: https://www.signavio.com/news/signavio-joins-sap/.

Skinner, D., R. R. Nelson, W. W. Chin, and L. Land (2015). "The Delphi Method Research Strategy in Studies of Information Systems". In: *Communications of the Association for Information Systems* 37, pp. 31–63. DOI: 10.17705/1CAIS.03702.

Song, M. and W. M. P. van der Aalst (2008). "Towards Comprehensive Support for Organizational Mining". In: *Decision Support Systems* 46.1, pp. 300–317. DOI: 10.1016/j.dss.2008.07.002.

Sonnenberg, C. and J. vom Brocke (2012). "Evaluations in the Science of the Artificial – Reconsidering the Build-Evaluate Pattern in Design Science Research". In: *Design Science Research in Information Systems. Advances in Theory and Practice*. Ed. by K. Peffers, M. Rothenberger, and B. Kuechler. Berlin, Heidelberg: Springer, pp. 381–397. DOI: 10.1007/978-3-642-29863-9_28.

Stettina, C. J. and J. Hörz (2015). "Agile Portfolio Management: An Empirical Perspective on the Practice in Use". In: *International Journal of Project Management* 33.1, pp. 140–152. DOI: 10.1016/j.ijproman.2014.03.008.

Suriadi, S., R. Andrews, A. H. M. ter Hofstede, and M. T. Wynn (2017). "Event Log Imperfection Patterns for Process Mining: Towards a Systematic Approach to Cleaning Event Logs". In: *Information Systems* 64, pp. 132–150. DOI: 10.1016/j.is.2016.07.011.

Teinemaa, I., M. Dumas, M. La Rosa, and F. M. Maggi (2019). "Outcome-Oriented Predictive Process Monitoring: Review and Benchmark". In: *ACM Transactions on Knowledge Discovery from Data* 13.2, pp. 1–57. DOI: 10.1145/3301300.

Thiede, M., D. Fuerstenau, and A. P. Bezerra Barquet (2018). "How Is Process Mining Technology Used by Organizations? A Systematic Literature Review of Empirical Studies". In: *Business Process Management Journal* 24.4, pp. 900–922. DOI: 10.1108/BPMJ-06-2017-0148.

van der Aalst, W. M. P. (2016). *Process Mining: Data Science in Action*. Second. Berlin, Heidelberg: Springer. DOI: 10.1007/978-3-662-49851-4.

van der Aalst, W. M. P., A. Adriansyah, et al. (2012). "Process Mining Manifesto". In: *Business Process Management Workshops*. Ed. by F. Daniel, K. Barkaoui, and S. Dustdar. Vol. 99. Lecture Notes in Business Information Processing. Berlin, Heidelberg: Springer, pp. 169–194. DOI: 10.1007/978-3-642-28108-2_19.

van der Aalst, W. M. P., M. Bichler, and A. Heinzl (2017). “Responsible Data Science”. In: *Business & Information Systems Engineering* 59.5, pp. 311–313. DOI: 10.1007/s12599-017-0487-z.

van Eck, M. L., X. Lu, S. J. J. Leemans, and W. M. P. van der Aalst (2015). “PM$\hat{2}$: A Process Mining Project Methodology”. In: *Advanced Information Systems Engineering*. Ed. by J. Zdravkovic, M. Kirikova, and P. Johannesson. Vol. 9097. Lecture Notes in Computer Science. Cham: Springer, pp. 297–313. DOI: 10.1007/978-3-319-19069-3_19.

Venable, J., J. Pries-Heje, and R. Baskerville (2012). “A Comprehensive Framework for Evaluation in Design Science Research”. In: *Design Science Research in Information Systems. Advances in Theory and Practice*. Ed. by K. Peffers, M. A. Rothenberger, and B. Kuechler. Vol. 7286. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, pp. 423–438. DOI: 10.1007/978-3-642-29863-9_31.

Venable, J., J. Pries-Heje, and R. Baskerville (2016). “FEDS: A Framework for Evaluation in Design Science Research”. In: *European Journal of Information Systems* 25.1, pp. 77–89. DOI: 10.1057/ejis.2014.36.

Verbeek, H. M. W., J. C. A. M. Buijs, B. van Dongen, and W. M. P. van der Aalst (2011). “XES, XESame, and ProM 6”. In: *Information Systems Evolution*. Ed. by P. Soffer and E. Proper. Berlin, Heidelberg: Springer, pp. 60–75. DOI: 10.1007/978-3-642-17722-4_5.

vom Brocke, J., M. Jans, J. Mendling, and H. A. Reijers (2021). “A Five-Level Framework for Research on Process Mining”. In: *Business & Information Systems Engineering* 63.5, pp. 483–490. DOI: 10.1007/s12599-021-00718-8.

von der Gracht, H. A. (2012). “Consensus Measurement in Delphi Studies: Review and Implications for Future Quality Assurance”. In: *Technological Forecasting and Social Change* 79.8, pp. 1525–1536. DOI: 10.1016/j.techfore.2012.04.013.

Wynn, M. T. and S. W. Sadiq (2019). “Responsible Process Mining - A Data Quality Perspective”. In: *Business Process Management*. Ed. by T. T. Hildebrandt, B. van Dongen, M. Röglinger, and J. Mendling. Vol. 11675. Lecture Notes in Computer Science. Cham: Springer, pp. 10–15. DOI: 10.1007/978-3-030-26619-6_2.

Zellner, G. (2011). “A Structured Evaluation of Business Process Improvement Approaches”. In: *Business Process Management Journal* 17.2, pp. 203–237. DOI: 10.1108/14637151111122329.

Zuhaira, B. and N. Ahmad (2021). “Business Process Modeling, Implementation, Analysis, and Management: The Case of Business Process Management Tools”. In: *Business Process Management Journal* 27.1, pp. 145–183. DOI: 10.1108/BPMJ-06-2018-0168.

VII Appendix

VII.1 Index of Research Articles

Research Paper 1: Opportunities and Challenges for Process Mining in Organisations – Results of a Delphi Study

Martin N, Fischer DA, Kerpedzhiev GD, Goel K, Leemans SJJ, Röglinger M, van der Aalst WMP, Dumas M, La Rosa M, Wynn MT (2021). Opportunities and Challenges for Process Mining in Organisations – Results of a Delphi Study. In: *Business & Information Systems Engineering*. DOI: `10.1007/s12599-021-00720-0`
(VHB-JQ3[4]: B, ABDC[5]: A, SJR[6]: Q1, IF[7]: 5.675)

Research Paper 2: A Portfolio Management Method for Process Mining-enabled Business Process Improvement Projects

Fischer DA, Marcus L, Röglinger M (2023). A Portfolio Management Method for Process Mining-enabled Business Process Improvement Projects. Submitted to: *Outlet hidden due to the double-blind review process of the journal*

Research Paper 3: PraeclarusPDQ: A Framework for Process Data Quality Management

Sadeghiansl S, Fischer DA, ter Hofstede AHM, Adams MJ, Andrews R, Comuzzi A, Ko J, Koschmider A, Wynn MT, Ziolkowski T (2023). PraeclarusPDQ: A Framework for Process Data Quality Management. Submitted to: *Big Data Research*.
(VHB-JQ3: n.a., ABDC: n.a., SJR: Q1, IF: 3.578)

Research Paper 4: Towards Interactive Event Log Forensics: Detecting and Quantifying Timestamp Imperfections

Fischer DA, Goel K, Andrews R, van Dun CGJ, Wynn MT, Röglinger M (2022). Towards Interactive Event Log Forensics: Detecting and Quantifying Timestamp Imperfections. In: *Information Systems*. DOI: `10.1016/j.is.2022.102039`
(VHB-JQ3: B, ABDC: n.a., SJR: Q1, IF: 3.18)
Earlier version published in *Business Process Management*, 2020. DOI: `10.1007/978-3-030-58666-9_18`.

---
[4]VHB-JQ3: VHB-Jourqual 3
[5]ABDC: Australian Business Deans Council Journal Quality List
[6]SJR: Scimago Journal & Country Rank
[7]IF: Impact Factor

Research Paper 5: An Assisted Approach to Business Process Redesign

Fehrer T, Fischer DA, Leemans SJJ, Röglinger M, Wynn MT (2022). An Assisted Approach to Business Process Redesign. In: *Decision Support Systems*. DOI: 10.1016/j.dss.2022.113749

(VHB-JQ3: B, ABDC: A*, SJR: Q1, IF: 6.969, Senior Scholars' List of Premier Journals)

Throughout the dissertation, I also co-authored the following research papers. These papers are not part of this dissertation.

M. Röglinger, C. van Dun, et al. (2021). "Automated Process (Re-)Design". In: *Proceedings of the International Workshop on BPM Problems to Solve Before We Die (PROBLEMS 2021)*. Ed. by I. Beerepoot et al. Vol. 2938. CEUR Workshop Proceedings. CEUR, pp. 28–33. URL: https://ceur-ws.org/Vol-2938/paper-PROBLEMS-28.pdf

I. Beerepoot et al. (2023). "The Biggest Business Process Management Problems to Solve before We Die". In: *Computers in Industry* 146, p. 103837. DOI: 10.1016/j.compind.2022.103837

T. Fehrer et al. (2022). "A Tool for Assisted Business Process Redesign". In: *Proceedings of the Best Dissertation Award, Doctoral Consortium, and Demonstration & Resources Track at BPM 2022*. Ed. by C. Janiesch et al. Vol. 3216. CEUR Workshop Proceedings. CEUR, pp. 97–101. URL: https://ceur-ws.org/Vol-3216/paper_243.pdf

M. Röglinger, D. Fischer, et al. (2021). *Prozessdigitalisierung für das „New Normal" - Branchenübergreifende Studie zu Herausforderungen und Chancen der Prozessoptimierung*. URL: https://digital.abbyy.com/fraunhofer-studie.html

VII.2 Individual Contribution to the Included Research Articles

This dissertation is cumulative and includes five research papers. All research papers were written in teams with multiple co-authors. This section outlines the settings and describes my contribution to the five papers. The descriptions follow the Contributor Roles Taxonomy (CRediT) by Allen et al. (2019).

Research Paper P1 entitled "*Opportunities and Challenges for Process Mining in Organisations – Results of a Delphi Study*" (Martin et al. 2021; Section VII.3) was written by a team of ten authors. In line with my role as the second author, I held a crucial role in most parts of the research project. I contributed significantly to the design of the research methodology. I also took a leading role in the Delphi study's iterative investigation and data curation process. In addition, I was responsible for developing meaningful visualizations based on the data. In terms of writing, I was responsible for the original drafting of individual sections and was involved in reviewing and editing the entire paper.

Research Paper P2 entitled "*A Portfolio Management Method for Process Mining-enabled Business Process Improvement Projects*" (Fischer et al. 2023; Section VII.4) was written by a team of three authors. In line with my role as the first author, I held a crucial role in all parts and administered the research. I contributed significantly to conceptualizing the research objectives and the design of the research methodology. I also led the iterative investigation and validation process of a DSR project. Furthermore, I developed the associated software prototype. In addition, I was responsible for developing meaningful visualizations based on the data. In terms of writing, I was responsible for the original drafting of most sections and was involved in reviewing and editing the entire paper.

Research Paper P3 entitled "*PraeclarusPDQ: A Framework for Process Data Quality Management*" (Sadeghianasl et al. 2023; Section VII.5) was written by a team of ten authors. In line with my role as the second author, I held a crucial role in most parts and administered the research project. I contributed significantly to the design of the DSR methodology. I also led the validation of the research outputs. In terms of writing, I was responsible for the original drafting of individual sections and was involved in reviewing and editing the entire paper.

Research Paper P4 entitled "*Towards Interactive Event Log Forensics: Detecting and Quantifying Timestamp Imperfections in Event Logs*" (Fischer et al. 2022; Section VII.6) was written by a team of six authors. In line with my role as the first author, I held a crucial role in all parts and administered the research project. I contributed significantly to conceptualizing the overarching research objectives and the design of the research methodology. I also led the iterative investigation and validation process of a DSR project. Furthermore, I developed the associated software prototype. In addition, I was responsible for developing meaningful visualizations based on the data. Regarding writing, I was responsible for the original drafting of the entire paper and stayed involved in the review and editing process.

Research Paper P5 entitled "*An Assisted Approach to Business Process Redesign*" (Fehrer et al. 2022; Section VII.7) was written by a team of five authors. In line with my role as the second author, I held a crucial role in most parts of the research project. I contributed significantly to conceptualizing the overarching research aims and the design of the research methodology. I also led the iterative investigation and validation process of a DSR project. In terms of writing, I was responsible for the original drafting of individual sections and was involved in reviewing and editing the entire paper.

VII.3 Research Paper 1: Opportunities and Challenges for Process Mining in Organisations – Results of a Delphi Study

Authors:

Niels Martin, Dominik A. Fischer, Georgi D. Kerpedzhiev, Kanika Goel, Sander J.J. Leemans, Maximilian Röglinger, Wil M.P. van der Aalst, Marlon Dumas, Marcello La Rosa, Moe T. Wynn

Published in:

Business & Information Systems Engineering 63, 511-527 (2021). DOI: 10.1007/s12599-021-00720-0

Abstract:

Process Mining is an active research domain and has been applied to understand and improve business processes. While significant research has been conducted on the development and improvement of algorithms, evidence on the application of Process Mining in organisations has been far more limited. In particular, there is limited understanding of the opportunities and challenges of using Process Mining in organisations. Such an understanding has the potential to guide research by highlighting barriers for Process Mining adoption and, thus, can contribute to successful Process Mining initiatives in practice. In this respect, this paper provides a holistic view of opportunities and challenges for Process Mining in organisations identified in a Delphi study with 40 international experts from academia and industry. Besides proposing a set of 30 opportunities and 32 challenges, the paper conveys insights into the comparative relevance of individual items, as well as differences in the perceived relevance between academics and practitioners. Therefore, the study contributes to the future development of Process Mining, both as a research field and regarding its application in organisations.

Keywords:

Process Mining, Opportunities, Challenges, Barriers, Delphi Study, Process Mining Adoption, Process Mining Use, Business Process Management

VII.4 Research Paper 2: A Portfolio Management Method for Process Mining-enabled Business Process Improvement Projects

Authors:
Dominik A. Fischer, Laura Marcus, Maximilian Röglinger

Submitted to:
Outlet hidden due to the double-blind review process of the journal

Extended Abstract:
Process mining has received tremendous attention from research and industry and established itself as an in-demand technology. While the technological maturity of process mining solutions is considered high due to extensive research and development investments over the past decades, organizations still face the challenge of elusive value when systematically adopting process mining. When establishing process mining, knowing which use cases to start with is often difficult. From a project selection perspective, Rott and Böhm (2022) provide a first approach to identifying suitable processes for process mining pilot projects. From a project execution perspective, van Eck et al. (2015) provide a methodology for the end-to-end execution of an individual process mining project. However, there is a lack of support beyond piloting when scaling and dealing with process mining project portfolios. Therefore, this paper addresses the following research question: *How can organizations manage process mining project portfolios?*

Based on design science research with situational method engineering, we propose a method for managing portfolios of so-called process mining value cases, which we define as process mining-enabled business process improvement projects (Gregor and Hevner, 2013; Ralyté et al., 2003). The overall goal of this management method is to support organizations in determining portfolios of process mining projects that generate business value by improving business processes. In line with design science research principles, the method is based on justification knowledge about project portfolio management and the business value of process mining. It is built around method requirements and three design objectives: structured guidance, consideration of process and context factors, and comparability of process mining value cases. The developed method consists of five activities that outline techniques, roles, and tools: *strategize*, *identify*, *select*, *implement*, and *monitor*. These activities are derived from the literature and supported by a panel of twelve experts from research and practice involved throughout the design and evaluation

of the method. Overall, the method is structured in a decision-support and user-guiding manner that requires humans in the loop. The method guides the user through the actions and decision points for each activity. A software prototype complements the method.

Overall, design science research projects should aim to design useful artifacts (Gregor and Hevner, 2013). To achieve this goal, the evaluation of the method was designed based on three phases according to Venable et al. (2012). During these phases, we evaluated the applicability and real-world fidelity by involving an expert panel of academics and practitioners. Furthermore, we substantiate the usefulness of the artifact through a real-world case study in a naturalistic setting.

Keywords:
Process mining, Business process improvement, Project selection, Portfolio management

References:

Gregor, S. and A. R. Hevner (2013). "Positioning and Presenting Design Science Research for Maximum Impact". In: *MIS Quarterly* 37.2, pp. 337–355. DOI: 10.25300/MISQ/2013/37.2.01.

Ralyté, J., R. Deneckère, and C. Rolland (2003). "Towards a Generic Model for Situational Method Engineering". In: *Advanced Information Systems Engineering*. Ed. by J. Eder and M. Missikoff. Vol. 2681. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, pp. 95–110. DOI: 10.1007/3-540-45017-3_9.

Rott, J. and M. Böhm (2022). "Mine the Right Process – towards a Method for Selecting a Suitable Use Case for Process Mining Adoption". In: *ECIS 2022 Research Papers*. URL: https://aisel.aisnet.org/ecis2022_rp/45.

van Eck, M. L., X. Lu, S. J. J. Leemans, and W. M. P. van der Aalst (2015). "PM2̂: A Process Mining Project Methodology". In: *Advanced Information Systems Engineering*. Ed. by J. Zdravkovic, M. Kirikova, and P. Johannesson. Vol. 9097. Lecture Notes in Computer Science. Cham: Springer, pp. 297–313. DOI: 10.1007/978-3-319-19069-3_19.

Venable, J., J. Pries-Heje, and R. Baskerville (2012). "A Comprehensive Framework for Evaluation in Design Science Research". In: *Design Science Research in Information Systems. Advances in Theory and Practice*. Ed. by K. Peffers, M. A. Rothenberger, and B. Kuechler. Vol. 7286. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, pp. 423–438. DOI: 10.1007/978-3-642-29863-9_31.

VII.5 Research Paper 3: PraeclarusPDQ: A Framework for Process Data Quality Management

Authors:

Sareh Sadeghianasl, Dominik A. Fischer, Arthur H.M. ter Hofstede, Michael Adams, Robert Andrews, Marco Comuzzi, Jonghyeon Ko, Agnes Koschmider, Moe Thandar Wynn, Tobias Ziolkowski

Submitted to:

Big Data Research

Extended Abstract:

After its emergence over two decades ago, process mining flourished as a discipline. There have been many contributions to its theory, it is widely applied in practice, and mature commercial environments support it. However, its potential for significant organizational impact is constrained by contemporary consideration and treatment of (poor) quality of event data. In practice, event logs tend to suffer from significant data quality problems that need to be recognized and resolved effectively for analysis results to be meaningful. Despite its importance, the topic of data quality in process mining has received limited attention. The complex and rich nature of process data quality management means that technological solutions should be able to evolve as our understanding of the field increases, and new strategies should not require a complete overhaul of the architecture of such solutions. However, a dedicated environment focusing on detecting, measuring, and repairing data quality problems is a sine qua non for the next generation of process data quality management. Therefore, this paper aims for the following research question: *How to design a reference architecture for process data quality management?*

Based on DSR and guidelines for empirically based reference architectures, this paper proposes PraeclarusPDQ, a reference architecture for process data quality management (Gregor and Hevner, 2013; Galster and Avgeriou, 2011). The reference architecture provides two interfaces. The first is a plugin interface, an abstraction that (currently) allows for four types of plugins: readers, writers, patterns, and actions. The second interface is a public-facing interface that facilitates user interaction. It also provides an API for external applications and services. The architecture is designed to accommodate new developments in data quality management strategies, as well as future developments in mitigation and prevention. It is built to anticipate new methods and strategies beyond current algorithms

and techniques. Therefore, the key underlying principle is abstraction and extensibility, not only for data quality improvement but also for visualization techniques that allow analysts to interact with data in new ways. The reference architecture has also been instantiated as an open-source software environment aiming to galvanize the process mining community and lead to significant future breakthroughs in the theory and practice of process data quality management.

In line with DSR and reference architecture development principles, multiple evaluation phases were integrated into the design and development process. The reference architecture and associated software framework have been tested and evaluated for their applicability, usefulness, and ease of use adopting the FEDS by Venable et al. (2016). The Praeclarus-PDQ framework is publicly available, which can serve as a rallying point for process data quality researchers. The framework will help them develop their research contributions in process data quality management and make them publicly available for feedback. Based on the open-source software approach, it is the ambition that the PraeclarusPDQ framework will evolve into an artifact that facilitates research collaboration in process data quality management and will become the hub for software contributions in this area.

Keywords:
Process mining, Process data quality, Event log, Process data governance

References:

Galster, M. and P. Avgeriou (2011). “Empirically-Grounded Reference Architectures: A Proposal”. In: *Proceedings of the Joint ACM SIGSOFT Conference – QoSA and ACM SIGSOFT Symposium – ISARCS on Quality of Software Architectures – QoSA and Architecting Critical Systems – ISARCS*. New York: Association for Computing Machinery, pp. 153–158. DOI: 10.1145/2000259.2000285.

Gregor, S. and A. R. Hevner (2013). “Positioning and Presenting Design Science Research for Maximum Impact”. In: *MIS Quarterly* 37.2, pp. 337–355. DOI: 10.25300/MISQ/2013/37.2.01.

Venable, J., J. Pries-Heje, and R. Baskerville (2016). “FEDS: A Framework for Evaluation in Design Science Research”. In: *European Journal of Information Systems* 25.1, pp. 77–89. DOI: 10.1057/ejis.2014.36.

VII.6 Research Paper 4: Towards Interactive Event Log Forensics: Detecting and Quantifying Timestamp Imperfections in Event Logs

Authors:

Dominik A. Fischer, Kanika Goel, Robert Andrews, Christopher G.J. van Dun, Moe T. Wynn, Maximilian Röglinger

Published in:

Information Systems 106, 102039 (2022). DOI: `10.1016/j.is.2022.102039`

Abstract:

Timestamp information recorded in event logs plays a crucial role in uncovering meaningful insights into business process performance and behaviour via Process Mining techniques. Inaccurate or incomplete timestamps may cause activities in a business process to be ordered incorrectly, leading to unrepresentative process models and incorrect process performance analyses. Thus, the quality of timestamps in an event log should be evaluated thoroughly before the event log is used for any Process Mining activity. To the best of our knowledge, research on the quality assessment of event logs remains scarce. Our work presents a user-guided and semi-automated approach for detecting and quantifying timestamp-related issues in event logs. We define 15 metrics related to timestamp quality across two axes: four levels of abstraction (event, activity, trace, log) and four quality dimensions (accuracy, completeness, consistency, uniqueness). The approach has been implemented as a prototype and evaluated regarding its design specification, instantiation, and usefulness in artificial and naturalistic settings by including experts from research and practice. Overall, our approach paves the way for a systematic and interactive enhancement of event log quality during the data preprocessing phase of Process Mining projects.

Keywords:

Process Mining, Event log, Data quality, Timestamps, Quality assessment

VII.7 Research Paper 5: An Assisted Approach to Business Process Redesign

Authors:
Tobias Fehrer, Dominik A. Fischer, Sander J.J. Leemans, Maximilian Röglinger, Moe T. Wynn

Published in:
Decision Support Systems 156, 113749 (2022). DOI: `10.1016/j.dss.2022.113749`

Abstract:
For many organizations, the continuous optimization of their business processes has become a critical success factor. Several related methods exist that enable the step-by-step redesign of business processes. However, these methods are mainly performed manually and require both creativity and business process expertise, which is often hard to combine in practice. To enhance the quality and effectiveness of business process redesign, this paper presents a conceptualization of assisted business process redesign (aBPR). The aBPR concept guides users in improving business processes based on redesign patterns. Depending on the data at hand, the aBPR concept classifies four types of recommendations that differ in their level of automation. Further, this paper proposes a reference architecture that provides operational support for implementing aBPR tools. The reference architecture has been instantiated as a prototype and evaluated regarding its applicability and usefulness in artificial and naturalistic settings by performing an extensive real-world case study at KUKA and interviewing experts from research and practice.

Keywords:
Business Process Redesign, Reference Architecture, User Guidance, Business Process Management