A Cross-Organizational Process Mining Framework

Ünal Aksu

SIKS Dissertation Series No. 2022-4
The research reported in this thesis has been carried out under auspices of SIKS, the Dutch Research School for Information and Knowledge Systems.

Promotiecommissie:

Prof. dr. Antonio Ruiz Cortés	(University of Seville, Spain)
Prof. dr. J.F.M. (Frans) Feldberg	(Vrije Universiteit Amsterdam, The Netherlands)
Prof. dr. Thomas Troels Hildebrandt	(University of Copenhagen, Denmark)
Prof. dr. Yannis Velegrakis	(Utrecht University, The Netherlands)
Dr. ir. Irene Vanderfeesten	(Eindhoven University of Technology, The Netherlands)

This thesis was accomplished with financial support from NWO in the *AMUSE* project [code: 628.006.001].

ISBN: 978-94-6423-649-1
NUR: 801, 983
DOI: 10.33540/1007

Cover design by Ünal Aksu
Printed by ProefschriftMaken | www.proefschriftmaken.nl

A Cross-Organizational Process Mining Framework

Een Organisatieoverschrijdend Kader Voor Process Mining

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof.dr. H. R. B. M. Kummeling, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op woensdag 2 februari 2022 des middags te 2.15 uur door

Ünal Aksu

geboren te Artvin, Turkije

Promotoren:	Prof. dr. ir. H.A. Reijers
	Prof. dr. S. Brinkkemper
Co-promotor:	Dr. ir. D.M.M. Schunselaar

Dedicated to

my family

&

all orphans.

Acknowledgments

First and foremost, I am immensely grateful to my supervisors, Hajo Reijers and Dennis Schunselaar. Hajo: thanks for the life-changing experience and for supporting me to grow in many ways. Dennis: thanks for guiding me in this journey and for all the suggestions and feedback you provided.

It certainly was a great pleasure to take part in the AMUSE project. I would like to express my gratitude to Sjaak Brinkkemper and Slinger Jansen for this opportunity. In this project, I had the chance to work with wonderful people of the dynamic company Afas to whom I owe a special thank you: Rolf de Jong, Erik van de Ven, Arjan van Hooff, Henk van der Schuur, and Dennis van Velzen. I also would like to thank Machiel de Graaf, Ariaan Hoogendijk, and the wonderful people of the NEXT and ANTA teams. For their countless effort, I owe a special thank you to Wilfred, Bouke, Paul, and Rik. I am also thankful to Guru, Michiel, and Siamak for their support and help.

Next, a big thank you goes out to Marzieh for helping me and for bringing me in contact with domain experts from ING. Many thanks, Maartje, Praveen, and Joost, for your help. It was great to have closer contact with domain experts to test the practical value of my research. Special thanks to Cocky Vlaar and Menno Rasch for giving me the chance to work together with great people at the Information and Technology Services department of Utrecht University. Thanks a lot, Juriaan, Niels, Karim, Jacqueline, Jan-Willem, Laurens, Corrie, and Arjan.

Clearly, when you collaborate with other researchers, you gain more insight and knowledge on your work. In this regard, I would like to express my sincere gratitude to the co-authors of the publications included in this thesis: to Adela del-Río-Ortega and Manuel Resinas, kind researchers of the beautiful city of Seville. Also, a special thank you goes out to Aleksandra for the wonderful guidance and company in that great city.

From the bottom of my heart, I would like to say a big thank you to the most adventurous couple ever, Banu & Deniz, for their support and encouragement.

I am grateful to all my former and present colleagues for helping me and making this journey more enjoyable. I cannot thank enough the members of the former VU Business Informatics group, with special mentions to Han, Ermeson, Isaac, Henrik, Fakhra, and Marie. In addition, thanks go out to the VU administrative staff, in particular to Mojca and Caroline, for helping me. All members of the UU BPM&A group have my deepest thanks, with special mentions for Inge, Jelmer, Xixi, Iris, Vinicius, Jan Martijn, Nico, Sietse, and Jens. Also, I would like to thank Marcelo and Sara for their collaboration and for making the office cozy with fun discussions.

Furthermore, I am grateful to the reading committee members for taking the time and assessing the work presented in this thesis.

Outside of the academic world, I am thankful to my friends for the wonderful times we shared and for the uplifting social atmosphere. In particular, thanks a lot, Cemil (kaptan), Enisio & Rita, Ediz, Bilge, and Sasha.

Finally, I would like to say a few words of thanks in Turkish. Doktora çalışmam boyunca destekleriyle beni hiçbir zaman yalnız bırakmayan aileme sonsuz teşekkürler ederim. Afacan yeğenim Eda'ya da neşe kaynağı olduğu için teşekkür borçluyum. Özel bir teşekkürü de tüm destekleri için Nilgün&Ataşen Altun'a yolluyorum. Başta Alpay Karagöz ve Onur Kaynak olmak üzere, Innova bünyesinde çalışma fırsatı bulduğum tüm iş arkadaşlarıma da destekleri için müteşekkirim. Birlikte büyüdüğüm ve cok şey paylaştığım kuzenlerim özelinde tüm akrabalarıma destekleri için minnettarım. Son olarak, tüm öğretmenlerime emekleri için teşekkür ediyorum.

Ünal Aksu,
Amsterdam, 1 August 2021

Summary

Groups of people collaborate within organizations to deliver value to their customers. To establish such collaborations, which can lead to valuable outcomes for customers, a set of coordinated activities, events, and decision points are orchestrated in business processes. However, due to the diverse contexts of organizations, the meaning of terms used in business processes often differ from one organization to another. Moreover, the way a business process is executed may differ across organizations, in some branches of the same organization, or for certain products or services offered to customers, or even across organizational units. Those differences, furthermore, affect both the performance of business processes and the relevance of performance indicators used for measuring that performance. Therefore, when such points are not taken into account, efforts to enable organizations to learn from each other for performance improvement can lead to three main issues, namely unfairness, inaccuracy, and inadequacy.

The aforementioned issues inspire the work in this thesis. In particular, the thesis focuses on providing relevant insights for organizations to improve their performance by learning from each other. With the theoretical Cross-Organizational Process Mining Framework that we developed, that learning becomes possible for organizations. The main contributions of this thesis are five approaches. The approach for automatically deriving Key Performance Indicators (KPIs) from Ontological Enterprise Models deals with the unfairness issue. Inaccuracy is the main focus of the approach for predicting relevant KPIs for organizations. Identifying the perspectives that can be adequate for organizations to learn from each other is the main focus of the two other approaches: the approach for the automated generation of engaging dashboards and the approach for interactive process performance dashboard generation. The last approach is devoted to building process benchmarks for performance improvement such that organizations can benefit from each others' best practices. Each approach is applied in a real-life setting to show its usefulness and practical value.

Overall, the work presented in this thesis provides important contributions to perform cross-organizational process mining in a fair, accurate, and adequate fashion.

Samenvatting

Groepen mensen werken samen binnen organisaties om waarde te leveren aan hun klanten. Om dergelijke samenwerkingsverbanden tot stand te brengen die kunnen leiden tot waardevolle resultaten voor klanten, wordt een reeks gecoördineerde activiteiten, gebeurtenissen en beslissingsmomenten georkestreerd in bedrijfsprocessen. Door de uiteenlopende context van organisaties verschilt de betekenis van termen die in bedrijfsprocessen worden gebruikt echter vaak van organisatie tot organisatie. Bovendien kan de wijze waarop een bedrijfsproces wordt uitgevoerd verschillen tussen organisaties, in sommige takken van dezelfde organisatie, of voor bepaalde producten of diensten die aan klanten worden aangeboden, of zelfs tussen organisatorische eenheden. Die verschillen zijn bovendien van invloed op zowel de prestaties van bedrijfsprocessen als de relevantie van prestatie-indicatoren die worden gebruikt om die prestaties te meten. Wanneer met dergelijke punten geen rekening wordt gehouden, kunnen inspanningen om organisaties in staat te stellen van elkaar te leren met het oog op prestatieverbetering leiden tot drie belangrijke problemen, namelijk oneerlijkheid, onnauwkeurigheid, en ontoereikendheid.

Voornoemd problemen die we hebben geïdentificeerd inspireren het werk in deze dissertatie. In het bijzonder richt het proefschrift zich op het verschaffen van relevante inzichten voor organisaties om hun prestaties te verbeteren door van elkaar te leren. Met het theoretische Organisatieoverschrijdend Kader voor Process Mining dat we hebben ontwikkeld, wordt dat leren mogelijk voor organisaties. De belangrijkste bijdragen van dit proefschrift zijn vijf benaderingen die door het raamwerk worden gebruikt om oplossingen te bieden voor de eerder genoemde drie hoofdproblemen. De aanpak voor het automatisch afleiden van Kritieke Prestatie-Indicatoren (KPI's) uit Ontologische Ondernemingsmodellen pakt het onnauwkeurigheidsprobleem aan. Onnauwkeurigheid is het belangrijkste aandachtspunt van de aanpak voor het voorspellen van relevante KPI's voor organisaties. Het identificeren van de perspectieven die geschikt kunnen zijn voor organisaties om van elkaar te leren is de belangrijkste focus van de volgende twee benaderingen: de benadering voor het geautomatiseerd genereren van boeiende dashboards en de benadering voor het interactief genereren van procesprestatiedashboards. De laatste benadering is gewijd aan het bouwen van procesbenchmarks voor prestatieverbetering, zodat organisaties kunnen profiteren van elkaars best practices. Elke benadering wordt toegepast in een real-life setting om het nut en de praktische waarde ervan aan te tonen.

Alles bij elkaar, levert het werk dat in deze dissertatie wordt gepresenteerd belangrijke bijdragen aan het uitvoeren van cross-organisationele process mining op een eerlijke, accurate en adequate manier.

ACRONYMS

AI	Artificial Intelligence
ANP	Analytic Network Process
B2B	Business-to-Business
B2C	Business-to-Customer
B2G	Business-to-Government
BI	Business Intelligence
BPM	Business Process Management
BPMN	Business Process Model and Notation
CA	Condition Action
CEO	Chief Executive Officer
CIO	Chief Information Officer
CM	Change Management
CRM	Customer Relationship Management
DBMS	Database Management System
DKUL	Dashboard and Key Performance Indicator Usage Logs
ERP	Enterprise Resource Planning
HRM	Human Resource Management
IM	Incident Management
IS	Information Systems
ISO	International Organization for Standardization
ITSM	Information Technology Service Management
KPI	Key Performance Indicator
KRD	Key Performance Indicator Relevance Data
OEM	Ontological Enterprise Model
PM	Problem Management
PPI	Process Performance Indicator
SBPM	Semantic Business Process Management
SGT	Sequence Graph Transform
SLA	Service Level Agreement
SME	Small-Medium Sized Enterprise
UML	Unified Modeling Language
UX	User Experience

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Organizations are entities established for a particular purpose. To accomplish their purpose with an intended impact, groups of people collaborate in organizations. Such collaborations are joined forces to deliver value to customers and stakeholders. Often, that value delivery is done through products and services. Importantly, the only way any organization can perform its operations for providing products and services is via its business processes. More specifically, with the involvement of a number of organizational functions, a set of coordinated activities, events, and decision points are instrumented in any business process to lead to an outcome that is valuable for customers [40]. Accordingly, business processes are among the key assets of organizations. Hence, the management and continuous improvement of business processes should be the primary focus of every organization [133].

Measurement is essential for process management and improvement. If it is not known whether a process produces the desired outcome, what adjustments are needed for it cannot be identified [39]. In other words, the success of a process cannot be defined and tracked unless its progress is quantified, i.e., measured. Without a clear measurement, the role of luck in improvement will become much bigger than it should [14]. Furthermore, measurement enables organizations to have a better understanding of whether their processes achieve the performance to attain their goals.

Key Performance Indicators (KPIs) are used by organizations as a means to measure and monitor their business performance. A KPI is simply a metric that indicates how well an organization is performing in relation to its goals. As a means to monitor KPIs, organizations use *dashboards*. Dashboards provide organizations visibility into their performance by consolidating and integrating required information [42, 54, 158]. The information displayed on dashboards supports decision-makers in organizations in making informed decisions to take actions on improving the business processes in their organization. Simply put, data-driven evidence is vital for organizations to gain resilience and to be more adaptive to changes [133]. Instead of relying on intuition and assumptions, high-performance organizations focus on the discovery and analysis of the inefficiencies in their processes using the data produced during the execution of those processes [14].

Information systems are intensively used by today's organizations to both support their business and execute their business processes. For example, an Enterprise Resource Planning (ERP) system is one of the most used information systems to support manufacturing, delivery, and financial processes. Similarly, marketing, sales, and customer service processes are typically supported by Customer Relationship Management (CRM) systems. These systems enable organizations to record vast amounts of data, which is a very valuable source to analyze and provide insights into understanding and improving processes.

Process Mining is a collection of techniques that aims to analyze how processes are executed in reality [139]. It allows organizations to understand the performance and conformance of processes based on the data produced during their execution. Thus, valuable insights into and recommendations on how to improve business processes can be obtained. Data recorded by information systems need to be extracted in the form of event logs [139] to obtain such insights and recommendations. An *event log* is a collection of events, each of which refers to an activity performed in a process. The events performed within the execution of a single process instance refer to a case. For example, in a bank, each mortgage application can be a case, which goes through a loan application process. The sequence of the events that are related to a particular case is called a trace.

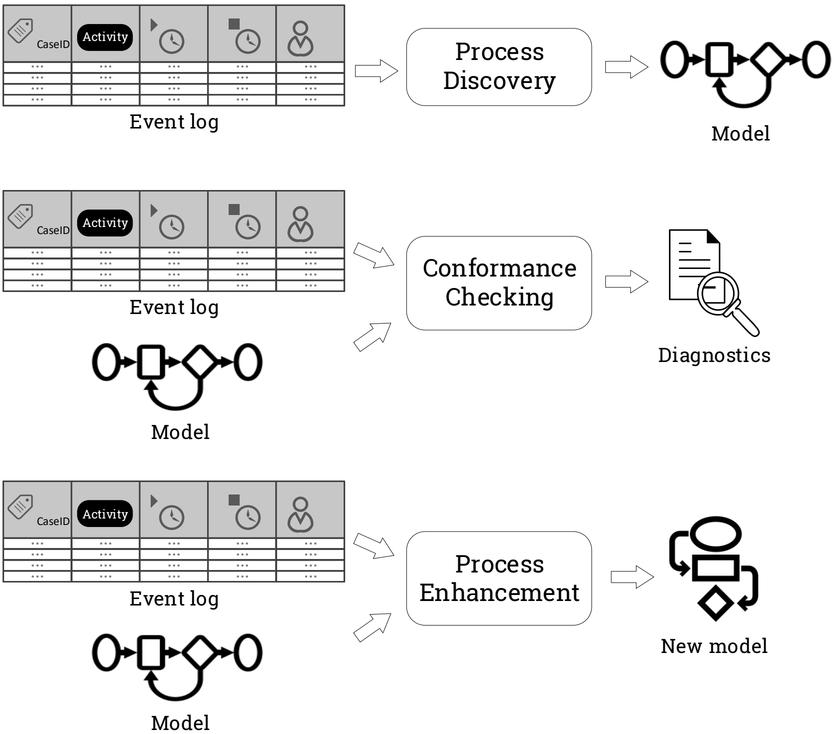

Figure 1.1: An overview of process mining tasks

There are three main tasks (see Figure 1.1) that process mining techniques focus: *process discovery*, *conformance checking*, and *process enhancement* [139]:

- *process discovery*: the construction of an accurate process model that can precisely represent the observed behavior in the given event log of a process.
- *conformance checking*: the detection of differences between the modeled behavior and the observed behavior of a process in terms of compliance and correspondence.
- *process enhancement*: the derivation of recommendations to repair or extend a process using the information about its actual execution.

How process mining is applied in cross-organizational settings and what issues we identified in such applications are described below.

Process Mining in Cross-Organizational Settings

Process mining is often used to gain data-driven insights into increasing the efficiency of processes in a single organization. For this purpose, there exist several goals that are commonly set when processes mining is used in single organizations. Understanding actual process executions, detecting bottlenecks in processes, finding the process paths that are either frequently or rarely taken, and spotting compliance issues in processes are the goals, to name some of the more popular ones.

However, process mining is principally not limited to use in a single organization. It can help to analyze any processes that: (1) differ within the same organization or (2) differ across multiple organizations, or (3) are distributed over multiple organizations. Accordingly, the use of process mining in such settings is called *cross-organizational process mining*. Due to various reasons (e.g., regulations, location, environmental conditions, etc.), the activity sequence of a business process execution may differ across organizations, or in some branches of the same organization, or for certain products or services offered to customers, or even across organizational units. Each unique activity sequence of the process execution is called *process variant*. Analyzing the differences and commonalities between process variants can enable organizations to learn from each other. For example, by exploiting the best practices that yield better performance, worse-performing organizations can improve their processes.

Horizontal and vertical are the two types of cross-organizational process mining [45, 138]:

- *horizontal*: focuses on the analysis of a process that goes through multiple organizations, i.e., multiple organizations collaborate to handle the instances of a single process. Each organization performs a fragment (like jigsaw puzzle parts) of the same process.
- *vertical*: focuses on the analysis of a process of which executions may differ across multiple organizations, although it serves essentially the same purpose within these organizations. Since such process execution variations can be observed between the organizational units or branches of a single organization, those sub-units can be abstracted as organizations to benefit from vertical cross-organizational process mining in a single organization.

For conducting horizontal cross-organizational process mining, it is required to collect and merge the event logs for a process that passes the boundaries of multiple organizations. As each fragment of that process is the responsibility of a different organization, event logs of the end-to-end process are distributed. Hence, having a merged event log for such processes and analyzing them in-depth are not trivial tasks [142]. Similarly, there are challenges related to vertical cross-organizational process mining [142]. For instance, the granularity of activities and the logging details may differ between organizations for the same process. Importantly, the labels used for activities and their meanings may vary from one organization to another. Furthermore, ensuring privacy is rather difficult in multi-organizational settings.

Despite the importance and relevance of cross-organizational process mining, current cross-organizational process approaches (e.g., [22, 23, 103, 136]) do not provide generic and comprehensive solutions to meet the aforementioned challenges. Considering the deficiencies of those approaches, we identified three main issues, which inspire the work in this dissertation.

Issues

1. **unfairness**: Current approaches do not take organizational characteristics into account in order to determine whether organizations are comparable. This is not fair, especially when organizations are benchmarked based on their performance without considering their specific characteristics (e.g., number of employees, legal form, and location) and context. Moreover, each organization can have expertise in different subjects, e.g., depending on its goals and knowledge. For instance, an ERP software vendor mostly focuses on production,

order management, and invoicing, whereas risk management is typically the main focus of an insurance company. Furthermore, the number of employees of an organization can determine its operational scope, e.g., an organization with only five employees may not handle the same amount of requests compared to an organization that has five hundred employees active in multiple locations.

2. **inaccuracy**: Current approaches generally match the text labels of activities in processes, not the meaning behind them. This leads to inaccurate comparisons. For instance, the meaning of the term *loan* is different for a bank and a library. Moreover, the same term might incur different regulations for a bank than for a library. Making comparisons without considering the meaning of terms and their influence within organizations might generate undesired and ambiguous insights, which may then trigger irrelevant improvement actions and changes in processes.
3. **inadequacy**: Current approaches mostly consider generic process model metrics or event log metrics to provide insights into understanding and improving processes of organizations. Such metrics are often not adequate because organizations make decisions with respect to the KPIs that are specifically of their interest and aligned with their strategic goals.

On the basis of the issues explained above, we specify our considerations to deal with these issues. These considerations provide the foundation for the work in this dissertation.

Considerations

1. **fairness**: The business environment in which an organization is active, i.e., context, and the characteristics of the organization influence both the processes the organization performs and the performance of these processes. For example, there can be a non-profit organization that can be compared with another typical revenue-oriented organization in terms of customer loyalty. However, such a comparison may not yield any insight into improving the processes of a non-profit organization. When context and organizational characteristics are taken into account, the factors that affect processes and their performance, i.e., the selection of organizations by the framework, can be better captured in a sense-making manner. Thus, the conclusions that will be drawn through process mining to provide insights will become fair in cross-organizational settings.
2. **accuracy**: It is reasonable for organizations to use the same terminology for processes having similar purposes. However, the meaning of the terms in a shared terminology may vary for the same process and reflect different concepts in some organizations. For example, the term *delivery* may encompass different activities, e.g., packing, transportation, or customer's receive, in various organizations. Therefore, the meaning behind the terms used in activity labels in processes is essential for accurately detecting the similarities and differences between the processes of multiple organizations.
3. **adequacy**: Organizations are defined as decision factories [34]. Accordingly, they use the KPIs that are relevant for assessing their performance and making decisions. Hence, insights provided through process mining and derived using the common process metrics (e.g., average execution duration) are often not adequate for organizations to make informed decisions. In the case of average execution duration of a process, outliers will highly influence the calculated value, e.g., the number of executions completed in a very short time or executions took an extremely long time. In this regard, to provide adequate insights, similarities and differences between the processes of multiple organizations need to be measured and interpreted using the KPIs relevant for these organizations.

Scope

In accordance with the challenges, issues, and considerations explained above, we set the scope of our framework vertical cross-organizational process mining and focus on the challenges that are related to it. In particular, our aim is to enable organizations to improve performance by benefiting from the similarities in and differences between their processes. To obtain such similarities and differences from process executions and derive recommendations for performance improvement, we place discovery and enhancement process mining tasks at the center of our scope.

1.1 Research Approach

1.1.1 Research Questions

Based on the discussion of the scope, issues (unfairness, inaccuracy, and inadequacy), and considerations (fairness, accuracy, and adequacy), the main research question of this dissertation is stated as follows:

MRQ: *How can cross-organizational process mining be performed in a fair, accurate, and adequate fashion to provide relevant insights for organizations to improve their performance?*

We derived four sub research questions on the basis of the main research question:

RQ1: *How can organizations be selected for fair cross-organizational process mining?*

RQ2: *How can organizations be compared for accurate cross-organizational process mining?*

RQ3: *How can adequate perspectives be determined for cross-organizational process mining?*

RQ4: *How can relevant insights for organizations be provided using cross-organizational process mining?*

It is important to note that the first three sub research questions form the basis for the last sub research question (RQ4).

1.1.2 Research Methods and Techniques

The research in this dissertation is carried out by using several research methods and techniques, namely case study research, focus group, expert interview, literature review, and conceptual modeling. First, these methods and techniques will be explained in this section. Second, their applications per chapter, the related framework components, and the corresponding dissertation theme will be listed (see Table 1.1).

Case Study Research

Case study research is a method that involves the study of a single case or multiple cases to investigate the particularity and complexity of a phenomenon in its real-life context [113, 126]. In case study research, it is essential to conduct a close examination of design artifacts for the purposes of understanding, theory building, and testing. Furthermore, initiative, pragmatism, optimism, and persistence are needed to successfully complete case study research within the field of information systems [35]. In this dissertation research, case studies are performed to build and evaluate the design of artifacts, such as the implementation of the approaches that fulfill the purpose of each component of the framework (e.g., see Chapter 3, Chapter 4, and Chapter 5). The case study protocols that are followed in this dissertation research are the ones defined by Yin [161].

Focus Group

A focus group is a research method that is defined as a moderated discussion on a topic determined by the researcher [95, 113]. It is very useful for understanding the range of opinions of people across groups as there is a much more natural environment than personal interviews [80, 127]. That is achieved by allowing people to interact, influence, and be influenced by others. The questions in focus groups are mostly open-ended but carefully determined. In this dissertation, focus groups composed of experts from the industry were used to evaluate the automated dashboard generation approach (see Chapter 5).

Expert Interview

An expert interview is a type of research method that has been broadly discussed and used in many research areas. It is described as a qualitative interview based on a topical guide that focuses on the knowledge of a field expert [94, 113]. The aim of this method is to explore and collect data from the field experts by means of interactions, which are typically driven by questions. The structure of expert interviews depends on the way of forming questions. More specifically, questions can be very specific and structured, semi-structured, or open-ended. To encourage experts and enrich the discussion with them, semi-structured interviews are mostly preferred in this dissertation research. This method is used almost in all chapters.

Literature Review

A literature review is a research method that is essential for conducting research [32, 77, 123]. It creates a robust foundation for advancing knowledge by identifying relevant theories, methods, and gaps in any discipline [41]. Hence, to collect and assess existing knowledge, every literature review needs to be done accurately. Importantly, having a systematic in doing literature review provides a reliable knowledge acquisition process. Specifying research questions for motivating the literature review, building search criteria, searching existing studies, assessing the collected studies to extract knowledge are the steps that such a process contains [32, 77, 123]. Accordingly, for acquiring the relevant knowledge for the work in this dissertation, the steps and guidelines described in [77] are followed. Related work and background are the primary sections in which the obtained relevant knowledge is explained within the chapters of this dissertation.

Conceptual Modeling

Conceptual modeling is a technique that is often used to denote several aspects of a given domain in a particular form [154]. In general, a graphical representation or a mathematical formulation is used to describe selected phenomena. To do so, entities in the phenomena and the possible relationships between them are determined and expressed [49, 98]. For building a representation both for the framework and its components, this technique is used. Another usage of this technique can be seen in formulating patterns for deriving KPIs (see Chapter 3).

The overview of this dissertation research in terms of the mapping between research questions, research methods and techniques, framework components, and dissertation theme is shown in Table 1.1. In particular, research methods and techniques that have been employed in each chapter and used as the main tool to answer the mapped research question are highlighted using ticks (✓) in black. Similarly, secondary research methods and techniques are denoted using ticks (✓) in gray. In the same table, the framework components column displays the corresponding components of the

Table 1.1: Overview of this dissertation research

Chapter	Research Question	Research Methods and Techniques					Framework Components	Dissertation Theme
		Case Study Research	*Focus Group*	*Expert Interview*	*Conceptual Modeling*	*Literature Review*		
2	MRQ	✓		✓	✓	✓	•All	•All
3	RQ1	✓		✓	✓	✓	•Metric Comparison Catalog	•Performance •Process Mining
4	RQ2	✓	✓	✓		✓	•Metric Comparison Catalog •Metric-Based Improvement Catalog	•Cross-Organizational •Performance
5	RQ3	✓		✓	✓	✓	•Metric-Based Improvement Catalog	•Performance
6	RQ3	✓	✓	✓	✓	✓	•Metric-Based Improvement Catalog	•Performance •Process Mining
7	RQ4	✓	✓	✓	✓	✓	•Metric-Based Improvement Catalog •Cross-Organizational Process Mining Techniques	•Performance •Cross-Organizational •Process Mining

cross-organizational process mining framework for the research described in each chapter. There are three main themes in this dissertation research: *performance*, *process mining*, and *cross-organizational*. Each of these is elaborated at the beginning of this chapter. In this regard, the concerned theme in each chapter is given in the dissertation theme column in Table 1.1. The themes that are the secondary concerns of the research described in a chapter are shown in gray text.

1.2 Contributions

This thesis focuses on providing relevant insights for organizations to improve their performance by learning from each other. Such learning becomes possible for organizations with the Cross-Organizational Process Mining Framework that we developed. The main contributions of this thesis revolve around the five novel approaches employed by the framework. The framework orchestrates three components that provide solutions to the aforementioned three main issues (i.e., unfairness, inaccuracy, and inadequacy) in cross-organizational process mining applications. These components are *Metric Comparison Catalog*, *Metric-Based Improvement Catalog*, and *Cross-Organizational Process Mining Techniques*. The relation between these components and the approaches is shown in Table 1.1.

For addressing the issues explained above, we introduce three particular inputs: *Organizational Context*, *Business Semantics*, and *Metrics Semantics*. In the framework, these inputs are combined with event logs. With the use of Organizational Context, it becomes possible to determine what to compare, i.e., what organizations are comparable. The Business Semantics input is aimed at the determination of how to compare, i.e., how organizations can be compared. Although similar terms are used in the business processes within organizations, their meaning may differ from one organization to another. To capture the actual meaning of the elements in business processes, Business Semantics play an important role. Metrics Semantics is the input that is used to determine which metrics can be compared and how for providing relevant insights into understanding and improving processes in organizations. With this input, it becomes possible for the framework to focus on the performance measurement means, i.e., KPIs that are relevant for organizations while providing insights for them. The framework with its components is shown in Figure 1.2. In Chapter

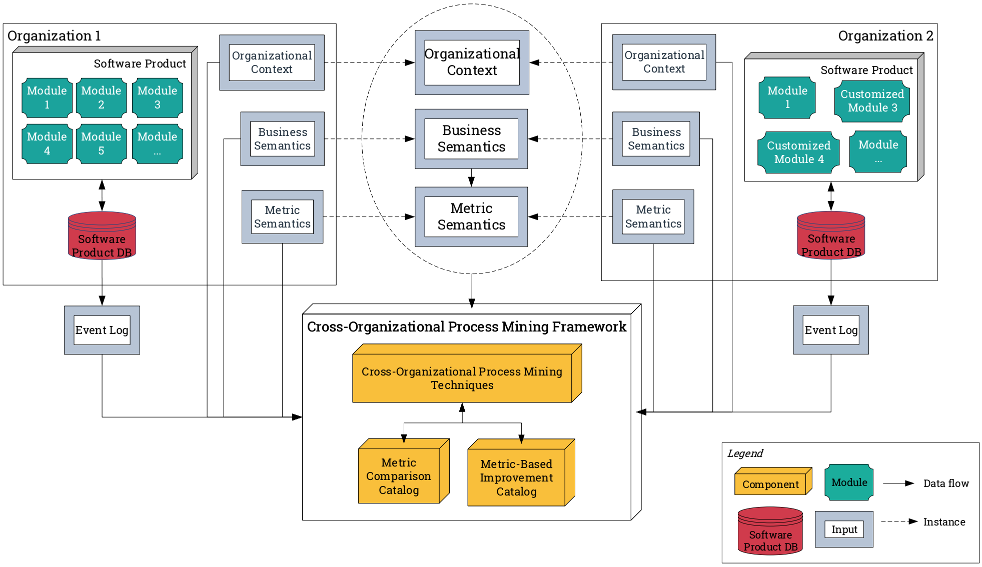

Figure 1.2: Cross-organizational process mining framework

2, we give the details of the framework.

The five novel approaches employed by the framework are as follows:

1. *KPI derivation from Ontological Enterprise Models*: The main focus of this approach is on identifying organizations that are in the same context. In this approach, we tackle the problem of capturing real-world enterprise concepts of organizations. For this, we use ontological enterprise models that allow us to obtain the meaning of enterprise concepts in a machine-readable form. Without requiring any technical knowledge from organizations, custom KPIs to their needs are made available using such machine-readable inputs. With this, a fair comparison of organizations can be performed in cross-organizational process mining. This approach is explained in Chapter 3.

2. *Predicting relevant KPIs for organizations*: This approach focuses on the automated selection of KPIs using machine learning techniques. Our findings show that the relevance of KPIs is linked to the properties of organizations, and further, these links are useful to automate the identification of relevant KPIs. By identifying the KPIs in which organizations have a shared interest, it becomes possible to establish an accurate comparison of those organizations. In Chapter 4, this approach is elaborated.

3. *Automated generation of engaging dashboards*: Finding the perspectives that can be adequate for organizations to learn from each other is the main focus of this approach. Such perspectives are determined by generating the dashboards that organizations will be engaged in. We developed a reusable model that captures the knowledge of dashboard design that is inherent dashboard design principles, which are commonly expressed in the natural language texts. How the approach achieves that goal is given in Chapter 5.

4. *Interactive generation of process performance dashboards*: This approach is an extension of the previous approach. This extension is developed using chatbot technologies to facilitate dashboard generation. In particular, it enables decision makers to interactively build the

process performance dashboards for monitoring the performance of the processes in their organization. With this approach, we tackle the problem of creating engaging dashboards in real-time by providing guidance to users. This approach is explained in Chapter 6.

5. *Building process benchmarks for performance improvement*: This approach is developed to build process benchmarks such that organizations can improve performance by learning from each other using the automatically detected significant differences in their process executions. In this approach, we show what practices applied in organizations affect performance and how such practices can be exploited for improving performance in other organizations. The details of the approach are given in Chapter 7.

How each of these approaches is related to the framework components and dissertation themes can be observed in Table 1.1. As each approach is explained in a separate chapter, using the chapter column in the table, one can see the relations between the approaches and framework and the themes of the dissertation. For example, the approach for KPI derivation from Ontological Enterprise Models is elaborated in Chapter 3, and it is related to the Metric Comparison Catalog component of the framework. Similarly, that approach is associated with the performance and process mining themes of the dissertation.

1.3 Publications

This thesis presents research on the applications of cross-organizational process mining in a fair, accurate, and adequate fashion to provide relevant insights for organizations to improve their performance. Parts of this research led to peer-reviewed scientific publications; 5 conference papers and 1 in progress journal article. The list below provides an overview of them:

- [4] Ü. Aksu, D. M. M. Schunselaar, and H. A. Reijers. A cross-organizational process mining framework for obtaining insights from software products: Accurate comparison challenges. In *2016 IEEE 18th Conference on Business Informatics (CBI)*, pages 153–162, 2016.
- [5] Ü. Aksu, D. M. M. Schunselaar, and H. A. Reijers. An approach for automatically deriving key performance indicators from ontological enterprise models. In *International Symposium on Data-driven Process Discovery and Analysis (SIMPDA)*, pages 38–53, 2017.
- [7] Ü. Aksu, D. M. M. Schunselaar, and H. A. Reijers. Automated prediction of relevant key performance indicators for organizations. In *International Conference on Business Information Systems (BIS)*, pages 283–299, 2019.
  This publication received the best paper award in that conference, i.e., BIS 2019.
- [6] Ü. Aksu, A. del-Río-Ortega, M. Resinas, and H. A. Reijers. An approach for the automated generation of engaging dashboards. In *OTM Confederated International Conferences" On the Move to Meaningful Internet Systems*", pages 363–384, 2019.
- [3] Ü. Aksu and H. A.Reijers. How business process benchmarks enable organizations to improve performance. In *2020 IEEE 24th International Enterprise Distributed Object Computing Conference (EDOC)*, pages 197–208, 2020.
- [8] Ü. Aksu, A. del-Río-Ortega, M. Resinas, and H. A. Reijers. Interactive generation of process performance dashboards. This manuscript is submitted to a journal, 2021.

The author of this thesis has also contributed to research projects beyond the scope of this thesis.

Other publications:

- [18] E. S. Borges, M. Fantinato, Ü. Aksu, H. A. Reijers, and L. H. Thom. Monitoring of non-functional requirements of business processes based on quality of service attributes of web services. In *International Conference on Enterprise Information Systems (ICEIS), Proceedings*, pages 588–595, 2019.
- [29] C. F. Castro, M. Fantinato, Ü. Aksu, H. A. Reijers, and L. H. Thom. Towards a conceptual framework for decomposing non-functional requirements of business process into quality of service attributes. In *International Conference on Enterprise Information Systems (ICEIS(2)), Proceedings*, pages 481–492, 2019.
- [30] C. F. Castro, M. Fantinato, Ü. Aksu, H. A. Reijers, and L. H. Thom. Systematizing the relationship between business processes' and web services' non-functional requirements. In *International Conference on Enterprise Information Systems (ICEIS), Proceedings*, pages 473–497, 2019.
- [115] A. Revina, Ü. Aksu, and V. G. Meister. Method to address complexity in organizations based on a comprehensive overview. *Information*, 12(10):423, 2021.
- [114] A. Revina and Ü. Aksu. Towards a business process complexity analysis framework based on textual data and event logs. In *17th International Conference on Wirtschaftsinformatik (WI 2022), February 21-23, 2022, Nürnberg, Germany, Proceedings*, 2022 in press.

1.4 Thesis Outline

The remainder of this dissertation is divided into the following seven chapters:

Chapter 2 — *Cross-Organizational Process Mining Framework*. This chapter introduces the theoretical framework on cross-organizational process mining. The framework consists of three components: Metric Comparison Catalog, Metric-Based Improvement Catalog, and Cross-Organizational Process Mining Techniques. Each component aims to provide solutions for the issues discussed above (see Subsection Issues). Furthermore, the framework is exemplified both from a software vendor perspective and a client organization perspective.

Chapter 3 — *Automatically Deriving KPIs from Ontological Enterprise Models*. This chapter presents our approach for the automated derivation of tailored KPIs for organizations. In particular, the approach takes an *Ontological Enterprise Model (OEM)* as input to automatically retrieve the knowledge on enterprise concepts in organizations. By means of the KPI derivation patterns expressed in terms of enterprise concepts, the approach generates the KPIs that are customized to each organization. The approach is demonstrated to show its real-life use, and the obtained results are presented.

Chapter 4 — *Predicting Relevant Key Performance Indicators for Organizations*. In this chapter, we present our approach that is aimed at the automated selection of relevant KPIs for organizations. To automate the selection, the approach employs prediction models that are aimed at predicting the relevance of KPIs. Such prediction models are trained for determining the factors that make certain KPIs relevant for particular organizations. The approach is evaluated by checking its prediction quality in a real-life setting. Furthermore, the obtained results that set the approach apart from state-of-the-art are discussed.

Chapter 5 — *Automated Generation of Engaging Dashboards*. This chapter presents our approach for the automated generation of *engaging* dashboards for organizations. The approach employs a decision model for visualizing KPIs. The decision model is developed by collecting and analyzing the dashboard design principles available in the literature. The common usability of the decision model is first tested. Then, it is incorporated into the implementation of the approach. After that, the approach is evaluated in a real-life setting to show its practical value.

Chapter 6 — *Interactive Generation of Process Performance Dashboards*. In this chapter, we present our approach for the interactive generation of process performance dashboards. The approach employs a chatbot that interacts with decision makers to build process performance dashboards by collecting their interests on Process Performance Indicators (PPIs). Moreover, the chatbot enables decision makers to view and decide engaging visualizations of PPIs in real-time.

Chapter 7 — *Performance Improvement through Process Benchmarks*. This chapter presents a holistic approach that is aimed at helping organizations to improve performance by enabling them to learn from each other. In particular, the approach provides process benchmarks that show the significant differences in the executions of processes in multiple organizations. Such differences are determined and highlighted with respect to the KPIs, which are specifically relevant for the organizations involved. Thus, what practices yield better or worse performance in organizations become visible. Organizations can use this information as the basis to take actions that can either increase their performance or prevent situations that may result in an undesired performance. Furthermore, the approach implemented and it is applied in a case study following its performance evaluation.

Chapter 8 — *Conclusion*. In the last chapter, research questions are answered by summarizing the contributions of this dissertation research. Moreover, both research and practical implications are discussed. Finally, we provide a reflection on the limitations of this dissertation research and list the potential research directions as future work, which other researchers pick and work on.

Chapter 2

Cross-Organizational Process Mining Framework

Abstract
Software vendors offer various software products to large numbers of enterprises to support their organization, in particular, Enterprise Resource Planning (ERP) software. Each of these enterprises uses the same product for similar goals, albeit with different processes and configurations. Therefore, software vendors want to obtain insights into how the enterprises use the software product, what the differences are in usage between enterprises, and the reasons behind these differences. Cross-organizational process mining is a possible solution to address these needs, as it aims at comparing enterprises based on their usage.

In this chapter, we present a novel Cross-Organizational Process Mining Framework which takes as input, besides event log, semantics (the meaning of terms in an enterprise), and organizational context (characteristics of an enterprise). The framework provides reasoning capabilities to determine what to compare and how. Besides, the framework enables one to create a catalog of metrics by deducing diagnostics from the usage. By using this catalog, the framework can monitor the (positive) effects of changes on processes. An enterprise operating in a similar context might also benefit from the same changes. To accommodate these *improvement suggestions*, the framework creates an improvement catalog of observed changes. Later, we provide a set of challenges that have to be met in order to obtain the inputs from current products to show the feasibility of the framework. Next to this, we provide preliminary results showing they can be met and illustrate an example application of the framework in cooperation with an ERP software vendor.

keywords- cross-organizational process mining, framework, enterprise resource planning

This chapter is based on the following publication:

[4] Ü. Aksu, D. M. M. Schunselaar, and H. A. Reijers. A cross-organizational process mining framework for obtaining insights from software products: Accurate comparison challenges. In *2016 IEEE 18th Conference on Business Informatics (CBI)*, pages 153–162, 2016.

2.1 Introduction

Software vendors offer various enterprise software products (hereinafter product) to large numbers of *enterprises* (hereinafter organizations) to support them in their business processes. For instance, organizations use Enterprise Resource Planning (ERP) software for order management, delivery, invoicing, stock management, keeping customer records, and human resource management. However, each organization uses the same product for similar goals but different settings. Due to the varying context of each organization, configuration files and workflows need to be modified to adjust the generic software to the needs of each organization [121]. Furthermore, organizations use the same terms and functions in different ways in the same product. For example, the term *loan* has different semantics within a bank compared to a library. In accordance with this information, cross-organizational process mining is a possible solution, which aims at comparing multiple organizations using process mining.

Cross-organizational process mining enables one to compare organizations from different perspectives, e.g., the ordering of activities or subdivision of work amongst resources. Current cross-organizational mining approaches ([22, 23, 136]) mainly use event logs, which are a collection of events as the sole input. At the same time, the current approaches primarily use text matching techniques to relate terms between organizations to find out differences in software product usage. This might hinder accurate comparisons. For instance, in the aforementioned example, the current cross-organizational process mining approaches match the term *loan* inside a bank to the term *loan* inside a library. However, the semantics of the term *loan* is different for a bank and a library. The approaches do not take into account these semantics. Making comparisons without semantics might generate inadequate and ambiguous insights, which may trigger irrelevant improvement actions and changes. Eventually, these irrelevant improvement actions and changes may cause unexpected results. Next to the semantics, the approaches do not consider the organizational context, e.g., the term *loan* might bring different regulations or preconditions for a bank than for a library.

In order to address the aforementioned issues and make a more accurate comparison between organizations, we present a novel Cross-Organizational Process Mining Framework that takes semantics and organizational context, in addition to event logs, as inputs. With these inputs, the framework provides reasoning capabilities that determine what to compare and how. For instance, with the semantics and the organizational context, the framework can determine that the term *loan* has a different semantic for a bank than for a library. By using this semantic difference, the framework can determine that the loan process in the bank is not comparable to the loan process in the library. Moreover, by having clear semantics and the organizational context, we can monitor changes in the processes and the (positive) effects they may bring within a given context. An organization operating in a similar context might also benefit from the same changes. In order to accommodate these *improvement suggestions*, the framework will create a catalog of observed improvements.

Organizations may be sensitive in case of using their data explicitly for comparison. There may be security regulations across the organizations, which may complicate the comparison. Although privacy is an important concern, there are various approaches [25] to ensure the anonymity of the inputs and the results considering the privacy-related concerns. Next to this, specific approaches can be developed to anonymize the semantics and organizational context of an organization.

However, the majority of the software products do not provide semantics and an organizational context that can be used by the framework. For example, *delivery* is a common term used in many software products and what it implies may not be applicable for every organization. Simply put, terms used in software products are not necessarily shared among all organizations [121]. In order to show the feasibility of the framework, we need to provide the inputs the framework expects. To this end, we provide a set of challenges, which have to be met in order to obtain these inputs from current software products. Next to presenting these challenges, we sketched some solutions in meeting the challenges in cooperation with a Dutch ERP software vendor (our industrial partner) for showing these challenges can be met. Later, we give an example application of the framework.

The structure of the chapter is as follows. In Section 2.2, we present the Cross-Organizational Process Mining Framework. In Section 2.3, we show the feasibility of the framework by using the ERP product, which is developed by our industrial partner, and we discuss the challenges by providing the solutions to them in order to be able to show the feasibility of the framework. In Section 2.4, we illustrate an example application of the framework. We conclude the chapter in Section 2.5 and list future work in Section 2.6.

2.2 Cross-Organizational Process Mining Framework

In this section, we introduce the Cross-Organizational Process Mining Framework and discuss its architecture. Figure 2.1 shows the architecture of the framework. The framework consists of three components, Cross-Organizational Process Mining Techniques, Metric Comparison Catalog, and Metric-Based Improvement Catalog. The first one focuses on giving insights into usage commonalities and differences between various organizations in the usage of software products, and execution of their business processes. The second one contains information on how to compare the metrics. The third one enables the framework to determine and store the possible changes that will have positive effects on a metric. This can be used to propose the same changes for other processes in the same context in another organization. In the remaining of this section, we give the details of the components.

2.2.1 Cross-Organizational Process Mining Techniques

The Cross-Organizational Process Mining Framework contains cross-organizational process mining techniques to compare organizations. However, each organization can have expertise on different subjects, e.g., depending on its goals and knowledge. For instance, while an ERP software vendor focuses on order management, human resources, product and stock management, and invoicing, an insurance company focuses on risk management. Beyond this, the number of employees of an organization can determine its operational scope, e.g., an organization with only five employees cannot handle the same amount of issues compared to an organization that has five hundred employees working in different locations. For this reason, the techniques need input that defines the characteristics of an organization to make more accurate comparisons. We name this *Organizational Context*.

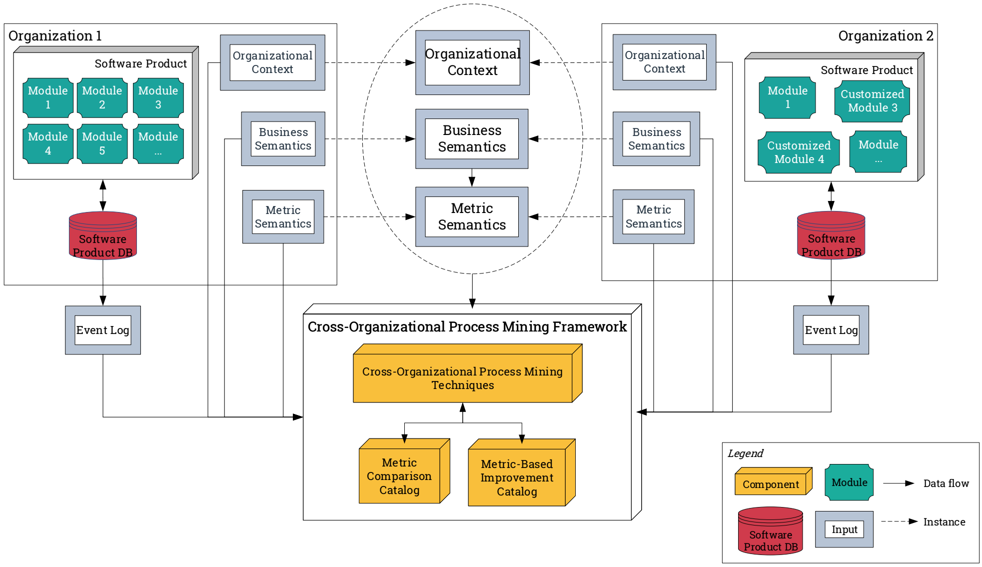

Figure 2.1: Architecture of the cross-organizational process mining framework

The Organizational Context contains information about an organization, e.g., number of employees, specialties, locations, and operated area like space or underground in a country. Further, it enables the techniques to determine the comparability of organizations and perform a fair comparison among them. An organization can operate a process differently under various circumstances. For example, the delivery process of an organization operating in a flat country can have a smaller delivery duration than an organization that operates in a highland area. Moreover, there can be local laws and regulations that affect organizations differently. As a result, while determining the comparability, the techniques need to consider the organizational context.

In addition to the comparability of organizations, the framework needs to determine how to compare organizations such that it can enable organizations to learn from each other. For this purpose, differences in and similarities between organizations in terms of their usage of a software product play an important role. For example, which functionalities are used and how they are used to perform processes in organizations are relevant inputs to spot differences and similarities between them. In this regard, the framework contains cross-organizational process mining techniques. These techniques discover commonalities and differences in usage among organizations. However, the terms inside the commonalities can carry different meanings across organizations. Therefore, the techniques need inputs defining the semantics of the usage, *Business Semantics*, which are common for organizations.

The Business Semantics works like an *ontology* that defines the semantics of things that are common for all organizations. For example, Figure 2.2 shows how two organizations are handling an issue reported by the customer. On the one hand, the first organization checks if the issue is already known or not. In case the problem is known, the organization shares the solution for the known problem and closes it. If the problem is not known, then the organization checks whether it is reproducible or not. If it is not reproducible, then the organization rejects it. On the other hand, the second organization does these two checks in one step called *Evaluate*. The first organization defines *Evaluate* as a two steps activity that starts with *Check Known Problem* and ends with *Reproduce Problem*. Based on this information, the techniques can determine that the semantics are different for the same terms. In addition, in this example, the techniques determine that *Reject* has a different meaning in two organizations, i.e., in the first one, *Reject* reflects a problem that is not reproducible, while in the second one is reflecting the rejection of a problem due to being invalid. As a result, Business Semantics enables the framework to determine comparable things more accurately.

2.2.2 Metric Comparison Catalog

Metrics are required to measure the aspects that will be compared. In our case, organizations will be compared based on their software product usage. Such usage is generally dependent on the strategy and goals of organizations. In other words, how an organization monitors its goals is the key to determine what means are relevant for comparing software product usage. Accordingly, organizations use metrics to monitor their progress in using software products. In this regard, the Metric Comparison Catalog enables one to compare the metrics of the observed software product usage of organizations. For instance, the throughput of two processes in the same organizational context can be compared. In order to know which metrics are comparable, the framework needs semantics for the metrics. We name this *Metric Semantics*. The Metric Semantics is comprised of a metric's context, its definition, how and when to calculate it, and an optional threshold. On the one hand, the metric's context is a reference to the organizational context via the enterprise semantics. Thus, the framework can identify a metric's meaning. In order to compare the metrics, this component contains metric classifiers. A metric classifier gives information about the metric, how to compare it, and information about the operations that must be applied to the metric before comparison. For instance, summation and division operations must be done for an average metric after a unit conversion. On the other hand, the threshold enables the framework to compare deviation boundaries for metrics. In particular, there can be a deviation from predefined boundaries for metrics arising from legal regulations or agreements. For instance, in a country, there can be a legal upper limit for a process' completion time or average response time for a specific step in the process. Let's assume that in a country, there is a legal upper limit for the loan process completion time for banks. By defining this upper limit as a threshold in the framework, both the software vendor developing the software for banks and the banks themselves can benefit from it. The framework can enable the software vendor to do benchmarking based on deviations from the threshold. From the bank's perspective, the framework can show the distance from the threshold among others and how it changes over time. Eventually, the threshold can be a trigger for organizations to take improvement actions.

Figure 2.2 shows an example scenario in which a software vendor wants to obtain insights from two organizations' problem resolution based on a throughput metric *(throughput-resolved metric is defined as the total number of resolved incidents in an hour)*. On the one

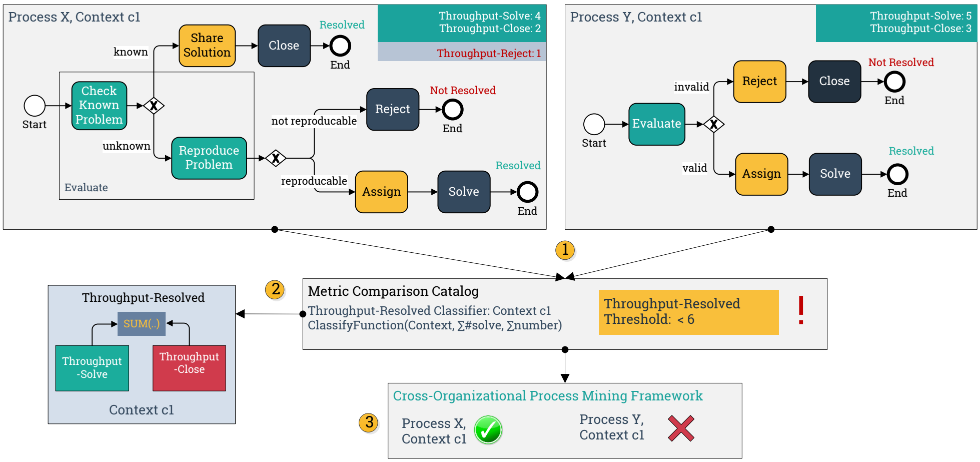

Figure 2.2: An example scenario showing the usage of metric comparison catalog

hand, there is Process X in Context c1, which reflects an incident management process. The process starts with checking the incident, whether it is a known problem or not. If it is a known problem, then the *Share Solution* branch is activated. Otherwise, the *Verifying Reproducibility* branch is executed. On the other hand, there is another process (Process Y), which reflects an incident management process in the same context as the previous. In Process Y, an incident is firstly evaluated, and then it is solved or closed. As shown in the figure, there is a *Close* step in both processes, but the semantic is different. In Process X, *Close* reflects that the incident is *resolved*. Conversely, in Process Y, it means it is *not resolved*. In addition, there is a throughput metric for both solve and close. During the cross-organizational process mining, the Metric Comparison Catalog checks the metrics' context and compares them. In our example, the Metric Comparison Catalog determines that in Process X, the step *Close* has the same semantic from a throughput metric perspective as the step *Solve*. Therefore, the total throughput for resolved in Process X is the summation of throughput-solve and throughput-close ((4 2) 6 incidents).

As shown in the figure, there is a threshold definition for the throughput metric, such that the framework can generate a warning using the threshold definition in order to trigger improvement in real-time. In the example, the threshold value is defined as a minimum number of resolved issues. Based on this, the framework checks if there is any throughput number value for problem resolution less than the threshold. If so, the framework generates a warning. With this, one can be notified in real-time to investigate the process execution. Next, the semantics and organizational context provide a base to create derived metrics. For instance, an organization can have different branches around the world in different cities, countries, and also on continents. Using this information, the framework can enable the Metric Comparison Catalog to generate derived metrics for each level in the organization and compare them with other branches that have a similar organizational context. As a result, this structure of the Metric Comparison Catalog enables one to do real-time process mining as a next step.

2.2.3 Metric-Based Improvement Catalog

As the main focus of the framework is to enable organizations to learn from each other, it should provide recommendations based on the differences between and the similarities in the software product usage of organizations. Thus, it can become possible for organizations to identify what changes may yield better or worse performance in their business. For this purpose, the framework has a Metric-Based Improvement Catalog component aimed at providing improvement proposals. This component stores the information about the changes done in a process and their effects on the metrics. With this information, the framework can determine the possible changes that can be applied to another process that operates in the same context.

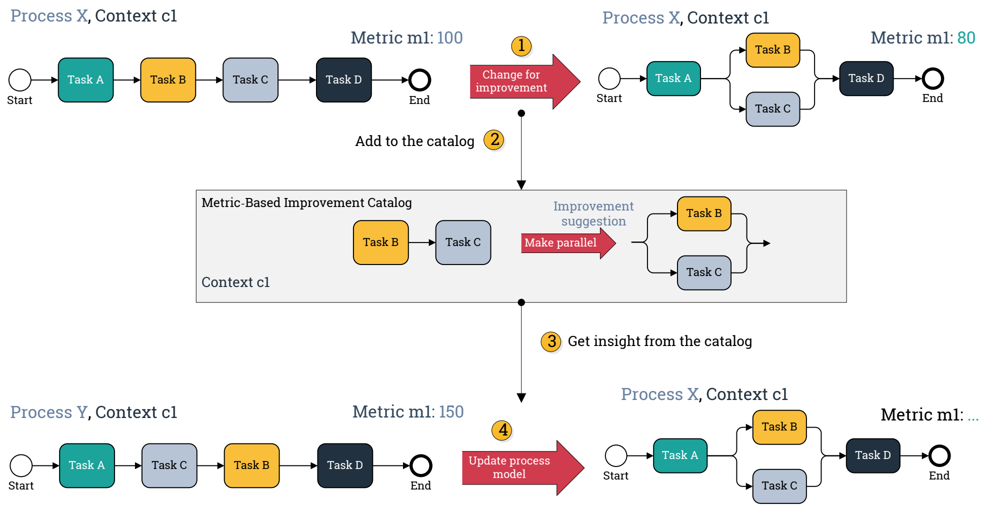

Figure 2.3: An example scenario shows the usage of metric-based improvement catalog

Figure 2.3 shows an example scenario. In the figure, there is Process X in Context c1 that starts with Task A. Task A is followed by B, C, and D. Afterwards, Process X ends. This process has metric *m1* with a value of *100*. After a while, the organization using this process changes its model without changing its goal and context. The organization decides to parallelize tasks B and C (Process XE). After the change, the metric's value improves by *20%* percent (from 100 to 80). The Metric-Based Improvement Catalog stores this information. The framework knows there is another process (Process Y) that is used by a different organization in the same context. Furthermore, the organization uses the same metric for the process, but the process has a different model. With this knowledge, the framework provides the improvement proposal to *parallelize B and C for Process Y*.

In this section, we discussed the Cross-Organizational Process Mining Framework and its components by explaining the opportunities that they bring with the example scenarios. In the following section, we list a set of challenges at transforming a software product's data into the inputs required by the framework. Next, we sketch solutions to meet these challenges and illustrate an example application.

2.3 Challenges and Solutions

As shown by the architecture of the framework in Figure 2.1, the framework needs four inputs; organizational context, business semantics, metric semantics, and event logs. For this reason, we describe challenges that need to be met in order to make the inputs available for the framework. Next, most software products do not provide these inputs in general. Therefore, we sketch solutions for some of the challenges to turn a software product's data into the inputs expected by the framework. Next, to show the applicability of our solutions, we have applied our solutions to the ERP product developed by our industrial partner and used by more than *10.000* customers. In the ERP product, the inputs that are required by the framework are not readily available. Therefore, we follow the steps shown in figure 2.4 to develop solutions for transforming the ERP product's data into the required inputs. In order to show the feasibility of the sketched solutions, we use two sample organizations' data that is provided by our industrial partner.

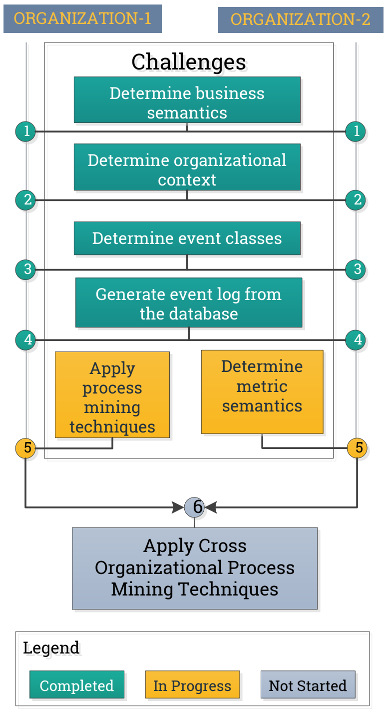

Figure 2.4: Steps to transform product data into framework inputs

2.3.1 Business Semantics

The framework needs clean and unambiguous business semantics of the product and each of its elements. With the semantics, the framework can determine which data is comparable and how it can be compared.

- **Challenge 1.1**: Determine the business semantics of the current product.

Having clean and unambiguous semantics is the most important challenge for us because it is the precondition to gain information and create knowledge. Without understanding the meaning of the data, we cannot use it properly. However, the business logic and semantics behind the elements are currently hard-coded in the application. Next, the current product's database contains partial information related to this. In order to use the semantics, first, we need to uncover them. At the company, we worked with architects to uncover the semantics. Besides this, we talked with product owners and database administrators to discover the business logic inside the current product. As a result of these discussions, we created a *constrained class model* (see Figure 2.5).

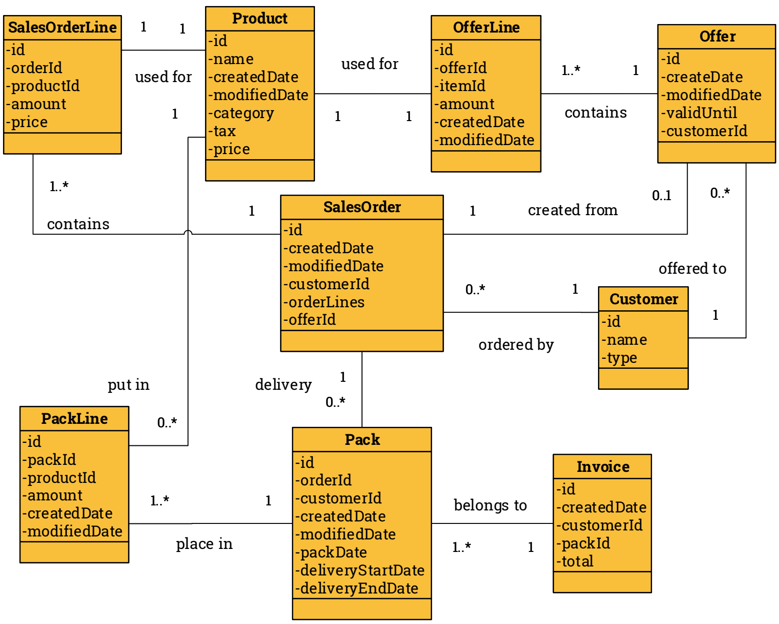

Figure 2.5: Constrained class model of the ERP product

The model shows the elements of the current product's business logic. With this, we were able to get the general semantics for organizations that use the current product. Later, we created a *class model* for each organization to get organization specific semantics.

We mapped one class model to the other (see Figure 2.6) in order to determine which elements are comparable. We did the mapping by following three steps; selection of an element from a model, checking its meaning inside the organization, and matching an element from another model. With this, we were able to give the proper semantics as an input to the framework. For now, we skip non-matching elements, but we will provide solutions in the future.

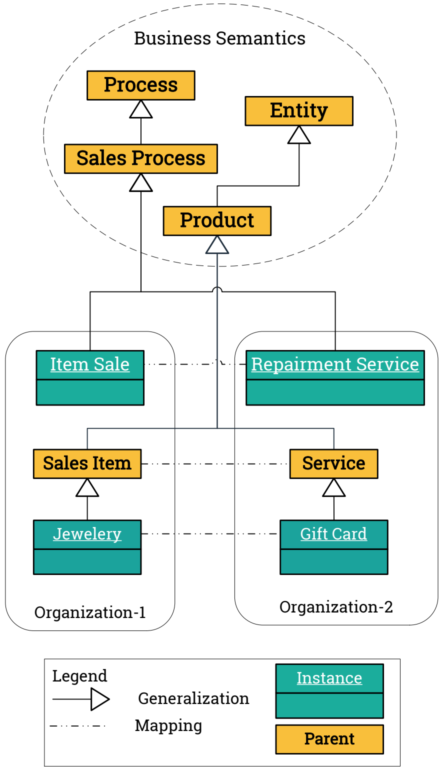

Figure 2.6: Constrained class model mapping using business semantics

2.3.2 Organizational Context

To eliminate mis-comparison of organizations having different characteristics, the framework needs the organizational context as an input. The organizations adjust their processes depending on their characteristics and circumstances arising from these characteristics to reach their goals. In particular, the number of employees, the working domain, legal regulations, and the working environment can be the main reasons for adjusting the processes [121]. Depending on the circumstances of the organization, there can be steps or checks that the organization can skip or merge to gain efficiency when executing a process from a reference library [40]. Next to this, these adjustments affect the process metrics. For instance, there can be a throughput metric that has not in the same value range in two organizations. In addition, there can be time-dependent behavior among organizations. With the organizational context input, the framework can determine how to compare the organizations showing time-dependent behavior.

- **Challenge 2.1**: Determine organizational contexts.

In order to make a fair comparison between organizations, we also need to identify whether they are in the same context. Our industrial partner provided us data that shows which organization uses which modules of the current product, how many entities are stored in specific tables for each organization, and characteristic attributes (e.g., name, location, customer target audience, and number of employees) of each organization. This data was adequate as a starting point to identify how to compare the organizations. In particular, based on the values stored in specific tables, product owners are able to interpret and explain the context of the organizations. For example, invoice types or product types were good indicators to determine what processes are executed and how in the organizations. Based on this information, we discussed with the product owners to select the relevant parts of the data that can help us to compare organizations. Specifically, we performed exploratory data analysis to identify what characteristics of organizations are useful for dividing organizations into groups. At this point, our main consideration was to keep each group as homogeneous as possible so that comparable organizations remain in the same group. In the end, we selected the target customer audience and location as characteristic attributes. Target customer audience attribute value can have one of these values; B2B (Business-to-Business), B2C (Business-to-Customer) or B2G (Business-to-Government). Later, we filtered the organizations based on these attributes to select two of them as the sample for an example application.

2.3.3 Event Log

The event log, a footprint of a process' execution, is the essential input for process mining. An event log is a collection of recorded actions for a particular process instance. Each action belongs to an event class, i.e., event type. For example, a process comprising of create, update, save, and archive steps may have ⟨create, update, save, update, save, update, save, archive⟩ as a possible trace. In this trace, *create*, *update*, *save*, and *archive* represents event classes. However, most software products do not record these actions explicitly. Therefore, it is required to derive an event log from software products' data.

- **Challenge 3.1**: Determine event classes.

The constrained class model that we created while determining business semantics contains relations between objects. These relations represent the interaction between objects in the current product. Next to this, the current product's help document contains the actions one can do while using the product. By using the constrained class model and the help document, we can determine the event classes. The event classes enable one to determine which part of the product's data can be used for event log generation. To this end, we listed possible event classes by using the constrained class model and the help document. For example, one may want to generate event log for the delivery process and can identify the delivery-related event classes from the constrained class model. Thus, all delivery-related events in the data can be retrieved. Later, we discussed with the product owners to select suitable event classes in the list.

- **Challenge 3.2**: Generate an event log from the database.

The current product records particular actions that are done by the user as operations in the database, e.g., item creation, item update, item packing, and item delivery. Furthermore, there are some date columns in the database that store the time that specific action was executed. For example, the update action time is stored in a *modifiedDate* column, the delivery start is stored in *deliveryStartDate* column, and the delivery complete time is stored in *deliveryEndDate* column. Using these date columns, we can create events. However, to do process mining, we need to build the traces that depict the sequence of events. In order to build the traces, we chose an approach that uses Redo logs as our reference to generate an event log [61]. Database schema usage and process instance identification parts of the approach helped us to determine the process identifiers in order to build traces. However, relationships for the current product are not stored at DBMS (Database Management System) level, but they are stored in the database at particular tables. For example, tables related to delivery have no direct relations defined in the database, but these relations are stored in a separate relationships table. Therefore, we developed a software module that reads the relationship information stored at particular tables in the database and creates the database schema. Then, we matched the tables in the schema with the event classes that we created in the previous step. After that, we prepared custom database queries and extracted the event log.

2.3.4 Metric Semantics

In order to know which metrics are comparable and how to compare them, the framework needs semantics for the metrics.

- **Challenge 4.1**: Determine metric semantics.

The current product has business intelligence features for tracking Key Performance Indicators (KPI) that are defined as a standard set. As a starting point, we discussed these KPIs with product owners in order to determine whether they are related to the processes or the operational data, e.g., disk usage and database size. After listing process related metrics, we will determine their contexts and comparison methods in order to compare them in an accurate way.

2.4 An Example Application of the Framework

In this section, we give an example application of the cross-organizational process mining framework. We use the inputs that we transform from an ERP software product's data in the previous section. In the example, we illustrate how the framework generates insights both from a software vendor's (in this example, it is our industrial partner) and an organization's perspective. Figure 2.7 illustrates the interaction between the user and our framework. To obtain insights, first, one needs to have process mining questions. The questions help to determine which cross-organizational process mining techniques can be applied. In this example, we use order management data of two sample organizations (Organization-1 and Organization-2). They are located in different countries in Europe. Both organizations are B2B. While Organization-1 is selling office materials, Organization-2 offers prepaid phone repair services. Both use the same software product, developed by our software vendor, for their sales processes. In the current product, a sales process can start with or without an *offer*. The former continues the same as the latter if the customer accepts the offer.

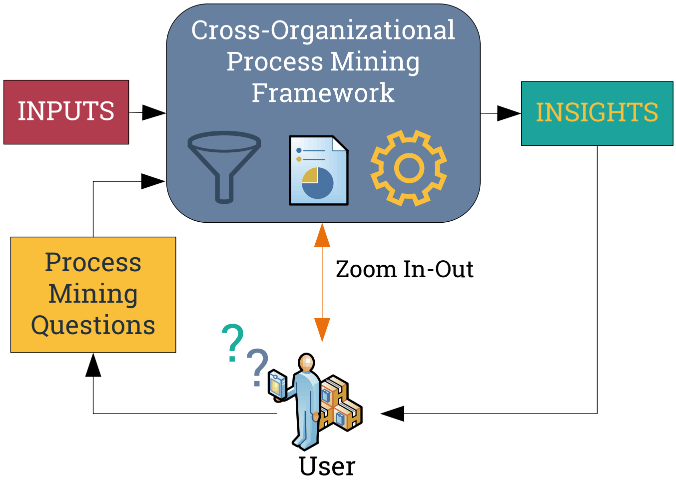

Figure 2.7: An illustration showing the interaction between the user and our framework

2.4.1 Software Vendor Perspective

Here, we explain the interaction between our industrial partner and our framework. As a software vendor, our industrial partner wants to see how organizations use their software product. Regarding this, the software vendor has the following questions.

Process Mining Question 1 (PMQ1): *Does the sales process start with an offer or without an offer in organizations based on an organization's location?*
To generate insights for this question, the framework firstly determines the semantic for the term *sales process*. The framework looks into the Business Semantics and determines what a sales process is and what are the characteristics of the sales process. Then, it checks the event logs if organizations execute a sales process. Furthermore, the framework checks how a sales process can start. After determining the starting points, the framework searches for starting points which are matching with the defined criteria in the question. After the search, the framework shows the generated insight. In this example, the framework generates: "All the companies which are located in Europe start the sales process without an offer.". Regarding the generated insight, the software vendor may provide new process mining questions to zoom in and obtain more specific insights. Let's assume that the software vendor created a new process mining question, namely PMQ2.

PMQ2: *What is the most frequent start activity for the sales process in organizations which are located in Europe?*
During the generation of the insights for this question, the framework determines the meaning of *most frequent* in addition to the steps in the previous question. As a metric, *most frequent* is defined in metric semantics. Therefore, the framework checks the metric semantics and the information about the term, *most frequent*, how and when to calculate it. In this example, the framework determines the meaning of the "most frequent" as a *case frequency metric* and how to calculate it. At the end of the metric calculation, the

framework generates the insight: “The most frequent start activity is CreateOrder. 90% of the cases in Organization-1 and 89% in Organization-2 start with CreateOrder activity.”.

Again, using the obtained insight, the software vendor can create new process mining questions. For instance, the next new question can be “What is the most frequent start and end activities for completed sales processes in organizations which are located in Europe?”. For this question, the framework needs to determine the meaning of the term *completed* for a sales process. Then, the framework can generate insights.

In addition, the software vendor can define a metric using the Metric Comparison Catalog. Let’s assume that the software vendor wants to compare the most frequent path in sales processes across organizations. The software vendor wants to see the result as a similarity metric (is defined by the vendor as the ratio of total common activity types inside the most frequent path of a sales process to the total activity types inside the most frequent path’s of all sales processes). Based on the similarity metric, the framework checks activity types that have the same semantic. The framework determines that only *NewOrder* and *CreateOrder* activities have the same semantic. Based on this, the framework generates the insight: “The similarity is 1/5 for the most frequent variant of the sales processes. The most frequent variant for Organization-1 is: 58%, [CreateOrder, PreparePack, Deliver] and the most frequent variant for Organization-2 is: 66%, [NewOrder, CheckStock, CreateInvoice]”.

2.4.2 Client Organization Perspective

A client organization also can obtain insights and benefit from cross-organizational process mining. For instance, an organization that wants to compare itself with others can obtain insights that can help to determine how to improve its processes. Let’s assume that an organization wants to compare itself with others in terms of process completion time, in particular, the mean case duration. To this end, the organization defines a process mining question, *What is the difference between our organization and the others based on the average duration of performing a sales process instance?*.

To generate insights for this question, the framework needs to determine the meaning of *average*. Let’s assume that average is by default defined in the framework as *the difference from the mean case duration metric*. With this metric definition, the framework generates an insight, “A sales process instance is performed on average 40.25 hours in all organizations. Your organization performs a sales process by spending 70% more time than the average.”. As we discussed before, based on this insight, one can create more specific process mining questions. For example, the organization may also want to compare itself with others based on location, number of employees, or based on another characteristic.

The examples from a software vendor’s perspective and an organization’s perspective indicate that one can obtain more specific insights by executing more interaction cycles with our framework. In each cycle, one can define more granular process mining questions using organizational contexts, metric semantics, or business semantics.

2.5 Conclusion

In this chapter, we present a generic Cross-Organizational Process Mining Framework that is aimed at comparing organizations based on the usage of a software product. The framework uses organizational context, business semantics, and metric semantics, apart from current approaches ([22, 23, 136]), in order to compare the organizations more accurately. As processes are not static, i.e., they have to be adapted to changes [40], they will evolve and vary over time. Such evolution often results in differences in the performance of processes. Therefore, such changes on processes need to be captured appropriately. In this regard, the inputs allow the framework to monitor the *concept drift*[1], i.e., the same process variant may operate differently under various circumstances, possibly depending on the season. For example, there may be seasonal or environmental circumstances affecting the features of a process. The same process may have longer execution times in summer than winter. Also, the same process may have less delivery duration at in flat area than in a highland area. Furthermore, the framework has a Metric Comparison Catalog component that enables the framework to determine how to compare the metrics. Moreover, by having clear semantics and the organizational context, the framework can track changes in the processes and their (positive) effects. An organization operating in a similar context might benefit from the same changes. In order to accommodate these *improvement suggestions*, the framework creates a Metric-Based Improvement Catalog of observed suggestions.

In order to be able to show the feasibility of the framework, we used the ERP product developed by our industrial partner. We first checked the availability of the inputs which our framework needs. In the product, the inputs were not readily available. Therefore, we listed challenges how to transform the product's data into the inputs required by the framework. Then, to meet these challenges, we sketched solutions that can be applied for any other products used by other organizations. On the one hand, the challenges related to the semantics can be resolved with the help of experts who are directly involved in product development. On the other hand, the challenges related to the event log extraction from a database can be met by extending different approaches ([27, 59, 61, 87, 110, 152]). However, there is still not a generic solution that can be applied to any kind of database which has no redo logs or just reflecting the current state of the data.

In addition, we gave an example application of the framework both from a software vendor and an organizational perspective. The example shows how a software vendor and an organization can benefit from the framework in order to obtain more insights. In the example, we also discussed the user interaction with our framework.

Other software vendors, who focus on comparing processes within different organizations, can apply our framework to their products by meeting the challenges that we presented. Next to this, organizations can benefit from by comparing themselves with other organizations. In particular, the Metric-Based Improvement Catalog provides improvement suggestions to the organizations by capturing changes inside other organizations that have positive effects on the metrics.

[1] In machine learning, concept drift means that a model loses its accuracy in predicting a target variable over time due to unforeseen changes in the statistical properties of that variable. In the context of process mining , instead of a variable, the complete process is investigated. This makes it complicated to define the notion.

2.6 Future Work

Cross-Organizational Process Mining is a novel and emerging area. However, there are related approaches to cross-organizational process mining from that we can benefit. These approaches are mainly focusing on process model similarity and comparison based on syntax. The approach in [140] presents process equivalence, including *fitness*[2] and *precision*[3] notions. This approach can help us to develop a generic method to determine the similarity of process models among different organizations.

In addition, the approaches listed in [21] and [24] are beneficial to sketch a solution to make accurate comparisons. Moreover, the meta-model presented in [37] uses a semantic mapping that allows the design-time analysis of process performance indicators. Based on this, one can develop enhanced analysis techniques. And also, developed techniques can be integrated with machine learning techniques in order to propose better improvement suggestions.

Furthermore, the case study explained in [24] uses the Process Tree approach, presented in [120], to illustrate the process model comparison from a control-flow perspective. We can benefit from the Process Tree approach in order to visualize and emphasize the differences and commonalities between processes. Furthermore, we can extend this approach to reflect performance metrics that we are going to discover from the event log.

[2]The discovered model should allow for the behavior seen in the event log [137], i.e., a measure reflects how much the model can reproduce the traces in the event log.

[3]The discovered model should not allow for behavior that is completely unrelated to what was seen in the event log [137].

Chapter 3

Automatically Deriving KPIs from Ontological Enterprise Models

Abstract
Organizations use Key Performance Indicators (KPIs) to monitor whether they attain their goals. Software vendors that supply generic software provide predefined KPIs in their software products for these organizations. However, each organization wants KPIs to be tailored to its specific goals. Therefore, software vendors spend significant efforts on tailoring KPIs to organizations. That tailoring process is time-consuming and costly due to differences in the real-world phenomena of these organizations. In this context, we present our novel Automated KPI Derivation Approach. To automate the derivation of KPIs, our approach obtains the exact meaning of the terms in the real-world phenomena of an organization that is modeled in the form Ontological Enterprise Models (OEMs). As a proof-of-concept we implemented our approach. We demonstrate its use in a real-life setting and present preliminary results.

keywords- key performance indicators, ontological enterprise modeling, enterprise resource planning

This chapter is based on the following publication:

[5] Ü. Aksu, D. M. M. Schunselaar, and H. A. Reijers. An approach for automatically deriving key performance indicators from ontological enterprise models. In *International Symposium on Data-driven Process Discovery and Analysis (SIMPDA)*, pages 38–53, 2017.

3.1 Introduction

Organizations use generic software, e.g., Enterprise Resource Planning (ERP) software, to support their business processes. Within this software, measuring the performance of business processes is essential for organizations [101] while they are progressing towards their goals. To this end, organizations use Key Performance Indicators (KPIs). For instance, *average duration of product delivery* is a KPI that organizations use to monitor their product delivery processes. By tracking this KPI, organizations can predict how much staff must be assigned to their product delivery processes to keep the duration of a product delivery below a certain threshold, e.g., on average 3 days.

Software vendors that supply generic software products typically provide predefined KPIs for organizations. However, predefined KPIs will not work successfully in all organizations because organizations want KPIs tailored to their specific organizational goals [101]. For instance, while *average duration of product delivery* is a relevant KPI for an organization that operates in the retail domain, for a library, it may not be relevant. Currently, to deal with providing tailored KPIs, software vendors either customize their software products for each organization based on the KPI definitions of organizations or they include Business Intelligence (BI) functionality into their software products to let organizations design custom KPIs and dashboards. However, both of these solutions require a significant effort of software vendors and organizations [101]. To this end, a large number of approaches have been proposed for defining, modeling, and customizing KPIs [23, 37, 71, 104, 105, 109]. These approaches either provide a set of KPIs with the assumption that they will be directly used by numerous organizations or introduce a structure for defining and modeling KPIs such that organizations can customize KPIs for themselves. However, if an organization wants to use the KPIs that are defined using one of these approaches, then, first, the organization needs to obtain the knowledge on the exact definitions of these KPIs, and secondly, the organization needs to identify the related activities in its business processes to calculate the value of these KPIs. For example, *The percentage of on-time product delivery* can be defined as a KPI via these approaches to monitor a product delivery process of various organizations, e.g., retailers, rental agencies, and banks. On the one hand, organizations need to manually determine what are the related real-world business activities in their product delivery process to derive that KPI, e.g., packing, shipping, and physical delivery. On the other hand, organizations need to identify the operations that are required to calculate the value of that KPI. All but one [37] do not provide a solution for determining the related activities and operations in business processes to derive KPIs and calculate their values. To determine the related activities and operations in business processes for deriving a KPI and calculating its value, del Río-Ortega et al. use annotations [37] in business processes as a solution. But manually annotating the business processes in organizations requires a significant effort due to the required technical knowledge from organizations. Moreover, the term *delivery* in the same KPI, *The percentage of on-time product delivery* might indicate different concepts for a retailer and a rental agency. For a retailer, it might indicate a set of activities that end when the customer receives the product and becomes the owner of it. But the same term in a rental agency might indicate an activity that results with the delivery of the key of a house to a customer where the customer neither becomes the owner of the house nor the key of the house. However, these approaches do not consider the differences in the meanings of real-world business activities in the business processes in organizations. Therefore, due to the deficiencies of these approaches that we exemplify above, tailoring KPIs to organizations using these approaches becomes time-consuming, costly, and error-prone ①.

Within this chapter, we present a novel approach to automatically derive tailored KPIs for organizations. In organizations, mostly end-users define the requirements for tailoring KPIs who often do not consider technical implementation details. ② End-users express these requirements in terms of real-world enterprise concepts such as business processes, the activities in business processes, products, or services. ③ Although software vendors cope with tailoring KPIs by developing custom-made solutions for organizations, customization possibilities are limited and require technical knowledge from organizations.

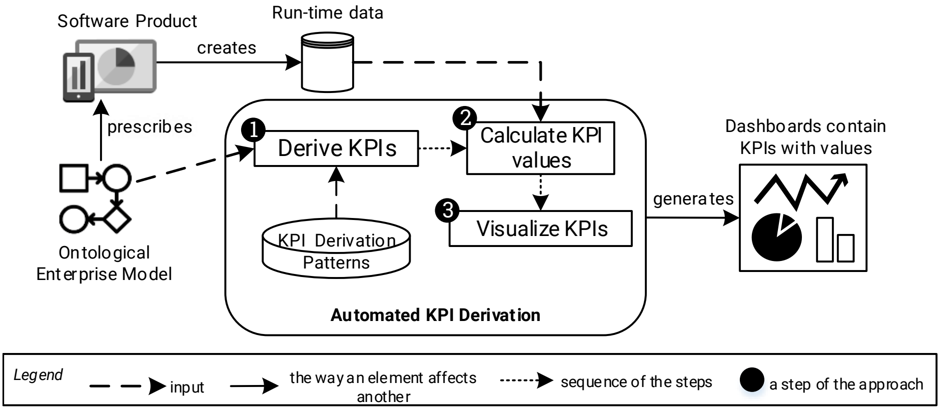

Figure 3.1: Our automated KPI derivation approach

To meet the aforementioned challenges (①, ②,and ③), we develop a novel approach that derives tailored KPIs for organizations using the knowledge on real-world enterprise concepts in organizations. To automatically obtain that knowledge, our approach takes an *Ontological Enterprise Model (OEM)* as input. Using an OEM, our approach derives KPIs automatically by means of KPI derivation patterns (see ❶ in Figure 3.1). A KPI derivation pattern denotes the criteria when and how KPIs will be derived. If our approach determines that a given OEM satisfies the criteria in a KPI derivation pattern, it will derive KPIs with respect to these criteria. Subsequently, our approach calculates the values of the derived KPIs (see ❷ in Figure 3.1) using the *run-time data* for that given OEM. Finally, our approach visualizes the derived KPIs (see ❸ in Figure 3.1) in the form of dashboards. In short, our approach provides tailored KPIs to organizations by automatically deriving the KPIs using real-world enterprise concepts of these organizations without requiring any technical knowledge from them. To validate our approach, we apply it in a case study; we will present preliminary results in this context.

The chapter is structured as follows. We describe our research approach, which is an instance of the design-science paradigm, in Section 3.2. After that, in Section 3.3, we present our approach aimed at automatically deriving KPIs from an OEM by means of KPI derivation patterns. In Section 3.4, we describe the case study context where we validate our approach. In Section 3.5, we describe how we define KPI derivation patterns to derive

KPIs in real-world phenomena of organizations. Afterwards, in Section 3.6, we show the applicability of our approach. In Section 3.7, we provide an overview of related work on deriving KPIs and tailoring them to organizations. Finally, in Section 3.8, we present our conclusions and the potential directions for future work.

3.2 Research Method

In this research, we propose a novel approach to automatically derive tailored KPIs for organizations from an OEM by means of KPI derivation patterns. By applying the design-science methodology, we develop the design artifact of our approach. Then, we evaluate it on a case study at a Dutch ERP software vendor by applying the four case study techniques that are defined by Yin [161] as sources of case study evidence. In the following paragraph, we elaborate on how we applied these techniques.

Firstly, in the beginning of this research, a training session is conducted to get familiar with an OEM language (called NEXT OEM Language) that the case study company is developing to model an organization's business processes in the form of OEMs [144]. Secondly, to see how real-world phenomena, essential to an organization, are modeled in the form of OEMs, we examined a set of OEMs that are modeled by the case study company. During that examination, we modeled a sample OEM and observed the generated software from that OEM to see how real-world phenomena essential to an organization are realized. Thirdly, to elicit the knowledge on the exact definitions of KPIs, which are offered by the case study company in its current ERP software product, we reverse engineered and analyzed the documentation of that ERP software product. To verify the knowledge on KPIs that we obtained, we conducted unsupervised interviews with the dashboard team of the case study company that maintains the KPIs, which are offered in the current ERP software product of the case study company. Finally, based on that knowledge, we defined KPI derivation patterns via the NEXT OEM Language such that one should be able to derive KPIs when a given OEM matches the criteria expressed in these patterns. These KPI derivation patterns were reviewed and evaluated by the dashboard team of the case study company and also by the architects who work on defining the NEXT OEM Language at the case study company. In the following section, we explain how our approach derives KPIs automatically by means of KPI derivation patterns.

3.3 Automated KPI Derivation from an Ontological Enterprise Model

In this section, we explain the details of our Automated KPI Derivation Approach. Which inputs our approach takes and in which steps these inputs are used are depicted in Figure 3.1. In short, our approach takes an OEM for deriving tailored KPIs for organizations using KPI derivation patterns. Then, it calculates the values of the derived KPIs using the run-time data for the given OEM and visualizes the KPIs in the form of dashboards. To determine *when* and *how* KPIs can be derived, our approach contains KPI derivation patterns.

A KPI derivation pattern is composed of two parts: *when* and *how*. The *when* part of a KPI derivation pattern specifies the criteria that a given OEM needs to meet for deriving a set of KPIs. The actions that need to be applied while deriving a set of KPIs are expressed in the *how* part of a KPI derivation pattern. For the representation of KPI derivation patterns, we use Condition-Action (CA) rules [88]. A CA rule is activated when its condition becomes true, i.e., if the criteria expressed in the *when* part of a KPI derivation is fulfilled, then the *how* part of it will be executed.

Definition 4.1. A KPI derivation pattern is encoded in terms of a Condition-Action rule (c, a), where **c** is an expression that returns a value of true or false for an element of a given OEM. **a**=a1,...,an is a list of actions (an action can be included multiple times) that will be applied while deriving a set of KPIs from a KPI derivation pattern. If **c** returns true for an element of a given OEM, a KPI will be derived from the given KPI derivation pattern. Then, the actions in the list **a** will be applied on the derived KPI to calculate its value using the run-time data of the given OEM. The last action returns the calculated value.

Step 1-Derive KPIs: In this step, our approach derives a set of KPIs when a given OEM (the precondition for our approach) satisfies the criteria expressed in the *when* part of a KPI derivation pattern. Then, our approach names the derived KPIs with respect to a *naming rule,* which corresponds to a particular KPI derivation pattern. A naming rule prescribes how the KPIs that are derived from a particular KPI derivation pattern will be named, i.e., which elements in the given OEM will be used for naming and how. For example, a set of derived KPIs will be named by the concatenation of the following texts: "total value of" and the text in the name of a particular OEM element that is contained in the when part of a specific KPI derivation pattern. Below, in Algorithm 1, we show how our approach derives KPIs.

Algorithm 1: KPI derivation. The algorithm derives KPIs from a given *OEM* OEM^G by means of *KPI derivation patterns* KPI^P and names the derived KPIs KPI^D with respect to *naming rules NR*.

forall ***KPI pattern*** $kp \in KPI^P$ **do**
    $c \leftarrow$ GETWHENPARTOFPATTERN(kp)
    $r \leftarrow$ GETNAMINGRULE(kp, NR)
    forall ***OEM element*** $e \in OEM^G$ ***that meets the critera expressed in*** c **do**
        $k \leftarrow$ CREATEKPI(kp, e)
        APPLYNAMINGONKPI(k, r)
        add created KPI k into KPI^D
    end
end

Step 2-Calculate KPI Values: The run-time data for a given OEM contain references to corresponding elements in an OEM. Using the references, one can obtain the related run-time data for a particular OEM element. In this step, to calculate the value of a derived KPI, our approach gets the referenced OEM element in the derived KPI. After that, our approach gets the actions (defined in the *how* part of the referenced KPI derivation pattern) that are required to calculate the value of the derived KPI. Then, our approach obtains the related run-time data for the referenced OEM element using the function named *getInstancesOf*. This function returns the related run-time data for a given OEM element from the run-time data of the given OEM using the references from the run-time data to the corresponding elements in the given OEM. If there is no related run-time data for the referenced OEM element, our approach indicates that KPI by a *no-value* marker. The calculation of the KPI values is depicted in Algorithm 2.

Algorithm 2: KPI value calculation. This algorithm calculates the values of the derived KPIs KPI^D using the run-time data OEM^R of OEM^G that is used at deriving KPIs.

forall ***Derived KPI*** $k \in KPI^D$ **do**
  $e \leftarrow$ getReferencedOEMElement(k)
  $kp \leftarrow$ getReferencedKPIDerivationPattern(k)
  $Acts \leftarrow$ getHowPartOfPattern(kp)
  $eIns \leftarrow$ getInstancesOf(e, OEM^R)
  $v \leftarrow$ calculateKPIValue($eIns$, $Acts$)
  setKPIValue(k, v)
  add KPI k into KPI^{DV}
end

Step 3-Visualize KPIs: Derived KPIs with their values will be visualized using *visualization rules* and *filtering rules*. *A visualization rule* prescribes the presentation style for derived KPIs, i.e., which visual element placement scheme, chart type, or coloring scheme needs to be used. Based on which dimensions derived KPIs can be filtered is prescribed by *Filtering rules*. To show a set of derived KPIs from different perspectives, KPI dimensions can be combined with these KPIs. Related to that, our approach shows derived KPIs from a time perspective.

In this chapter, we presented our Automated KPI Derivation Approach that is a part of our Cross-Organizational Process Mining Framework, which we introduced in [4]. We considered the implementation of our approach[1] as a plug-in of *ProM* [141], which is an extensible process mining framework that supports a large number of process mining techniques by means of plug-ins. Our considerations for *ProM* are two-fold: (1) *ProM* handles the import and export of run-time data, (2) we aim to benefit from a wide variety of process mining techniques in the future while developing our framework. We have implemented our approach (except some filtering functionality). We can load an OEM with its run-time data and show the dashboards for the automatically derived KPIs. In the following section, we define the context of the case study that we conduct to evaluate our approach.

3.4 Case Study Context

As we introduced in Section 3.1, our Automated KPI Derivation Approach automatically derives KPIs from OEM. In the case study that we conduct, an OEM was provided by the case study company. The case study company is developing a novel model-driven software generation approach [144] to aid automatically generating ERP software from a model. As part of NEXT, a declarative modeling language, namely the NEXT OEM Language is being developed by the case study company. This language will contain modeling elements

[1]The implementation of our Automated KPI Derivation Approach is available at `https://amuse-project.org/software/`

such that they will reveal and reflect the meaning of the real-world enterprise concepts of an organization (e.g., business processes, the activities in business processes, products, or services) and the interaction between these concepts.

The NEXT OEM Language provides a holistic perspective (i.e., not separate perspectives for data, interactions, time, and conditions as in the UML) for modeling these concepts in an organization. Thus, the information inside an organization, which is required to generate the tailored ERP software for that organization can be captured by means of the model, which is created via the NEXT OEM Language. Moreover, the language is aimed at modeling that required information in the form of OEMs. Since an OEM is a precondition for our approach, we use the NEXT OEM Language to define KPI derivation patterns. Thus, we can determine whether the given OEM meets the criteria expressed in these KPI derivation patterns. Although the NEXT OEM Language is under development, it does not bring any limitations to our approach because it provides sufficient elements for deriving KPIs from the given OEM by applying our approach. In this context, we introduce the case study company that provides the OEM for our approach. Then, we describe the relevant aspects of the NEXT OEM Language for our approach while expressing KPI derivation patterns in terms of the elements of the given OEM, which is created using that language.

Company Identification: Our case study company is an ERP software vendor that develops and distributes its ERP software product (called MyERPSuite). MyERPSuite is used by more than 1.2 million end-users of 10,000 customers from various domains, e.g., retail, production, and accountancy. Furthermore, to provide relevant insights for these customers, the dashboard team of the case study company created a rich set of dashboards in MyERPSuite based on customer experience and expert knowledge. These dashboards consist of numerous KPIs for various areas such as sales, purchase, finance, and product development.

NEXT OEM Language: We start with a simple OEM to familiarize the reader with the NEXT OEM Language. We suppose there is an organization called NRetailCorp, which operates in the retail domain. Its main business is selling products to its customers. Figure 3.2 depicts a simple OEM that NRetailCorp created to model its business. As shown in Figure 3.2, NRetailCorp sometimes makes sales offers ❶ to its customers. It is also possible that a customer may order ❷ products without a sales offer being issued. After a sales order has been placed by a customer, NRetailCorp delivers ❸ the products to the customer. Subsequently, NRetailCorp creates a sales invoice ❹ and sends it to the customer to inform the customer.

To whom NRetailCorp delivers are denoted with the *person* and *organization* blocks in the OEM (see Figure 3.2). From the perspective of NRetailCorp, these blocks perform the role of *customer*, which can be either people or organizations in real-world. In addition, the products that NRetailCorp delivers are denoted by the block that is *good*. The type of the good, person, and organization blocks is **entity**. An entity represents a specific thing that exists independently in the real world, e.g., a person, an organization, a service, time, or goods. In the NEXT OEM Language, the difference between entities is captured by their type, which can have a value of organization, person, service, time, or good.

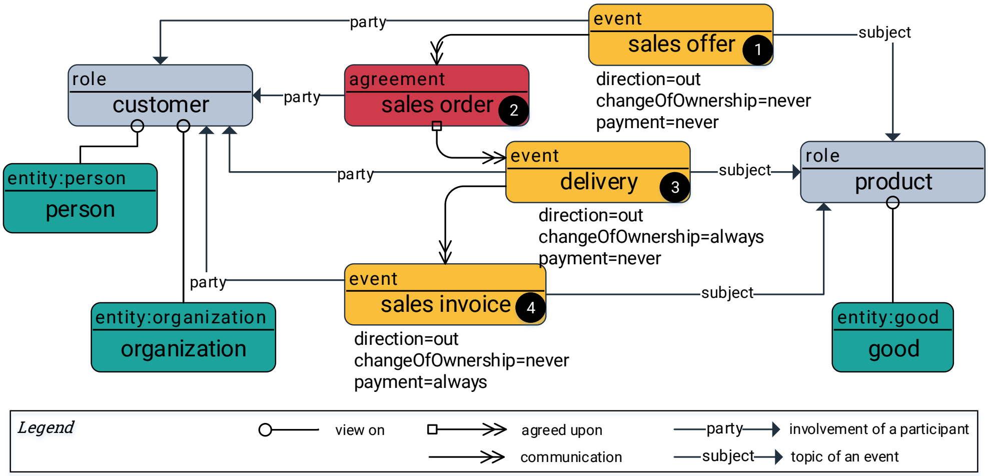

Figure 3.2: A simple OEM created by NRetailCorp to model its business

The customer (a view on the *person* or *organization* block) and product blocks (a view on the *good* block) are both of type **role**. Roles are views on entities that describe in what way NRetailCorp considers the entities. For example, a person might be both a customer and a supplier of NRetailCorp.

Real-world business activities (e.g., making a sales offer, delivering products, or invoicing) that are relevant for NRetailCorp to administer are denoted by **event** blocks (see ❶, ❸, and ❹ in Figure 3.2). Since not a single organization is exactly identical to another, these activities vary from one organization to another. Therefore, to capture the variability between organizations, each event block has a set of *characteristics*. For instance, the interactions between NRetailCorp and external entities and what the interactions are about are encoded in the *party* and *subject* characteristics, respectively. Something of value that leaves NRetailCorp is indicated by setting the *direction* characteristic to *out*. Using the *changeOfOwnership* characteristic, the delivery event indicates whether the good is sold or rented out (see ❸ in Figure 3.2). The necessity of payment at an event is encoded in the *payment* characteristic: as the delivery of products is not free of charge, the *payment* characteristic of an invoicing event is set to *always* (see ❹ in Figure 3.2).

As mentioned above, events are used to abstract real-world business activities in an organization. However, an event is stateless and timeless while modeling, i.e., at design-time, it will be executed at a specific moment in time and will move into a state. Therefore, the state of an event is encoded in the *lifecycle* characteristic, which will get a value (*done* or *todo*) at its execution. Moreover, time related information of an event, which will get values during execution, are encoded in the following characteristics: *plannedStartDate*, *plannedEndDate*, *actualStartDate*, and *actualEndDate*.

An **agreement** (see ❷ in Figure 3.2) expresses an obligation between NRetailCorp and a customer to make certain things happen for the sale of products: goods have to be delivered to the customer. The edges between agreements and events indicate the events which can be performed based on an agreement. Moreover, the edges between events indicate the communication between real-world business activities. Using edges, one can traverse between events and agreements both at design time and run-time. To this end, the event has the following characteristics: *incomingConnectedEvents*, *outgoingConnectedEvents*, *incomingConnectedAgreements*, and *outgoingConnectedAgreements*. The contents of the agreement are represented by the events that can be performed based on this agreement, e.g., delivery of the products for NRetailCorp. The monetary value of the contents of the agreement is encoded in the *value* characteristic of the agreement. The *value* of an agreement will be available at run-time.

Modelers and end-users in organizations are not necessarily familiar with technical concepts such as database management systems, programming languages, or development environments. By using minimal but sufficient elements with their characteristics, the NEXT OEM Language enables modelers and end-users to reason and model real-world phenomena, essential to organizations. Thus, a modeler or an end-user can easily comprehend OEMs independently from technical details and limitations. How could it be generalized to any process of the same organizations. Moreover, by being independent of technical details, OEM becomes reusable for various organizations, e.g., the OEM that NRetailCorp created (see Figure 3.2) can be applicable also for other organizations in the retail domain, real-estate agencies, and construction companies. Since the language is under development, over time, more and more concepts will be added to the language by the case study company to capture much more real-world phenomena in OEMs, e.g., planning, production, accounting, and resource management. In the following section, we define KPI derivation patterns using the NEXT OEM Language.

3.5 KPI Derivation Patterns

In Section 3.3, we presented our approach that is aimed at automatically deriving KPIs from an OEM by means of KPI derivation patterns. To enable our approach to determine whether a given OEM meets any criteria for deriving KPIs, we select a set of KPIs in real-world phenomena essential to organizations and define KPI derivation patterns in terms of OEM elements based on the similarities of the selected KPIs. To do so, first, we obtain the meanings of the selected KPIs and the operations required to calculate the values of these KPIs. Secondly, with this knowledge, we formulate the selected KPIs using the NEXT OEM Language. Finally, we analyze the similarities and differences between these formalizations. Accordingly, in the following first subsection, we describe how we obtain the meanings and calculation logic of the selected KPIs and formulate these KPIs using the NEXT OEM Language. In the second subsection, we explain how we define KPI derivation patterns using these formalizations.

3.5.1 Obtaining the Knowledge on KPIs

As a set of KPIs in real-world phenomena essential to organizations, we took the KPIs that are offered by the case study company to its customers in its current ERP product (MyERPSuite). The majority of the KPIs that are frequently used by a significant amount of customers are located in the dashboards for Sales, Purchase, and HRM. Therefore, together with the dashboard team, we selected a set of KPIs from these areas. Afterwards, we acquired the knowledge on the exact definition of these KPIs and the calculation of their values. With this knowledge, we formalized each KPI using the NEXT OEM Language. Below, we list the formalizations of the KPIs mostly used from the dashboard for the Sales area.

KPI-Total value of sales orders: In this KPI, sales orders are *agreements* between an organization and its customers for the delivery of products. To derive this KPI, we need to determine the *delivery events* (see ❸ in Figure 3.2) that are in the context of agreements (see ❷ in Figure 3.2) and not yet executed, i.e., the *lifecycle* characteristic has the value *todo*. Then, using the edges between the deliveries and agreements, we can determine the sales orders. Afterwards, we can derive this KPI and calculate its value by summing up the values of the determined agreements. In the formalizations, we use **EVENTS**, **AGREEMENTS**, **ROLES**, and **ENTITIES** to denote the set of run-time instances of *Events*, *Agreements*, *Roles*, and *Entities*, respectively.

$$
\begin{aligned}
\mathit{Goods} &= \{g \mid g \in \textsc{Entities} \wedge g.\text{type}=\text{good}\} \\
\mathit{DeliveriesToDo} &= \{d \mid d \in \textsc{Events} \wedge d.\text{changeOfOwnership}=\text{always} \wedge d.\text{direction}=\text{out} \\
&\qquad \wedge d.\text{payment}=\text{never} \wedge d.\text{lifecycle}=\text{todo} \wedge d.\text{subject} \in \mathit{Goods}\} \\
\mathit{Orders} &= \{o \mid o \in \textsc{Agreements} \wedge o.\text{outgoingConnectedEvents} \cap \mathit{DeliveriesToDo} \neq \emptyset\} \\
\text{KPI SalesOrdersValue} &= \sum_{o \in \mathit{Orders}} o.\text{value}
\end{aligned}
$$

KPI-Average delivery duration: In this KPI, the delivery duration of a product is defined as the difference between the end and start time of a delivery event. In order to derive this KPI, we need to determine the *delivery events* (see ❸ in Figure 3.2) that are completed, i.e., the *lifecycle* characteristic has the value *done*. Thereafter, we can calculate the value of this KPI by applying a function that calculates the difference between the *actualEndDate* and *actualStartDate* of an event. To this end, we use a function named *ActualDuration*, which is formalized below.

$$
\begin{aligned}
\mathit{DeliveriesDone} &= \{d \mid d \in \textsc{Events} \wedge d.\text{changeOfOwnership}=\text{always} \wedge d.\text{direction}=\text{out} \\
&\qquad \wedge d.\text{payment}=\text{never} \wedge d.\text{lifecycle}=\text{done} \wedge d.\text{subject} \in \mathit{Goods}\} \\
\text{ActualDuration}(e) &= e.\text{actualEndDate} - e.\text{actualStartDate} \text{ where } e \in \textsc{Events} \\
\text{KPI DeliveryAverageDuration} &= \frac{\sum_{d \in \mathit{DeliveriesDone}} \text{ActualDuration}(d)}{\mid \mathit{DeliveriesDone} \mid}
\end{aligned}
$$

KPI-Percentage from offer to order: In this KPI, acceptance means that a sales offer is accepted by a customer, and a sales order is created for that offer. Therefore, we need to determine the sales offers that are followed by a sales order to derive this KPI. We formalized sales orders in the KPI-*Total value of sales orders*. Now, we need to obtain sales offers (see ❶ in Figure 3.2) that precede the determined sales orders. To this end, we need to check the events that can precede an agreement. Then, we can calculate the value of this KPI.

$$Offers = \{e \mid e \in \textsc{Events} \wedge e.\text{subject} \in Goods \wedge e.\text{direction}=\text{out} \wedge e.\text{payment}=\text{never} \wedge e.\text{changeOfOwnership}=\text{never}\}$$

$$Offer2Order = \{o \mid o \in Offers \wedge o.\text{outgoingConnectedAgreements} \cap Orders \neq \emptyset\}$$

$$\text{KPI Offer2OrderPercentage} = \frac{\mid Offer2Order \mid}{\mid Offers \mid} \times 100$$

The formalized KPIs above are not exclusive to the Sales area. Table 3.1 lists some of the KPIs from the dashboards for Sales, Purchase, and HRM that show resemblances with each other with respect to OEM elements that are used for formulating them. For example, both *Average delivery duration* and *Average receive duration* are the KPIs that show the average duration of an event, which will be performed as the result of an agreement. While in *Average delivery duration* the event is a delivery and the agreement is a sales order, in *Average receive duration* the event is a receive, and the agreement is a purchase order. Furthermore, both delivery and receive events change the ownership of goods.

Table 3.1: KPIs expressed using the NEXT OEM Language

Area	**KPI**	**Formula**
Sales	*Total value of sales orders*	$\sum_{s \in Orders} s.\text{value}$
Purchase	*Total value of purchase orders*	$\sum_{p \in PurchaseOrders} p.\text{value}$
HRM	*Total value of the personnel costs at employment*	$\sum_{e \in Employments} e.\text{value}$
Sales	*Average delivery duration*	$\frac{\sum_{d \in DeliveriesDone} \text{ActualDuration(d)}}{\lvert DeliveriesDone \rvert}$
Purchase	*Average receive duration*	$\frac{\sum_{g \in GoodsReceiptsDone} \text{ActualDuration(g)}}{\lvert GoodsReceiptsDone \rvert}$
Sales	*Percentage from offer to order*	$\frac{\lvert Offer2Order \rvert}{\lvert Offers \rvert} \times 100$
HRM	*Percentage from vacancy to employment*	$\frac{\lvert Vacancy2Employment \rvert}{\lvert Vacancies \rvert} \times 100$

3.5.2 Defining KPI Derivation Patterns

By analyzing the formalizations of the KPIs that are depicted in Table 3.1, we express KPI derivation patterns using Condition-Action (CA) rules that we explained in Section 3.3. Then, together with the dashboard team of our case study company, we check whether new KPIs can be derived using these KPI derivation patterns. For the cases that new KPIs can be derived using a KPI derivation pattern, together with the same team, we have evaluated that whether these new KPIs are relevant for organizations. The first row of Table 3.1 contains the following KPIs: *Total value of sales orders*, *Total value of purchase orders*, and *Total value of the personnel costs at employment*. By analyzing the formalizations of these KPIs, we note that there is a pattern: each KPI is related to an *agreement*, e.g., sales order in Sales, purchase order in Purchase, and employment in HRM. Therefore, we define a KPI derivation pattern, namely *Agreement Totality*, and express it as a CA rule according to the definition of a KPI derivation pattern (see Definition 1 in Section 3.3). The existence of *agreement* elements is the *c* part of the CA rule that corresponds to the when part of KPI Pattern-1. In the *a* part of the CA rule, there are two actions. The first action (a1) gets the related run-time data for an agreement from the run-time data of an OEM (OEM^R) using the function named *getInstancesOf* (see Step- 2 in Section 3.3). The second one (a2) sums the values in the related data and calculates the value of a KPI derived from *Agreement Totality*.

(c) if a $\in$ OEM $\wedge$ a is an agreement
then
(a1) $E1$ getInstancesOf(a, OEM^R)
(a2) $\sum_{el \in \mathrm{E1}} el.value$

Figure 3.3: *Agreement Totality* is expressed using CA rules

In the second row of Table 3.1, there are two different KPIs: *Average delivery duration* and *Average receive duration*. By analyzing the formalizations of these KPIs, we note that these KPIs show the average duration for an *event* on which an organization and an external entity agree, e.g., the delivery event on which an organization and its customer agree at a sales order. Based on that, we specified a new pattern, namely *average duration of an event that is executed as a result of an agreement*. While checking the new pattern together with the dashboard team, we noticed that the new pattern enables one to derive a new KPI for the HRM area. Because there is an *agreement* in the HRM area, namely *employment* that can be used with the pattern to derive a new KPI that shows average work duration. We discussed the relevance (i.e., whether a valuable insight is provided for organizations via new KPI) of that new KPI with the dashboard team. The team mentioned that from an organization's perspective, only tracking the average duration of something of value that enters or leaves an organization is important and relevant for the organization, e.g., goods come into or leave the organization. However, in the KPI, *Average work duration*, there is no change of ownership because in that KPI, *time* is the entity, and it is not owned by anyone in the real-world phenomena of organizations. Therefore, we need to update the new pattern such that it addresses the events that change the ownership of *goods*. The second KPI derivation pattern that we defined, namely *Ownership Duration* is shown below.

By checking the third row of Table 3.1, we note that the two KPIs in that row show the percentage of the *events that preceded an agreement* to the *events that can precede an agreement*. The resemblance in the formalizations of these two KPIs seems to be a pattern, i.e., in both KPIs, we see an event that precedes an agreement. Based on this resemblance, we specified a new pattern, namely *Continuation Percentage*.

(c) if e $\in$ OEM $\wedge$ e is an event $\wedge$ e.changeofOwnership=always $\wedge$ e.payment=never $\wedge$ e.subject $\in$ Goods
then
(a1) $E1$ getInstancesOf(e, OEM^R)
(a2) $E2$ $\{el \mid el \in E1 \wedge el.\text{lifecycle=done}\}$
(a3) $\frac{\sum_{el \in E2} \text{ActualDuration}(el)}{|E2|}$

Figure 3.4: *Ownership Duration* is expressed using CA rules

In the condition part of that pattern, we determine whether a given event can precede an agreement using its *outgoingConnectedAgreements* characteristic at design time. After getting the related run-time data for the given event, we need to filter when the given event preceded an agreement in the run-time. For that, we need to check the *lifecycle* characteristic of an event. If an event's *lifecycle* characteristic has the *done* value, this means that event preceded an agreement in the run-time.

(c) if e $\in$ OEM $\wedge$ e is an event $\wedge$ e.changeofOwnership=never $\wedge$ e.payment=never $\wedge$ e.subject $\in$ Goods $\wedge$ e.outgoingConnectedAgreements $\cap$ Agreements $\neq \emptyset$
then
(a1) $E1$ getInstancesOf(e, OEM^R)
(a2) $E2$ $\{el \mid el \in E1 \wedge el.\text{lifecycle=done}\}$
(a3) $\frac{|E2|}{|E1|} \times 100$

Figure 3.5: *Continuation Percentage* is expressed using CA rules

After defining this new pattern, we started verifying it together with the dashboard team. We found that there is an *event*, namely *purchase offer* that can precede an *agreement*, namely *purchase order* in the Purchase area. This means that a new KPI, *Percentage from purchase offer to purchase order*, can be derived using the new pattern. With the dashboard team, we discussed whether the new KPI forms a basis for decisions to improve the purchasing process in organizations. The team stated that by tracking the new KPI, organizations can determine the actions to increase the chance that purchase offers lead to purchase orders. Furthermore, the team mentioned that the new KPI might be used as an indicator by organizations to compare their suppliers. This means that the new KPI seems quite relevant, and it provides specific and valuable insights for the Purchase area.

Since the defined KPI derivation patterns are expressed using the NEXT OEM Language, the changes in the NEXT OEM Language will require modifying the KPI derivation patterns. For example, if the event element in the NEXT OEM Language becomes insufficient to capture particular real-world business activities, then it will be updated by the case study company, and subsequently, we will modify the defined KPI derivation patterns based on that update. Due to the changes in the NEXT OEM Language, maintenance of the OEMs that were created with it will require some effort from the case study company. In addition, over time, more concepts of the ERP applications, for example, finance, accounting, project management, and customer relationship management, will be able to be modeled in the form of OEMs by developing new elements for the NEXT OEM Language. In accordance with that, we will be defining new KPI derivation patterns for deriving the KPIs related to these concepts. In the following section, we demonstrate the use of our approach in a real-life setting and present preliminary results.

3.6 Validation

In this section, we show the applicability of our approach to a case study. As we described in Section 3.3, our approach takes two inputs: an OEM and the run-time data for that OEM. In the case study, an OEM will be provided by the case study company. The OEM that the case study company provided us is an extended version of the simple OEM that we presented in Section 3.4. In the extended OEM, there is a purchasing part in addition to the sales part in the simple OEM. Unfortunately, the run-time data for the extended OEM is not operational yet because the language that the company develops for modeling real-world phenomena in organizations in the form of OEMs is still under development. However, this does not prevent us to evaluate our approach because the business processes that are modeled in the extended OEM are offered by the case study company to its customers in its current ERP product (recall MyERPSuite in Section 3.4). Therefore, we use MyERPSuite as a source for obtaining the run-time data, which captures the same information. In the following subsections, we describe how to obtain the run-time data and elaborate on the application of our approach, respectively.

3.6.1 Obtaining Run-Time OEM Instances

Together with an expert in the case study company, we selected two real organizations (Organization A and Organization B) from the customers of the case study company. The selected organizations operate in the retail domain, and both are Business-to-Business (B2B) organizations. Moreover, these organizations execute sales and purchasing processes in MyERPSuite with respect to the extended OEM. In order to obtain the run-time data for that OEM, together with three domain experts in the case study company, we created mappings from the database schema of MyERPSuite to the elements inside the given OEM. These mappings are embedded in a set of database queries to extract the execution history from the database of MyERPSuite for each element in the OEM. Each mapping specifies which table in the database and which fields in a table are related to a mapped OEM element. We executed these database queries for the selected organizations and extracted the run-time data as an Event Log for each organization. Both of these event logs contain 6 months of data (January 2014 to June 2014). The event log for Organization A contains 108K cases and 159K events. The event log for Organization B contains 9K cases and 27K events. In both of the event logs, there are 5 types of events: *Sales Order*, *Delivery*, *Sales Invoice*, *Purchase Order*, and *Goods Receipt*.

3.6.2 Applying the Automated KPI Derivation Approach

In this subsection, we discuss the results that we obtained by the application of our Automated KPI Derivation Approach on the case study, which we explained its context in Section 3.4. In Figure 3.6, the derived KPIs for the two organizations in the case study are shown. As we mentioned while elaborating on the steps of our approach in Section 3.3, our approach shows the derived KPIs from a time perspective, i.e., the values of the derived KPIs are shown using a particular time dimension in a timeline. Accordingly, in the same figure (in Figure 3.6), time series charts are used to show the values of the derived KPIs for each month in a timeline.

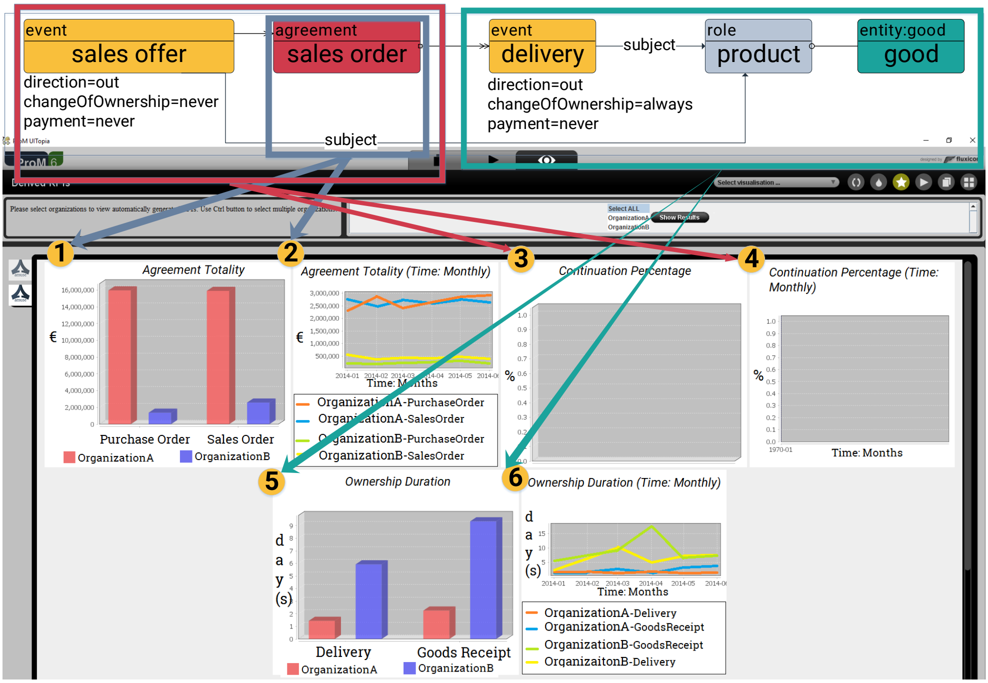

Figure 3.6: Derived KPIs and their mapping to the fragments of the given OEM

The KPIs that are derived using the KPI derivation pattern, *Agreement Totality* are depicted in the first chart (see ❶ in Figure 3.6). The second chart (see ❷ in Figure 3.6) shows these KPIs on a monthly basis. Our approach determined two different agreements by applying KPI Pattern-1: sales order and purchase order. This means that there are two concrete instances of KPI Pattern-1 in the given OEM. Based on that, our approach derived two KPIs, namely *Total value of sales orders* and *Total value of purchase orders*. These two KPIs consist of the total value of agreements for an organization. By analyzing the values of the derived KPIs, one can note that the total value of the agreements of Organization A is higher than Organization B.

The third chart (see ❸ in Figure 3.6) shows the derived KPIs based on the KPI derivation pattern, *Ownership Duration*. The same KPIs are shown in the fourth chart (see ❹ in Figure 3.6) on a monthly basis. In these charts, we expected to see two KPIs: *Percentage* *from offer to order* and *Percentage from purchase offer to purchase order* with respect to the given OEM. However, the charts are empty. This means that there are references from the run-time data to the OEM elements specified in KPI Pattern-2. Regarding that, we discussed the empty chart with two product managers from the case study company. The two product managers noted that the selected two organizations are not using sales offer and purchase offer events, which are not mandatory to execute in MyERPSuite.

The KPIs that are derived from the KPI derivation pattern, *Continuation Percentage* are shown in the fifth (see ❺ in Figure 3.6). In addition, the derived KPIs are depicted on a monthly basis in the sixth chart (see ❻ in Figure 3.6). By applying KPI Pattern-3 to the given OEM, our approach determined two events that change the ownership of goods and follow an agreement, namely delivery, and goods receipt. This means that there are two concrete instances of KPI Pattern-3 in the given OEM. In accordance with that, our approach derived two KPIs using KPI Pattern-3, namely *Average delivery duration* and *Average receive duration*. By analyzing the fifth chart, one can note that the average delivery duration KPI for the two organizations is nearly the same; however, the average duration of goods receipts for Organization A is different than for Organization B.

Above, we presented the results that we obtained by applying our approach in a real-life case study. The results indicate that our approach automatically derives and captures relevant KPIs for organizations. Moreover, we showed that using the derived KPIs, relevant insights can be provided to organizations.

3.7 Related Work

Much work has been conducted on measuring, tracking, and evaluating the performance of business processes using KPIs. Moreover, various approaches have been proposed for defining, modeling, and customizing KPIs. In this section, we list some of the approaches, which touch upon the aspects that we are interested in: deriving KPIs and tailoring them to organizations.

Jansen-Vullers et al. present a set of KPIs [71] for each dimension of the Devil's Quadrangle [20] to evaluate the effectiveness of a redesign. Despite the fact that some of the KPIs are formalized, most of them are manually defined in natural language. However, on the one hand, it is required to obtain the meaning of each KPI to determine whether they are relevant for numerous organizations. On the other hand, organizations have to implement custom solutions to calculate the values of these KPIs. Therefore, a significant effort is required to customize the provided KPIs for organizations.

A framework for modeling KPIs within organizations has been presented by Popova and Sharpanskykh [109]. In this framework, the authors formulate KPIs and relationships between them. However, all these formalizations are in natural language and specified manually. In addition, while defining KPIs, the meanings of the terms have not been taken into account; only the values assigned to the set of attributes of the KPIs have been taken into account. For example, PI20-efficiency of the planning process is defined as a KPI, but it is not explained when this KPI can be derived and what is the meaning of efficiency in this KPI.

Pedrinaci and Domingue present a Metrics Ontology [104] that is aimed at specifying domain-independent KPIs and computing their values. Subsequently, by using semantic technologies on top of that Metrics Ontology to support automated reasoning for computing KPIs, Pedrinaci et al. present a tool called SENTINEL [105]. On the one hand, the authors manually define a set of KPIs. On the other hand, it is not clear whether there are explicit relations between the computed KPIs and the activities in business processes. Unfortunately, it requires technical knowledge of ontological query building to define KPIs by end-users. As a result, it becomes a highly technical challenge to specify when and how these KPIs can be derived for organizations. Unlike that, in our approach, we aim to meet this challenge by defining KPI derivation patterns to automatically derive KPIs from an OEM such that it requires minimal technical knowledge from end-users to reason and model.

To enable one to define KPIs while modeling business processes, del Río-Ortega et al. present a metamodel [37]. With this metamodel one can define various KPIs; however, the KPIs need to be manually defined and associated with the model elements inside process models for each organization. Furthermore, the metamodel does not consider the meaning of the activities in the process models. Therefore, the KPIs defined for an organization using this metamodel cannot be applied in another organization to derive the KPIs automatically, i.e., the metamodel is a reference model for identifying the KPIs in the scope of an organization.

For comparing organizations using cross-organizational process mining, Buijs et al. present an approach [23]. Within this approach, a set of metrics is used to compare organizations. For example, precision, cost-based fitness, and behavioral appropriateness. However, these metrics do not serve as the same goals with KPIs, as they are not directly related to the business processes performed by organizations.

The aforementioned approaches define KPIs either in natural language or in a way such that it is not possible to automatically derive and tailor them to organizations. Although the approach [23] automatically derives a limited number of KPIs, only some of them are directly related to business processes.

3.8 Conclusion and Future Work

In this chapter, we presented our Automated KPI Derivation Approach aimed at automatically deriving tailored KPIs for organizations from *Ontological Enterprise Models* (OEMs). An OEM gives the opportunity to know which aspect of an organization a modeler tries to model. In other words, we can gain the meaning of each modeling element in an OEM in terms of the business of an organization using the characteristics specific to each modeling element. For example, with an agreement element, we can obtain the meaning of an agreement between a customer and an organization in the real-world phenomena of that organization. Moreover, different from other approaches, by automatically obtaining the meanings of the elements in an OEM, our approach can deduce the knowledge for automatically deriving KPIs that are related to the modeled aspect of an organization and relevant for the organization.

We believe that our approach proposes a better way than current approaches by having the following advantages to meet the challenges of the process of deriving tailored KPIs such as time-consuming, costly, and being error-proneness. ① By obtaining the meanings of the elements in an OEM, our approach automatically copes with ambiguity in the terms in business processes and reduces the complexity at the process of deriving tailored KPIs. ② Moreover, by automatically obtaining the exact meaning of the terms in business processes in various organizations that are required for customizing KPIs, our approach lowers the efforts of software vendors and organizations on customizing KPIs. ③ In addition, we provide a uniform view for KPIs such that the KPIs that are derived from the same KPI derivation pattern will be automatically named and visualized accordingly. Although the number of KPI derivation patterns that are currently supported in our approach is limited, we are still analyzing the KPIs related to various concepts in organizations (for example, finance and customer relationship management) to define new patterns for deriving more KPIs.

We have validated our approach by means of two steps. Firstly, as a proof-of-concept, we implemented our approach, which shows its feasibility. Secondly, we applied our approach in a case study and derived KPIs for two selected organizations whose business processes are modeled in the OEM that we illustrated in this chapter, and discussed how these organizations can be compared using these derived KPIs. Other software vendors, who focus on deriving KPIs automatically and work in the ERP domain, can apply our approach. To this end, software vendors need to make the required inputs available for our approach. First, they need to model the business processes that they offered in their products in the form of an OEM using the NEXT OEM Language. Second, they need to obtain the run-time data for that OEM and define references from the run-time data to the corresponding elements in the OEM.

In future work, we will focus on the aspects that make KPIs relevant for organizations. For instance, a set of KPIs in the HRM area that is relevant for an organization might not be relevant for another that has outsourced the business processes for that area. This means that there can be factors that determine the relevance of KPIs for organizations, such as domain, location, targeted customer audience, or the number of employees. Therefore, we will focus on identifying the factors that can affect the relevance of KPIs for organizations. Moreover, we can do benchmarking between organizations using KPIs so that they can see how they perform in comparison to each other.

Chapter 4

Predicting Relevant Key Performance Indicators for Organizations

Abstract
Organizations utilize Key Performance Indicators (KPIs) to monitor whether they attain their goals. For this, software vendors offer predefined KPIs in their enterprise software. However, the predefined KPIs will not be relevant for all organizations due to their varying needs. Therefore, software vendors spend significant efforts on offering relevant KPIs. That relevance determination process is time-consuming and costly. We show that the relevance of KPIs may be tied to the specific properties of organizations, e.g., domain and size. In this context, we present our novel approach for the automated prediction of which KPIs are relevant for organizations. We implemented our approach and evaluated its prediction quality in an industrial setting.

keywords- key performance indicators, prediction, relevance

This chapter is based on the following publication [received the best paper award]:

[7] Ü. Aksu, D. M. M. Schunselaar, and H. A. Reijers. Automated prediction of relevant key performance indicators for organizations. In *International Conference on Business Information Systems (BIS)*, pages 283–299, 2019.

4.1 Introduction

Organizations measure the performance of their business processes to determine whether they attain their goals. As a means for that, Key Performance Indicators (KPIs) are used [101]. *Average duration of product delivery* is a KPI that organizations use to monitor their product delivery processes. By tracking this KPI, organizations can predict how much staff must be assigned to their product delivery processes to keep the duration of a product delivery below a certain threshold, e.g., on average 3 days.

To support organizations in process performance measurement, software vendors offer predefined KPIs in their software products. With this, they aim to provide the maximal set of KPIs that may be relevant for most organizations. However, predefined KPIs will not work successfully in all organizations because they want relevant KPIs aligned to their specific goals [101]. For example, *taken leave per day* is a relevant KPI for a production organization, whereas it may not be relevant for a university, which has a similar number of employees. For this reason, software vendors include Business Intelligence (BI) functionality into their software products and let organizations develop custom KPIs. Although organizations may do this, custom development of KPIs still requires a significant effort both from software vendors and organizations [101].

Numerous studies have been conducted for determining relevant KPIs for organizations [28, 37, 72, 74, 146]. In these studies, relevant KPIs are either defined from scratch or selected from a set of KPIs (e.g., a KPI library) for each organization. Moreover, in these studies, the identified reasons that make certain KPIs relevant for one organization are not usually reusable at determining the KPIs for another. Therefore, for current approaches tailoring KPIs is a manual endeavor that needs to be repeated for each organization and requires a significant effort both from software vendors and organizations.

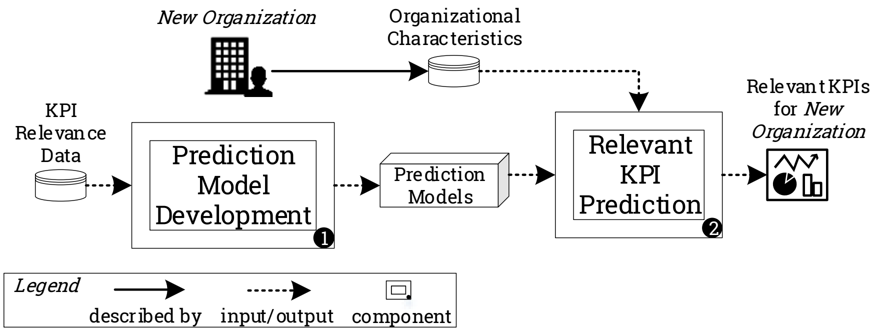

Figure 4.1: Our approach for predicting relevant KPIs for organizations

Within this chapter, we propose a novel approach for the automated prediction of relevant KPIs for organizations. The approach takes a set of prediction models aimed at predicting the relevance of KPIs and the characteristics of a new organization, e.g., domain, location, and the number of employees. By checking the given organizational characteristics of that organization against the relevance factors of KPIs encoded in the prediction models, the approach predicts which KPIs are relevant to the organization (see

❷ in Figure 4.1). To determine the relevance factors of KPIs and develop prediction models (see ❶ in Figure 4.1), the approach uses the known relevance values of a set of KPIs for a number of organizations, which needs to be given in the form of a specific input, KPI Relevance Data. By means of the automatically determined relevance factors of KPIs, we automate the prediction of relevant KPIs for organizations, which is manually repeated for every organization in current approaches. Thus, our approach sets itself apart from state-of-the-art. We evaluate the prediction quality of our approach by applying it in a real-life setting at a Dutch ERP software vendor. In this context, we discuss the results that we obtained.

In Section 4.2, we present our approach aimed at the automated prediction of relevant KPIs for organizations. The details of the implementation of the approach are given in Section 4.3. Afterwards, in Section 4.4, we evaluate the prediction quality of our approach by applying it in a real-life, industrial setting and present the results obtained in the application. Section 4.5 is devoted to the discussions on the implications of the obtained results. In Section 4.6, we provide an overview of related work on providing relevant KPIs to organizations. Finally, we present our conclusions and potential directions for future work in Section 4.7.

4.2 Approach

In this section, we explain the details of our approach on the automated prediction of relevant KPIs for organizations. As introduced in Section 4.1, there are two tasks: predicting relevant KPIs and developing prediction models. They are taken care of by the components Prediction Model Development and Relevant KPI Prediction. The former uses prediction models and the organizational characteristics of a new organization as inputs; the latter uses KPI Relevance Data as the only input. For the sake of simplicity, first, the definitions of organizational characteristics, prediction models, and KPI Relevance Data are listed below. Then, the details of each component are given.

Definition 1 ***Organizational Characteristics*** *contain the values of a set of characteristics (e.g., domain, location, and number of employees) by which organizations can be characterized.*
Example: Organization o1 = {domain Retail ∧ location Amsterdam ∧ numberOfEmployees [10-19] ∧ doesExport Yes ∧ industryClassification-MainGroup 47 ∧ industryClassification-SubGroup 8109 }.

Definition 2 ***Prediction Models*** *are aimed at predicting the relevance of KPIs. Each prediction model encodes a KPI, the factors that are the determinants to what extent the KPI will be relevant for organizations, and a prediction modeling technique, which outperforms predicting the relevance value of the KPI for organizations using those relevant factors.*

Definition 3 ***KPI Relevance Data*** *is a 2-tuple: (1) the relevance values of a set of KPIs for a number of organizations where that KPI set is considered as the comprehensive set from which a sub-set will be selected and (2) the key characteristics of these organizations with their values, i.e., Organizational Characteristics. For example, in Figure 4.3, an excerpt from a sample KPI Relevance Data is depicted.*

In our approach, the relevance value of a KPI can be a numeric value from a scale, namely KPI Relevance Scale. As the KPI Relevance Scale, we use a five-points Likert-type scale: $[1, 5]$, where a higher value denotes a higher relevance. The reason for using a five-points scale is that it has been recommended by many researchers [43, 58, 129] as the optimal number of categories for relevance.

Relevant KPI Prediction: To predict which KPIs are relevant for a new organization, two inputs are required: the organizational characteristics of that organization and prediction models. The prediction modeling technique encoded in each prediction model is executed with the given organizational characteristics. Thus, a predicted relevance value will be obtained for the KPI. In the output, the obtained relevance values are sorted from highest to lowest. Afterwards, the KPIs that have the highest predicted relevance value, a value of 5 in the KPI Relevance Scale used in our approach (see in Fig 4.2), are marked as the set of relevant KPIs for the new organization. However, this marking is flexible, and one can say that a value of either 4 or 5 may be presented as the set of relevant KPIs for the new organization. For this, to what extent a KPI is used for making decisions about the related business process in an organization may be a reason.

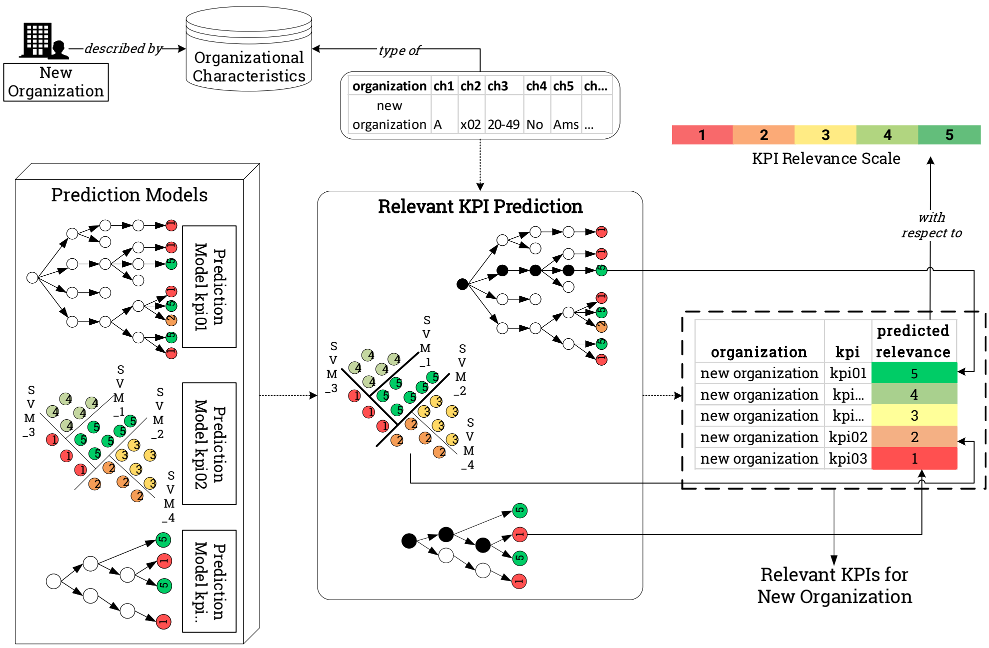

Figure 4.2: Predicting relevant KPIs for a new organization

Prediction Model Development: This component takes KPI Relevance Data as input. For each KPI in the input, an analysis task is performed to determine what organizational characteristics are the determinants of the relevance value of a KPI for organizations. The reason for performing the task per KPI is that relevance factors may vary from one KPI to another. For example, "the number of employees" may be the only factor that makes a KPI relevant for organizations, whereas, for those organizations, the relevance of another KPI may be dependent on both "number of employees" and "organization type", e.g., whether it is a non-profit organization.

Since the organizational characteristics in the given KPI Relevance Data are raw data, they need to be transformed into features to better represent the underlying patterns in the given KPI Relevance Data. That transformation is done by the encoders within this component. More specifically, a one-hot encoder is used for each feature. Then, a feature subset is selected for each prediction modeling technique employed within the component—employed techniques are listed in the implementation documentation of the approach[1]. This feature subset selection helps to make sense of the features for prediction modeling techniques. To keep the best performing features in subsets, the worst performing feature at each iteration is eliminated, and then the dependencies between features are uncorrelated by a dimensionality reduction.

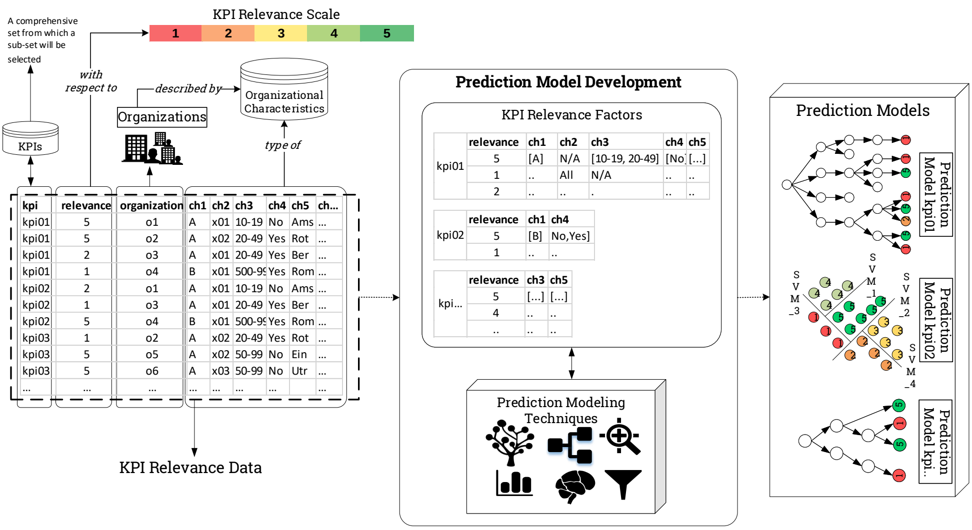

Figure 4.3: Creation of the prediction models for predicting relevant KPIs

Afterwards, the component trains and tests each prediction modeling technique to find out the best performing prediction modeling technique for each KPI. The reason for that is a prediction modeling technique may not outperform for predicting the relevance values of all KPIs since the relevance values in a given KPI Relevance Data may not be the same for all KPIs. Moreover, each prediction modeling technique has its own noise handling mechanism. For example, while Random Forest may be the best for an imbalanced set

[1]The implementation of our approach is available at `https://amuse-project.org/software/`

of relevance values, other prediction modeling techniques, e.g., Ada Boost may perform poorly. While training a prediction modeling technique, the component chooses a set of appropriate hyperparameters to discover the parameters that may result in more accurate predictions. To do so, the component uses a cross-validated grid-search algorithm. As a result of the train and test, the component identifies the best performing prediction modeling technique at finding the relevance factors of each KPI. For this, the balanced accuracy metric [57, 124] is used. By doing so, we aim to deal with the relevance values of KPIs that may have an imbalanced distribution in a given KPI Relevance Data. When the relevance factors and best performing prediction modeling techniques are determined for all KPIs, the component creates the prediction models for the KPIs. In particular, the relevance factors of a KPI, the selected prediction modeling technique for identifying them, and the KPI itself are encoded in the form of a prediction model.

In the next section, we give the details of the implementation of the approach.

4.3 Implementation

In this section, we give the details of the implementation[1] of our approach. On the one hand, the implementation is a constructive proof of the approach. On the other hand, it shows the applicability of the approach. In Figure 4.4, the technical details of the implementation are depicted.

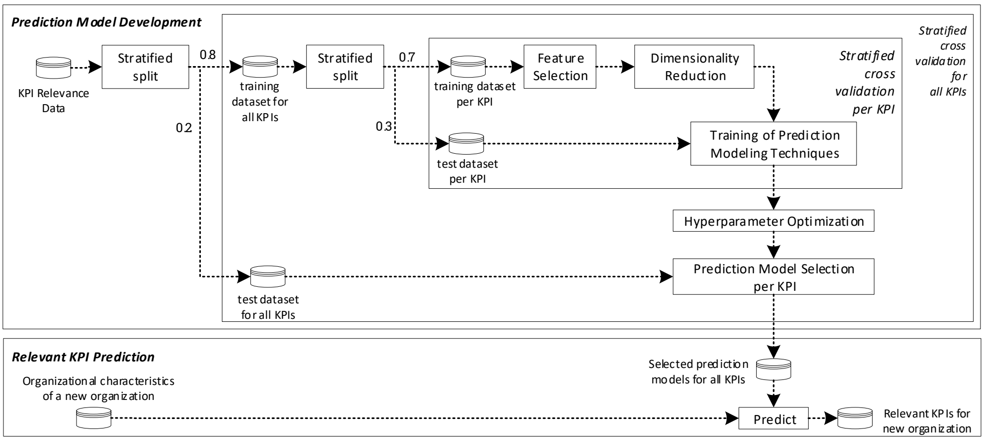

Figure 4.4: Technical details of the implementation of our approach

As explained, to predict relevant KPIs, the approach requires prediction models as input. This is taken care of the Prediction Model Development component within the approach. It takes KPI Relevance Data as input and develops prediction models. To accurately capture the knowledge in the given KPI Relevance Data, as shown in Fig 4.4, a nested (two-level) stratified cross-validation is used: (1) for all KPIs and (2) per KPI. More specifically, both model development and testing will be carried out n-times, which is specified in each stratified cross-validation block, using a different sample dataset of the given KPI Relevance Data. By doing so, we aim to develop prediction models that both capture the patterns in the given KPI Relevance Data, but also generalizes well to unseen organizational characteristics of new organizations.

Within the *stratified cross-validation block for all KPIs*, prediction model development and testing for all KPIs will be done in 5-folds. This means that in each fold, an 80%/20% stratified split [102] is done to divide the given input into training and test datasets. By doing so, the approach can develop prediction models and test them using the sample datasets that are preserving the percentage of the data points for each KPI and class (i.e., relevance values). As a result, the approach creates two different datasets in each fold: training dataset for all KPIs and test dataset for all KPIs. Similarly, within the *stratified cross-validation block per KPI*, the approach does the prediction model development and testing for each KPI in 5-folds. A training dataset per KPI and a test dataset per KPI will be generated in each fold. The aforementioned prediction modeling techniques will be trained and tested using these datasets. In order to avoid over-fitting to the training data, a 70%/30% split is preferred [63].

The approach utilizes feature selection and dimensionality reduction [99] to select the organizational characteristics in the training dataset per KPI that contribute most to the prediction variable (the relevance value of a KPI). Then, the approach applies feature scaling to have a standardized range of the values of the selected organizational characteristics. By doing so, we aim to make that each selected organizational characteristic may equally influence the prediction variable.

Afterwards, the approach trains each prediction modeling technique contained in it. Meanwhile, the approach tunes the parameters of each prediction modeling technique to determine the parameter set with which each trained prediction modeling technique performs best. When all prediction modeling techniques are trained, the approach tests them using the test dataset per KPI mentioned above. Using the balanced accuracy metric, the approach selects the best performing prediction model for a KPI from the set of the prediction models that are created in all folds of the *cross-validation block per KPI* and tested. When the best performing prediction model for each KPI is selected, then the approach completes the prediction model generation process. In other words, the selected prediction models are ready for predicting relevant KPIs for organizations.

To achieve high quality at predicting relevant KPIs for organizations while developing prediction models within the approach, we use 3 types of meta-algorithms: stacking, boosting, and bagging [163]. In stacking, a meta-technique tries to learn the best combination of the prediction models of the primary prediction modeling techniques, which are combined as a stack. In boosting, the same prediction modeling technique is applied in a chain to learn and fix the prediction errors of prior prediction models developed in the chain. Different sub-samples of the training dataset are taken, and multiple prediction models are generated in bagging. Then, these models are aggregated to form a final prediction model, which has a better accuracy value.

Since most properties of organizations are categorical data types and the scale we used for relevance values (KPI Relevance Scale) has multiple points, it is required to support multi-class classification [63] prediction modeling techniques from the machine learning discipline within our approach. Accordingly, our approach employs the prediction modeling techniques that are listed in the implementation documentation of the approach[1]. However, our approach is flexible to support continuous (numeric) data types. This can be indicated in the configuration where the approach learns the organizational characteristics contained in a given KPI Relevance Data. Moreover, our approach is extensible to support regression prediction modeling techniques in the case that one may want to predict a decimal value for the relevance value of KPIs instead of a numeric value from a KPI Relevance Scale.

In addition, to obtain better predictive performance, the approach reduces the problem of multi-class classification to multiple binary classification problems while developing prediction models. In this regard, we apply the following strategies: one-vs-rest and one-vs-one [63]. The former involves training a single prediction modeling technique per class, whereas in the latter, a particular prediction modeling technique is trained for each different pair of classes.

In the following section, we describe how we evaluate our approach in a practical use of its implementation.

4.4 Evaluation

In this section, we demonstrate the use of the proposed approach in an industrial setting and evaluate how accurately it predicts relevant KPIs for organizations. In this regard, in the following subsections, we describe how we develop prediction models and use them at predicting relevant KPIs for organizations in a case study.

4.4.1 Data Collection

In order to develop prediction models, the approach requires KPI Relevance Data. However, although software vendors usually know the key characteristics of the organizations that they deliver their software products, they are typically not aware of the relevance of the KPIs that they offer to those organizations. Therefore, we investigated whether we could identify a proxy for this type of data. KPI Usage Logs are typical data sources in which software vendors typically keep track of how KPIs are being used by organizations. In general, software vendors either record these logs using the software product in which they offer KPIs for organizations or using a third-party BI tool (e.g., Qlik Sense and Microsoft Power BI), which they use as a means for enabling organizations to develop custom KPIs. KPI Usage Logs is a data source from which one can obtain information on the usage of KPIs. For example, how many times a particular KPI is used in an organization, when that KPI is used, and how much time has spent using the KPI can be obtained from KPI Usage Logs. The obtained information on the usage of KPIs can be seen as the interest of organizations in KPIs. Moreover, one can interpret the usage of KPIs in organizations as the interest of the organizations on KPIs, i.e., the relevance of KPIs for the organizations. Thus, as a primary proxy for known relevance values of KPIs for organizations, KPI Usage Logs are determined.

A Dutch ERP software vendor, the case study company, records KPI Usage Logs for the KPIs that the company offers to its customers. In the company, we had a training session on the KPIs, which are offered to its customers within its ERP software product. In particular, we examined the KPIs in the Human Resource Management (HRM) area. The reason for that is that human resources form a key asset in any organization. As such, the availability of the employees in an organization is essential in performing its business processes to attain its goals. However, due to various reasons, employees may not always be available, for example, sickness, injury, maternity leave, or vacations. Absence and leave are the two sub areas in HRM that concern the unavailability of employees in organizations. The former deals with the unexpected reasons for the unavailability of employees. The latter focuses on the unavailability of employees resulting from statutory rights as granted by labor laws. In this regard, together with two experts who manage the KPIs in the company, we selected 13 KPIs from the absence and 6 KPIs from the leave sub area. While selecting the KPIs, our main consideration was the wide usage of the KPIs by organizations to get sufficient data points such that our approach can predict relevant KPIs for a new organization accurately. The selected 19 KPIs are commonly used by more than 2000 client organizations of the software vendor. Afterwards, the experts defined a set of metrics for transforming the usage of the KPIs into relevance values. Since the defined metrics require a minimum of one-year usage of the KPIs, the relevance values of the selected 19 KPIs are obtained for approximately 1100 organizations, which use those KPIs at least for a year.

In addition to the obtained relevance values, the characteristics of those 1100 organizations were required to create KPI Relevance Data and develop the prediction models for the selected KPIs. To determine which characteristics and their values for those organizations are available in the company, we arranged three meetings with various experts. The first meeting was with the following experts: a director—the CIO (Chief Information Officer) of the company—who has knowledge on scoping the KPIs offered to the organizations and a senior product developer who is an expert on designing and developing KPIs. These experts explained the characteristics of the organizations that they consider while scoping and developing the KPIs offered to the organizations. A marketing manager participated in the second meeting and described the characteristics of the organizations that often request adjustments for the offered KPIs. In the last meeting, together with a product manager, we analyzed the data about the organizations to identify what characteristics and their values are available within the company. As a result, we selected the characteristics shown in Table 4.1. The characteristics shown in the table go beyond the characteristics considered for identifying relevant KPIs for organizations in the related literature. Specifically, in recent studies [68, 73, 107, 108], a small number of characteristics are often taken into account, for example, often, sector of organizations and type of business processes in organizations. In that sense, our approach has a broad coverage of characteristics.

All the characteristics we selected are categorical, and the values of them were available for 750 out of the previously selected 1100 organizations. Moreover, the names of these characteristics are translated from their original definitions in Dutch.

By combining the obtained relevance values of the selected KPIs 19 with the organizational characteristics of the aforementioned 750 organizations, we created the KPI Relevance Data for our evaluation. This means that the required input for developing the prediction models for the KPIs is ready. Accordingly, the approach created 19 prediction models for predicting the relevance of the selected 19 KPIs.

Table 4.1: Selected characteristics of organizations

Characteristic	Explanation with some example values
Legal Form - Main Group	The main group of the legal form of an organization, e.g., with or without legal entity.
Legal Form	The legal form of an organization, e.g., private limited company, foundation or association.
Non-Profit	A non-profit organization uses the money it earns to help people. However, a profit organization invests the money it earns on developing new products or services to sell them and make more money.
Industry Classification - Main Group	The main group to which the organization is assigned by the Chamber of Commerce within the Netherlands. For example, construction is the main group to which general civil and utility construction organizations are assigned. Transportation and storage is another example for the main group.
Industry Classification - Sub Group	The sub group to which the organization is assigned by the Chamber of Commerce within the Netherlands. Construction of residential buildings and construction of railways are two example sub groups of the construction main group. Similarly, passenger transport and freight transport are two example sub groups of the transport and storage main group.
Province	The province where the organization is registered
Number of - Employee Range	The range of the total number of employees in the organization. For example, 10-19, 20-49, and 50-99.
Import	Whether the organization does import
Export	Whether the organization does export
Has Subsidiary Organizations	Whether the organization has subsidiary organizations

4.4.2 Applying the Automated Relevant KPI Determination Approach

We predicted relevant KPIs for a set of organizations using the developed prediction models. Since the developed prediction models are for the KPIs in the absence and leave sub areas, the set of the organizations for which relevant KPIs are predicted are accordingly selected. In particular, the organizations are selected from the client organizations of the case study company that are not purchased and not use to the selected 19 KPIs but use the related functionalities, i.e., absence and leave within the ERP software product of the company. As a result, we selected 261 organizations and predicted relevant KPIs for them.

To determine the prediction accuracy of our approach in the case study, we collaborated with the CIO of the company and a senior product manager in the case study company. The reason for collaborating with these two experts is that these experts have extensive knowledge both on the organizations for which relevant KPIs are predicted and on the expected relevance values of the selected KPIs for those organizations. Then, we calculated the prediction accuracy of our approach by comparing the predicted relevance values of the KPIs for the organizations with the expected relevance values of the KPIs for these organizations, which are provided by the aforementioned two experts. In Figure 4.5, the prediction accuracy of our approach in the case study is depicted.

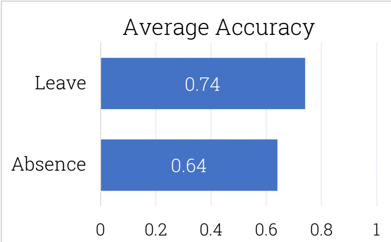

Figure 4.5: Prediction accuracy of our approach in the case study

As depicted in Figure 4.5, in the case study, our approach achieves a 74% balanced accuracy at predicting the relevance of 6 KPIs in the leave sub area for 261 organizations. Similarly, a 64% balanced accuracy is achieved at predicting the relevance of 13 KPIs in the absence sub area for the same organizations. The weighted average of the prediction quality values for these two sub areas will show the prediction quality of our approach for the HRM area, which is 67%. In the following section, we discuss the implications of those results.

4.5 Discussion

To consider a certain prediction quality as good, one should look into the context of the application where that quality is measured [56, 117, 128]. In this regard, we discuss the prediction accuracy of our approach. Since no other study has so far focused on the automated prediction of relevant KPIs for organizations, there is no exact reference to compare the prediction quality of our approach with. However, we think that the prediction quality of our approach shown in Fig 4.5 is reasonable [128].

As mentioned in Section 4.3, the problem that our approach tries to solve is a multi-class classification. Having a balanced distribution of each class has a significant effect on prediction quality in multi-class classification. However, this was not the case for the KPIs that we used in the case study. Notably, for the lower relevance values (e.g., 1 and 2 with respect to the used KPI Relevance Scale), there were fewer data points than for the higher relevance values (e.g., 4 and 5 with respect to the used KPI Relevance Scale). As a result, the approach was inclined to predict higher relevance values for some KPIs, which are expected to have lower relevance values by the experts in the case study company. This indicates that the prediction quality of the approach has been negatively affected due to lacking data points.

One of the possible reasons for a lower prediction quality value in the absence sub area is that there were fewer data points in the used KPI Relevance Data for each KPI in the absence sub area than for each KPI in the leave sub area. In particular, there was a limited variety of small organizations in the known relevance values of the KPIs in the absence sub area. This was because in small organizations, the management of absence-related data is ad-hoc, i.e., absence-related data may not be stored day-to-day. Therefore, there was missing usage information in the KPI Usage Logs for those KPIs for small organizations. By contrast, leave operations in organizations are mostly recorded day-to-day since there are regulations defined by law to keep the data related to these leave operations up-to-date. In addition, leave is a type of operation that can be planned ahead, whereas absence has a more unpredictable nature.

Table 4.2: Outperformed predicting modeling techniques for the KPIs in the HRM area

	Logistic Regression	Support Vector Machines (SVMs)	Decision Tree	Random Forest	Stacked (Decision Tree & SVMs)
Absence	7	3	2	1	0
Leave	0	0	2	2	2

As a result of having fewer data points for the KPIs in the absence sub area, as shown in Table 4.2, the prediction modeling techniques that use the linear separation method for input data outperformed at predicting the relevance values of the KPIs in this sub area. However, tree/forest based prediction modeling techniques were the majority in the leave sub area since there were more data points for the KPIs in this sub area.

We also had the idea to apply our approach in the finance area. Using the KPI Usage Logs of 109 KPIs in the finance area, we created KPI Relevance Data. Then, we analyzed this data before applying the approach on it. We found out that for more than 60% of the KPIs, there are fewer data points for 3 out of 5 known relevance values. We decided against actually using the approach although the data was not good enough, i.e., containing fewer data points. Unfortunately, the approach performed worse than predicting the relevance of the KPIs in the HRM area. We examined the failing predictions for the KPIs in the finance area with the two experts together with whom we determined the prediction accuracy of the approach in the HRM case—a director and a senior product developer. These experts pointed out that the expected relevance of the KPIs in the finance area is mostly dependent on various financial characteristics of organizations such as debt, revenue, payment periods of both the customers and suppliers of these organizations, and how the products and services are sold by these organizations. Although our approach is extensible to new organizational characteristics, however; unfortunately, these organizational characteristics are not available in the case study company since these are mostly sensitive data about organizations.

Software vendors that focus on automatically predicting relevant KPIs for their customers and operate various domains can apply our approach. However, if these software vendors may want to predict relevant KPIs for their customers using a different set of KPIs and organizational characteristics than we demonstrated in the case study, they need to provide their KPI Relevance Data to our approach and develop prediction models using the approach. Then, these software vendors can predict relevant KPIs for their customers by executing the approach with the developed prediction models.

4.6 Related Work

Due to the high interest in both academia and business, there is a broad literature in the field of organizational performance measurement. Notably, researchers proposed various approaches dealing with determining relevant KPIs for organizations since KPIs are widely used as a means for measuring the performance of organizations. Within these approaches, creating relevant KPIs afresh for any organization or choosing KPIs from a reference set of KPIs (e.g., a KPI library) as the relevant set for a particular organization are the two common ways of determining relevant KPIs. In this section, we list some of the works, which cover the following question that we are interested in: how are relevant KPIs determined for organizations?

Much work has been conducted on defining relevant KPIs from scratch for organizations in various domains. Granberg and Munoz develop KPIs for airport managing organizations [62]. Similarly, to monitor the performance of airports, a set of KPIs is proposed in [46]. Kaganski et al. [74] describe the development of KPIs for small and medium-sized enterprises (SMEs). While a set of KPIs for the organizations that have highly diverse product families are defined in [119], Elliot et al. [44] specify a set of KPIs for a large pediatric healthcare organization. Since the development of KPIs in the aforementioned works is from scratch and manual, in each work, it is required to have an intensive technical knowledge of the organization to which relevant KPIs are determined. Thus, a significant effort is required to obtain that knowledge.

Apart from the aforementioned works, del Río-Ortega et al. present a metamodel [37] as a basis for working with KPIs. Using the language proposed as part of the metamodel, one can model KPIs within the process models of the processes in an organization. Then, the values of the modeled KPIs can be derived from the execution logs of the process models. However, this still requires each organization to determine relevant KPIs for itself and model them using the proposed metamodel. Therefore, this will require a significant effort of each organization.

In some studies [68, 73, 107, 108], researchers focus on selecting a subset from a set of KPIs to determine the relevant set of KPIs for organizations. Within that selection process, researchers mostly consider the sector of an organization or a set of business processes in an organization. However, due to the varying needs of organizations, a KPI subset that is selected as the relevant set for one organization may not be relevant for all other organizations, which are in the same sector or perform similar business processes with that organization. Therefore, that KPI subset selection process needs to be repeated for many organizations. To deal with that, Analytic Network Process (ANP) is utilized [28, 72, 81, 146]. In particular, certain characteristics of KPIs such as reliability, comparability, and understandability are taken into account to determine the priorities of a set of existing KPIs in organizations. This is mostly done together with specific experts in organizations. Then, the KPIs that have the highest priorities are selected as the relevant KPIs for organizations. However, on the one hand, since the considered characteristics of KPIs are subjective to experts, the priority of a KPI may vary from one organization to another. On the other hand, ANP is a time-consuming and complex multi-criteria decision-making method, and therefore requires a significant effort from organizations.

4.7 Conclusion and Future Work

In this chapter, we presented a novel approach aimed at the automated prediction of relevant KPIs for organizations. A set of prediction models aimed at predicting the relevance of KPIs and the organizational characteristics of a new organization are the required inputs by the approach. The approach determines which of the KPIs that are encoded in the prediction models are relevant for that new organization using the relevance factors of the KPIs. To identify these factors automatically and develop prediction models, the approach employs prediction modeling techniques and applies them on the known relevance values of KPIs for organizations, which should be given in the form of a specific input, KPI Relevance Data.

To show the accuracy of our approach, we implemented it and demonstrated it in a case study at a Dutch ERP software vendor. Within the case study, together with experts in the company, we selected 19 KPIs from the HRM area that are offered to organizations by the company in its ERP software product. The known relevance values of the selected KPIs were not available in the company. Therefore, we identified KPI Usage Logs as a proxy for known relevance values of KPIs, and subsequently we created KPI Relevance Data and developed the prediction models for the selected 19 KPIs together with the experts in the company. Afterwards, the relevance values of the KPIs were predicted for 261 organizations, which are new to those KPIs. Finally, we evaluated the prediction quality of the approach by comparing the predicted relevance values of the KPIs against the expected relevance values of those KPIs, which are provided by two experts in the company. The prediction quality of the approach was of sufficient quality to show the practical usage of the approach. As a result, we automate the selection of relevant KPIs for every organization. For current approaches, this is a manual endeavor that needs to be repeated for every single organization. Thus, we believe that our approach lowers the efforts of software vendors for determining relevant KPIs for their client organizations or the efforts of these organizations doing this themselves.

In future work, we want to extend our approach for determining relevant KPIs for different roles in organizations since the relevance of a KPI might vary from one role to another in organizations. For example, there may be a significant difference in the relevance value of a KPI on daily stock changes between for a CEO and for a warehouse employee. Furthermore, sales, purchasing, and logistics are the areas to which we envision extending our approach since their commonality among organizations in addition to the less sensitivity of data for organizations in these areas in comparison to other areas, e.g., finance and accounting. Besides, we plan to develop a decision graph aimed at identifying which visualization best suits for particular KPIs. Thus, engaging dashboards comprising relevant KPIs can be built automatically. Moreover, the approach in this chapter and the approach that we presented in [5] are part of our Cross-Organizational Process Mining Framework, which we introduced in [4], and will be incorporated together into the framework. With this, we aim to provide recommendations for organizations using the benchmarks that are developed utilizing relevant KPIs.

Chapter 5

Automated Generation of Engaging Dashboards

Abstract
Organizations use Key Performance Indicators (KPIs) to monitor whether they attain their goals. To support organizations at tracking the performance of their business, software vendors offer dashboards to these organizations. For the development of the dashboards that will engage organizations and enable them to make informed decisions, software vendors leverage dashboard design principles. However, the dashboard design principles available in the literature are expressed as natural language texts. Therefore, software vendors and organizations either do not use them or spend significant efforts to internalize and apply them literally in every *engaging dashboard* development process. We show that engaging dashboards for organizations can be automatically generated by means of automatically visualized KPIs. In this context, we present our novel approach for the automated generation of engaging dashboards for organizations. The approach employs the decision model for visualizing KPIs that is developed based on the dashboard design principles in the literature. We implemented our approach and evaluated its quality in a case study.

keywords- key performance indicators, dashboard, visualization.

This chapter is based on the following publication:

[6] Ü. Aksu, A. del-Río-Ortega, M. Resinas, and H. A. Reijers. An approach for the automated generation of engaging dashboards. In *OTM Confederated International Conferences" On the Move to Meaningful Internet Systems*", pages 363–384, 2019.

5.1 Introduction

To determine whether they attain their goals, organizations measure the performance of their business execution. To do so, they use Key Performance Indicators (KPIs). As a means to monitor KPIs, organizations use dashboards that are either developed by themselves or offered by software vendors. A typical dashboard aims to inform decision makers by displaying the information that they need to improve the business processes in their organization. In particular, such information is displayed mostly as a table or a graph. By doing so, it adds visual attractiveness to grab the attention of decision makers and enable them to make informed decisions at a glance.

However, most dashboards are poorly designed displays, although adequate technology is used while developing them. Therefore, most dashboards fail to communicate efficiently and effectively since they mainly focus on decoration rather than substance [42, 50, 54, 97, 158]. For example, the dashboard depicted in Figure 5.1 is an incident management dashboard of an organization[1]. By analyzing this dashboard, one can see that the dashboard goes against the dashboard design principles in the literature. For example, pie charts have many slices that make them unreadable; also, their colors are distracting, which causes misleading associations. More importantly, this is a cluttered design that does not reflect the overall status of the related business processes in that organization. Since there is an overload of the information displayed as a cluttered view, decision makers need to spend substantial effort to identify the messages that the dashboard is designed to convey. As a result, the dashboard is “not engaging” decision makers to take relevant decisions for improving the performance of their organization.

Instead of providing only a fraction of the insight that is needed to monitor business, engaging dashboards communicate in a manner that enlightens decisions makers for informed decisions [14, 50]. More specifically, dashboards engage decision makers if the available dashboard design principles in the literature [42, 52, 54, 79, 92, 93, 97, 134, 158] are used when creating them. Moreover, engaging dashboards enable decision makers to sense and process the displayed information rapidly through the visualization elements, which can be quickly examined and understood without requiring any further interpretation. Not to distract decision makers with overloaded information, the right context for KPIs is visualized in such a way that it inspires actions. Furthermore, “Are we on track?” and “How well is our organization performing its business?” are such questions in organizations to which complete answers can be obtained at a glance in engaging dashboards. Simply put, engaging dashboards do not require any investigation, analysis, or aggregation of the information, which is a must for informed decisions and is distributed inside an organization.

To overcome these issues in the field of dashboard development, several approaches are available in the literature. Within these approaches, mostly dashboards are either developed from scratch for each organization, or a template is created and customized for organizations depending on their specific needs. This customization process is carried out by software vendors or by their client organizations. Although organizations may perform this customization process, it still requires a significant effort both from software vendors and organizations [42, 50–52, 54]. To deal with that, some approaches [31,

[1] The example dashboard is taken from `https://adniasolutions.com/dashboard-design-principles/introduction-to-dashboards/`, Last accessed Aug. 2021

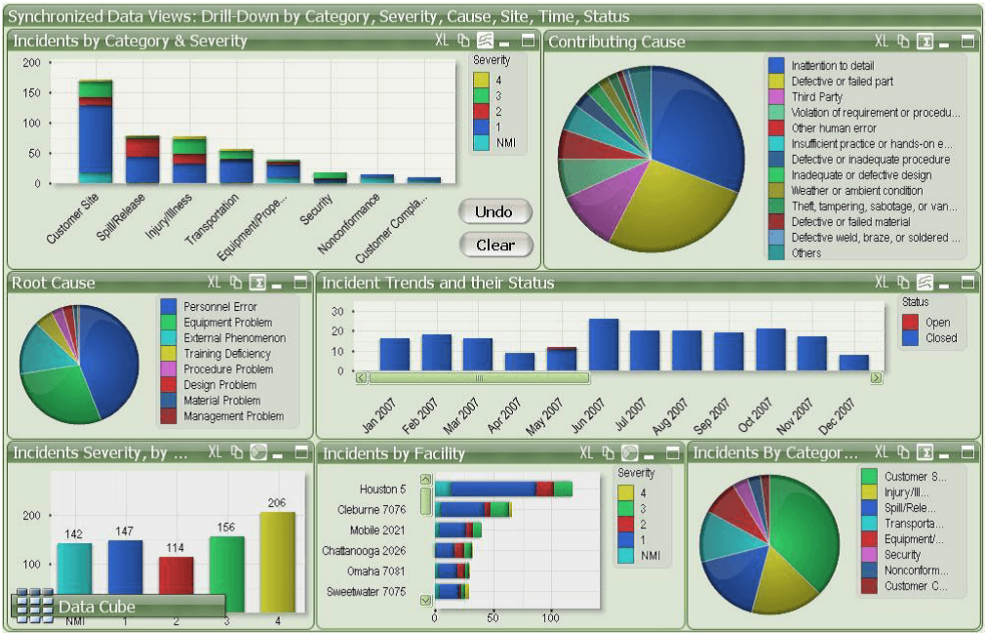

Figure 5.1: An example of a non-engaging dashboard

66, 75, 100] focus on the automation of dashboard development. Creating a dashboard template and expressing the structure of a dashboard in terms of the elements of a descriptive dashboard design language are the two prominent ways in these approaches. However, these approaches either cover only a few of the state-of-the-art dashboard design principles [42, 52, 54, 79, 92, 93, 97, 134, 158] or require human intervention to incorporate each dashboard design principle consistently. Therefore, the KPIs that are visualized using these approaches still lead to misinterpretations.

Furthermore, there are several studies [31, 66, 75, 76, 100, 148] in the related literature aimed at automatically generating dashboards. In [31, 100], a notation is introduced such that a dashboard with its KPIs can be created based on a user-created model with that notation. However, it still remains a challenge for the user to obtain the knowledge to identify appropriate visualizations for KPIs. In another study [148], a selection is made from existing dashboards for a user based on the comparison of the requirements of users on dashboards. Since existing dashboards are used without tailoring, identifying what visualizations are appropriate for KPIs and employing dashboard design principles are neglected. Although a mapping-based visualization of KPIs was proposed in a recent study [76], it lacks essential relationships in quantitative information (e.g., correlation, ranking). To sum up, there is no ready-to-use decision mechanism that exploits available dashboard design principles for visualizing KPIs and generating dashboards.

With this chapter, we propose a novel approach for the automated generation of *engaging* dashboards for organizations (See Figure 5.2). The approach takes a set of KPIs with their attributes and values, as well as a decision model that is developed for visualizing those KPIs as inputs. As such decision models for visualizing KPIs are not readily available,

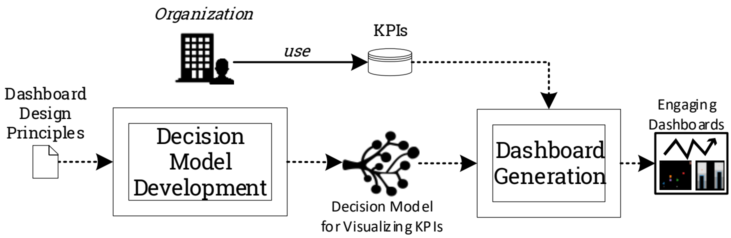

Figure 5.2: Our approach for the automated generation of engaging dashboards

we developed a decision model for visualizing KPIs, which is our second contribution in addition to the approach. The decision model that our approach uses is developed by analyzing the prominent dashboard design principles in the literature and evaluated to show its common usability. Using the decision model, the approach determines which visualization element will be used to display each KPI on a dashboard. Depending on the attributes and the values of a KPI, a particular table or graph will be chosen as the visualization element. By means of the automatically determined visualization elements for each KPI, we automate the generation of engaging dashboards for organizations. Our approach sets itself apart from the state-of-the-art in conveying relevant messages to decision makers via automatically generated engaging dashboards. Thus, decision makers can make informed decisions to improve the performance of their organizations.

In the evaluation of the approach, first, we check the common usability of the decision model developed for visualizing KPIs in two organizations with experts. Then, in one of the organizations, we execute the approach, and together with experts in that organization, we compare the newly created dashboard with an existing dashboard to see how our approach helps them at making informed decisions. The results that we obtained indicate that this new approach is able to fulfill the needs of organizations for improving their business.

We provide the background on dashboard design principles in Section 5.2. In Section 5.3, we present our approach for the automated generation of engaging dashboards. In Section 5.4, we evaluate the decision model for visualizing KPIs that our approach employs, and then present the results obtained while evaluating the dashboard generated using our approach in a case study. Section 5.5 is devoted to the discussion of the obtained results. In Section 5.6, an overview of the related work on developing dashboards for organizations is given. Finally, we present our conclusions and directions for future work in Section 5.7.

5.2 Theoretical Background

Dashboards are pervasive means to display important information at a glance, as needed to achieve objectives. Accordingly, much work has been conducted on developing the dashboards that are communicating important information and engaging. Notably, to facilitate dashboard development researchers and practitioners developed several guidelines [1, 42, 51, 52, 54, 79, 91–93, 97, 134, 158, 162]. Within these guidelines, principles for visualizing quantitative information in dashboards, i.e., dashboard design principles are described. Simply put, dashboard design principles describe what visual representations (e.g., various graphs) should be used and how they should be used. In this context, we list the dashboard design principles available in the literature.

Dashboard Design Principles

In the literature, numerous researchers and practitioners provide various dashboard design principles [1, 42, 51, 52, 54, 79, 91–93, 97, 134, 158, 162] to develop dashboards that are communicating important information visually in the most informative way such that organizations can make informed decisions to improve their business. These dashboard design principles are mostly expressed as natural language texts in the form of rules and best practices. Some researchers follow a more structured approach [1, 42, 51, 52, 54, 79, 91, 92, 162] and provide mechanisms (e.g., a table consists of rules for selecting graphs or a diagram shows which graphs should be used in which condition) such that organizations can determine appropriate visualization elements while visualizing certain quantitative information in dashboards. In this context, we identify which of these dashboard design principles for visualizing quantitative information are more suitable to determine sense-making visual elements for displaying KPIs. In this regard, what properties [42, 54, 97, 106, 134, 158] should a dashboard have to be considered "engaging" are listed below.

- **Consistent view**: Visual aspects of the dashboard (e.g., visualizations, colors, fonts, alerts, naming conventions, etc.) should be consistent and free from conflicts and ambiguity.
- **Contain what is important**: The dashboard should only contain relevant and timely information. Redundant information should be eliminated.
- **Deliver actionable information**: The messages conveyed via the dashboard should be suitable to its users for making decisions. The dashboard should provide clear signals about the monitored performance, e.g., proper alerts based on the predefined targets.
- **Content over decoration**: Visual attractiveness should not be the main concern in a dashboard.

Furthermore, we take the dashboard design principles that provide comprehensive guidance explained in [42, 51, 52, 54, 79, 92, 93, 97, 134, 158] as the sources for developing a decision model for visualizing KPIs. In these sources, eight typical relationships that can be encoded in quantitative information are discussed. We explain each relationship by specifying the visualization elements, which are mostly recommended and used for visualizing that relationship.

Time series: This relationship is about how a set of values change over time based on particular time units (intervals), e.g., by year, month, day, or hour. Line Graph is the graph that is well-know and mostly used for displaying this relationship. Bar Graphs and Area Graphs are also often used for displaying a time series relationship.

Ranking: How a set of values relate to each other in a particular order is described in ranking relationships. Since bars in graphs can be easily understood by any audience and best encode the values in a ranking relationship, Bar Graphs are mostly used to display a ranking relationship in dashboards.

Part-to-whole: This relationship is about how much the parts of a whole contribute to the whole, i.e., expressing the proportions of a whole. As a common practice, Pie Graphs are used to display a part-to-whole relationship.
Deviation: In this relationship, the focus is on how one or more values in a set of values vary from a reference, e.g., forecast. This is achieved by comparing values with a reference and displaying the degree of that difference. The values that divert from a reference are represented as bars in graphs and displayed in Diverging Bar Graphs, i.e., Variance Graph in most of the time.
Distribution: This relationship expresses the way how a set of values are distributed across a particular range that is from lowest to highest. Histogram and Box-Plots are well-known graphs that are usually used for displaying a distribution relationship.
Correlation: How a set of values affect each other is expressed in a correlation relationship. Mostly, two-paired, i.e., categorized set of values, are analyzed to see how they relate to each other: whether the values in one set increase or decrease based on the values in another set. Scatter Plot is the most used graph to display a correlation relationship.
Nominal comparison: This relationship describes a set of values based on a categorical scale without an order. For instance, the revenue of each department in an organization. Bars in graphs best encode values on a categorical scale, and therefore Bar Graphs are the most common visualization elements used to display a nominal comparison in dashboards.
Geospatial: The values in a geospatial relationship are located based on their geographical location. Spatial Maps are always used for visualizing this relationship.

Although there are many types of graphs for visualizing quantitative information, most of them are not recommended and listed as the graph types to avoid [42, 52, 54, 79, 92, 93, 97, 134, 158], such as Pie, Donut, Radar, Funnel, Circle, Area Graphs, or 3D Graphs. The main reason for that is these graphs fail effectively communicating quantitative information and causing misinterpretations. Overlapping shapes, missing scales, hidden values, distracting decoration, and cluttered view are the problems these graphs commonly have.

Within the dashboard design principles available in the literature, researchers provide guidance on using colors, resizing visualization elements, and placing them in dashboards in addition to determining visualization elements. To decide how visualization elements should be placed in dashboards, layout patterns are devised. The most common layout pattern is the Z-diagram layout [10], where readers follow the shape of the letter z while scanning quantitative information. In this regard, we define our visualization element placement strategy in our approach. Moreover, to achieve consistency in dashboards using colors and resizing visualization elements, there are guidelines in the literature [42, 52, 54, 92, 97, 134]. Since these visual aspects of dashboard development are not the main focus of our approach, we use an embedded mechanism, i.e., a fixed set of colors and size values.

5.3 Approach

This section elaborates on our approach for the automated generation of engaging dashboards. The procedure to automatically generate engaging dashboards consists of two tasks, as introduced in Section 5.1: (1) developing the decision model for visualizing

KPIs and (2) generating dashboards automatically using the decision model. The second task is automated and takes a set of KPIs with attributes and the values of these KPIs as inputs in addition to the decision model itself. KPIs with attributes and values are taken as input in a "machine-readable" format. For this, human involvement is required. To reduce that human involvement, KPIs with attributes and values are desired to be defined such a "machine-readable" format that enables their automated analysis and computation as proposed in [37].

Unlike the second task, the first task is not automated in our approach. The reason for that is the available dashboard design principles in the literature are in the form of natural language texts. Thus, human interpretation is required to develop a decision model for visualizing KPIs using those principles [42, 51, 52, 54, 79, 92, 93, 97, 134, 158]. However, this task only needs to be performed once. The decision model created as its output, and presented in this chapter, can be re-used in any scenario for the automated generation of engaging dashboards. In particular, it is possible to prune or extend the decision model for a given set of KPIs of a certain organization, which is part of the second step, the automated dashboard generation. In this sense, the amount of human involvement required will highly depend on the way KPIs are defined, i.e., the amount of information provided for them in their definition and its correspondence with the attributes required by our approach. In this context, we now explain how we developed the decision model for visualizing KPIs, and then give the details of the automated dashboard generation task.

5.3.1 Developing the Decision Model for Visualizing KPIs

As explained in Section 5.2, we identified the most prominent sources [42, 51, 52, 54, 79, 92, 93, 97, 134, 158] for dashboard design principles. Using these sources, we construct a decision model for visualizing KPIs, which is shown in Fig 5.3 and encoded as such in our approach. We explain how we construct the decision model by listing our considerations below.

A typical KPI may have a single value or a set of values as quantitative information. For example, the total revenue of an organization or the total revenue of each department within an organization. We take this attribute of KPIs as the top decision point of the decision model (see ① in Figure 5.3). Then, we determine how a KPI with a single value and a KPI with a set of values should be visualized using the most common types of visualization elements, namely tables and graphs.

When a KPI has a single value, bar graphs better convey the message of that KPI [42, 51, 52, 54, 92, 97, 158]. Since a KPI must have a "*target*", that target needs to be displayed in a graph together with the value of the KPI. This can be achieved in the most informative way using a Bullet Graph [42, 51, 53, 54, 92, 93] since it is a special, simplified bar graph and is designed for visualizing a value along with a comparative measure to enrich the meaning of the value.

If a KPI has a set a of values, then we need to determine the "*purpose*" of the KPI. That purpose can be taking the attention of a decision maker to "*look up*" the values or "*revealing the relationship*" between the values for the decision maker (see ② in Figure 5.3). Tables are the visual elements that are designed to look up values [42, 51, 52, 54, 92, 97, 158]. While constructing a table, it is important to emphasize how individual values in a table relate to the target of the KPI, which is visualized. It is recommended to use a table where the values

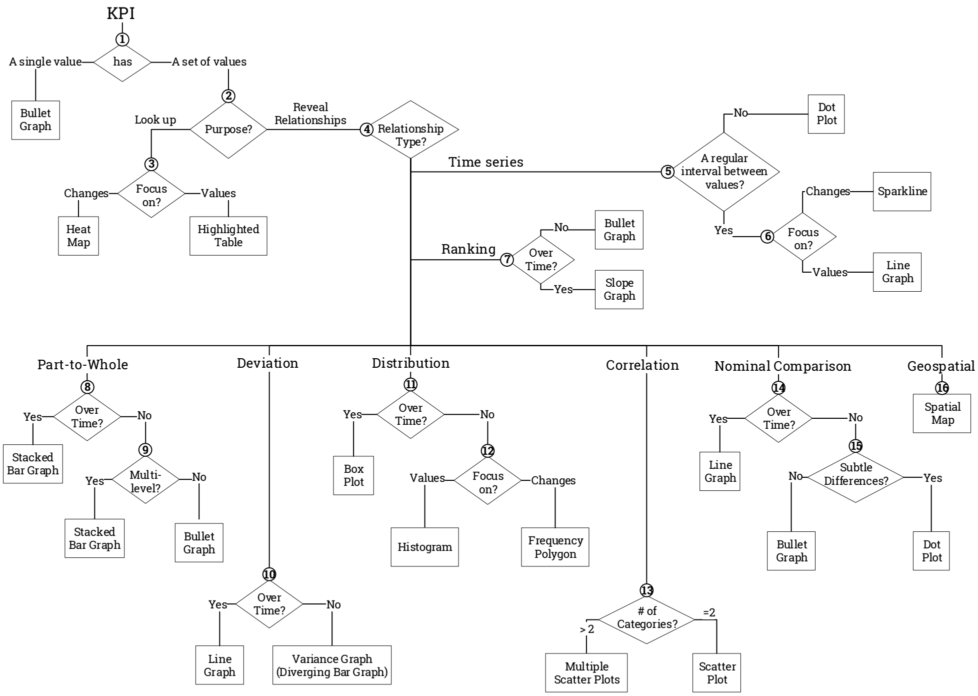

Figure 5.3: Decision model used for visualizing KPIs

of the KPI represented are highlighted with colors according to the fulfillment of its target value, e.g., red if not fulfilled and green if fulfilled [42, 51, 54, 92, 97, 158]. However, in addition to purpose, while looking up values, a KPI may require decision makers to focus on the "*changes of values*" rather than "*individual values*" (see ③ in Figure 5.3). This can be achieved by using a Heat Map. In a Heat Map, values are represented by colors, and one can easily determine precise individual values using the color scheme if needed.

When the purpose of a KPI is to reveal the relationship between its values to decision makers, graphs are used. To determine what graphs are particularly useful for specific relationships (see ④ in Figure 5.3), we take the relationships as the base that we listed in Section 5.3, and then describe how each relationship can be visualized such that decision makers will be engaged in.

Time series: Although Line Graph is the commonly used graph for visualizing a time-series relationship [42, 51, 54, 79, 92, 93, 97, 134, 158], connecting the data points representing values as a line will cause a misleading communication when the values are not collected "*at a regular interval*". Dot Plot deals with that problem by displaying a time-series relationship in the form of points in which missing values are not displayed. Furthermore, a KPI may aim decision makers to "*focus on the history of changes in values*" over time instead of the values over time (see ⑤ and ⑥ in Figure 5.3). To this end, there is a special graph, namely Sparkline [51, 52, 54], which provides a simple and quick view of the history of changes in values at a glance to determine whether there is anything unexpected. Although bar graphs and area graphs are commonly used for this relationship, they miserably fail to show changes over time [51, 52, 54], and they especially clutter the display when values are categorized or benchmarked against various comparative measures, e.g., target or forecast.

Ranking: Since a KPI must have a target, a Bullet Graph will perform better than classical Bar Graphs [42, 51, 53, 54, 92, 93] at displaying values along with a comparative measure. If "*changes in rankings over time*" are important, Slope Graph outperforms among other graphs [51, 52, 54, 134] since it focuses on the evaluation of rankings between two or more points in time (see ⑦ in Figure 5.3).

Part-to-whole: Although Pie Graphs are quite often used for visualizing part-to-whole relationships, they are listed in the graphs to avoid [42, 51, 52, 54, 92, 93, 97, 134, 158] due to several reasons. One of the reasons for that [51, 52, 54, 92, 97] is that view will be cluttered when there are many slices, and many of them have similar sizes. Another reason is the common practice of creating a slice named as "others," which mostly causes misleading interpretations. In addition, as seen in many examples [42, 51, 52, 54, 92], the total of slices is not checked correctly, e.g., total does not add up to 100. To overcome these problems, bars are recommended [42, 51, 52, 54, 92, 93, 97, 134, 158] to encode values. While displaying a "*part-whole-relationship over time*" Stacked Bar Graphs outperform among other bar graphs since they will not require the duplication of each proportion for each time unit, which causes a cluttered view [51, 52, 54]. When a KPI is solely about a part-to-whole relationship with no time involvement, it is required to check whether a "*multi-level*" hierarchy exists between values (see ⑧ and ⑨ in Figure 5.3). For example, the revenue of an organization may be the aggregation of the revenues of its branches, and even the revenue of each branch may be the total of the revenue of various departments. When there is a multi-level hierarchy in a part-to-whole relationship, we select Stacked Bar Graphs. Otherwise, Bullet Graph outperforms than classical Bar Graph displaying a KPI with its target. In Stacked Bar Graphs, the target of a KPI can be displayed using lines with a secondary axis.

Deviation: Since Diverging Bar Graphs perform well in visualizing a deviation relationship [42, 52, 54, 97, 158], and there are no competing alternatives, we select them for visualizing the KPIs that reveal a deviation relationship. However, to see a "*deviation relationship from a time-perspective*", i.e., how deviations evolve, the Line Graph stands out as the best option [51, 52, 54, 93, 97, 158] due to its power of showing things over time in a simple way (see ⑩ in Figure 5.3).

Distribution: Although Box-Plots are common in visualizing distributions, interpreting Box-Plots requires specific statistic knowledge [51, 52, 54]. To take actions based on the displayed relationship, decision makers will prefer simple graphs that require less effort [42, 51, 52, 54, 92, 134]. As the Histogram is a special type of bar graphs and bars are easy to understand by everyone, we select the Histogram (see ⑫ in Figure 5.3) as the graph to visualize KPIs when the focus is on values across the range of distribution. If the "*changes of the shape*" of distribution are the main focus, Frequency Polygon (see ⑫ in Figure 5.3) outperforms than Histogram [42, 51, 52, 54, 92, 93, 97, 134, 158]. Moreover, when a distribution relationship needs to be displayed "*over time*", we select to use Box-Plots (see 11 in Figure 5.3) since others will cause a cluttered view due to the duplication of the range of the distribution [42, 51, 52, 54, 92, 93, 97, 134, 158].

Correlation: Although Scatter plots perform quite well visualizing a correlation relationship, an increase in the number of categories will make a Scatter Plot very complex. As the biggest negative effect of this increase, the readability and interoperability of a Scatter Plot will dramatically decrease since adding categories will hinder some values beyond others. Although circles are used to support the added categories in scatter plots [51, 52, 93, 97, 158], they overlap and decrease the understandability when values are closer to each other. Multiple Scatter Plots can be used. In this way, when a correlation between two categories is to be displayed, scatter plots are used. Otherwise, we propose multiple scatter plots to be used. A Multiple Scatter Plot consists of a number of scatter plots where each scatter plot displays the correlation in the values of two categories (see 13 in Figure 5.3). In each scatter plot, the target of a KPI can be visualized using lines.

Nominal comparison: As discussed in ranking relationship, instead of Bar Graphs, we select the Bullet Graph due to its simplified and beneficial view where bars best encode a particular relationship. However, if bars become similar in length, detecting the subtle differences between them can become difficult. To capture these subtle differences, a Dot Plot is the most effective alternative in which the scale has no longer need to start at zero, which is a must [42, 51, 52, 54] for Bar Graphs (see 14 in Figure 5.3). When the aim is to display a set of values on a categorical scale over time, bars fail since they cause a cluttered view [42, 51, 52, 54, 92, 93, 97, 134, 158] by duplicating each discrete value for each time point. For that reason, we select the Line Graph to show a nominal comparison relationship over time, where a separate line represents each discrete value (see 15 in Figure 5.3).

Geospatial: The de-facto way of displaying a geospatial relationship is using a map called Spatial Map (see 16 in Figure 5.3) and no criticism has been found in this regard [42, 51, 52, 54, 79, 92, 93, 97, 134, 158].

To execute the developed decision model, it is required to provide the KPI attributes that map to the decision points of the decision model and identify which visualization element needs to be used. In this regard, first, we identified what attributes of KPIs are taken into account while visualizing KPIs within the described dashboard design principles in the literature. Then, we transformed the identified KPI attributes into a single set. Finally, we checked the completeness of the identified KPI attributes against the developed decision model. This check is conducted by controlling the existence of mapping both from each KPI attribute to the decision points in the decision model and vice versa. The identified KPI attributes are listed in Table 5.1.

5.3.2 Generating Dashboards Automatically

To generate dashboards automatically in our approach, a set of KPIs with attributes, their values, and the decision model for visualizing KPIs are needed as inputs. These attributes

Table 5.1: KPI attributes required by the decision model

KPI Attribute	Definition
Relationship type of values	Describes how the values of the KPI are related. For example, time series, correlation, ranking, part-to-whole, nominal comparison, or distribution.
Purpose	Whether the KPI is about looking up its values or revealing the relationship between its values.
Focus	Describes what is the focus of the KPI with respect to its purpose attribute. Example values: look up-changes, look up-values, relationship-changes in a time series, relationship-values in a distribution.
Time interval	Whether the KPI needs to be displayed over time.
KPI values	Describe the quantitative information of the KPI.
Categories	The discrete groups in which one or more values exist. For example, the total revenue is a KPI that has a single category, which contains a single value. However, the revenue per department is a KPI that will have a category for each department.
Sort direction	Describes how the categories or the values in a category will be ordered, e.g., ascending or descending. This is especially important in ranking and distribution relationships.
Multi-level hierarchy	Whether there is a hierarchy or main-sub grouping in the categories attribute of the KPI. For example, main group: region and sub-group: county.
Regular interval between values	This will be determined using the attribute time interval. If there is any missing value in the values of the KPI based on its time interval, the branch “No” will be selected in the related decision point of the decision model.
Subtle difference threshold for the values of the KPI	Describes the limit of the difference between the values of the KPI that should be clearly detectable at a nominal comparison.

are common attributes described when defining KPIs [37]. The approach determines what kind of visualization element will be utilized for each KPI by applying the given KPIs with their attributes on the given decision model. In particular, a mapping from the decision points in the given decision model is searched for the given KPIs with attributes. The approach completes this search when a visualization element for each KPI is determined. Then, each determined visualization element is created and placed in dashboards, as shown in Fig 5.4.

In addition to the KPI attributes listed above as required by the decision model, the approach uses a set of KPI attributes while creating dashboards and displaying them according to the values of KPIs. Those KPI attributes are listed in Table 5.2.

To determine how many dashboards need to be created, we defined a strategy so-called Dashboard Creation Strategy in the approach. The strategy is based on the relations between the KPIs of an organization. More specifically, the KPIs that are related to the business processes in a particular process area will be grouped and placed onto a particular dashboard. For example, the KPIs related to the sales process and the KPIs involved in the purchasing process of an organization are combined into the dashboard, Order Management. In addition, the KPIs about creditors and debtors are grouped into the dashboard, namely the Finance dashboard.

The creation of each determined visualization element for a KPI consists of four tasks:

Table 5.2: Additional KPI attributes required for visualizing dashboards

KPI Attribute	Definition
Process area	The category of the business process that is related to the KPI. This attribute is used for determining the number of dashboards that will be created. Example values: Order Management, Finance.
Target	The value or value-range that needs to be achieved with respect to the related strategic goals of the organization. Example target values: zero, at minimum €50K, a reduction of 10%, precisely 7 days, or within $1-3$ days.
Target Thresholds	The set of value-range that shows to what extent the target of the KPI is achieved. Each threshold has a lower and upper bound value. For example, good: [KPI target-10K, KPI target-30K], bad:[KPI target-30K, KPI target-50K]
Human resource interest	Represents the interest of the human resources in the KPI. It can have a value of Responsible or Informed. This attribute is used within the dashboard split strategy of the approach.
Name	The textual description used to define the KPI.
Unit	The quantity used as the standard for the measurement of the KPI's values. Although this is not important at determining visualization elements, it is essential to convey an informative message via KPIs.

(1) creating the visualization parts for the target thresholds of the KPI on the visualization element, (2) creating the visualization parts for the KPI values, (3) creating the visualization parts for the target of the KPI, and finally (4) combining all parts as a single visual element, e.g., graph or table. In the first task, the approach creates bars and arranges them with respect to the boundaries of target thresholds. Then, in the second task, the approach creates the visualization part in the form of bars, dots, or lines using the KPI values. These forms depend on the determined visualization element. Depending on the type of the target of the KPI (e.g., achievement, reduction, absolute, zero, or min-max), the visual signs that indicate the target are created as a visualization part in the third task. In addition, in the third task, noticeable alerts that indicate whether the KPI and its categories are on target or not (see the cross-mark, check-mark, and warning signs used as alerts in Fig 5.5) are created. In the last task, the approach combines all these visualization parts as a single visualization element considering the embedded coloring [2], orientation, and resizing rules for visualization in it. How many categories should be visualized in graphs is determined using an implicit, configurable parameter in the approach. The reason for that is to determine the orientation (horizontal or vertical), which increases the readability. For example, a ranking relationship better reads when it is horizontal and has a maximum of 10 categories where the rest is grouped as "others."

Similarly, to determine how created visual elements will be placed on dashboards is determined using the strategy, Dashboard Split Strategy, that we defined in the approach. In this strategy, the approach creates a flow through a combination of visual weight and

[2] `http://colorbrewer2.org` is used as the source for color selection.

visual direction to take advantage of how people read through a design. The created flow splits a dashboard into two areas: top and bottom. By applying the most common layout pattern (the Z-diagram layout [10]), which is recommended for simple designs, the approach defines the route that the human eye travels on these areas: left to right and top to bottom. To determine the order of the KPIs that the human eye should read in this travel, the approach uses the KPI attribute "human resource interest." The KPIs that have the value "Responsible" in their "human resource interest" attribute will be placed to the top area of dashboards. Then, the KPIs that have the value "Informed" in their "human resource interest" attribute will be placed to the bottom area of dashboards. The KPIs will be ordered in an ascending order based on their names in each area unless there is a particular ordering for displaying KPIs, such as the relevance of KPIs for decision makers.

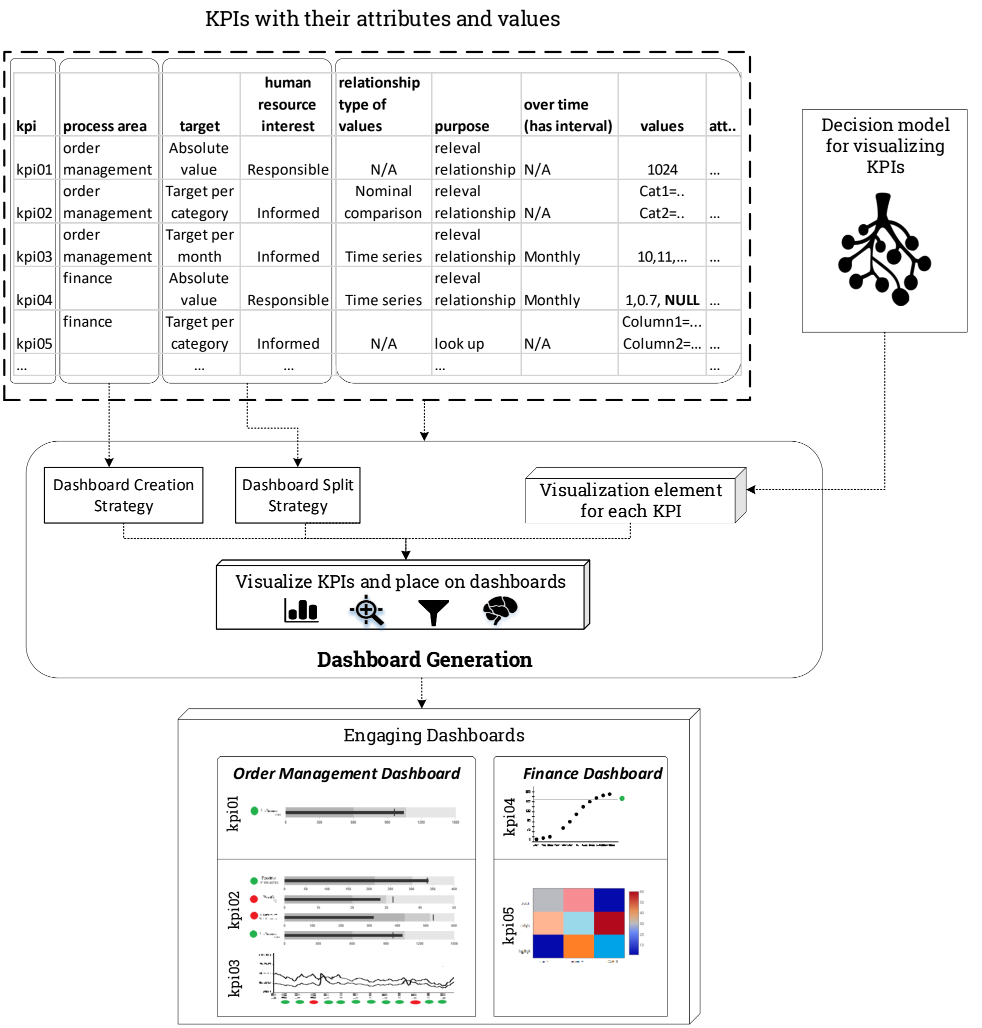

Figure 5.4: Generating engaging dashboards for organizations

In the next section, we give the details of the evaluation of both the approach[3] and the decision model for visualizing KPIs.

5.4 Evaluation

In this section, first, we explain how we evaluated the decision model for visualizing KPIs, which was described in Section 5.4. Then, we elaborate on the evaluation of the proposed approach in a case study.

5.4.1 Evaluation of the Decision Model

We evaluated the decision model within two organizations, *A* and *B*, for confidentiality reasons. We did so by discussing our considerations and walking through each path in the decision model together with the experts in these organizations (for more details on their background, see Table 5.3 and Table 5.4). The experts with whom we worked together are actively involved in the dashboard development process in their organizations. We collected their opinions about the decision model using the three-points Likert-type scale (agree, somewhat agree, and disagree). While collecting experts' opinions in each organization, we had an open discussion meeting on the usefulness of the decision model to the needs of each organization at dashboard development. In particular, we gathered opinions related to two aspects of the decision model: 1) decision points and 2 visualization elements. Then, an average value is calculated for each organization using the collected expert opinions. That average value shows to what extent the decision model is useful to the needs of an organization at visualizing KPIs.

Table 5.3: Evaluation of the decision model in organization *A*

Expert	**Area of expertise**	**Years of expertise**	**Meeting duration (hours)**	**To what extent agree on the decision points in the decision model**	**To what extent agree on the visualization elements in the decision model**
Software Architect-1	Dashboard development	>5	1	Agree:13 Somewhat agree:2 Not agree:0	Agree:20 Somewhat agree:3 Not agree:0
Software Architect-2	Product management and Information visualization	>20	1	Agree:13 Somewhat agree:2 Not agree:0	Agree:20 Somewhat agree:3 Not agree:0
Manager	Product management and Dashboard development	>15	1	Agree:12 Somewhat agree:3 Not agree:0	Agree:19 Somewhat agree:4 Not agree:0

[3]The implementation of our approach for the automated generation of engaging dashboard is available at `http://amuse-project.org/software/` . In the implementation, two *Python* libraries are required: *Plotly* is for visualizing KPIs and *Dash* is for creating dashboards.

Organization *A*: In order to automatically generate ERP software from a model, a Dutch ERP software vendor is developing a novel model-driven software generation approach. As part of that approach, a declarative modeling language is being developed that is aimed at modeling an organization's business in the form of an ontological enterprise model. In order to build dashboards automatically for its client organizations with the power of that declarative modeling language, this company is currently investigating how KPIs can be automatically visualized. Since this is highly related to the approach that we propose, in this company, we evaluated the decision model. The details of the evaluation of the decision model in this organization are listed in Table 5.3. As shown in the table, the experts, on average *agree* on the decision points and also on the visualization elements in the decision model. Only for a minority of the decision points and the visualization elements in the decision elements they indicated their partial agreement, i.e., somewhat agree. In particular, the experts shared their partial agreement for the following decision points in the decision model: 6, 12, and 15. Similarly, for the visualization elements at these decision points, the experts shared their partial agreement. The experts did not mention any disagreement for the decision points or the visualization elements in the decision model.
Organization *B*: To monitor the usage of physical resources, the IT department of a Dutch bank uses a dashboard. This dashboard consists of a set of KPIs in which particular psychical resources are monitored with respect to their response rates. A performance management expert maintains that dashboard in accordance with the change requests coming from the performance monitoring chapter lead of the IT department. We followed the same procedure that we explained above for the evaluation of the decision model in this organization. The details of the evaluation of the decision model in this organization are listed in Table 5.4. As depicted in the table, the experts agreed on average more than 70% of the decision points and also the visualization elements in the decision model. In addition, only for one visualization element in the decision model a disagreement is mentioned. This was for the Box Plot graph that is identified for visualizing a distribution relationship over time. As in the evaluation in Organization A, we received partial agreement feedback in organization *B* for the decision points 6, 12, and 15 in the decision model.

Table 5.4: Evaluation of the decision model in organization *B*

Expert	**Area of expertise**	**Years of expertise**	**Meeting duration (hours)**	**To what extent agree on the decision points in the decision model**	**To what extent agree on the visualization elements in the decision model**
Performance Management Expert	Dashboard development	>20	1.5	Agree:12 Somewhat agree:3 Not agree:0	Agree:18 Somewhat agree:4 Not agree:1
Performance Monitoring Chapter Lead	Dashboard design and monitoring	>15	1.5	Agree:12 Somewhat agree:3 Not agree:0	Agree:18 Somewhat agree:4 Not agree:1

In both organizations that we evaluated the decision model, *agree* is the calculated average value of the usefulness of the decision model for the needs of organizations.

5.4.2 Evaluation of the Approach

The proposed approach was evaluated in the first organization, which we evaluated the decision model, namely organization *A*. In that organization, the three aforementioned experts developed a finance dashboard template to create dashboards for the client organizations of the company. Together with the same three experts whom we worked on in the evaluation of the decision model, we created a sample dashboard using that template. After that, the created sample dashboard was used as the *existing dashboard* in the evaluation of the approach. We executed our approach for the KPIs, in total 8, contained in the existing dashboard, and created a new dashboard. Then, together with the aforementioned three experts, we evaluated our approach by comparing the existing dashboard with the *newly generated dashboard*. The results that we obtained are explained below.

As shown in Fig 5.5, the KPIs that have implicit target values on the existing dashboard, namely KPI-1, KPI-2, KPI-6, and KPI-7, are displayed differently in the newly generated dashboard than the existing dashboard. These KPIs are visualized as bullet graphs in the newly generated dashboard since each of them has a single value. Moreover, the target values of these KPIs are also highlighted using a diamond shape; corresponding alerts are added next to each KPI based on those target values. Besides, the approach visualized KPI-4 as a dot plot as that KPI aims to display a nominal comparison relationship. To do so, whether the differences between values are subtle is checked with respect to the given subtle difference threshold by the experts.

Furthermore, the approach visualized 3 out of 8 KPIs, namely KPI-3, KPI-5, and KPI-8, slightly different than the existing dashboard. These KPIs are visualized in the newly generated dashboard in a way such that each KPI conveys its intended message clearly. More specifically, since KPI-3 is about a nominal comparison and has no subtle differences between its values, this KPI is visualized as a bullet graph. As KPI-5 presents values over time, the visualization element for that KPI has not changed. The only change is the addition of its target. KPI-8 reveals a ranking relationship, and it is visualized as bullet graphs accordingly. For each of these three KPIs, the approach added a noticeable alert next to each graph to indicate target achievement.

Since there were no defined values in the human resource interest attribute of the displayed 8 KPIs, the KPIs are displayed in the top area of the newly generated dashboard and placed in orange boxes, which is the assigned color for the value “responsible” of human resource interest attribute.

To determine the implications of the differences between the existing dashboard and the newly generated dashboard in the evaluation of the approach, we had an open discussion meeting with the experts who were involved in the evaluation of the approach. The experts confirmed that visualizing KPI-1, KPI-2, KPI-6, and KPI-7 as bullet graphs in the newly created dashboard helps them to observe a KPI along with its comparative measure, e.g., target. Similarly, the experts agreed on visualizing KPI-4 as a dot plot as the subtle difference become visible to make precise decisions. In addition, as KPI-1 and KPI-6 are visualized differently in the newly generated dashboard than the existing dashboard and become noticeable, the experts decided to check the need for these KPIs. Besides, the experts agreed both on the usefulness of the displayed target thresholds (e.g., perfect, good) and the alerts (cross-mark, check-mark, and warning signs) that display whether

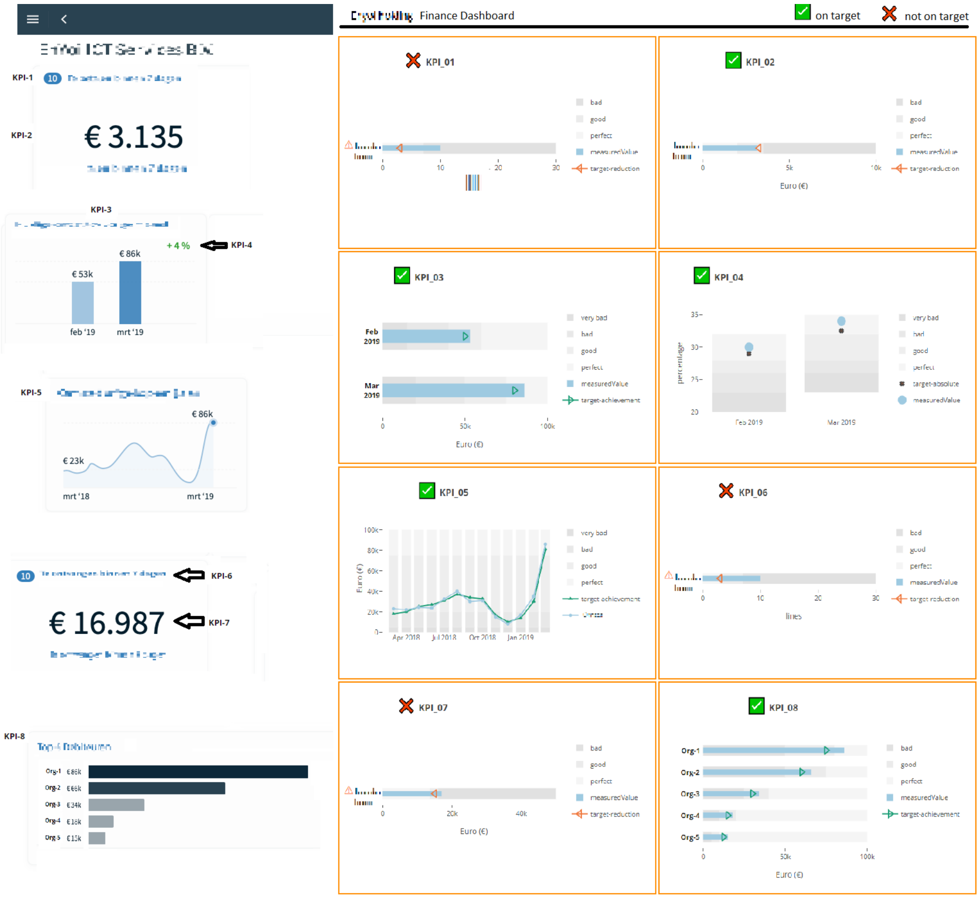

Figure 5.5: Existing dashboard (left) vs. newly generated dashboard (right)

the KPIs are on target and to what extent at a glance.

In the following section, we discuss the results that we obtained in the evaluation of both the decision model and the approach.

5.5 Discussion

Regarding the evaluation of the decision model we proposed, as listed in Table 5.3 and Table 5.4, we obtained a partial agreement for the following decision points in the decision model: ⑥, ⑫, and ⑮. Similarly, for the visualization elements at these decision points, we obtained a partial agreement. The experts who are involved in the evaluation of the decision model expressed that the difference between the visualization elements at these decision points is not big since these decision points are rarely investigated in their organizations while developing dashboards. This shows that our decision model has a wide coverage of decision points, considering even the less common scenarios. Moreover, the experts mentioned that the visualization elements in those decision points, namely Slope Graph, Frequency Polygon, and Bullet Graph, are very simple and useful. However, the experts noted that these graphs are not completely supported in most business intelligence software products, although they are not very new. This means that our decision model helps organizations to determine simple and useful visualization elements for creating engaging dashboards.

Furthermore, there was no decision point that the experts neither in Organization A nor in Organization B disagree. However, only for one visualization element in the decision model, we received a disagreement in Organization B. This was about the Box Plot graph that is used for visualizing a distribution relationship over time. Since the Box Plot graph requires particular knowledge of statistics to interpret the message of it, especially for decision makers who are not familiar with the Box Plot graph grasping the conveyed message with it may not be easy. However, although there are alternatives, they have more drawbacks

As for the evaluation of the automated approach itself, since the newly generated dashboard enabled organizations to check and eliminate the KPIs that are not often a source for decision making in their organization, in that respect, the approach helps organizations to focus on the KPIs that are relevant to their business.

The results of the evaluation confirm that the approach proposed in this paper is of sufficient quality to show its practical usage. On the one hand, as we observed in the evaluation, the newly generated dashboard by the approach can help organizations to clearly observe KPIs along with their comparative measures. On the other hand, the approach enables organizations to focus on the message conveyed via KPIs with engaging visualizations, which is the main substance for wise decisions. Moreover, organizations can detect whether any KPI is not a good source for decision making and avoid misleading communications.

Software vendors that focus on automatically generating dashboards for their client organizations can apply our approach. To do so, these software vendors need to provide KPIs with attributes and the values of these KPIs to our approach in addition to the decision model, which is already encoded in the approach. For obtaining these required inputs, organizations may leverage formal notations for defining KPIs, such as PPINOT [37]. Using them, on the one hand, organizations can reduce their management efforts on KPIs since these formal KPI definitions enable their automated analysis and computation. On the other hand, formally defined KPIs can be integrated into our approach for facilitating the automated generation of dashboards.

One of the limitations of the approach is the decision model development task since it is not automated. In addition, our approach is limited to visualizing KPIs as tables and graphs. However, in the literature, there are principles on how to combine multiple visual elements as a single visual element, e.g., multi-panes for visualizing a KPI. Additionally, the visual aspects of visual elements such as fonts, coloring, responsiveness are not covered so far.

5.6 Related Work

In this section, we list some of the works that relate to the approach we proposed for the automated generation of engaging dashboards.

A model-driven dashboard development approach is proposed in [31] to automatically create dashboards with the code necessary for their deployment. To create a dashboard automatically, the approach requires a dashboard user to model both the dashboard and the related KPIs. Then, an engine executes the model and creates the dashboard with its code. To handle the change management of dashboards, the approach is enriched with the observers [100] who manage the maintenance of dashboards. However, to derive dashboards automatically using this approach, organizations need to have the intensive knowledge required to visualize KPIs in an effective way and should model each dashboard using the notation in the approach. Since these tasks are manual, the approach will require a significant effort of every organization that wants to apply it.

To create customized dashboards automatically considering the requirements of different users, Vázquez-Ingelmo et al. have proposed an approach that uses domain engineering practices [148]. Based on the analysis of the similarities and differences between users' requirements and existing dashboards, a feature model is constructed. The constructed feature model specifies what visualization will be created within a dashboard. Since the approach uses existing dashboards as a base, for each existing dashboard users' requirements need to be obtained. Therefore, using this approach, each organization will need to spend a great effort in addition to the effort for internalizing dashboard design principles.

To visualize KPIs for production planning on a BPMN model in the manufacturing domain, Heidema et al. [66] present an approach. Based on the dashboard design principles defined in [52], the applicability of visualizations is determined in the context of BPMN. Then, a set of KPIs for the manufacturing domain, which is listed in ISO 22400 is automatically visualized on a BPMN model. However, the visualization elements that are supported by the approach are limited. For example, the values in a time series relationship cannot be seen since only sparklines are used, which do not contain values. Moreover, some relationship types are not covered, such as part-to-whole, distribution, correlation, and geospatial. Displaying the values over time in a various relationship is not addressed comprehensively within the approach. Since the approach is dependent on BPMN models, adding visualization elements on the relatively large BPMN models with numerous elements will clutter the view and distract decision makers.

Koetter and Kochanowski [78] proposed a modeling language, namely ProGoalML, to build a monitoring infrastructure automatically for KPIs. The language enables organizations to model their KPIs in their business process models, i.e., annotate the business process models using the language components proposed. Kintz [75] has proposed a dashboard design methodology that can transform the inputs created by ProGoalML to formal KPI definitions, which are required as inputs to derive dashboards automatically by the dashboard engine—a component of the proposed methodology. Kintz et al. have extended the proposed methodology by adding support to create the dashboards that are customized to users [76]. To determine how KPIs should be visualized, from data types to visualizations 4 mappings are employed within the aforementioned methodology. However, in these mappings, it is unclear what the data types imply, i.e., how the data type of a KPI can be determined is not explained. Furthermore, some important relationships in quantitative information, e.g., correlation, ranking, are not covered. Additionally, a data type is mapped to two visualizations in a mapping, which causes ambiguity.

As explained above, to develop the dashboards that are communicating important information and are engaging, each organization, first, has to internalize dashboard design principles, and then apply them. However, this is time-consuming and costly. To provide a solution to these problems, we proposed an approach for the automated generation of engaging dashboards.

5.7 Conclusion and Future Work

In this chapter, we presented a novel approach aimed at the automated generating of engaging dashboards for organizations by means of automatically visualizing KPIs. A set of KPIs with their attributes and values and a decision model developed for visualizing those KPIs are the required inputs by the approach. The approach determines which visualization elements (a table or a graph) will be used to visualize each KPI using the given decision model. The approach creates the dashboards based on the dashboard creating strategy encoded in it, and then places the built visualization elements on dashboards using the dashboard split strategy, which is also encoded in it. Since the available dashboard design principles are not in the machine readable form, we described how a decision model for visualizing KPIs can be devised.

To evaluate our approach, we conducted two tasks: an evaluation of the developed decision model and the evaluation of the created dashboards using our approach. The former was carried out in two organizations: an ERP software vendor and a bank. The latter was done with the ERP software vendor. In both tasks, we conducted the evaluation by informal interviews with the experts in the organizations who are actively involved in dashboard development. As a result, we showed that the approach enables organizations focusing on the messages conveyed via KPIs with engaging visualizations to make informed decisions for improving the performance of their organizations. In most recent approaches, visualizing KPIs is a manual endeavor and needs to be carried out in every single organization. Thus, we feel confident to that our approach lowers the efforts of software vendors for developing engaging dashboards for their client organizations and the efforts of these organizations doing this themselves.

In future work, we want to extend our approach by adding predictive technologies. In particular, we want to predict the values of KPIs such that decision makers can take preventive actions instead of corrective actions, which costs more to improve the business processes in their organizations. Moreover, to provide insights for organizations by means of the benchmarks that are developed using the relevant KPIs for them, we plan to integrate this approach with the approach we presented in [5]. In addition, we will add the support for automatically providing the consistency of visual aspects (e.g., colors, size, and spacing) in dashboards.

Chapter 6

Interactive Generation of Process Performance Dashboards

Abstract
Process performance measurement is essential for any organization that wants resilience in its business process management and aims to make process improvement meaningful. Regarding that, organizations use *Process Performance Indicators* (PPIs) to track and identify the success of their business processes. To provide automated support for the management of PPIs, there exists a well-established solution in the related literature, namely the *PPINOT* metamodel. The metamodel allows advanced definition and instrumentation of PPIs as well as automated analysis of PPIs. Nonetheless, significant efforts both from software vendors and organizations are required for visualizing PPIs to develop *engaging* process performance dashboards. The reason for that is the dashboard design principles, i.e., best practices, available in the literature are expressed as natural language texts and they need to be internalized and applied literally in every dashboard development process. Moreover, there is a prominent lack of support in current Business Intelligence (BI) systems in consolidating such best practices. Hence, to build such dashboards by means of automatically visualized PPIs, we propose an approach for the interactive generation of engaging dashboards. The PPINOT metamodel is integrated into the approach for the automated analysis of PPIs such that they can then be automatically visualized. For this purpose, the approach uses a decision model that we developed based on the dashboard design principles in the literature. Moreover, the approach employs a chatbot to interactively generate dashboards. The chatbot interacts with decision makers to collect the required inputs for creating dashboards. Thus, decision makers can interactively build dashboards without needing substantial knowledge on dashboard development and visualizing PPIs. We applied our approach in a case study and evaluated its usefulness. Our findings show that the approach can assist experts in making informed decisions through interactively created engaging dashboards.

keywords- chatbot, dashboard, PPINOT, process performance indicators (PPIs), visualization

This chapter is based on the following manuscript:

[8] Ü. Aksu, A. del-Río-Ortega, M. Resinas, and H. A. Reijers. Interactive generation of process performance dashboards. This manuscript is submitted to a journal, 2021.

6.1 Introduction

Process performance measurement is vital for business process management and improvement. What has to change in a process can only be identified if it is known whether a process performs as expected [39, 40, 133]. Simply put, the progress of a process towards its goals cannot be tracked unless that progress is measured, i.e., quantified [14]. With measurement, organizations can have a better understanding of the performance of their processes in attaining their goals. To determine whether their processes perform as desired, organizations use Process Performance Indicators (PPIs) that provide, at a glance, an overview of their business performance. A PPI is a metric [37] that reflects to what extent a business process reaches its goals. For example, the percentage of unresolved problems within a period of time in comparison to a defined target is a typical PPI used for tracking the performance of a problem management process.

Organizations use process performance dashboards to monitor their PPIs [37, 118, 160]. Such dashboards are generally developed either from scratch or by customizing the dashboards offered by software vendors. Within a typical process performance dashboard, PPIs are consolidated into a single visual display to assist in decision making [42, 54]. In this regard, each PPI that provides the information needed by decision makers is often displayed as a table or graph to grab their attention through visual attractiveness.

However, most dashboards are poorly designed displays and not useful despite the sophisticated data management and graphical abilities of the Business Intelligence (BI) systems that are used while developing them [50, 54, 106, 158]. Moreover, to impress its users, plenty of visualizations are often combined in dashboards; in turn, it becomes difficult for its users to concentrate on the essential information. Simply put, decorations hinder the substance that is the key to making informed decisions. Hence, most dashboards fail to communicate efficiently and effectively [42, 50, 54, 70, 97, 106, 158]. An example dashboard[1] for monitoring the performance of the incident management process in an organization is depicted in Figure 6.1. Clearly, one can see that the dashboard goes against the dashboard design principles in the literature [6, 106]. The dashboard has several significant issues. For example, pie charts have many slices that make them unreadable. Inconsistent and distracting colors cause misleading associations. More importantly, that dashboard does not reflect the overall status of the related process because of the cluttered design in which an overload of information is squeezed. Therefore, decision makers need to spend substantial effort to identify the messages that the dashboard is designed to convey. As a result, this dashboard is not "engaging" decision makers to make informed decisions that can yield improvement [14, 50].

Much work has been conducted on developing engaging dashboards that can communicate important information efficiently and effectively. Notably, numerous researchers developed guidelines [1, 42, 51, 52, 54, 79, 91–93, 97, 132, 134, 158, 162]. Within these guidelines, principles for visualizing quantitative information in dashboards, the so-called dashboard design principles are described. Specifically, what visual representations (e.g., various graphs) should be used and how they should be used are expressed in such dashboard design principles. Dashboards engage decision makers if the available dashboard design principles are used when creating them. The main reason for

[1] The example dashboard is taken from `https://adniasolutions.com/dashboard-design-principles/introduction-to-dashboards/`, Last accessed Aug. 2021

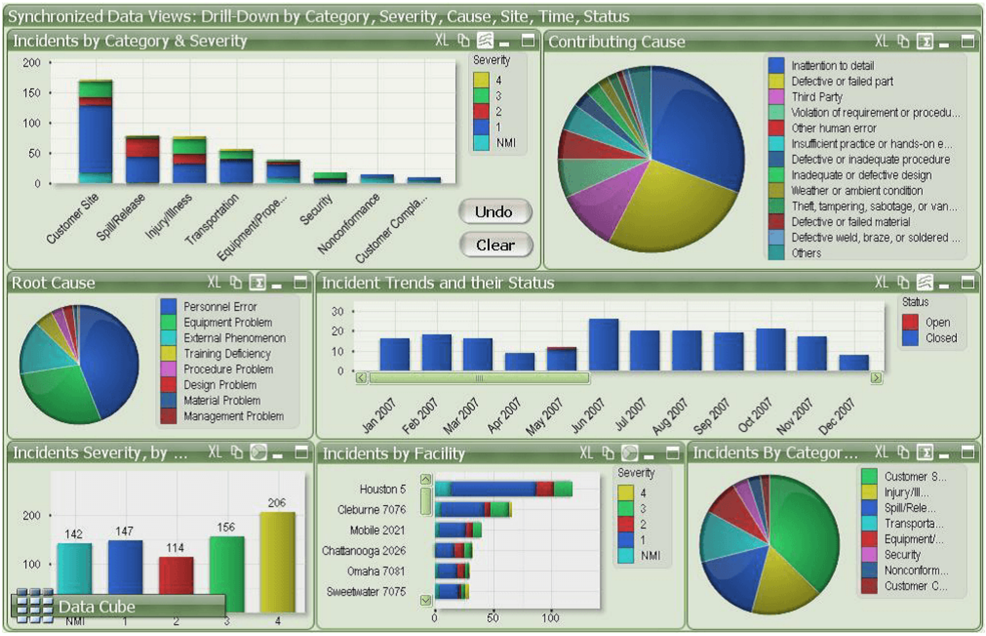

Figure 6.1: An example of a *non-engaging* dashboard

that is that engaging dashboards enable decision makers to sense and process the displayed information rapidly through visualization elements. Moreover, performance indicators on engaging dashboards are visualized in such a way that it inspires actions without overloaded information that causes distractions. Furthermore, "Are we on track?" and "How well is our organization performing its business?" are such questions in organizations to which complete answers can be obtained intuitively in engaging dashboards. Simply put, engaging dashboards do not require any investigation, analysis, or aggregation of the information, which is a *must* for informed decisions.

Although the literature on dashboard design principles is broad and comprehensive, in the field of dashboard development, mostly dashboards are either created from scratch for each organization or a dashboard template is tailored to the needs of organizations. Software vendors or their client organizations often carry out this customization process. However, it requires a significant effort both from software vendors and organizations [42, 50–52, 54]. To deal with that, several approaches in the literature [31, 66, 75, 76, 78, 82–84, 100, 135, 147–150] focus on developing dashboards automatically. Reusing parts of existing dashboards and expressing the structure of a dashboard or a dashboard template using an introduced notation (i.e., model-driven) are the two prominent ways in these approaches. Due to the limited coverage of state-of-the-art dashboard design principles or not completely addressing the relevance of visualized indicators for decision makers, the dashboards created using these approaches have shortcomings [50, 106]. For example, the approach intended to visualize performance indicators on BPMN models will generate a complicated view when multiple performance indicators are related to a single task in a BPMN model. Hence, decision makers will be distracted.

In this context, we propose a novel approach for the *interactive* generation of *engaging*

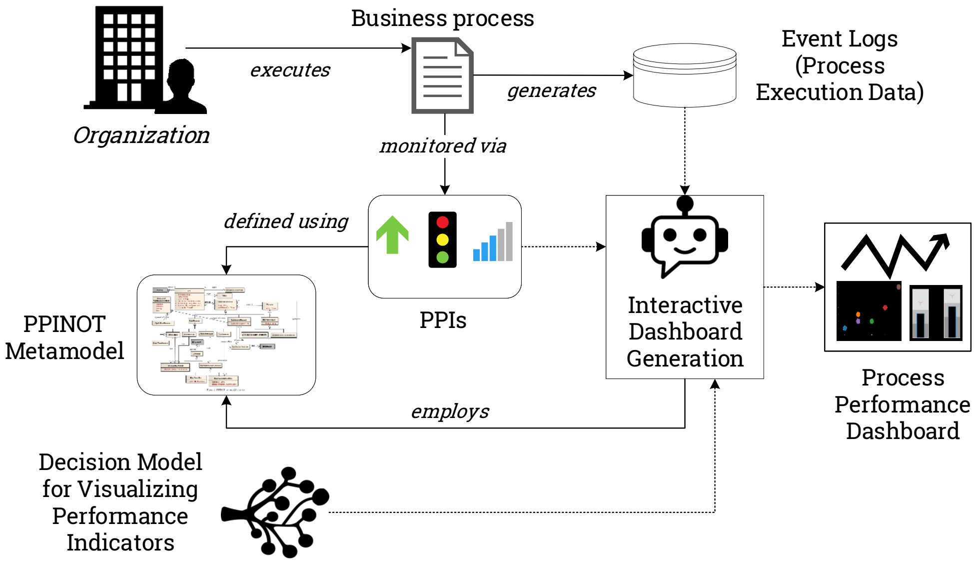

Figure 6.2: Our approach for interactive process performance dashboard generation

process performance dashboards. As shown in 6.2, The approach employs a chatbot to generate dashboards. For a given set of PPIs, the chatbot determines what visualizations are appropriate to engage decision makers. To do so, it uses a decision model that we previously developed based on the dashboard design principles in the literature. To provide automated support for the management of PPIs, del-Río-Ortega et al. proposed a well-established solution, i.e., the *PPINOT* metamodel [37]. Accordingly, we integrate the PPINOT metamodel into our approach for the automated analysis of PPIs and computing their values automatically. Hence, for an interacting decision maker, the chatbot can generate engaging dashboards by means of the automatically visualized PPIs together with their values. Thus, the approach enables decision makers to focus on the relevant messages conveyed via engaging dashboards without needing substantial knowledge to develop such dashboards.

In the evaluation of the approach, first, we check the common usability of the decision model in a case study in an organization. It is important to note that the decision model was evaluated in two organizations in our previous work [6]. Apart from the scenarios in which the decision model was evaluated in that work, we used here a different case from a different domain, and also evaluated the decision model in this context. Then, we execute the approach, and together with experts in that organization, we compare the newly created dashboards with existing dashboards to see how our approach helps them at making informed decisions. The results that we obtained indicate that our approach is useful and valuable to assist organizations in improving their business.

In Section 6.2, we provide the background on both organizational performance measurement and the PPINOT metamodel. We present our approach for the interactive generation of engaging process performance dashboards in Section 6.3. Section 6.4 is devoted to the evaluation of both our approach and the decision model, which is employed by our approach for visualizing PPIs. In the same section, we present the results obtained while evaluating the process performance dashboards that we created. In Section 6.5, we discuss the obtained results. An overview of the related work on developing process performance dashboards for organizations is given in Section 6.6. Finally, in Section 6.7, we present our conclusions and directions for future work.

6.2 Theoretical Background

In this section, first, we provide a summary of the terminology often used in organizational performance measurement and monitoring. After that, we explain the PPINOT metamodel that we integrate into our approach.

6.2.1 Performance Measurement in Organizations

Organizations measure their performance to determine how well they are doing to attain their goals. To figure out what changes, if any, to make, it is vital for organizations to know their actual performance. Furthermore, it is mostly not possible to know whether there is any improvement without measurement [133]. Metrics, process performance indicators, and key performance indicators are the common means for performance measurement [40]. Often, they are considered synonyms and used interchangeably. However, each has a specific meaning and purpose. Therefore, their proper use is essential to avoid ambiguity and to establish clear communication within organizations. In this regard, we explain these common performance measurement means below.

- **Metric**: A metric is a quantifiable measure to gauge performance or progress [12]. For example, the count of items (e.g., number of incidents) or the end time of an event (e.g., when the order is delivered). It is important to note that metrics do not provide complete information to establish a basis for a decision. To support decision making, additional information is required. With additional information, metrics can provide a clear indication that can be interpreted to make a decision.

- **Process Performance Indicator (PPI)**: A PPI is a metric that provides a reflection of a business process [37] to evaluate its performance based on its critical success factors, e.g., efficiency and effectiveness. For example, the percentage of solved incidents within a month in comparison to the performance target defined in the related business process is a common PPI of the incident management process. The exclusive focus of PPIs is the performance of business processes. PPIs can give a clear indication of what extent a business process achieves its goals. Thus, PPIs enable decision makers to make informed decisions.

- **Key Performance Indicator (KPI)**: A KPI is a metric that embeds performance targets so organizations can chart progress toward organizational strategic goals [42, 101]. For instance, the average duration of delivery is an essential KPI for an organization that operates in the logistics domain because that KPI shows to what extent the organization attains its strategic goal, i.e., fast delivery. KPIs have a broad focus on capturing the performance of the overall organization. In general, KPIs provide an aggregated overview of multiple business processes that are performed to reach a particular strategic goal in an organization. Accordingly, decisions that are made based on KPIs can affect the entire organization or parts of it. It is important to mention that a single process can solely be the contributor to a strategic goal in an organization. Thus, the PPIs used for monitoring that single process are also KPIs for such an organization. Simply put, it is likely to have overlapping PPIs and KPIs in organizations [12, 89]. Thus, PPIs can be considered as a particular case of KPIs, e.g., aggregation of some PPIs.

6.2.2 Performance Monitoring in Organizations

Dashboards provide organizations visibility into the performance of their business by consolidating and integrating required information to monitor the performance and make decisions [42, 54, 158]. To display such information on dashboards, performance indicators are visualized using graphs or tables. With this, the aim is to exploit visual elements for supporting decision makers in accessing, analyzing, and acting on the information needed to achieve organizational goals. Hence, appropriate visualizations are necessary to convey suitable messages that will engage decision makers. In this regard, what properties [42, 54, 97, 106, 134, 158] should a dashboard have to be considered "engaging" are listed below.

- **Consistent view**: Visual aspects of the dashboard (e.g., visualizations, colors, fonts, alerts, naming conventions, etc.) should be consistent and free from conflicts and ambiguity.
- **Contain what is important**: The dashboard should only contain relevant and timely information. Redundant information should be eliminated.
- **Deliver actionable information**: The messages conveyed via the dashboard should be suitable to its users for making decisions. The dashboard should provide clear signals about the monitored performance, e.g., proper alerts based on the predefined targets.
- **Content over decoration**: Visual attractiveness should not be the main concern in a dashboard. It should be adjusted to the content not to cause any distraction.

6.2.3 PPINOT metamodel

PPIs are commonly described as natural language texts in organizations. Such informal and ad-hoc way of defining them often causes ambiguity and coherence issues. Furthermore, the traceability between these textual expressions and processes disappears over time due to the dynamic nature of processes. The level of granularity that is chosen expressing PPIs affects their correct interpretation. For instance, PPIs defined from a technical perspective are less comprehensible for managers to figure out what is actually measured in the related process at a glance. To deal with these issues and provide automated support for the management of PPIs, del-Río-Ortega et al. proposed a well-established solution, namely the *PPINOT* metamodel [37]. Using the constructs proposed as part of the metamodel, one can easily define PPIs. Moreover, with the support of a visual notation [38], PPIs can be attached to process models independently from the process modeling standard used.

Furthermore, PPINOT enables the design-time analysis of the relationships between PPIs and process elements. Such analysis establishes the basis for the run-time analysis of PPIs, i.e., the values of PPIs can be automatically calculated from the execution logs of processes. With conditions that are linked to a business process model, traceability is established. In particular, using the linked process model elements, measurement is done. Figure 6.3 shows the details of the PPINOT metamodel.

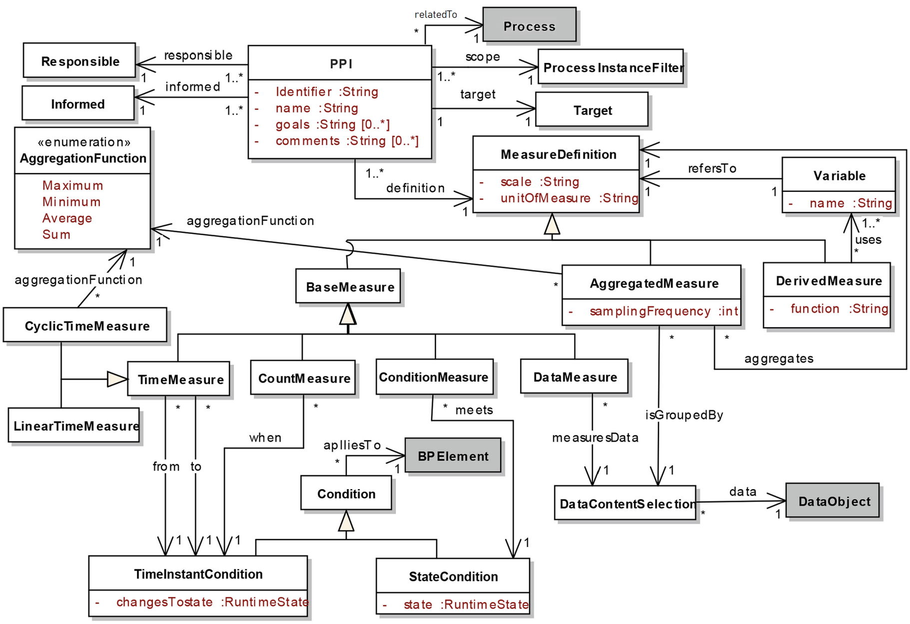

Figure 6.3: Excerpt of the PPINOT metamodel

What information is necessary to define PPIs using the PPINOT metamodel is given in Table 6.1.

PPINOT allows defining an extensive range of measures. For example, base measures, aggregated measures, and derived measures. A base measure represents a single-instance measure that is used for measuring the values of time, count, condition, or data. By aggregating base measures using an aggregation function, an aggregated measure is created that measures multiple process instances. Derived measures employ a mathematical function and use it over other measures, i.e., base measures and aggregated measures. Thus, a derived measure may represent either a single-instance or multi-instance measure. Each type of Base Measures is described in Table 6.2.

Considering the aforementioned strengths of the PPINOT metamodel, we integrated it into our approach. In particular, the chatbot employed in the approach exploits an implementation of the PPINOT metamodel to analyze PPIs and determine appropriate visualizations for them. Moreover, that implementation enables the chatbot to automatically compute PPI values. The implementation that we use in our approach is the

Table 6.1: PPI attributes in PPINOT

Attribute	Definition
Identifier	The unique identifier of the PPI.
Name	The name that describes the PPI.
Goals	The goals (e.g., strategic, operational, process) that the PPI is related to.
Comments	Additional information that cannot fit in other attributes.
RelatedTo	For which process the PPI is defined.
Scope	Limits the subset of instances of the process that must be considered to compute the value of the PPI.
Definition	How the PPI is measured based on the linked measure definition.
Target	The value or value range to be reached to indicate the accomplishment of defined goals. The target (i.e., simple target) is defined using a lower and/or an upper bound. When only the upper bound is set, a maximum is defined as the target, whereas a minimum is defined as the target if only the lower bound is set. In case both bounds are set, a range becomes the target. If the PPI has multiple values (e.g., value per incident type), several targets can be composed, i.e., the composed target. It is possible to define a restriction on the PPI value using the custom target in which functions or formulas denote the restriction.
Responsible	The human resource (e.g., person, role, or organizational unit) who is in charge of the state of the PPI.
Informed	The human resource (e.g., person, role, or organizational unit) has an interest in the PPI and needs to be informed about the state of the PPI.

Table 6.2: Sub-types of the base measure

Base Measure	Purpose
Time Measure	It measures the duration between two time instants, i.e., from the start time to the end time. In the Linear Time Measure, the first occurrence of the *from* time instant and the last occurrence of the *to* time instant condition are taken into account. The duration between the two time instants is aggregated considering every iteration in more than one process instance for the Cyclic Time Measure.
Count Measure	It measures how many times the certain instant condition is met, e.g., the number of times that the resolve incident activity happens.
Condition Measure	It indicates the measured fulfillment of a certain condition.
Data Measure	A certain part of a data object is measured.

Python implementation of the PPINOT metamodel, namely PPINot4Py[2]. In the following section, we explain our approach.

6.3 Approach

The details of our approach for the interactive generation of process performance dashboards are given in this section. Within the approach, a chatbot is used for handling interactions. PPIs denoted using the PPINOT metamodel are provided as a

[2]PPINot4Py is available at: `https://github.com/isa-group/ppinot4py`

"machine-readable" input to the chatbot. To compute the values of PPIs, event logs are used. For determining what visualizations are appropriate for PPIs, a decision model is implemented by the chatbot. Based on a dashboard split strategy, visualized PPIs are then combined in the form of a dashboard, the so-called process performance dashboard. These details are depicted in Figure 6.4 and explained below.

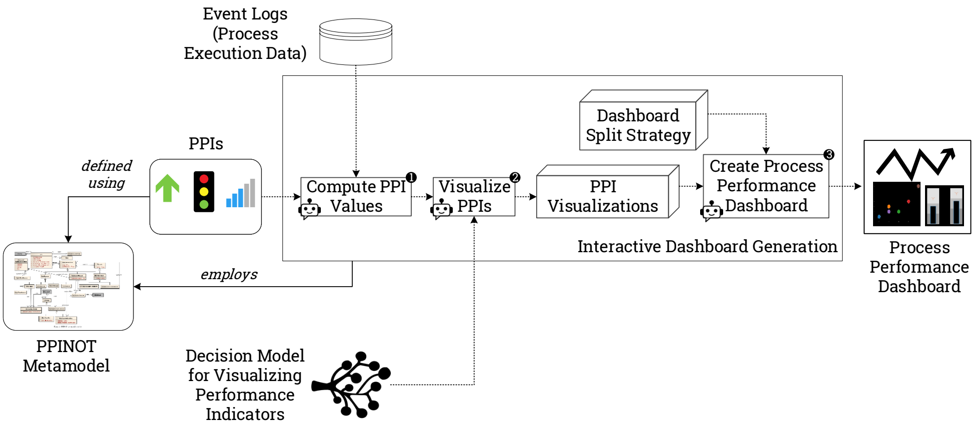

Figure 6.4: Interactive generation of process performance dashboards

By carrying out three subsequent tasks, the approach creates a process performance dashboard. These tasks are: (1) computing the values of each PPI, (2) interactive visualization of PPIs based on a decision model, and (3) blending visualized PPIs together to form a process performance dashboard. Each task is explained below.

6.3.1 Task 1-Compute PPI Values

The chatbot employed in the approach uses event logs as input to automatically compute PPI values. Event logs contain process execution data, i.e., historical records that are created during the execution of a process. As PPIs are defined using the PPINOT metamodel, the relationships between process elements and PPIs are encoded in PPIs. After obtaining these relationships, the chatbot determines corresponding records within event logs. Then, the chatbot calculates PPI values with respect to the *measure definition* in each PPI.

6.3.2 Task 2-Visualize PPIs

Appropriate visualizations for PPIs are determined by the chatbot as the result of a conversation with an interacting user. In the conversation, the chatbot supports the user to decide on PPI visualizations. This is achieved using a particular input, namely a decision model. Since PPIs are a particular case of KPIs (See Section 6.2.1), we provide the decision model [6] that we previously developed for visualizing KPIs as input to the chatbot. The decision model is shown in Figure 6.5. For the development of the decision model, we refer to Subsection 5.3.1 Developing the Decision Model for Visualizing KPIs in Chapter 5.

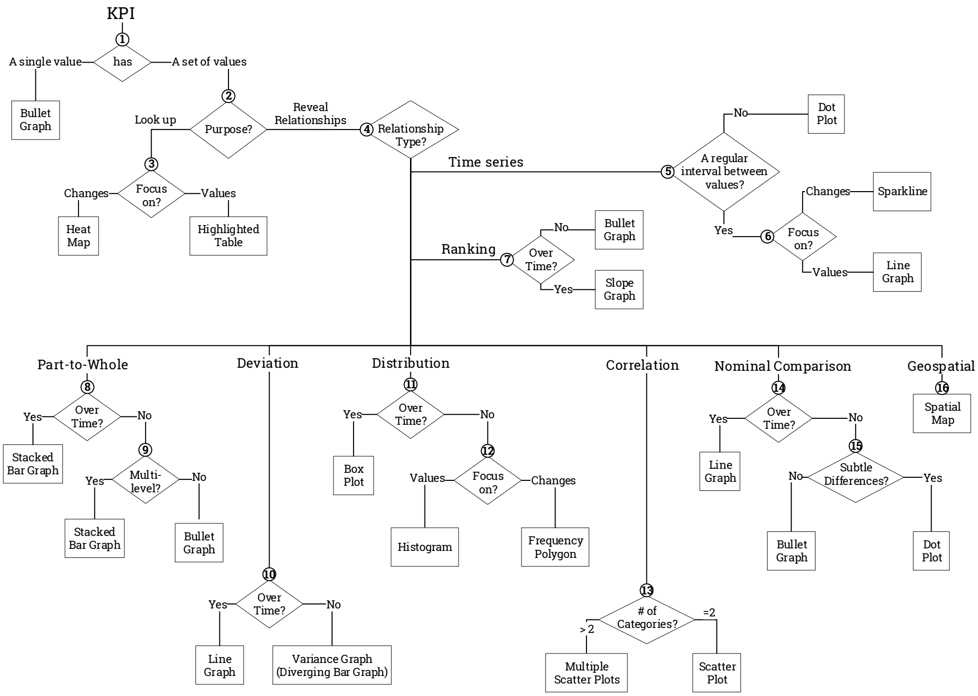

Figure 6.5: Decision model for visualizing KPIs

As can be seen in Figure 6.5, there are several decision points that require certain inputs, i.e., KPI attributes. Based on the provided KPI attributes, the decision model is traversed, and the appropriate visualization is determined for a given KPI. Similarly, appropriate visualizations for PPIs can be determined using the decision model when those attributes are present in PPIs. In this context, we made a mapping between the PPINOT metamodel elements and our decision model. More specifically, we checked the purpose and definition of each PPINOT metamodel element to identify which of them can provide the required information for KPI attributes. The mapping is given in Table 6.3 below. In the table, for each KPI attribute, mapped PPINOT metamodel elements are indicated together with the degree of the mapping.

Clearly, not all attributes required by the decision model can be obtained from PPIs, as indicated in the table. Therefore, the missing attributes in PPIs need to be provided to the chatbot via human interaction. To obtain such information via human interaction, we add a loop into the conversation flow (see Figure 6.6) of the chatbot. In particular, the chatbot will determine and collect required additional information by asking questions to the interacting user. Until the chatbot obtains the complete information required for visualizing a particular PPI, the interaction will continue with the user. After the loop, the chatbot creates visualizations for each PPI. Important to note that, in the loop, the chatbot can suggest visualizations using three basic heuristics:

1. **Time dimension**: If the scope of the PPI contains time information, the chatbot

checks the existence of a regular interval in the previous values of the PPI. A Line Graph will be suggested when a regular interval is detected.

2. **Dependency**: In case the PPI is an Aggregated Measure and its values comprise two categories, a Scatter Plot will be suggested additionally.

3. **Location**: When the PPI is an Aggregated Measure and the grouping is done based on a geographical term (e.g., country, county, city, etc.), the chatbot will suggest a Spatial Map as an additional visualization.

Building a visualization for a PPI consists of four steps: (i) creating the visualization parts for the *target thresholds* of the PPI, (ii) creating the visualization parts for the *values of PPI*, (iii) creating the visualization parts for the *target* of the PPI, and finally (iv) combining all parts as a single visual element, e.g., graph or table. Based on the defined thresholds for the target of a PPI, the approach creates bars and arranges them accordingly in the first step. Then, bars, dots, or lines are used depending on the determined visualization element for showing the values of a PPI in the second step. A visual sign, i.e., a noticeable alert (a cross-mark, a check-mark, or a warning sign), is put on the generated visualization element to indicate the achievement of the target of the PPI in the third step. In the last step, the approach combines the parts as a single visualization element considering the embedded coloring [3], orientation, and resizing rules for visualization in it. The reason for that is to determine the best orientation (horizontal or vertical) for readability. For example, a ranking relationship better reads when it is horizontal and has a maximum of 10 categories where the rest is grouped as "others." All but one of the required attributes for building a visualization element are available in the PPINOT metamodel. The only missing attribute is the target threshold. Since this attribute is not mandatory in our approach, we can conclude that the PPINOT metamodel provides the required minimum for building a visual element for a PPI.

To provide a relevant set of visualizations for an interacting user, the chatbot then asks for feedback on each visualization generated for every PPI. Thus, we ensure that the chatbot creates engaging visualizations for its users. Moreover, the obtained feedback is stored in a database for later use. For example, for users that have a similar profile, relevant visualizations can be determined using that database. How the visualized PPIs are combined in a process performance dashboard is explained in the following subsection.

6.3.3 Task 3-Create Process Performance Dashboard

To determine how visualized PPIs will be placed on a dashboard, the approach employs a particular strategy, namely the Dashboard Split Strategy. This strategy is used for creating a flow through a combination of visual weight and visual direction to take advantage of how people read through a design. More specifically, the created flow splits a dashboard into two areas: top and bottom. By applying the most common layout pattern (the Z-diagram layout [10]), which is recommended for simple designs, the approach defines the route that the human eye travels on these areas: left to right and top to bottom. The PPI attributes "responsible" and "informed" provide the information for determining the order of the PPIs that the human eye should read in this travel. The PPIs that are the

[3] `http://colorbrewer2.org` is used as the source for color selection.

Table 6.3: Mapping between the decision model and the PPINOT metamodel

KPI Attribute	Definition	**PPINOT Mapping** (✓: Fully, ✓: Partially, ✗: None)
Relationship type of values	Describes how the values of the KPI are related. For example, time series, correlation, ranking, part-to-whole, nominal comparison, or distribution.	✓ Time series via scope: Process Instance Filter and Measure Definition: Time Measure.
Purpose	Whether the KPI is about looking up its values or revealing the relationship between its values.	✗ Chatbot collects it from the user during interaction
Focus	Describes what is the focus of the KPI with respect to its purpose attribute. Example values: look up-changes, look up-values, relationship-changes in a time series, relationship-values in a distribution.	✗ Chatbot collects it from the user during interaction
Time interval	Whether the KPI needs to be displayed over time.	✓ scope: Process Instance Filter and Measure Definition: Time Measure.
KPI values	Describe the quantitative information of the KPI.	✓ Measure Definition: Base Measure, Aggregated Measure, or Derived Measure
Categories	The discrete groups in which one or more values exist. For example, the total revenue is a KPI that has a single category, which contains a single value. However, the revenue per department is a KPI that will have a category for each department.	✓ Measure Definition: Aggregated Measure, i.e., Collection of Base Measure with isGroupedBy
Sort direction	Describes how the categories or the values in a category will be ordered, e.g., ascending or descending. This is especially important in ranking and distribution relationships.	✗ Chatbot collects it from the user during interaction
Multi-level hierarchy	Whether there is a hierarchy or main-sub grouping in the categories attribute of the KPI. For example, main group: region and sub-group: county.	✓ Measure Definition: Multi-level Aggregated Measure, i.e., Collection of Base Measure with isGroupedBy
Regular interval between values	This will be determined using the attribute time interval. If there is any missing value in the values of the KPI based on its time interval, the branch "No" will be selected in the related decision point of the decision model.	✓ scope: Process Instance Filter and Measure Definition: Time Measure or not
Subtle difference threshold for the values of the KPI	Describes the limit of the difference between the values of the KPI that should be clearly detectable at a nominal comparison.	✗ Chatbot collects it from the user during interaction

responsibility of the interacting user with the chatbot will be placed in the top area of dashboards. The bottom area of dashboards is reserved for the PPIs that the user is simply informed about. In which order PPIs will be placed on the dashboard is determined based on the interaction sequence of users, i.e., in the order that a user starts visualizing PPIs. For interacting with users, the chatbot exploits a conversation flow. This is depicted in Figure 6.6. The flow starts with the retrieval of the user profile of an interacting user. Based on the user profile, the chatbot obtains and shows the processes that the user is involved in with either a responsible or an informed role. For this, the chatbot checks the following attributes of PPIs: “responsible” and “informed”. Then, the chatbot retrieves the PPIs that are defined for the selected process using the “relatedTo” attribute of PPIs. Within a loop, the chatbot interacts with the user to generate the visualizations for each PPI that is relevant for the user. When all PPIs are handled, the chatbot blends all visualizations together and creates a process performance dashboard. The created dashboard is then shown to the user. This interaction loop can be repeated depending on the remaining processes or the reactions of the user. For the sake of simplicity, we exclude exceptional paths, such as going back to a prior point in an interaction or exiting at any time.

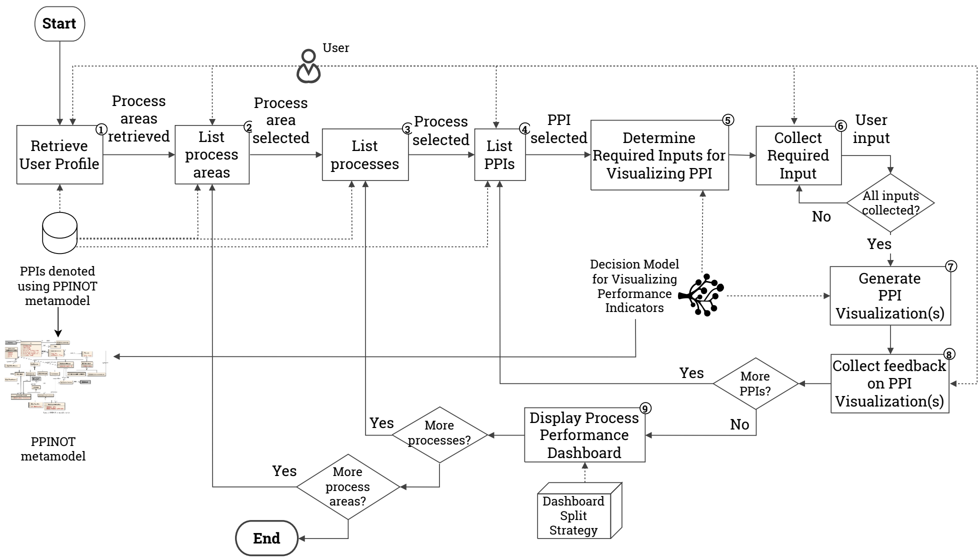

Figure 6.6: Chatbot conversation flow

We implemented our approach in Python. Our implementation[4] is comprised of two parts. The first part is the chatbot core. For this, we used Microsoft Bot Framework software development kit (SDK)[5]. The second part of our implementation is devoted to the visualization of PPIs and placing them on a performance dashboard. To accommodate this, we exploit two Python libraries: *Plotly* is for visualizing KPIs and *Dash* is for creating dashboards.

In the next section, we apply our approach in a case study and evaluate its usefulness.

[4]The implementation of our approach is available at `http://amuse-project.org/software/`

[5]Microsoft Bot Framework SDK is available at `https://dev.botframework.com/`

6.4 Evaluation

In this section, first, we explain how we evaluated the decision model that our approach exploits to visualize PPIs. Then, we elaborate on the evaluation of the proposed approach in a case study.

6.4.1 Case Study Organization

One of the biggest educational institutions in the Netherlands expressed its interest to us in learning from the data recorded about its IT Service Management (ITSM) processes. ITSM encompasses all processes that are essential for the planning and management of products and IT services offered to its customers in any organization [11]. Incident Management (IM), Change Management (CM), and Problem Management (PM) are typical processes within ITSM. Such processes are performed by the IT department of the institution. As there is a wide range of products and services offered to more than $25K$ customers ($20K$ students, $5K$ employees), multiple organizational units are involved in these processes. For instance, printing and workplace equipment (e.g., software and hardware) management are two separate services, which are managed by different organizational units.

To actively manage and monitor the performance of the aforementioned three processes, the IT department has eight essential PPIs. The description, target, and scope of those PPIs are shown in Table 6.4. The performance of the IM process is monitored using three PPIs that show the progress in managing questions, requests, and, most importantly, malfunctions about the offered products and services. Similarly, there are three PPIs to monitor the CM process, which is established to reduce or minimize the risk that modifications could affect the offered products and services. The last two PPIs shown in the table are used for monitoring the PM process of which purpose is to determine and analyze root causes to prevent issues from happening in the future. Due to privacy regulations, we adjusted the description and target values of those PPIs. Likewise, hereafter, organization *O* refers to the IT department of the institution because of confidentiality reasons.

Table 6.4: PPIs of organization *O*

PPI	**Description**	**Target**	**Scope**
IMPPI1	Percentage of type-1 incidents solved on-time	$\geq 80\%$	Monthly
IMPPI2	Percentage of type-2 incidents solved on-time	$\geq 80\%$	Monthly
IMPPI3	Percentage of type-2 incidents solved on-time per organizational unit	$\geq 80\%$	Monthly
CMPPI1	Percentage of standard RFCs completed on-time	$\geq 80\%$	Monthly
CMPPI2	Received standard RFCs compared to total changes	$\geq 50\%$	Monthly
CMPPI3	Received urgent RFCs compared to total changes	$\leq 10\%$	Monthly
PMPPI1	Percentage of problems resolved on-time	$\geq 80\%$	Monthly
PMPPI2	Percentage of problems resolved with a workaround	$\leq 20\%$	Monthly

To accomplish a data-driven service delivery, organization *O* has the initiative to improve the dashboards used for monitoring its ITSM processes. As part of that initiative, one of the authors of this work has been in close collaboration with experts in the organization for more than a year. Since the setting and needs of the organization are highly related to the approach that we propose in this chapter, its application in this context is highly relevant to determine its applicability and value.

Before applying our approach, we evaluate the decision model for visualizing KPIs, which is an input to the chatbot employed in our approach. This decision model was previously developed by us as part of our previous work on automated dashboard generation [6]. It is important to note that the decision model is evaluated in two organizations in that previous work. Since the institution operates in a different domain than the two organizations mentioned in that work, it is relevant to evaluate the decision model before using it in a different setting.

6.4.2 Evaluation of the Decision Model

The evaluation of the decision model is done by discussing our considerations and walking through each decision point and element in the decision model together with experts in organization *O*. The details on the background of the experts involved in the evaluation of the decision model are shown in Table 6.5. In particular, we worked together with three experts: one BI consultant and two process workers. The reason why we chose these experts is as follows: both of the process workers are involved in the development of the dashboards for the IM, CM, and PM processes, and moreover, they worked together with that particular BI consultant for dashboard development.

We collected the opinions of these experts about the decision model using the three-points Likert-type scale (agree, somewhat agree, and disagree). The opinion collection is done with a semi-structured discussion meeting. In the meeting, the experts shared their thoughts on the usefulness of the decision model based on their overall involvement in dashboard development. In doing so, we aimed not to limit the evaluation scope with the three ITSM processes. For each decision point and visualization element shown in the decision model, we gathered opinions.

The collected expert opinions are displayed in Table 6.5. From the table, one can observe to what extent the decision model is useful to the needs of organization *O* at visualizing KPIs. In particular, the experts agreed on average more than 80% of the decision points and also the visualization elements in the decision model.

6.4.3 Evaluation of the Approach

Data Collection

As explained in the previous section, there are three inputs required by our approach for interactively generating engaging process performance dashboards. These are: (1) PPIs denoted using the PPINOT metamodel, (2) event logs, and (3) a decision model for visualizing KPIs.

As for the first input, we use the PPIs listed in Table 6.4. Since these PPIs are defined as natural language texts, we collected the corresponding process models for them. Together with the aforementioned three experts in the evaluation of the decision model, we adjusted the process models based on the privacy restrictions in the organization. Moreover, the

Table 6.5: Evaluation of the decision model in organization O

Expert	**Area of expertise**	**Years of expertise**	**Meeting duration (hours)**	**To what extent agree on the decision points in the decision model**	**To what extent agree on the visualization elements in the decision model**
BI Consultant	Dashboard development	>15	1.5	Agree:13(87%) Somewhat agree:2 Not agree:0	Agree:20(91%) Somewhat agree:2 Not agree:0
Process Worker-1	Change management and Problem management	>10	1.5	Agree:13(87%) Somewhat agree:2 Not agree:0	Agree:21(95%) Somewhat agree:1 Not agree:0
Process Worker-2	Incident management and Change management	>10	1.5	Agree:13(87%) Somewhat agree:2 Not agree:0	Agree:21(95%) Somewhat agree:1 Not agree:0

experts explained to us the way each PPI is calculated by referring to the related elements in the process models. Using that information, we expressed the PPIs using the Python implementation[6] of the PPINOT metamodel. We plan to further develop our approach such that this input can be automatically extracted from process models annotated with PPIs, e.g., using Visual PPINOT.

For the second input, we obtained event logs for the three processes (i.e., IM, CM, and PM) from organization O. The collected event logs contain the execution history of these processes in 2020. Specifically, the number of events and cases that each event log consists of are as follows: IM, $100K$ events and $8K$ cases; CM, $12K$ events and $2K$ cases; and PM, $2K$ events and 130 cases.

Lastly, as the third input, we use the decision model for visualizing KPIs that we shared its details in this section.

Applying our Approach

The proposed approach was evaluated in organization O. A dashboard for each process is created by the same three experts. We collected the opinions of experts about the usefulness of the approach using the ten-points Likert-type scale (from strongly disagree-1 to strongly agree-10. For the assessment of usefulness, we adapted the work by Bhattacherjee in [16] and created a questionnaire, which is shown below.

a. Overall, the approach is useful for generating engaging process performance dashboards interactively.
b. The interaction points increase my productivity in identifying what matters to have in a process performance dashboard.

[6]The PPINOT metamodel Python implementation: `https://github.com/isa-group/ppinot4py`

c. Seeing the generated visualizations in real-time enhances my effectiveness in interpreting visualized information.
d. The generated dashboards improve my performance to make informed decisions.

Based on the feedback of the experts, we calculated the average score for each item in the questionnaire. Specifically, 9, 8.7, 9.3, and 8.7 are the averages for the questionnaire items, respectively.

Figures 6.7 and 6.8 show how a conversation evolves in a typical interaction with the chatbot for creating an engaging process performance dashboard. As can be seen in the first figure, the chatbot lists the processes in which the interacting user is involved. Then, based on the user's process selection, the chatbot lists the available PPIs. Once the user selects a PPI, the chatbot generates and shows the visualizations for the selected PPI. As indicated in the figure, depending on the PPIs' attributes, the chatbot can suggest visualizations. To place only the relevant visualizations in a dashboard that the interacting user will be engaged in, the chatbot collects feedback on the relevance of visualizations. As depicted in Figure 6.8, as soon as the visualizations for each PPI are generated and evaluated, the chatbot prepares the process performance dashboard and displays it. After that point, the user can continue creating process performance dashboards for other processes if needed.

To determine the implications of the differences between the existing dashboards and the newly generated dashboards for the three processes, we had a semi-structured discussion meeting with the same three experts. The newly generated three dashboards are shown in Figures 6.9, 6.10, and 6.11. Each visualization in the dashboards is surrounded by a rectangle. An orange rectangle indicates that the surrounded PPI is the responsibility of the interacting process worker. Whether a PPI is on target is indicated in each visualization using a *cross-mark* or a *check-mark*. Moreover, the *warning sign* where it is placed implies what to focus on reaching the target of a PPI.

For the three PPIs of the IM process (see Table 6.4), the generated dashboard is shown in Figure 6.9. As depicted in the figure, there are four visualizations for these three PPIs. In the top two visualizations, IMPPI1 and IMPPI2 are displayed as a Bullet graph (see ① in Figure 6.5) as these PPIs have a single value based on their defined scope, i.e., a percentage value for the current month. On the bottom left, multiple Bullet graphs are shown as the interacting user (Process Worker-2) was interested in the ranking of organizational units with respect to the achievement of the target defined in the PPI. The chatbot determined this visualization by following the following path in the decision model: ①, *PPI has a set of values*; ②, *purpose* is to *reveal relationships*, and ⑦, the selected relationship is *ranking* but not *over time* in Figure 6.5. The last visualization, on the bottom right, is a suggested visualization and placed on the dashboard as the interacting user mentioned that it is relevant. IMPPI2 is displayed from a time perspective in that visualization.

Figure 6.10 depicts the three PPIs for the CM process, namely CMPPI1, CMPPI2, and CMPPI3. As each of these PPIs has a single value based on their scope attribute, they are visualized as Bullet graphs (see ① in Figure 6.5). Although the chatbot suggested showing these PPIs over time, the interacting user (Process Worker-1) mentioned that the view on a time perspective for those PPIs is irrelevant. Therefore, no additional visualizations are placed on the CM process performance dashboard.

The created PM process performance dashboard for the interacting user (Process Worker-2) is displayed in Figure 6.11. As there are only two PPIs for this process and both

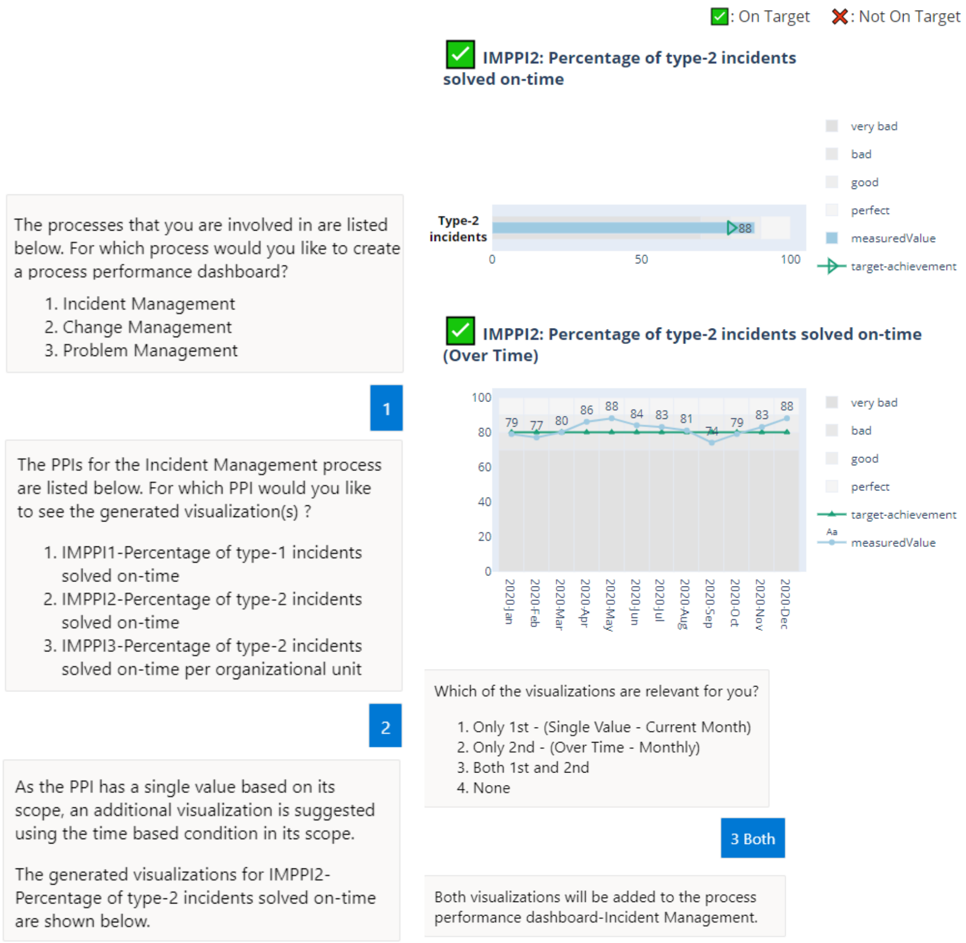

Figure 6.7: Chatbot can suggest additional visualizations based on PPI attributes

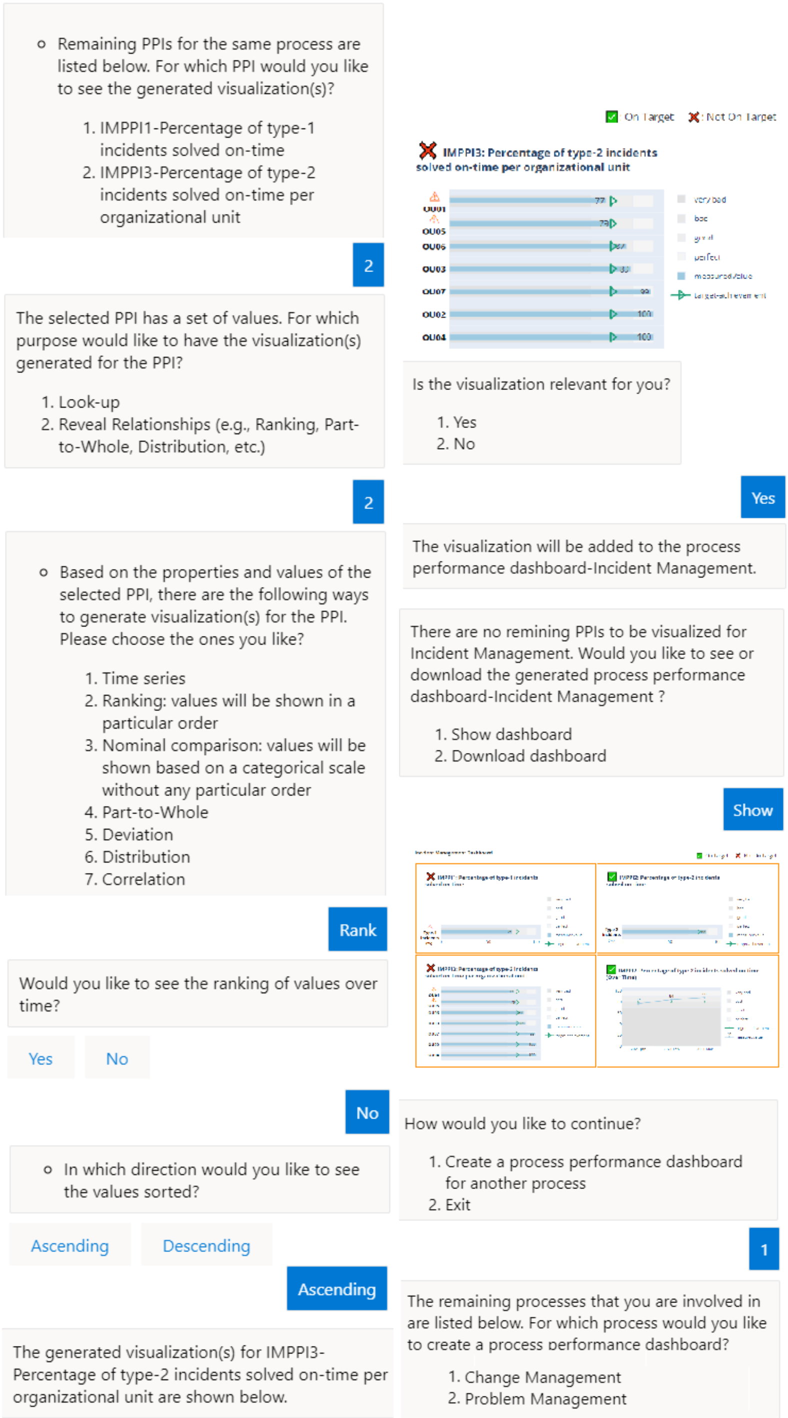

Figure 6.8: Evolution of conversation while interacting with the chatbot

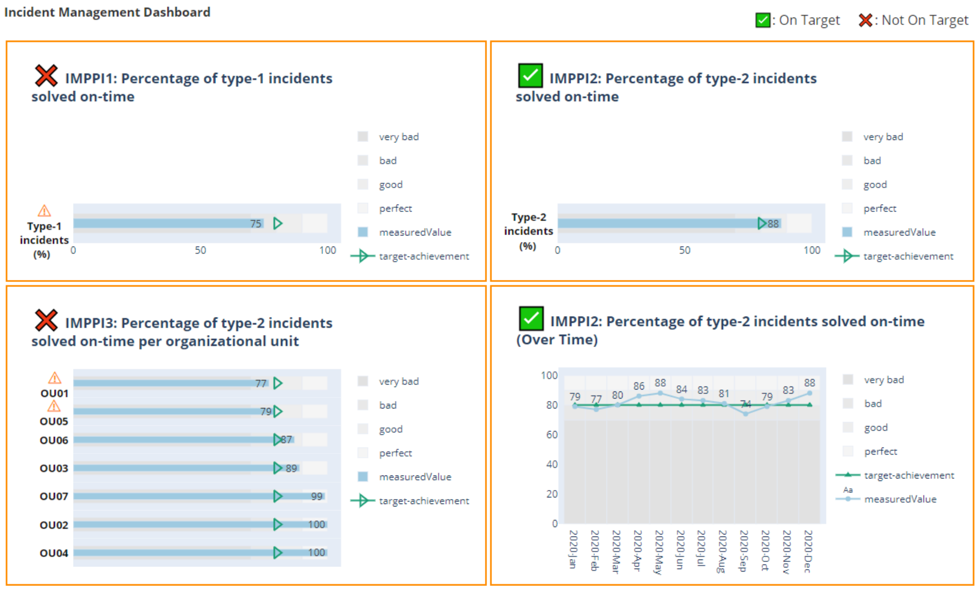

Figure 6.9: Generated process performance dashboard for incident management

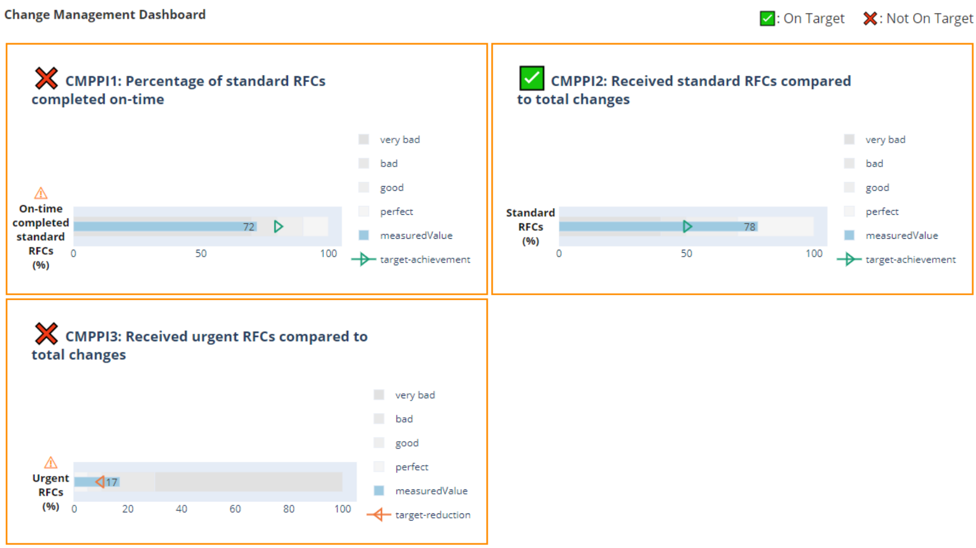

Figure 6.10: Generated process performance dashboard for change management

have a single value based on their scope, they are visualized in the form of Bullet graphs (see ① in Figure 6.5). Separate visual indicators are used due to the difference in the target type of each PPI (IMPPI1: achievement, IMPPI2:reduction).

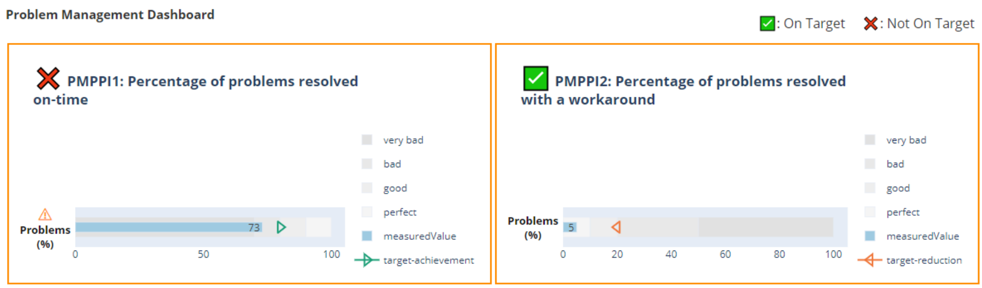

Figure 6.11: Generated process performance dashboard for problem management

The existing dashboards for the CM, PM, and IM processes in the case study organization are displayed in Figures 6.12, 6.13, and 6.14, respectively. As can be seen in these figures, a highlighted table with a smiley is the common means used for visualizing seven out of the eight PPIs. For the PPI IMPPI3, Bar graphs are used. In the visualization for that PPI, bars represent the organizational units that are stated in the PPI. Moreover, organizational units are placed alphabetically into the visualization. Besides, there is no target indicator on that visualization. The remaining seven PPIs are visualized using a highlighted table in which the value of each PPI is painted red or green. A PPI painted red means it is not on target, whereas on target, PPI is painted green. To show whether PPIs are on target in a different way, highlighted tables are supported with smileys, i.e., a happy face for on target and an unhappy face for not on target.

PPI	**PPI Description**	**Target**	**Current Month**	
CMPPI1	Percentage of standard RFCs completed on-time	>= 80%	72	☹
CMPPI2	Received standard RFCs compared to total changes	>= 50%	78	☺
CMPPI3	Received urgent RFCs compared to total changes	<= 10%	11	☹

Figure 6.12: Current process performance dashboard for change management

PPI	**PPI Description**	**Target**	**Current Month**	
PMPPI1	Percentage of problems resolved on-time	>= 80%	73	☹
PMPPI2	Percentage of problems resolved with a workaround	<= 20%	5	☺

Figure 6.13: Current process performance dashboard for problem management

In the discussion meeting we had with the same three experts, we collected their opinions. In particular, the experts shared their findings by comparing the newly generated dashboards with the existing dashboards for the three processes. For all the generated dashboards, the experts agreed on the value of using a consistent style, i.e., selected visualization elements, indicators, signs, and colors. The experts also mentioned that the displayed target thresholds (e.g., perfect, good) are useful for determining the severity of taking actions in case a PPI is not on target. Moreover, the experts emphasized that alerts

PPI	PPI Description	Target	Scope	Current Month	
IMPPI1	Percentage of type-1 incidents solved on-time	>= 80%	Monthly	75	☹
IMPPI2	Percentage of type-2 incidents solved on-time	>= 80%	Monthly	88	☺
IMPPI3	Percentage of type-2 incidents solved on-time per organizational unit	>= 80%	Monthly		

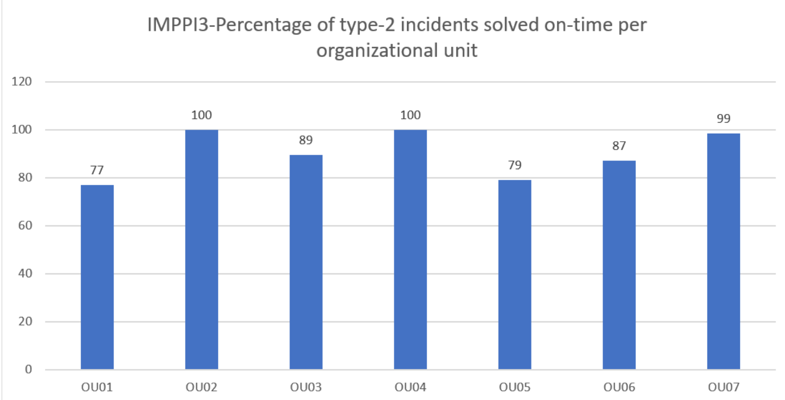

Figure 6.14: Current process performance dashboard for incident management

(cross-mark, check-mark, and warning signs) help them to detect what to focus on when monitoring the performance of processes.

Furthermore, the experts pointed out several benefits of using each newly created dashboard. Visualizing IMPPI3 to reveal the ranking relationships of organizational units is considered as highly relevant as the organizational units that could not attain their goals become noticeable. Similarly, best performing organizational units are now more visible. The experts noted that this distinction could assist them in revealing the best practices within organizational units. The experts confirmed that the additional visualization for IMPPI2 is very useful to observe changes over time. As the PPIs for the CM and PM processes are visualized as Bullet graphs, the experts indicated that this type of graph is clean, simple, and very powerful to capture the information that is needed for making to the point decisions. Moreover, the experts noticed that having tidy visualizations avoids cluttered views.

In the following section, we discuss the results that we obtained in the evaluation of both the decision model and the approach.

6.5 Discussion

In the evaluation of the decision model, for all but two decision points, we received full agreement feedback. The experts noted that the decision model provides a broad coverage of both decision points and visualization elements. These results show high similarity (i.e., overall more than 80% full agreement for both decision points and visualization elements) with the results we obtained in the prior two evaluations of the decision model—for the prior evaluations, we refer to our work on dashboard generation [6]. As each case study organization in these evaluations operates in a different domain, it is reasonable to infer that the decision model can be generalized to other domains.

As listed in Table 6.5, for two visualization elements in the decision model, the experts shared their partial agreements. These two are: the Bullet graph after the decision point ① and the Box Plot that is identified for visualizing a distribution relationship over time at the decision point ⑪. It is important to note that this feedback is aligned with our observations in the prior evaluations of the decision model. There are two reasons for this occurrence. First, although it is not a very new visualization [53, in 2013], experts in organizations are not very familiar with the Bullet graph visualization, which is often due to the lack of support for that particular visualization in BI software products. Second, the correct interpretation of Box Plots requires decision makers to have sufficient knowledge on statistics as that visualization is an advanced one.

Despite the partial agreements mentioned above, the Bullet graph visualization is considered rather useful by the experts. This attests that the decision model raises the awareness of various means on visualizing performance indicators.

The proposed approach enables decision makers to interactively create process performance dashboards in a few steps. In these steps, the employed chatbot provides real-time visualizations of PPIs and guides decision makers in creating tailored dashboards to their needs. Furthermore, by suggesting visualizations for PPIs, the chatbot assists decision makers in determining the needs for visualizing PPIs from different perspectives. Moreover, without requiring any substantial knowledge on information visualization and dashboard development, decision makers can generate dashboards tailored to their needs. Thus, organizations can avoid having outdated and inert dashboards.

With the interactively created engaging visualizations for PPIs, the approach enables organizations to focus on the messages conveyed via dashboards. Instead of placing all possible visualizations of PPIs on a dashboard, the approach puts only the relevant visualizations for decision makers on a dashboard. Thus, the approach can prevent misleading communications that often lead to unwise decisions in organizations.

The achievement of the aforementioned goals was measured through a questionnaire. In particular, the overall usefulness of the approach is checked in terms of productivity, effectiveness, and performance. The average scores obtained for these criteria are above 8.5 out of 10. This indicates that the approach is useful and can assist experts in making informed decisions.

The approach has the following limitation: denoting the PPIs using the PPINOT metamodel is not automated. However, del-Río-Ortega et al. developed a graphical notation, namely Visual PPINOT [38], to annotate business process models with PPIs expressed using the PPINOT metamodel. For this purpose, Visual PPINOT provides a set of diagramming conventions as a PPI extension to the BPMN standard. This extension is in our consideration to add the capability to our approach for automatically extracting PPIs from process models.

As the approach takes PPIs denoted using the PPINOT metamodel as input, organizations can reduce their management efforts on PPIs by leveraging such a formal notation. Moreover, that formal notation helps organizations to establish enduring traceability between PPIs and processes.

Organizations that focus on automatically generating dashboards tailored to their needs can apply our approach. To do so, the required inputs by the approach must be available. Since the decision model for visualizing KPIs is employed in our approach, the only particular input organizations need to collect is PPIs denoted using the PPINOT metamodel and the event logs to compute the PPI values. Similarly, software vendors can adapt our approach to provide automated dashboard generation features in their BI software products. In this way, they can focus on more intelligent activities regarding dashboards by reducing the burden of customizing dashboards for each client.

Threats to Validity In the context of the conducted case study and obtained results, the following types of validity threats can be considered: reliability, construct validity, internal validity, and external validity.

- *Construct validity*: To mitigate the threats to the construct validity, relevant measures from the related literature are chosen. Specifically, we used the three-points Likert-type scale to collect the opinions of experts on the generated dashboards and developed the questionnaire based on a commonly applied expectation model [16] to assess the usefulness of the approach. Moreover, with the long-term cooperation with the case study organization, the threats to the construct validity were lowered.
- *Reliability*: Regarding reliability, the size of the data used to perform the case study is considerably large. Moreover, the collected event logs contain cases handled in a year. With this, seasonal changes in processes are covered.
- *Internal validity*: This aspect of validity is about whether a third factor affects the studied causal relation and is caused by the researcher or another unknown source. Regarding internal validity, we looked for experts who have substantial knowledge of dashboards and processes inside the case study organization. With this, we identified the experts who can make a valid judgment on the decision model and unbiased comparisons between existing and generated dashboards.
- *External validity*: This aspect is concerned with to what extent the findings can be generalized. Important to note that the decision model evaluated in this work was previously evaluated in different scenarios in different domains. With this, we show that it is possible to generalize the use of the decision model in other real-life settings. In addition, the used PPIs are common PPIs in the ITSM area [11] and can be observed in most organizations. Despite the fact that a limited number of experts are involved, we believe that majority of the opinions of experts are generic and can be observed in many organizations performing ITSM processes. Apart from that, the two process workers work as the backup of each other. Therefore, one can assume that the opinions of experts represent a wider perspective. Finally, by taking the well-established and commonly applied solution on providing automated support for the management of PPIs, PPINOT, we minimize the threats to the external validity of our findings.

6.6 Related Work

The literature in the field of organizational performance measurement is quite extensive as there is a great interest in both academia and industry for this topic. Accordingly, much work has been conducted on visualizing information about organizational performance in the form of dashboards. Notably, researchers proposed various approaches that focus on visualizing performance indicators (e.g., KPI or PPI) and generating dashboards. In this section, we list some of the works that relate to the approach we proposed for the interactive generation of engaging process performance dashboards.

Heidema et al. present an approach in [66] that focuses on the visualization of KPIs on a BPMN model. What visualizations should be used in the context of BPMN is identified based on the dashboard design principles defined in [52]. The approach is tested using a set of KPIs for the manufacturing domain, which is listed in ISO 22400. However, the visualization elements that are supported by the approach are limited. For example, the values in a time series relationship cannot be seen since only sparklines are used, which do not contain values. Moreover, some relationship types are not covered, such as part-to-whole, distribution, correlation, and geospatial. Importantly, as the approach is dependent on BPMN models, visualizing multiple KPIs for a process that has a relatively large or a very small corresponding BPMN model will result in a cluttered view and distract decision makers.

To enable organizations to model their KPIs by annotating business process models, Koetter and Kochanowski [78] proposed a modeling language, namely ProGoalML. A process-oriented dashboard design methodology that can transform ProGoalML documents to formal KPI definitions, which are then used to automatically derive dashboards is proposed by Kintz [75]. That transformation is made possible with the semantic dashboard description language that is employed by the dashboard engine—a component of the proposed methodology. To enable the creation of user specific dashboards, by Kintz et al. [76] the methodology is extended. How KPIs are visualized is determined using the defined four mappings between data types and visual elements. However, what is encompassed by data types remains unclear. For instance, there is an apparent lack of explanation on how the data type of a KPI can be determined. Furthermore, some important relationships that are often observed in quantitative information, e.g., correlation and ranking, are left uncovered. In addition, the ambiguity that is led by the mapping from a data type to two visual elements needs to be addressed.

In a study, Vashisht and Dharia [147] proposed an approach that deals with the integration of a chatbot into BI software. In particular, a chatbot is developed to provide responses to its users regarding KPIs and their values, which are retrieved from the integrated BI software. Since the approach requires pre-developed dashboards, it does not go beyond implementation to retrieve KPIs and their values from a data source. More importantly, the dashboard design principles in the literature are not considered within the approach. Additionally, the approach is limited to only one representation to visualize KPIs: simple textual responses.

Chowdhary et al. [31] proposed an approach to generate dashboards together with the code required for their deployment. A dashboard user models a dashboard with its KPIs using a dashboard model editor. Then, the model is transformed into a metamodel representation and sent to an engine that executes it and creates a dashboard with its code. The approach is enriched with observers [100] to perform the maintenance of dashboards. However, the approach requires organizations to have substantial knowledge on both visualizing KPIs and dashboard design principles. Besides, the dashboard model creation is not automatic, and the notation used for it has to be internalized to apply the approach in an effective way.

To create customized dashboards automatically considering the requirements of different users, Vázquez-Ingelmo et al. leverage the Software Product Line (SPL) paradigm in their approach [148, 150]. What similarities and differences users have in using existing dashboards are determined using domain engineering practices. Then, a feature model [149] is constructed specifying what visualizations will be created within a dashboard. Based on a given user's goals, characteristics, and data access rights, a custom dashboard is provided to the user. Since the approach uses existing dashboards as a base, to what extent the dashboard design principles available in the literature were applied in these dashboards is not considered in the approach. Moreover, the meaning of indicators placed on existing dashboards and the way that their values are computed need to be incorporated into the feature model construction step of the approach.

Aimed at helping non-expert users in decision making by means of semi-automated visualizations, a methodology is proposed by Lavalle et al. [84]. Such users can generate dashboards to monitor their goals by following the steps described in the methodology. In particular, a User Requirements Model [83] is created considering the needs of users. Then, this model is combined with another model that captures the characteristics of data sources (i.e., Data Profiling Model [82]) to determine what type of visualization will be created. For determining visualizations, the SkyViz approach introduced in [60] is employed. With a Visualization Model [82], users are allowed to fine tune the visualization that is selected as suitable. Finally, all visualizations are grouped in a dashboard. Despite the usefulness of the methodology, some challenges are not completely addressed in it. For example, the link between user needs and data sources has to be constructed manually. Due to missing indicators (i.e., only measures used), to what extent a process attains its goals cannot be observed at a glance. This means that the message that a visualization conveys is subjective to its users. Therefore, the interpretation of the same visualization may differ between its users.

Apart from the aforementioned works, Van de Wiel et al. [135] studied several techniques to discover usage patterns of dashboard users such that a personalized dashboard for a given user can be built. Usage logs of the web pages of an ERP software, each of which contains a particular dashboard, are taken as the data source. Then, usage statistics of the components (e.g., various visualizations) contained in each dashboard are analyzed. By applying some pattern discovery techniques (e.g., process mining, clustering, and association rule mining), a model is developed that can suggest a dashboard tailored to the profile of a given user. However, it is not analyzed whether the components in existing dashboards are visualized using the dashboard design principles in the literature. Therefore, it is unlikely that the tailored dashboards will be engaging. Moreover, the mixed visualizations of metrics and indicators may cause a cluttered view that generally hinders the substance that is needed for making informed decisions.

With the approach that we proposed in this work, we provide a solution to the issues explained above.

6.7 Conclusion and Future Work

In this chapter, we presented our approach for the interactive generation of engaging process performance dashboards. This approach is developed by extending the automated dashboard generation approach that we previously proposed. The extended approach employs a chatbot that interacts with its users to build process performance dashboards by collecting their interests on Process Performance Indicators (PPIs). The chatbot provides real-time visualizations of PPIs for its users to view and decide on what visualizations should be included in a dashboard, which then becomes an engaging dashboard. How appropriate visualizations for PPIs are determined by the chatbot is through the employed decision model, which we previously developed. Based on the attributes of PPIs, that decision model is traversed by the chatbot, PPI values are computed, and then embedded into the chosen visualization. For the automated analysis of PPIs and computing their values automatically, the approach leverages the PPINOT metamodel, which is a powerful formal notation and allows advanced definition and instrumentation of PPIs.

We evaluated our approach by conducting two tasks: an evaluation of the previously developed decision model before using it in a different organizational setting and the evaluation of the interactively created dashboards using our approach. The two tasks were carried out in an educational institution. In both tasks, we conducted the evaluation by semi-structured interviews with experts who are actively involved in managing and monitoring process performance dashboards in the ITSM area. Moreover, through a questionnaire, we showed the usefulness of our approach.

As a result, we showed that the approach enables organizations to interactively generate dashboards tailored to their needs without requiring any substantial knowledge on information visualization and dashboard development. With suggested visualizations for PPIs, the approach increases the awareness in organizations on various means for visualizing information. Importantly, the approach can assist experts in making informed decisions by taking the focus on the messages conveyed via engaging visualizations. As no other approach available in the literature that enables the generation of engaging dashboards interactively, we conclude that our approach sets itself apart from the state-of-the-art in helping both software vendors and organizations to lower the significant efforts for developing engaging dashboards.

In future work, we want to extend our approach in two ways. First, we will add support for automatically extracting PPIs from process models. For this, we plan to incorporate Visual PPINOT into our approach. Second, we want to include predictive technologies such that decision makers can take preventive actions instead of rather costly corrective actions. Apart from that, we want to incorporate a qualitative analyses of transcriptions of interviews into the future case studies. Finally, user experience (UX) in dashboards [26, 86] is an interesting avenue for future work.

Chapter 7

Performance Improvement through Process Benchmarks

Abstract
The recurring but mutually distinct ways of executing a business process are referred to as process variants. There are approaches available in the literature aimed at finding such process variants and determining how they differ from each other. However, organizations are more interested in understanding the effect of these differences in terms of the performance of a business process. In this context, we propose a novel approach to enable organizations to learn from each other through business process benchmarks. To do so, the approach bins organizations based on what extent they achieve their performance targets in relation to their Key Performance Indicators (KPIs). Within each bin, process variants are identified using trace clustering. Then, significant differences among process variants are determined and highlighted. These differences help organizations to improve the performance of their processes. We implemented our approach, evaluated its performance, and applied it in a case study.

keywords- process variants, trace clustering, benchmarking, key performance indicators, process mining

This chapter is based on the following publication:

[3] Ü. Aksu and H. A.Reijers. How business process benchmarks enable organizations to improve performance. In *2020 IEEE 24th International Enterprise Distributed Object Computing Conference (EDOC)*, pages 197–208, 2020.

7.1 Introduction

The way a business process is executed may differ across organizations, or in some branches of the same organization, or for certain products or services offered to customers, or even across organizational units. The term *process variant* is used for the subset of the executions of a business process that distinguishes it from others due to various reasons. Identifying the similarities in and differences between process variants helps organizations to determine improvement opportunities and also actions to prevent undesired changes in the performance of the process. Furthermore, understanding the root causes of differences between process variants enables managers to make informed decisions for improving business processes. For example, one organizational unit may deliver very good performance by applying a best practice. An organization may want to disseminate that best practice across other units as well.

To determine and explain the similarities in and differences between two or more process variants, numerous techniques are available in the literature —in a recent survey [130] such techniques are discussed. In most of these techniques, the event logs that are produced during the execution of a process and corresponding to two or more variants of that process are compared. For this comparison, trace clustering methods are primarily preferred since they are devised to determine homogeneous groups of process instances from the executions of a given process. To do so, process instances, i.e., traces, are grouped based on their similarities. Such similarity is computed either using the alignment between the sequences of traces [19, 47, 48, 67, 90, 151] or from the feature vectors in each a trace is located and represented using its characteristics, e.g., frequency of activities [36, 64, 69, 125, 156]. However, these methods are complex and computationally expensive, especially when longer substrings are used for alignment. Similarly, feature vector-based methods have some limitations: either they only focus on the attributes of activities rather than the relations between activities or consider only proceed and succeed relations. Avoiding false positives while determining similar sequences in length-insensitive sequence problems is a challenge for some approaches, e.g., n-gram based [112]. Another challenge for some other approaches is dealing with noise in the given sequences [69, 90].

In this chapter, we propose a novel approach to enable organizations to learn from each other by means of *business process benchmarks*. A business process benchmark shows how the same business process is performed among the organizations that are in the same context. The approach takes the process performance of organizations in terms of the KPIs that are relevant to them. Organizations that have similar goals in terms of their KPIs, i.e., interested in the same set of KPIs, are grouped. Then, they are distributed into KPI bins based on what extent they achieve their process performance targets. The motivation for this is that organizations that have differences in their process performance are more likely to learn much from each other rather than from the organizations, which perform similarly with them.

Trace clustering is employed to identify the similarities in and differences between the process executions that yield better or worse performance. To address the aforementioned deficiencies of trace clustering methods, the approach adapts a sequence feature extraction technique (called SGT [111, 112]). By identifying and highlighting the significant differences between the obtained clusters in the previous step, the approach provides business process benchmarks. Organizations can use these benchmarks as the basis to identify both opportunities for improvement and actions to prevent undesired changes in the performance of their business processes. We implemented our approach and applied it in a real-life setting after evaluating its performance.

In Section 7.2, we provide the background on Sequence Graph Transform (SGT), which is the feature extraction technique that we employ in our approach for trace clustering. The approach is elaborated in Section 7.3. Implementation details of the approach are given in Section 7.4. Afterwards, in Section 7.5, we evaluate the performance of our approach, and then we apply it in a case study and discuss the obtained results in Section 7.6. In Section 7.7, we provide an overview of related work on trace clustering and organizational benchmarking. Finally, we present our conclusions and potential directions for future work in Section 7.8.

7.2 Theoretical Background

In the subsection below, we explain the feature extraction technique, Sequence Graph Transform (SGT), which we adapted from the area of data mining, and employ in our approach while determining trace clusters to create process benchmarks.

Sequence Graph Transform (SGT)

Depending on the associations between events, a sequence can be either *feed-forward* or *undirected*. In the former case, in a forward direction, events follow one another. In the latter, the order of events does not depend on each other. In this research, we focus on feed-forward sequences since the order of activities depends on one another in most business processes. In other words, the decisions that are made in a step determine the succeeding steps. For example, how an incident will be handled depends on all the decisions made regarding its categorization, prioritization, and severity determination that happens in the earlier steps of a typical incident management process.

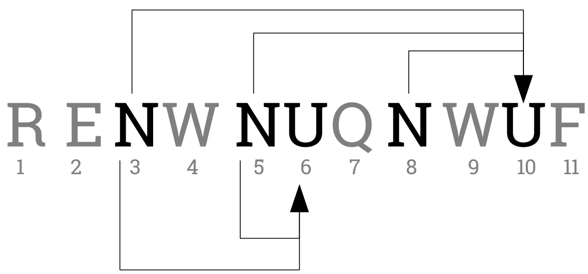

Figure 7.1: SGT: showing *effects* on events on each other.

An example of a feed-forward sequence is shown in Figure 7.1. The predecessors of *Event U* at positions 6 and 10 can be seen in the figure. The relative positions of one event at a time are taken to extract the sequence features. The relative positions for the event pair (N, U) are $\{(3,5),6\}$ and $\{(3,5,8),10\}$. From these positions, sequence features will be extracted. Note that only "*N* leading to *U*" can be interpreted using these positions. "*U* leading to *N*" requires a different set of relative positions, i.e., where *N* is preceded by *U*.

Let $\mathcal{V}$ denote a set of events in a given sequence. The associations between all events can be extracted to obtain sequence features in a $|\mathcal{V}|^2$ dimensional feature space. On the one hand, the similarity in sequences can be measured based on these features. On the other hand, these features can be used to visualize the sequence as a directed graph. For the sake of simplicity, we only give the definitions for SGT below. We refer to the original study on SGT done by Ranjan et al. [111] in which the step by step derivation of these definitions and graph generation are explained.

Let $\mathcal{S}$ denote a set of sequences, $s \in \mathcal{S}$ denotes a sequence that consists of events in $\mathcal{V}$. A sequence can be built using one or more instances of an event from $\mathcal{V}$. For instance, $\mathcal{S} = \{ABABCED, ABBE, ACEDD, \dots\}$ is built using the events in $\mathcal{V} = \{A, B, C, D, E\}$. $\mathcal{L}^s$ denotes the length of a sequence, s is equal to the number of events it contains. The event at position l in the sequence is denoted by s_l, where $l = 1, \dots, \mathcal{L}^s$ and $s_l \in \mathcal{V}$.

Let l and m be the positions of two events and $d(l,m)$ a distance measure. $\phi(d(l,m))$ denotes the quantification of the effect of the preceding event on the latter. Given that events u and v are at positions l and m, ψ^s_{uv} denotes the extraction of features of the sequence s in the form of associations between events, where $u, v \in \mathcal{V}$ and ψ is a function of an auxiliary function ϕ. A distance d and a tuning hyper-parameter κ are taken as the inputs by the function $\phi_\kappa(d)$.

Based on the assumptions above, the derived feature extraction formulas both for length-sensitive (7.1) and insensitive sequence (7.2) analysis problems are taken from [111, 112] and listed below.

$$\psi_{uv}(s) = \frac{\Sigma_{\forall (l,m) \in \mathcal{L}_{uv}(s)} e^{-\kappa|m-l|}}{|\mathcal{L}_{uv}(s)|} \tag{7.1}$$

$$= \frac{\Sigma_{\forall (l,m) \in \mathcal{L}_{uv}(s)} e^{-\kappa|m-l|}}{|\mathcal{L}_{uv}(s)| / \mathcal{L}(s)} \tag{7.2}$$

where $\Psi(s) = [\psi_{uv}(s)], u, v \in \mathcal{V}$ is the SGT feature representation of sequence s.

The SGT feature for the event pair (N, U) in the sequence in Figure 7.1 can be computed as (for κ=1 in length-sensitive SGT):

$$\mathcal{L}_{NU} = \{(3,6);(5,6);(3,10);(5,10);(8,10)\} \textit{ and}$$

$$\psi_{NU} = \frac{\Sigma_{\forall (l,m) \in \mathcal{L}_{NU}} e^{-|m-l|}}{|\mathcal{L}_{NU}|}$$

$$= \frac{e^{-|6-3|} \; e^{-|6-5|} \; e^{-|10-3|} \; e^{-|10-5|} \; e^{-|10-8|}}{5}$$

$$= 0.112$$

As shown above, one can obtain each sequence's SGT features for a given set of sequences. Vector representations can be used for these features to determine similar sequences, i.e., clustering.

Since SGT is good at capturing short and long-term dependencies in sequences, it enables avoiding increasing computation when extracting long-term similarity patterns.

Another advantage of SGT is related to the accurate comparison of sequences that have different lengths. Most subsequence matching techniques often lead to false positives for the given sequences, sq1:*REN*, sq2: *RENWRENREN*, and sq3: *RENWNUNQ*. In particular, in these techniques, sq1 will be consider similar to sq2 and sq3, which is due to the local alignment. To deal with, SGT considers mismatches inherently.

Moreover, SGT is robust to noise that is often observed in sequence problems. For the sake of simplicity, we will refer to the examples related to noise and discussed in [112]. Consider these sequences: a) *RENWFRE* and b) *RENWFRXE*. In b), a noise X appears between R and E. The Markov model transition probability for R and E in the two sequences will be as the following: 1.0 is for a) and 0.5 for b). However, SGT will handle the noise and calculate 0.5 for a) and 0.45 for b) when κ 5. Thus, the effect of stochasticity can be managed.

Considering the aforementioned strengths of SGT, we use it in our approach for trace clustering. In the following section, we explain how we use SGT at the clustering task of our approach.

7.3 Approach

In this section, we give the details of our approach for providing business process benchmarks to enable organizations to learn from each other. Our approach consists of four tasks: (1) selecting organizations and KPIs, (2) binning organizations, (3) clustering traces, and (4) benchmarking. Each task is detailed below in a separate subsection.

7.3.1 Selecting Organizations and KPIs

In order to determine which organizations may learn from each other, we need to identify whether they share the same context. To do so, we focus on Key Performance Indicators (KPIs) that organizations use to monitor whether they attain their performance goals. Our main motivation at this point is that the KPIs that are relevant for organizations indicate that these organizations have similar goals with respect to their business processes. In one of our previous works [7], we showed that organizations that have a similar goal indeed have a shared interest in certain KPIs. As such, we select organizations with respect to the KPIs that are relevant to them. From a set of assessments of the relevance of certain KPIs for organizations, the approach determines for which organizations business process benchmarks can be created and based on which KPIs. The selected organizations and KPIs will be taken as inputs by the next task, i.e., binning organizations.

7.3.2 Binning Organizations

Organizations that have a difference in their process performance are more likely to learn much from each other rather than from the organizations, which perform similarly. In addition, the similarities in the business processes of the organizations that perform similarly can be interpreted as the reasons for their similar process performance. Therefore, in this task, the organizations that are selected in the previous task are distributed into KPI bins. These bins are determined using the KPIs, which are also selected in the previous task. More specifically, the organizations are grouped based on what extent they achieve their process performance target, which is set in their KPIs.

To determine KPI bins, target thresholds of KPIs are used as input in this task. Target thresholds are defined as value-range scales. They are used to interpret to what extent the target of a KPI is achieved [6]. Each threshold has a lower and upper bound value that is used for building the value-range set of a KPI. For example, good: [KPI target-10K, KPI target-30K], bad:[KPI target-30K, KPI target-50K]. Target thresholds of KPIs may vary from one organization to another. Therefore, in this task, KPI bins are determined from the perspective of the organization, which is benchmarking its process performance against others. Thus, our approach deals with the subjectivity of the target thresholds of KPIs.

Furthermore, the approach can handle the variation among the target thresholds of multiple KPIs. For instance, one KPI may have three target thresholds, whereas another may have five. In this situation, the approach creates a KPI bin for each threshold combination. We will indeed involve end-users to select the threshold combinations that are relevant to them. If there are KPIs that are affected by multiple processes, it is required to combine these multiple processes as a meta-process considering their interdependencies. Thus, the selected process discovery algorithms can deal with the event logs of multiple processes. Aside from that, case attributes of process models can be incorporated into binning organizations. However, a scaling strategy is necessary in that case to interpret the meaning of smaller or higher values regarding case attributes.

Figure 7.2 depicts an example of binning organizations based on the target thresholds of a given KPI. In the figure, there are 4 KPI bins: *perfect*, *good*, *bad*, and *worse*. These bins are created based on the defined target thresholds for the given KPI. For example, for the bin *perfect*, 90 is the lower bound threshold value, and 100 is the upper bound value, which is the maximum for this KPI. Similarly, the lower and upper bound values for each target threshold are shown in the figure. The organizations that will be put into each bin are depicted in the figure as well. For example, *Org 02* and *Org 10* belong to the bin *Bad*.

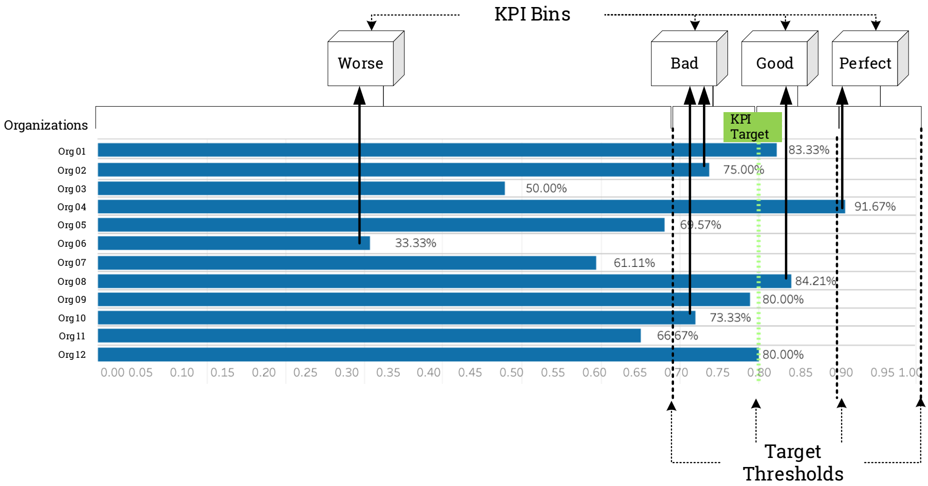

Figure 7.2: Binning organizations using KPI target thresholds

7.3.3 Clustering Traces

In the third task of our approach, the focus is on identifying the similarities in and differences between the business process executions of the organizations in each KPI bin. These bins were determined in the task we described earlier. A sequence feature extraction technique (SGT, see in Section 6.2) is adapted and applied on the business process executions, i.e., event log of the organizations in each KPI bin. In particular, the event log for the process that is monitored by the previously selected KPIs is used while clustering.

An essential decision in clustering is the number of clusters to choose. In the literature, several approaches are available to propose the optimal number of clusters for a given clustering problem. However, the quality of the process models that will be discovered for the clusters from a given event log is not taken into consideration in these approaches [156]. Therefore, we opt for using a range for the number of clusters. In our approach, we rely on process owners to determine that range.

Traces are compared by locating each unique trace in a feature space using the *k-means* clustering algorithm. Due to its simplicity, we opted for this algorithm. SGT extracts the short and long-term sequence features in the traces and embeds them in a finite-dimensional feature space. To facilitate the feature extraction, our approach shortens the activity names in the given event log. To do so, the approach employs a label encoder. Note that the approach does not convert activity names to numeric values: a string prefix is combined with the numeric value created by the label encoder, e.g., *Assign* is shortened as *T0*.

Although SGT promises [111] no increase in the computation to reduce the size of the clustering problem, we remove duplicate traces in the given event log. Thus, the approach assigns equal weight to all unique traces. Assigning equal weights on traces enables our approach to handle the varying absolute frequency of a trace that may be seen in the event logs of multiple organizations. Aside from that, without any domain knowledge on a given process, one cannot easily assign a weight based on the frequencies of traces, which mostly varies per organization.

Since the curse of dimensionality is a known phenomenon when analyzing data that have a high-dimensional space, for this clustering task, a dimensionality reduction technique is applied on the generated feature space by SGT. Hence, the essential parts that have more variation of the data are preserved.

The gap between clustering and evaluation is the main challenge of clustering techniques, as noted in [156]. In this context, there exist two ways to address this gap. The first way is evaluating the discovered process model for each cluster while clustering. However, this way is computationally expensive, especially for large event logs. Therefore, in our approach, the quality of the discovered process models is checked after clustering. This is done by using the common process mining metrics [139], namely replay fitness, precision, generalizability, and simplicity.

As a result of this task, we obtain the clusters for each KPI bin that best represent the variance in the behavior captured in a given event log. In the next and the last task of our approach, we focus on spotting significant differences between the discovered process models for these clusters.

7.3.4 Benchmarking

Organizations can benefit from determining the differences between and the similarities in their business processes that yield better or worse performance. For instance, what yields better performance in a business process can be interpreted as opportunities for improvement. Similarly, what yields a worse performance in a business process can become the basis for finding the actions to prevent any decline in the performance. In this context, the last task of our approach is devoted to identifying relevant differences between process models. They will be discovered from the event log of the clusters, which are created in the previous task.

In a recent survey on process variant analysis [130], approaches on determining differences between process executions are studied. We checked the applicability of these approaches for detecting statistically significant differences between process executions within our approach. Aligned with our goal, the approach proposed by Bolt et al. came forward since it provides an extensible basis for process specific metrics in addition to KPIs. Moreover, that approach is available as a plugin (called Process Comparator) in the process mining framework, ProM [145], which offers built-in features for handling event logs. As such, in this task, we execute that plugin for the event logs of each cluster pair. The cluster pairs are formed by picking the best performing cluster of two KPI bins based on fitness, precision, and generalizability. For the balancing of these process model metrics, we used the order in which they are listed here. This forming step ends when all KPI bin combinations are checked. Then, the event logs for a selected cluster pair are analyzed by the plugin, and the statistical differences between these event logs are projected onto a transition system. In the transition system, states and transitions are colored to spot those differences. Moreover, the thickness of each node's borders and arcs in the same transition system indicate the frequencies of the states and transitions for each cluster in the selected cluster pair. In addition, a set of metrics for highlighting the differences from the control-flow (frequency) and time perspective (elapsed time, remaining time) provided by the plugin are used for the selected cluster pairs.

Figure 7.3 illustrates how the significant differences between the event logs of a selected cluster pair are highlighted. The nodes in the figure represent activities, whereas arcs reflect the sequence of these activities in a process. The thickness of the arcs and nodes is determined based on the value of the selected process metric. Similarly, opposite colors, i.e., red and blue, are used to indicate significant differences between two event logs in terms of the selected process metric. For example, T5 is a significant activity in one cluster, whereas the activities T15 and T23 are significant in another cluster as the opposite colors are used for each (red vs. blue colors). In the next section, we give the details of the implementation of the approach.

7.4 Implementation

To demonstrate the feasibility of our approach and evaluate it in practice, we implemented it[1]. The implementation of the approach consists of two components. The first component covers the first three tasks of the approach (Selecting Organizations and KPIs, Binning

[1]The implementation of our approach is available at `http://amuse-project.org/software/`

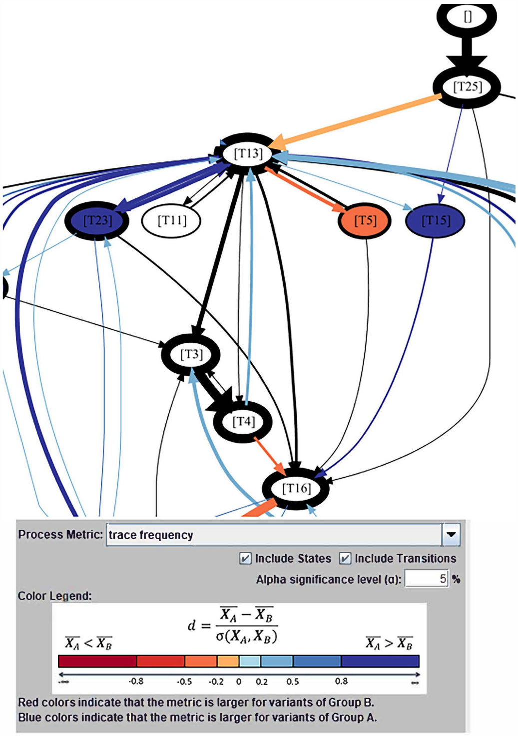

Figure 7.3: Detecting statistically significant differences between the event logs of a cluster pair for benchmarking

Organizations, and Clustering Traces). This component is developed in Python as a script that can be directly executed. The second component implements the last task of the approach, i.e., benchmarking. It is developed in the form of a plugin of the ProM process mining framework [145].

In the first component, we reused the SGT implementation in Python that is publicly available[2]. In the reused code, we configure two parameters that are relevant for our approach: *kappa* and *length-sensitivity*. The first parameter adjusts to what extent the long-term dependency will be captured while embedding the extracted sequence features in a finite-dimensional feature space. 5 is chosen as the value for this parameter to have a medium level dependency capture in the embedding. The second parameter is used to determine whether the length of sequences will be considered in embedding. The value of this parameter is set to *true* to capture the effects of trace lengths.

To deal with high dimensionality at feature extraction, we used the Principal Component Analysis (PCA) algorithm implementation in the *scikit-learn* Python machine learning library. We set the minimum variance coverage rate as 95% at dimensionality reduction since it is a commonly used value in similar problems.

[2]The SGT implementation in Python is available at `https://github.com/cran2367/sgt`

As mentioned, the second component of the implementation is devoted to the benchmarking task. The component calls the Flexible Heuristics Miner (FHM) plugin [157] of the ProM framework with default parameters. The reason for using FHM is that it is mostly used in other trace clustering methods [47, 156], and recommended as a robust discovery algorithm. The plugin reads the event log of each cluster and creates a heuristics net for each. Each heuristics net is converted to a Petri net to replay [116] them on the log and compute fitness using the PNetReplayer plugin [143]. Similarly, the PNetAlignmentAnalysis plugin [2] is employed to compute precision and generalizability. Based on these computed metrics, our approach determines the clusters for each KPI bin that best represent the variance in the behavior captured in a given event log.

In the next section, we evaluate the performance of the approach by comparing it with state-of-the-art trace clustering approaches. Then, in Section 6.6, we apply our approach in a case study.

7.5 Performance Evaluation

In this section, we explain how we evaluate the performance of our approach. Specifically, we compare the performance of the trace clustering task of our approach with state-of-the-art trace clustering methods. This is relevant because this task adapts an existing clustering method. For the performance evaluation, we followed the same scenario (Evaluation Scenario 2) described by Evermann et al. in the study [47] as it serves the same purpose: the method proposed in that study is compared with state-of-the-art trace clustering methods. The steps of that scenario are listed below.

1. Identify a set of configurations for the adapted trace clustering method based on the cluster range used in state-of-the-art trace clustering methods:
2. For each identified configuration:
   (a) Apply the configuration to the adapted trace clustering method
   (b) Perform clustering on the given event log
   (c) *Apply Flexible Heuristic Miner (FHM)* to discover a heuristics net
   (d) Convert the heuristics net to a Petri net
   (e) Compute performance metrics: simplicity (CN, CNC, and delta)
   (f) Replay the event log on the Petri net using *PNetReplayer*
   (g) Compute performance metrics: fitness
   (h) Replay the event log on the Petri net using *PNetAlignmentAnalysis*
   (i) Compute performance metrics: precision and generalizability
3. Compare the performance of the adapted trace clustering method with the performance of state-of-the-art trace clustering methods

We use the same set of performance metrics as well as the event log, which is taken from a loan application process and publicly available[3]. We compare the obtained results

[3]The event log is available at `https://joerg.evermann.ca/software.html`

with the performance measurements for state-of-the-art trace clustering methods, which are available in the aforementioned study. In Table 7.1, the abbreviations used in the column *Conf* refer to the following clustering methods respectively, *Active Trace Clustering* (ActiTraC) [156], *Disjunctive Workflow Schema Mining* [36], *Trace Clustering* [125], *Sequence Clustering* [151], and *AlignCluster* [47].

In these three configurations, we used several values for the number of clusters. The values that we used as the number of clusters are taken from the aforementioned study. Specifically, we used 3, 6, and 9 as the number of clusters. The results that we obtained are shown in Table 7.1. In the table, the best values for each quality criterion are shaded; the worst values are italicized.

As can be seen in the table, 4 out of the 6 criteria (CNC, CN, Fit, Prec), the last configuration that we used in our trace clustering method outperformed state-of-the-art trace clustering methods. In particular, our approach provides simpler process models with better precision and fitness. Moreover, for the remaining 2 criteria (Delta and Gen), in that configuration, we obtained values that are very close to the best value. However, the first configuration that we used has a lower precision but very good fitness and generalizability. In addition, in this configuration, our approach provides somewhat simpler process models. Although we used a limited set of configurations, this evaluation shows that our approach generally achieves a better performance than state-of-the-art trace clustering methods.

In the following section, we apply the technique in a real-life setting.

7.6 Case Study

In this section, first, we introduce the case study organization where we applied and evaluated our approach for creating process benchmarks. Second, we explain how we collect the data that are used in the evaluation. Third, we list the experts who interpreted the results that we obtained in the application of our approach. Fourth, we elaborate on the application of our approach to the collected data. Finally, we discuss the relevance of the results to the case study organization.

7.6.1 Case Study Organization

One of the biggest educational institutions in the Netherlands expressed its interest to us in learning from the data recorded about its IT Service Management (ITSM) processes. One of the vital processes among them is the Incident Management (IM) process. That process defines the way of managing questions, requests, and, most importantly, malfunctions about the products and services offered by the institution. Moreover, the process is executed through a third-party software. Since the institution offers a wide range of products and services to more than 25K customers (20K students, 5K employees), multiple organizational units are involved in this IM process. Each organizational unit is devoted to ensuring the quality of a particular set of offered products and services. For example, cloud-based email and printing are two services that are managed by separate organizational units.

Table 7.1: Performance of state-of-the-art trace clustering methods vs. our approach

Conf	**CNC**	**CN**	**Delta**	**Fit**	**Prec**	**Gen**
AT-3	1.1670	33.7120	0.0240	0.7000	0.4737	0.7322
AT-6	1.1198	26.196	*0.0326*	*0.6670*	0.5663	*0.6011*
AT-6-ICS95	1.1709	27.3087	0.0072	0.8529	0.3751	0.9661
DWS-Std	1.2275	30.8013	0.0103	0.8783	*0.3219*	0.9586
DWS-55510	1.1579	17.396	0.0208	0.7721	0.5459	0.9581
TC	1.1773	31.6533	0.0270	0.7823	0.4062	0.7419
TC-W1-H3	*1.2434*	*46.1867*	0.0129	0.8213	0.3232	0.8937
TC-W2-H3	1.1792	35.1787	0.0217	0.7235	0.4413	0.8037
TC-W3-H3	1.1542	31.6587	0.0279	0.6991	0.4686	0.7309
TC-W4-H3	1.1542	31.6587	0.0279	0.6991	0.4840	0.7332
SC-3	1.1976	32.9653	0.0071	0.8475	0.3631	0.9598
SC-6	1.1346	20.5973	0.0071	0.8644	0.5103	0.9905
SC-9	1.1273	17.7667	0.0078	0.8623	0.5408	0.9335
SC-12	1.1048	13.6540	0.0083	0.8607	0.5502	0.9628
SC-15	1.0974	12.0793	0.0104	0.8919	0.5760	0.9793
AC-maxPrec [mmP=-0.5, cGO=0.5, cGE=1, useSim=F, dim=log, c=9]	1.1201	15	0.0094	0.8036	0.5992	0.9957
AC-maxGen [mmP=-0.5, cGO=0.5, cGE=0.5, useSim=F, dim=log, c=3]	1.1507	22	0.0073	0.7775	0.5158	0.9988
AC-maxFit [mmP=-0.5, cGO=0.5, cGE=0.5, useSim=F, dim=log, c=9]	1.1209	14	0.0104	0.854	0.5679	0.9965
AC-minCNC [mmP=-1, cGO=1, cGE=1, useSim=F, dim=log, c=9]	1.1164	13	0.0108	0.8502	0.5872	0.9969
AC-minCN [mmP=-1, cGO=1, cGE=1, useSim=F, dim=log, c=9]	1.1164	13	0.0108	0.8502	0.5872	0.9969
AC-minDelta [mmP=-0.5, cGO=0.5, cGE=0.5, useSim=F, dim=sqrt, c=3]	1.1840	30	0.0072	0.7888	0.4767	0.9941
Our Approach [SGT_kappa=5, SGT_length-sensitive=T, c=3]	1.1765	20	0.0100	0.9823	0.3339	0.9860
Our Approach [SGT_kappa=5, SGT_length-sensitive=T, c=6]	1.1220	19	0.0069	0.9678	0.5213	0.9985
Our Approach [SGT_kappa=5, SGT_length-sensitive=T, c=9]	1.0909	11	0.0083	0.9837	0.6292	0.9968

To monitor customer satisfaction and provide the same level of quality for each product and service, the institution has a set of KPIs. These KPIs have the same target for every organizational unit that is involved in the same ITSM processes. Specifically, there are some KPIs for the IM process, and all of them are used to determine whether each organizational unit attains the shared goal of the IM process. Performance is tracked by monthly reports. In 2019, a few of the organizational units were not able to achieve their objectives about dealing with the malfunctioning of a number of products and services. Therefore, the institution wanted to investigate what the differences and similarities are in the execution of the IM process. Moreover, the managers involved in the IM process assumed that some organizational units may follow best practices, which would explain why they would provide a better performance.

Since both the setting and needs in the educational institution are highly related to the approach that we propose in this chapter, its application in this context is highly relevant to determine its applicability and value.

7.6.2 Data Collection

In accordance with the problem that the case study organization faced in 2019, we extracted the event log for the malfunctions processed in that year. Specifically, the event log consists of $150K$ events and $12K$ cases in which 26 unique activity and 40 unique organizational units exist. The time frame of the event log is one year. In addition to the required minimum attributes, the event log contains to what extent the resolution time of a malfunction adheres to the defined target of the single KPI for malfunctions. The KPI has a defined target value, which is 80% and there are four target thresholds: $[0-69]$ *worst*, $[70-79]$ *borderline*, $[80-89]$ *sufficient*, and $[90-100]$ *best*.

Before applying our approach to the collected data, together with the experts listed, we checked the number of cases per organizational unit. At this check, we found out some outliers. In other words, a very small number of malfunctions are received and handled by some organizational units that are specialized in occasionally used services and products. The experts suggested to filter out the organizational units that handled less than 50 malfunctions in a year to eliminate the potential bias that may be caused by these infrequent malfunctions. This filtering out cannot be considered as an issue for the applicability of our approach since more than 75% of the filtered out organizations received less than 10 malfunctions in a year. Moreover, such a low number of malfunctions is not adequate to determine relevant patterns in the process executions from which other organizational units can benefit and learn much. As a result of the filtering, in total, 24 organizational units remained for our analysis purposes.

7.6.3 Applying the Approach

We applied our approach using the collected data in the case study organization. As explained in the approach, in the first task, the organizations that may learn from each other will be selected based on the KPIs that are relevant to them. Accordingly, we selected the remained 24 organizational units since there is only one particular KPI used for monitoring the IM process for malfunctions.

Table 7.2: Experts involvement in the evaluation of the obtained results

Expert	Area of expertise	Years of expertise	Meeting duration (hours)
Process manager	Incident management	>10	3
Process manager-2	Incident management and Change management	>10	2
First-line support manager	Incident management	>20	2.5
Product and service manager	Educational services	>15	1.5

In the second task, our approach created 4 KPI bins based on the given 4 target thresholds for the given KPI. These are: KPI bin-worst, KPI bin-borderline, KPI bin-sufficient, and KPI bin-best. The selected 24 organizational units are distributed to these KPI bins based on what extent they achieved their process performance. The number of organizational units put into the created each bin is 12, 6, 5, and 1, respectively.

In the third task, trace clustering was performed. As the number of clusters, together with the experts in Table 7.2, we determined the value range from 3 to 9. Considering this input, for each KPI bin, clusters are created. Afterwards, for each KPI bin-cluster, a process model is discovered and replayed on the corresponding event log to compute the three process mining metrics, namely precision, fitness, and generalizability.

Finally, in the last task, the clusters that best capture the observed behavior in each KPI-bin are selected. This is done using the process mining metrics mentioned. Thus, we obtained all the cluster pairs that are relevant for checking the significant differences between them. Accordingly, we imported the event logs of each cluster pair separately into the Process Comparator plugin, and then obtained the highlighted transition system that is the projection of the differences between the event logs as the final results (i.e., business process benchmarks) of our approach. We interviewed with experts and asked each one to explain whether they see any performance improvement opportunities. As a result of their combined feedback, in 4 out of the 6 business process benchmarks, we observed relevant points that can enable the organizational units in the case study organization to learn from each other for improving the performance. In the remainder of this section, we discuss these.

7.6.4 Discussion

Together with the experts (see Table 7.2) in the case study organization, we analyzed the differences that are highlighted in the selected business process benchmarks. The business process benchmarks are determined by the experts as the ones that have the most potential to learn from for improving the IM process. In the first three figures, red colors are used to indicate the greater value in the frequency of the activity in the KPI bin-cluster that has a lower process performance achievement than the paired one with a bigger cluster value. Similarly, blue colors are used to indicate the greater value in the frequency of the activity in the KPI bin cluster that has a better process performance achievement than the paired one with a smaller cluster value. In the last figure, the colors serve similar purposes for the KPI bins but indicate duration values.

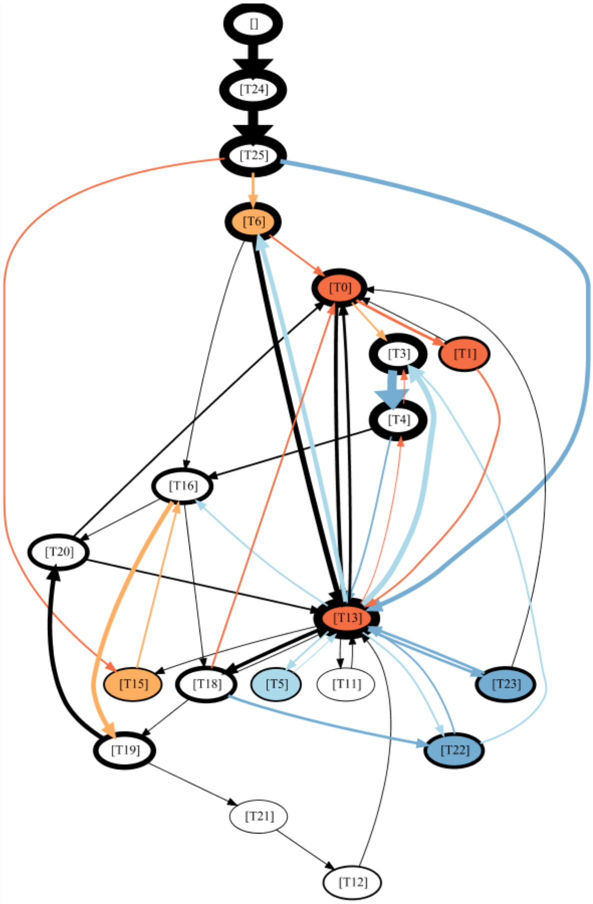

Figure 7.4: Borderline vs. Worst

Figure 7.4 depicts the highlighted statistically significant differences between the process execution of the organizational units (KPI bin-worst and KPI bin-borderline) that could not attain their KPI goals for the IM process. In the IM process, T0 and T1 are the activities that are related to hand-overs, whereas T22 and T23 correspond to interactions with callers. As shown in the figure, while T0 and T1 have red colors, T22 and T23 have blue colors. The experts interpreted this difference as the following: the organizational units in the KPI bin-borderline tend to add more information to the received malfunctions in terms of comments or questions. Then, they stop the SLA timer by moving malfunctions to the callers. However, the organizational units in the KPI bin-worst do hand-over the malfunctions either to other operators or organizational units. Therefore, the time spent in handovers increase and yield lower process performance. The experts suggest that these organizational units should use the comments area so that other operators will get notified and involved in the malfunction.

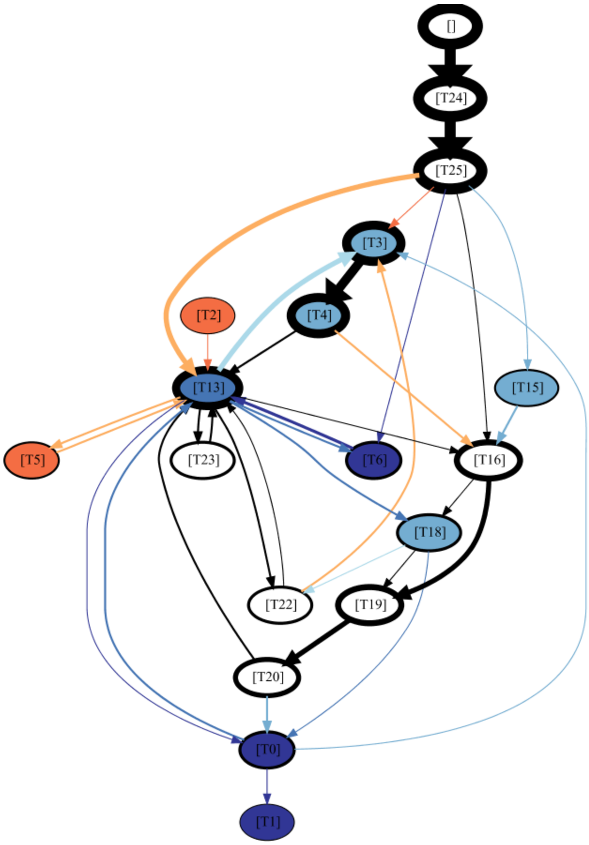

Figure 7.5: Sufficient vs. Borderline

Figure 7.5 shows the differences between the organizational units that were close to achieving the process performance and the organizational units that have already achieved it. With this, we aimed to find out what may be the quick wins to improve the process performance of the organizational units in the KPI bin-borderline. As depicted, T2 and T5 are the two red-colored activities. In the IM process, while activity T2 is related to supplier involvement, T5 is about send backs between the second-line support and the first-line support. The significant frequency of these activities grabbed the interest of experts. The experts interpret this situation as a ping-pong behavior for the malfunctions in which suppliers are involved. Moreover, this is mentioned as an unexpected situation since several big suppliers are already integrated into the IM software to avoid such send backs. To deal with this situation and identify new supplier integration needs, together with experts, we carry out the cluster analysis of the malfunctions in which suppliers are involved

The differences between the organizational units that have attained their KPI goals are shown in Figure 7.6. As can be seen in the figure, the light-blue colored activities happen less frequently in the best performing organizational units. The experts mentioned that T23 is a specific activity that is about the confirmation from callers for resolved malfunctions. We

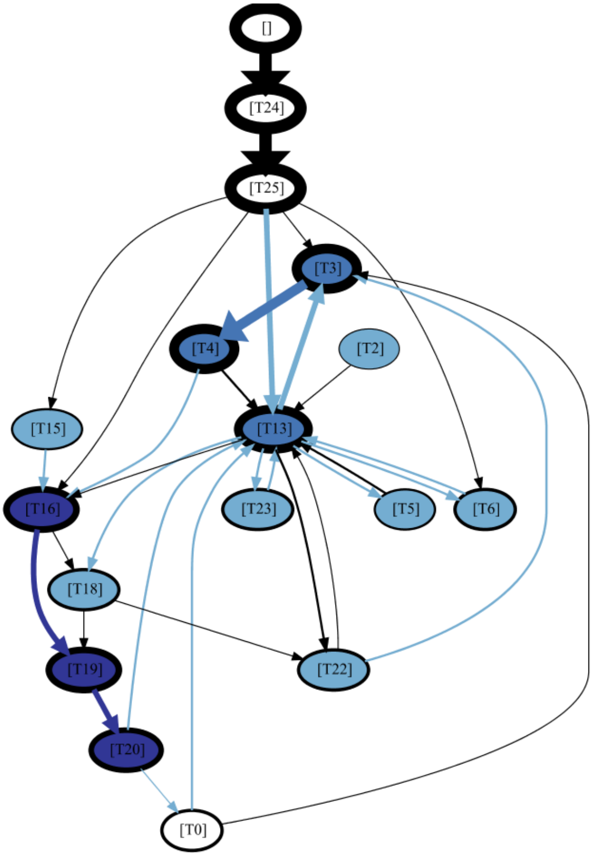

Figure 7.6: Best vs. Sufficient

checked that there is only one organizational unit in the KPI bin-best. This organizational unit has a face-to-face interface with callers. Therefore, the organizational unit gets confirmation from callers, meanwhile resolving malfunctions. Thus, the confirmation activity is mostly not necessary for this organizational unit. Moreover, the experts mentioned that although there is a confirmation collection step in the process, this activity is not frequent in the discovered process models as expected. Currently, the experts redesign the IM process to remove the necessity of a caller confirmation for resolved malfunctions in certain situations.

In Figure 7.7, a time perspective (i.e., duration) is chosen: the differences between the organizational units that were close to achieving the process performance and the organizational units that have already achieved it are highlighted. The experts indicated two important points based on the figure. Firstly, activity T5 shows the delay due to the send backs between the first-line support and the second-line support. Secondly, the organizational units that perform sufficiently spend more time (from T0 to T13) on the discussions for determining operators for resolving malfunctions. To find the reason behind the second point, together with experts, we analyzed the malfunctions of the KPI

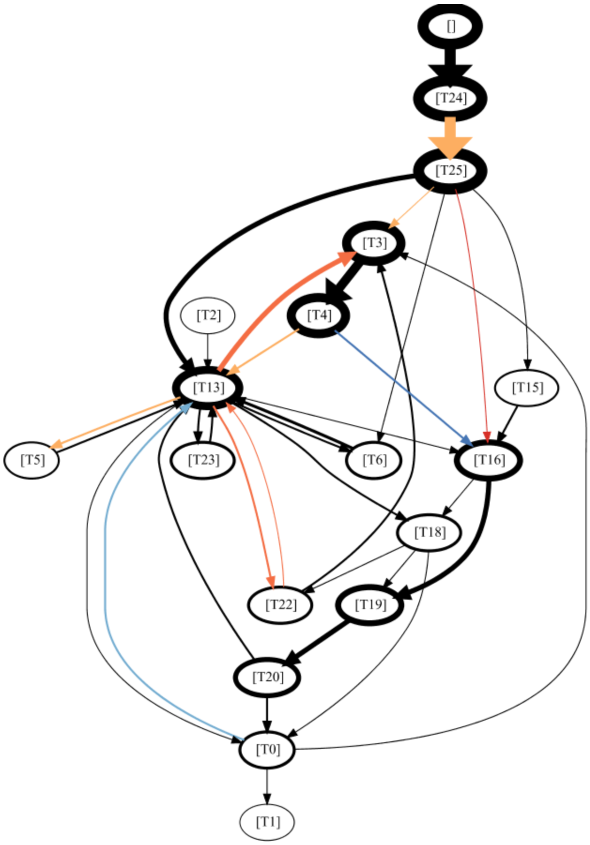

Figure 7.7: Sufficient vs. Borderline (time perspective)

bin-sufficient organizational units. In particular, we checked the organizational units that are involved in commenting on malfunctions for determining resolving operators. From these, we observed that some of the organizational units prefer face-to-face discussion if the involved operators are in their neighbor organizational units. The experts interpret this as a strategy to avoid send backs and potential delays. Based on this observation, together with experts, we investigate the correlation between non-neighbor organizational unit involvements and process performance achievement.

Based on the discussions and the results, the case study clearly confirms that the approach proposed in this chapter can enable different parts of an organization to learn from each other. Moreover, software vendors that offer software products to their client organizations may benefit by applying our approach. In particular, those clients can be benchmarked based on their process performances since they use the same software product offered by the same vendor. In addition, governmental bodies can learn much from each other through process benchmarks since the majority of the processes they perform often overlap.

Domain expert involvement is a key element in identifying improvement opportunities and prevention actions from process benchmarks. While clustering traces, equal weights are assigned on traces to capture all behaviors in the events logs of all organizations. Assigning varying weights on traces may reduce the dependency on domain experts. For example, building a drill-down mechanism on process benchmarks and focusing on mainstream behavior may help non-business users to interpret the significant differences between the mainstream behavior of organizations.

It is important to note that the labels and meanings of the activities in a business process, which is performed in several organizations may vary. This is an important challenge in addition to gathering and pre-processing the data of such a process from different information systems in those organizations. In our previous works [4, 5], we discussed such challenges by introducing the approaches, which we proposed to meet them. The proposed approach in this chapter concludes our framework introduced in [4]. Therefore, our approach relies on the internal outputs of our framework that is aimed at meeting the aforementioned challenges.

7.7 Related Work

In this section, we list some of the works that relate to the approach we proposed for enabling organizations to learn from each other using process benchmarks. The most recent survey [130] on process variant analysis served as the basis for the snowball search method that we used for finding the works.

An analytical approach proposed by Buijs and Reijers [22] compares the variants of a process based on the alignments between the executions and the model of that process, which is performed by multiple governmental bodies. Partington et al. [103] compared patient pathways across four hospitals to visually detect the differences in process variants from which the hospitals can learn from each other. Trace clustering is combined with data and text mining in [155] to find the patterns that drive certain control-flow behavior in process variants.

Furthermore, some works [9, 33, 65, 159] focused on advancing visualizations to ease the identification of differences between process variants.

Ballambettu et al. [13] annotated the process models of process variants with a set of metrics to detect the key differences between these variants. Nguyen et al. [96] introduced Perspective Graphs that are graph abstractions of event logs using a set of relations between entities (e.g., resource, location, etc.) in the event logs. These perspective graphs are then employed to compare process variants and identify the differences between them. Similarly, event logs for process variants are projected onto transition systems by Bolt et al. in [17]. After that, these transition systems are compared to detect the significant differences in their states and transitions in terms of commonly used metrics (i.e., frequency, duration, etc.). However, in these works, the main focus mostly is on where process variants differ and how. Which differences matter is only determined based on a limited set of metrics rather than relevant KPIs for organizations. Moreover, domain experts need to go through each process variant pair. The first two tasks in our approach aim to cope with these challenges.

7.8 Conclusion and Future Work

In this chapter, we presented a novel approach aimed at enabling organizations or organizational units to learn from each other through business process benchmarks. As the organizations that are in the same context will learn much from each other, for benchmarking, we used the KPIs that are relevant for them.

To identify which variations in the process executions among organizations yield a better or worse performance, trace clustering is employed in our approach. As the order of tasks mostly depends on one another in business processes, a sequence feature extraction technique proposed to address this challenge is used within the clustering traces task of the approach. Then, significant differences between the process variants in the obtained clusters are determined as benchmarks.

The evaluation of our approach consists of two steps. First, we tested the performance of the trace clustering method in our approach by comparing it with state-of-the-art trace clustering methods. The result of this test showed that our trace clustering method has a better performance than other methods. Secondly, we applied our approach in a case study in an educational institution. We discussed the results that we obtained in that case study. This discussion showed that our approach is very helpful in revealing useful insights for organizational units to learn from each other. As the proposed approach is generic in the way it is developed, there is no limitation for applying it across organizations. By doing so, organizations can learn from each other to improve their performance.

In future work, we want to extend our approach by making it goal-driven. For example, an organization may be interested in learning from other organizations to identify only preventive actions to eliminate the risk of performing badly. As such, that organization will want to observe what similarities are present in the process variants of inefficient organizations. Moreover, we would like to incorporate decision mining technologies in our approach since the decision points in a business process are very likely to be associated with its process variants. Analyzing the organizations that achieve similar performance with different process models is also an interesting future work direction. Lastly, the approach can be extended to capture dependencies between business processes in organizations to provide cross-benchmarks for organizations.

Chapter 8

Conclusion

This chapter concludes this dissertation research in which we developed a theoretical cross-organizational process mining framework. The ultimate goal of the framework is to provide relevant insights for organizations to improve performance. The framework orchestrates three components, each of which is devoted to providing solutions to the issues introduced in Chapter 1. First, we summarize the contributions of this dissertation research by reflecting on the way how they address the research questions, which are defined in the introduction chapter (see Section 1.1). After that, the implications of the presented framework for research and practice are summarized. We list some open issues and limitations of the presented approaches. Finally, an outlook on future research directions is provided.

8.1 Contributions

In this dissertation, the cross-organizational process mining framework that we developed is presented. The framework orchestrates three components that provide solutions to the three main issues (unfairness, inaccuracy, and inadequacy) in cross-organizational process mining applications. The main contributions of this thesis are on the basis of the developed five novel approaches that are employed by the three components. How each contribution provides answers to the research questions specified in the introduction chapter is elaborated below.

- The theoretical cross-organizational process mining framework: This contribution provides answers to the main research question, i.e., **MRQ** (*How can cross-organizational process mining be performed in a fair, accurate, and adequate fashion to provide relevant insights for organizations to improve their performance?*).
  To perform cross-organizational process mining in a fair, accurate, and adequate fashion, we introduced the theoretical cross-organizational process mining framework. The framework consists of three components: (1) Metric Comparison Catalog, (2) Metric-Based Improvement Catalog, and (3) Cross-Organizational Process Mining Techniques. Each component aims to address a particular issue. More specifically, the Metric Comparison Catalog deals with the unfairness issue. Similarly, the Metric-Based Improvement Catalog focuses on addressing the inadequacy issue. The inaccuracy issue was addressed by the Cross-Organizational Process Mining Techniques component.

- The approach for automatically deriving KPIs from Ontological Enterprise Models: This contribution provides answers to research question **RQ 1** (*How can organizations be selected for fair cross-organizational process mining?*).
  How organizations that are in the same context can be identified is the main focus of this approach. Thus, a fair comparison can be performed via cross-organizational process mining. An Ontological Enterprise Model (OEM) is the input that the approach requires to automatically obtain knowledge on the context of organizations. Furthermore, the approach employs KPI derivation patterns to derive tailored KPIs for organizations. With this, the approach ensures the fair comparison of organizations by determining the ones that share enterprise concepts.

- The approach for the automated prediction of relevant KPIs for organizations: This contribution finds answers to research question **RQ 2** (*How can organizations be compared for accurate cross-organizational process mining?*).
  To enable an accurate comparison of organizations, this approach focuses on the automated selection of KPIs for organizations. Thus, the organizations that have a shared interest in the same set of KPIs can be compared. To automate the selection, the approach employs prediction models that are trained on predicting relevant KPIs. Moreover, these models are able to determine the factors that make certain KPIs relevant for specific organizations.

- The approach for the automated generation of dashboards: This contribution provides answers to research question **RQ 3** (*How can adequate perspectives be determined for*

cross-organizational process mining?).
By generating the dashboards that organizations will be engaged in, the approach identifies what perspectives can be adequate for providing relevant insights for organizations. A decision model for visualizing KPIs is employed by the approach to find the best means to visualize KPIs. Using this knowledge, the Metric-Based Improvement Catalog of the framework can detect how differences in and commonalities between organizations should be converted to benchmarks.

- The approach for the interactive generation of process performance dashboards: As this approach is developed by extending the previous approach, similarly, it provides answers to research question **RQ 3**. In particular, the approach focuses on generating dashboards at the process level (i.e., process performance dashboards), unlike the previous approach, which focuses on the dashboards at the organization level. To interactively generate process performance dashboards by collecting the interests of decision makers on Process Performance Indicators (PPIs), a chatbot is devised. The chatbot enables decision makers to view visualizations in real-time and decide on which ones should be included in a dashboard. Thus, the approach assists experts in making informed decisions by taking the focus on the messages conveyed via engaging visualizations.

- The approach aimed at enabling organizations to learn from each other by means of business process benchmarks: This contribution finds answers to research question **RQ 4** (*How can relevant insights for organizations be provided using cross-organizational process mining?*).
The goal of this approach is to enable organizations to learn from each other. In particular, the approach is devoted to providing relevant insights for organizations in the form of process benchmarks. In these benchmarks, the significant differences in the executions of processes in multiple organizations are highlighted based on the relevant KPIs. These are automatically derived, predicted, and visualized by the approaches explained above. What practices yield better or worse performance in organizations can be obtained from the highlights on process benchmarks. Thus, this approach ensures that relevant insights will be provided for organizations.

To perform a fair comparison of organizations through cross-organizational process mining, we developed the approach for automatically deriving KPIs from Ontological Enterprise Models. That approach focuses on identifying organizations in the same context to tackle unfairness. Another approach that we developed (Predicting relevant KPIs for organizations) focuses on determining the set of KPIs in which organizations have a shared interest, i.e., relevant. With this, one can achieve an accurate comparison of organizations using their relevant KPIs. This approach employs machine learning techniques to predict the relevance of KPIs for organizations and deals with the inaccuracy issue. What perspectives are adequate for providing relevant insights for them are determined by generating the dashboards that organizations will be engaged in. In this regard, our third approach employs a decision model to find the best means to visualize KPIs and generate engaging dashboards automatically. Furthermore, we extended this approach using chatbot technologies. In particular, we integrated a state-of-the-art performance indicator meta-model framework (called PPINOT) into our approach such that process performance dashboards can be created interactively by decision-makers. Our last approach (Building process benchmarks for performance improvement) is devoted to providing relevant insights for organizations in the form of process benchmarks. In these benchmarks, the significant differences in the executions of processes in multiple organizations are highlighted based on the relevant KPIs, which are automatically derived, predicted, and visualized by the approaches explained above. What practices yield better or worse performance in organizations can be obtained from the highlights on process benchmarks. Thus, it becomes possible to provide relevant insights for organizations into improving performance.

8.2 Implications

In this section, we list the implications of this dissertation research for practice and research, respectively.

8.2.1 Implications for Practice

The framework presented in this dissertation has several implications both for organizations that strive to improve their performance and for software vendors that focus on understanding how their client organizations use the offered products and services.

Organizations that aim to improve their performance can obtain insights and benefit in many ways from the presented cross-organizational process mining framework. For instance, organizations can find answers to the following question, "What best practices present in other organizations that perform better than our organization and have a shared interest in the same set of KPIs?" This question can also be phrased to identify certain worse practices to avoid. Moreover, the framework can encourage organizations that perform similar processes to establish process standardization. For example, municipalities can learn from each other since the processes that they perform have much the same purpose. Similarly, organizations in the healthcare domain, e.g., hospitals and laboratories, can benefit from the known best practices as they play an important role in the public good. The exchange of best practices between legal bodies, e.g., police departments and courthouses, even can save lives.

Similarly, software vendors can benefit from the framework by exploiting the differences in and commonalities between the processes of their client organizations. Such commonalities and differences may be used by a software vendor to determine the features that would increase the intended and correct use of its products and services by the client organizations. Furthermore, software vendors can use these commonalities and differences to deduce potential product improvements. In addition, the capabilities of the framework can be incorporated into software products as novel features by software vendors. Thus, software vendors gain a competitive advantage and become more resilient in their operational context.

The approach for predicting relevant KPIs for organizations has great potential for applying it in various domains. Since the features that we used for prediction model development are generic and can be observed in many organizations, the developed predictions model can be beneficial for organizations in other domains to identify relevant KPIs. Moreover, with specific features from a certain domain, prediction models specifics to that domain can be easily developed using our approach. Related to that, Netherlands Organization for Applied Scientific Research (TNO) has recently contacted us to adapt our approach in the safety domain.

Aside from that, the decision model we presented for visualizing KPIs can be integrated into BI software products to guide users in determining appropriate visualizations for indicators. Thus, the knowledge required for developing engaging dashboards using state-of-the-art dashboard design principles in each organization can become transparent to multiple organizations in a more convenient way.

8.2.2 Implications for Research

The framework presented in this dissertation has some implications for research in the following fields: Information Systems (IS), Business Process Management (BPM), and Business Intelligence (BI).

The components of the framework are extensible. Due to the nature of each component, new techniques can be added. For example, researchers develop techniques to enrich process benchmarks and include these techniques in the Metric-Based Improvement Catalog. A technique for the feasibility check of best practices can be very useful for organizations for identifying the required changes with their potential costs. The effects of information systems in best practices can be analyzed to provide insights for the organizations that use the same information systems. Developing techniques that blend dynamic process redesign and process automation to improve performance is a relevant direction for the BPM field. Similarly, the Metric Comparison Catalog can be extended in a way that process benchmarks can be tailored to particular performance improvement directions. For instance, reducing costs or improving quality can be the strategic goals of an organization. Accordingly, that organization may want to learn from other organizations that achieve better performance from one of these dimensions. Furthermore, the last component, i.e., the Cross-Organizational Process Mining Techniques, can contain new and novel process discovery techniques as state-of-the-art process discovery techniques still have some limitations in capturing process reality. Since processes in organizations are interdependent, the performance issue in a process can diffuse into other processes. In this regard, techniques that are able to discover the reality of multiple processes simultaneously can be very useful to spot the weak links in process chains in organizations. Once such techniques become readily available, organizations can be benchmarked considering multiple processes at a glance.

In Chapter 4, we presented the approach aimed at the automated prediction of relevant KPIs for organizations. Simply put, the approach employs machine learning algorithms to solve a multi-class classification problem. To this end, it uses the characteristics of organizations as features. In this way, our approach provides a basis for the development of new techniques that can be drilled down to a role-level or even a user-level relevant KPI prediction. For example, considering the characteristics and responsibilities of a role, relevant KPIs for the role can be determined to detailed process performance measurement and monitoring. Similarly, how users interact with dashboards can be analyzed to understand the interest of users in KPIs for building user specific dashboards.

Furthermore, this approach also shows the potential value of the data generated by BI applications to develop further automated techniques. For example, the relevance of KPIs may change over time. Using such input, techniques that dynamically adjust dashboards to the needs of decision-makers can be developed in the future.

We presented a decision model for visualizing KPIs in Chapter 5. This decision model is derived from the dashboard design principles available in the literature. With this decision model, we transformed the knowledge on dashboard design to a reusable and visible form. Thus, researchers in the BI field can benefit from it while developing end-to-end automated BI techniques. In addition, the automated dashboard generation approach presented in the same chapter also shows potential to further enhance it considering User Experience (UX) principles such as coloring or shapes.

8.3 Limitations and Open Issues

In this section, we discuss the limitations and open issues of the approaches presented in this dissertation.

The KPI derivation approach that is presented in Chapter 4 uses KPI derivation patterns to generate KPIs from a given OEM. The patterns employed in this approach are determined by analyzing a set of KPIs manually. However, this can be improved by leveraging natural language processing techniques.

As there is no other study that we can take as reference to compare the quality of our automated KPI prediction approach (see Chapter 5), this approach is open for further evaluation by applying it in different operation domains.

While developing the decision model for visualizing KPIs (see Chapter 5), we excluded the dashboard design principles that deal with combining multiple visual elements, e.g., small multiple, which is a set of similar graphs that use the same scale and axes to easily compare them. The reason for that is building a single visual element using tables and graphs is highly dependent on the needs of roles and even end-users rather than organizational level needs. For example, a top manager may want to have a particular KPI visualized based on a certain dimension, e.g., location. However, a team leader may have no right to see the aggregated data that is visible to the top manager. Therefore, for the sake of simplicity, we opted for excluding the principles focus on combined visualizations.

It is important to note that in the proposed framework, we mostly focused on the challenges that are related to vertical cross-organizational process mining. Although multiple organizations are involved in the execution of processes that cross the boundaries of those organizations, not every organization may be interested in obtaining insights into understanding such processes. The main reason for that is each organization only executes a fragment of the process. Therefore, it is very unlikely that such distinct organizations can determine relevant improvement actions for other organizations.

8.4 Future Work

In each of the chapters, we included ideas for future research. On top of these, in this section, we sketch some of the interesting and promising directions for future research.

One of the future research directions that is interesting is to apply predictive techniques [55, 153] to predict the future values of KPIs such that decision makers can take preventive actions rather than corrective actions.

Another future research direction is to extend the approach that is aimed at the prediction of relevant KPIs such that relevant KPIs for roles or end-users can be selected automatically. Taking this extension as a basis, dashboards for roles and end-users can also be automatically generated by further developing the approach, which we presented in Chapter 5.

Semantic Business Process Management (SBPM) aims to combine BPM with semantic technologies such that automation capabilities in processes can be increased to provide more convenient features to process workers through information systems. Ontologies and semantic web services are the two common examples of such technologies. Incorporating such technologies into process mining in cross-organizational settings may be an enabler to automatically obtain and compare the enterprise concepts of organizations in terms of their performance.

Apart from that, decision mining techniques [15, 85, 122, 131] can be used to enrich the approach that provides business process benchmarks by analyzing the differences and commonalities between process variants. In addition, dependencies between processes in organizations can be considered for in-depth process benchmarks.

Furthermore, as a promising research field, Artificial Intelligence (AI) can be used for automated identification of the visual differences between the dashboards of organizations. Such differences may then be checked with state-of-the-art dashboard design principles and recommendations can be generated for improving the visual means used for monitoring the performance in organizations. Similarly, AI technologies can be used to train and spot the similarities and differences between the process models of organizations. This can be further advanced with the employment of semantic technologies to diagnose what (activities, fragment of the process, or decisions) yields in better or worse performance in executing similar processes in various organizations.

Bibliography

[1] A. V. Abela. *The Presentation: A Story About Communicating Successfully With Very Few Slides*. CreateSpace Independent Publishing Platform, 2010. (Cited on pages 71, 72, and 91.)

[2] A. Adriansyah, J. Munoz-Gama, J. Carmona, B. F. van Dongen, and W. M. P. van der Aalst. Alignment based precision checking. In *Business Process Management Workshops - BPM International Workshops, Tallinn, Estonia. Revised Papers*, pages 137–149, 2012. (Cited on page 127.)

[3] Ü. Aksu and H. A.Reijers. How business process benchmarks enable organizations to improve performance. In *2020 IEEE 24th International Enterprise Distributed Object Computing Conference (EDOC)*, pages 197–208, 2020. (Cited on pages 9 and 118.)

[4] Ü. Aksu, D. M. M. Schunselaar, and H. A. Reijers. A cross-organizational process mining framework for obtaining insights from software products: Accurate comparison challenges. In *2016 IEEE 18th Conference on Business Informatics (CBI)*, pages 153–162, 2016. (Cited on pages 9, 14, 37, 65, and 136.)

[5] Ü. Aksu, D. M. M. Schunselaar, and H. A. Reijers. An approach for automatically deriving key performance indicators from ontological enterprise models. In *International Symposium on Data-driven Process Discovery and Analysis (SIMPDA)*, pages 38–53, 2017. (Cited on pages 9, 32, 65, 87, and 136.)

[6] Ü. Aksu, A. del-Río-Ortega, M. Resinas, and H. A. Reijers. An approach for the automated generation of engaging dashboards. In *OTM Confederated International Conferences" On the Move to Meaningful Internet Systems"*, pages 363–384, 2019. (Cited on pages 9, 68, 91, 93, 98, 104, 111, and 123.)

[7] Ü. Aksu, D. M. M. Schunselaar, and H. A. Reijers. Automated prediction of relevant key performance indicators for organizations. In *International Conference on Business Information Systems (BIS)*, pages 283–299, 2019. (Cited on pages 9, 52, and 122.)

[8] Ü. Aksu, A. del-Río-Ortega, M. Resinas, and H. A. Reijers. Interactive generation of process performance dashboards. This manuscript is submitted to a journal, 2021. (Cited on pages 9 and 90.)

[9] R. Andrews, S. Suriadi, M. T. Wynn, A. H. M. ter Hofstede, A. Pika, H. H. Nguyen, and M. La Rosa. Comparing static and dynamic aspects of patient flows via process model visualisations. *Information and Software Technology*, 2016. (Cited on page 136.)

[10] E. Arnold. Graphics overview. *Journal of Applied Communications*, 62(3):6, 1979. (Cited on pages 73, 80, and 100.)

[11] AXELOS. *ITIL® Foundation, ITIL 4 edition*. TSO (The Stationery Office), 2019. (Cited on pages 103 and 113.)

[12] J. D. Baker. Language of improvement: metrics, key performance indicators, benchmarks, analytics, scorecards, and dashboards. *AORN journal*, 102(3):223–227, 2015. (Cited on pages 94 and 95.)

[13] N. P. Ballambettu, M. A. Suresh, and R. P. J. C. Bose. Analyzing process variants to understand differences in key performance indices. In *Advanced Information Systems Engineering - 29th International Conference, CAiSE, Essen, Germany. Proceedings*, pages 298–313, 2017. (Cited on page 136.)

[14] S. Barr. *Prove it!: How to create a high-performance culture and measurable success*. Wiley, 2017. (Cited on pages 1, 69, and 91.)

[15] E. Bazhenova, F. Zerbato, B. Oliboni, and M. Weske. From BPMN process models to DMN decision models. *Information Systems*, 83:69–88, 2019. (Cited on page 145.)

[16] A. Bhattacherjee. Understanding information systems continuance: An expectation-confirmation model. *MIS quarterly*, pages 351–370, 2001. (Cited on pages 105 and 113.)

[17] A. Bolt, M. de Leoni, and W. M. P. van der Aalst. Process variant comparison: Using event logs to detect differences in behavior and business rules. *Information Systems*, 74:53–66, 2018. (Cited on page 136.)

[18] E. S. Borges, M. Fantinato, Ü. Aksu, H. A. Reijers, and L. H. Thom. Monitoring of non-functional requirements of business processes based on quality of service attributes of web services. In *International Conference on Enterprise Information Systems (ICEIS), Proceedings*, pages 588–595, 2019. (Cited on page 10.)

[19] R. P. J. C. Bose and W. M. P. van der Aalst. Context aware trace clustering: Towards improving process mining results. In *Proceedings of the SIAM International Conference on Data Mining, SDM, Sparks, Nevada, USA*, pages 401–412, 2009. (Cited on page 119.)

[20] N. Brand and H. van der Kolk. Workflow analysis and design. *Deventer: Kluwer Bedrijfswetenschappen*, 1995. (Cited on page 47.)

[21] J. C. A. M. Buijs and H. A. Reijers. Comparing business process variants using models and event logs. In *Enterprise, Business-Process and Information Systems Modeling - 15th International Conference, BPMDS 2014, 19th International Conference, EMMSAD 2014, Held at CAiSE 2014, 2014. Proceedings*, pages 154–168, 2014. (Cited on page 29.)

[22] J. C. A. M. Buijs and H. A. Reijers. Comparing business process variants using models and event logs. In *Enterprise, Business-Process and Information Systems Modeling - 15th International Conference, BPMDS 2014, 19th International Conference, EMMSAD 2014, Held at CAiSE 2014, Thessaloniki, Greece, June 16-17, 2014. Proceedings*, pages 154–168, 2014. (Cited on pages 3, 15, 28, and 136.)

[23] J. C. A. M. Buijs, B. F. van Dongen, and W. M. P. van der Aalst. Towards cross-organizational process mining in collections of process models and their executions. In *Business Process Management Workshops - BPM 2011 International Workshops, Clermont-Ferrand, France, August 29, 2011, Revised Selected Papers, Part II*, pages 2–13, 2011. (Cited on pages 3, 15, 28, 33, and 48.)

[24] J. C. A. M. Buijs, B. F. van Dongen, and W. M. P. van der Aalst. Mining configurable process models from collections of event logs. In *Business Process Management - 11th International Conference, BPM 2013, Beijing, China, August 26-30, 2013. Proceedings*, pages 33–48, 2013. (Cited on page 29.)

[25] A. Burattin, M. Conti, and D. Turato. Toward an anonymous process mining. In *3rd International Conference on Future Internet of Things and Cloud, FiCloud 2015, Rome, Italy, August 24-26, 2015*, pages 58–63, 2015. (Cited on page 15.)

[26] C. Burnay, S. Bouraga, S. Faulkner, and I. Jureta. User-experience in business intelligence-a quality construct and model to design supportive bi dashboards. In *International Conference on Research Challenges in Information Science*, pages 174–190, 2020. (Cited on page 116.)

[27] D. Calvanese, M. Montali, A. Syamsiyah, and W. M. P. van der Aalst. Ontology-driven extraction of event logs from relational databases. In *Business Process Management Workshops - BPM 2015, 13th International Workshops, Innsbruck, Austria, 2015, Revised Papers*, pages 140–153, 2015. (Cited on page 28.)

[28] D. Carlucci. Evaluating and selecting key performance indicators: an anp-based model. *Measuring Business Excellence*, 14(2):66–76, 2010. (Cited on pages 53 and 64.)

[29] C. F. Castro, M. Fantinato, Ü. Aksu, H. A. Reijers, and L. H. Thom. Towards a conceptual framework for decomposing non-functional requirements of business process into quality of service attributes. In *International Conference on Enterprise Information Systems (ICEIS(2)), Proceedings*, pages 481–492, 2019. (Cited on page 10.)

[30] C. F. Castro, M. Fantinato, Ü. Aksu, H. A. Reijers, and L. H. Thom. Systematizing the relationship between business processes' and web services' non-functional requirements. In *International Conference on Enterprise Information Systems (ICEIS), Proceedings*, pages 473–497, 2019. (Cited on page 10.)

[31] P. Chowdhary, T. Palpanas, F. Pinel, S.-K. Chen, and F. Y. Wu. Model-driven dashboards for business performance reporting. In *Proceedings of the 10th International Enterprise Distributed Object Computing Conference IEEE EDOC*, pages 374–386, 2006. (Cited on pages 69, 70, 86, 92, and 114.)

[32] H. M. Cooper. Scientific guidelines for conducting integrative research reviews. *Review of educational research*, 52(2):291–302, 1982. (Cited on page 6.)

[33] C. Cordes, T. Vogelgesang, and H. Appelrath. A generic approach for calculating and visualizing differences between process models in multidimensional process mining. In *Business Process Management Workshops - BPM International Workshops, Eindhoven, The Netherlands. Revised Papers*, pages 383–394, 2014. (Cited on page 136.)

[34] K. Daniel. *Thinking, fast and slow*. Farrar, Straus and Giroux (FSG), 2011. (Cited on page 4.)

[35] P. Darke, G. Shanks, and M. Broadbent. Successfully completing case study research: combining rigour, relevance and pragmatism. *Information systems journal*, 8(4):273–289, 1998. (Cited on page 5.)

[36] A. K. A. de Medeiros, A. Guzzo, G. Greco, W. M. P. van der Aalst, A. J. M. M. Weijters, B. F. van Dongen, and D. Saccà. Process mining based on clustering: A quest for precision. In *Business Process Management Workshops, BPM International Workshops, BPI, BPD, CBP, ProHealth, RefMod, semantics4ws, Brisbane, Australia. Revised Selected Papers*, pages 17–29, 2007. (Cited on pages 119 and 128.)

[37] A. del-Río-Ortega, M. Resinas, C. Cabanillas, and A. R. Cortés. On the definition and design-time analysis of process performance indicators. *Information Systems*, 38(4), 2013. (Cited on pages 29, 33, 48, 53, 64, 74, 78, 85, 91, 93, 94, and 95.)

[38] A. del Río-Ortega, M. Resinas, A. Durán, B. Bernárdez, A. Ruiz-Cortés, and M. Toro. Visual ppinot: A graphical notation for process performance indicators. *Business & Information Systems Engineering: Vol. 61, No. 2*, 2019. (Cited on pages 95 and 112.)

[39] P. F. Drucker. *People and performance: The best of Peter Drucker on management*. Routledge, 1995. (Cited on pages 1 and 91.)

[40] M. Dumas, M. L. Rosa, J. Mendling, and H. A. Reijers. *Fundamentals of Business Process Management*. Springer, second edition, 2018. (Cited on pages 1, 23, 28, 91, and 94.)

[41] C. G. Durbin. How to read a scientific research paper. *Respiratory care*, 54(10):1366–1371, 2009. (Cited on page 6.)

[42] W. W. Eckerson. *Performance dashboards: measuring, monitoring, and managing your business*. Wiley, second edition, 2010. (Cited on pages 1, 69, 70, 71, 72, 73, 74, 75, 76, 77, 91, 92, 94, and 95.)

[43] M. Eisenberg and X. Hu. Dichotomous relevance judgments and the evaluation of information systems. *Proceedings of the American Society for Information Science*, 24:66–69, 1987. (Cited on page 55.)

[44] C. Elliot, C. Mcullagh, M. Brydon, and K. Zwi. Developing key performance indicators for a tertiary children's hospital network. *Australian Health Review*, 42(5):491–500, 2018. (Cited on page 64.)

[45] R. Engel, W. Krathu, M. Zapletal, C. Pichler, R. J. C. Bose, W. M. P. van der Aalst, H. Werthner, and C. Huemer. Analyzing inter-organizational business processes. *Information Systems and e-Business Management*, 14(3):577–612, 2016. (Cited on page 3.)

[46] M. Eshtaiwi, I. Badi, A. Abdulshahed, and T. E. Erkan. Determination of key performance indicators for measuring airport success: A case study in libya. *Journal of Air Transport Management*, 68:28–34, 2018. (Cited on page 64.)

[47] J. Evermann, T. Thaler, and P. Fettke. Clustering traces using sequence alignment. In *Business Process Management Workshops - BPM, 13th International Workshops, Innsbruck, Austria. Revised Papers*, pages 179–190, 2015. (Cited on pages 119, 127, and 128.)

[48] D. R. Ferreira, M. Zacarias, M. Malheiros, and P. Ferreira. Approaching process mining with sequence clustering: Experiments and findings. In *Business Process Management, 5th International Conference, BPM, Brisbane, Australia. Proceedings*, pages 360–374, 2007. (Cited on page 119.)

[49] P. Fettke. How conceptual modeling is used. *Communications of the Association for Information Systems*, 25(1):43, 2009. (Cited on page 6.)

[50] S. Few. Why most dashboards fail. Technical report, Perceptual Edge, 2007. Last accessed May 2021. (Cited on pages 69, 91, and 92.)

[51] S. Few. *Now you see it: simple visualization techniques for quantitative analysis*. Analytics Press, 2009. (Cited on pages 71, 72, 74, 75, 76, 77, and 91.)

[52] S. Few. *Show me the numbers*. Analytics Pres, second edition, 2012. (Cited on pages 69, 70, 71, 72, 73, 74, 75, 76, 77, 86, 91, 92, and 114.)

[53] S. Few. Bullet graph design specification. Technical report, Perceptual Edge, 2013. Last accessed May 2021. (Cited on pages 74, 76, and 112.)

[54] S. Few. *Information dashboard design*. Analytics Pres, 2013. (Cited on pages 1, 69, 70, 71, 72, 73, 74, 75, 76, 77, 91, 92, and 95.)

[55] C. D. Francescomarino, C. Ghidini, F. M. Maggi, W. Rizzi, and C. D. Persia. Incremental predictive process monitoring: How to deal with the variability of real environments. *CoRR*, abs/1804.03967, 2018. (Cited on page 145.)

[56] J. Frost. Making predictions with regression analysis. `https://statisticsbyjim.com/regression/predictions-regression`, May 2017. Last accessed May 2021. (Cited on page 62.)

[57] V. García, R. A. Mollineda, and J. S. Sánchez. Index of balanced accuracy: A performance measure for skewed class distributions. In *Pattern Recognition and Image Analysis, 4th Iberian Conference*, pages 441–448, 2009. (Cited on page 57.)

[58] M. Gluck. Exploring the relationship between user satisfaction and relevance in information systems. *Information Processing & Management*, 32(1), 1996. (Cited on page 55.)

[59] S. Goel, J. M. Bhat, and B. Weber. End-to-end process extraction in process unaware systems. In *Business Process Management Workshops - BPM 2012 International Workshops, Tallinn, Estonia, September 3, 2012. Revised Papers*, pages 162–173, 2012. (Cited on page 28.)

[60] M. Golfarelli and S. Rizzi. A model-driven approach to automate data visualization in big data analytics. *Information Visualization*, 19(1):24–47, 2020. (Cited on page 115.)

[61] E. González López de Murillas, W. M. P. van der Aalst, and H. A. Reijers. Process mining on databases: Unearthing historical data from redo logs. In *Business Process Management: 13th International Conference, BPM 2015, Innsbruck, Austria, 2015, Proceedings*, pages 367–385, 2015. (Cited on pages 25 and 28.)

[62] T. A. Granberg and A. O. Munoz. Developing key performance indicators for airports. In *3rd ENRI International Workshop on ATM/CNS*, 2013. (Cited on page 64.)

[63] F. J. Gravetter and L. B. Wallnau. *Essentials of statistics for the behavioral sciences*. Cengage Learning, 8 edition, 2013. (Cited on pages 58 and 59.)

[64] G. Greco, A. Guzzo, L. Pontieri, and D. Saccà. Discovering expressive process models by clustering log traces. *IEEE Transactions on Knowledge and Data Engineering*, 18(8):1010–1027, 2006. (Cited on page 119.)

[65] J. Gulden. Visually comparing process dynamics with rhythm-eye views. In *Business Process Management Workshops - BPM International Workshops, Rio de Janeiro, Brazil. Revised Papers*, pages 474–485, 2016. (Cited on page 136.)

[66] Y. K. Heidema, I. T. P. Vanderfeesten, J. Erasmus, R. Keulen, and K. Dizy. Visualizing performance indicators for production planning in bpmn 2.0 in the manufacturing domain. Master's thesis, Eindhoven University of Technology, 2018. (Cited on pages 70, 86, 92, and 114.)

[67] B. F. A. Hompes, J. C. A. M. Buijs, W. M. P. van der Aalst, P. M. Dixit, and H. J. Buurman. Discovering deviating cases and process variants using trace clustering. In *27th Benelux Conference on Artificial Intelligence, Hasselt, Belgium. Proceedings*, pages 5–6, 2015. (Cited on page 119.)

[68] B. Ioan, A. S. Nestian, and S.-M. Tita. Relevance of key performance indicators (kpis) in a hospital performance management model. *Journal of Eastern Europe Research in Business & Economics*, 2012. (Cited on pages 60 and 64.)

[69] S. Jablonski, M. Röglinger, S. Schönig, and K. M. Wyrtki. Multi-perspective clustering of process execution traces. *Enterprise Modelling and Information Systems Architectures (EMISAJ) International Journal of Conceptual Modeling*, 14:2:1–2:22, 2018. (Cited on page 119.)

[70] A. Janes, A. Sillitti, and G. Succi. Effective dashboard design. *Cutter IT Journal*, 26(1):17–24, 2013. (Cited on page 91.)

[71] M. H. Jansen-Vullers, M. W. N. C. Loosschilder, A. P. A. M. Kleingeld, and H. A. Reijers. Performance measures to evaluate the impact of best practices. In *Proceedings of Workshops and Doctoral Consortium of the 19th International Conference on Advanced Information Systems Engineering (BPMDS workshop)*, pages 359–368, 2007. (Cited on pages 33 and 47.)

[72] V. Kachitvichyanukul, H. T. Luong, and R. Pitakaso. A hybrid mcdm approach to kpi selection of the coordination problems of production and sales departments-an empirical study of iron and steel industry of china and taiwan. In *13th Asia Pacific Industrial Engineering and Management Systems Conference*, 2012. (Cited on pages 53 and 64.)

[73] S. Kaganski, A. Snatkin, M. Paavel, and K. Karjust. Selecting the right kpis for smes production with the support of pms and plm. *International Journal of Research in Social Sciences*, 1(3): 69–76, 2013. (Cited on pages 60 and 64.)

[74] S. Kaganski, M. Paavel, and J. Lavin. Selecting key performance indicators with support of enterprise analyse model. In *9th International DAAAM Baltic Conference "Industrial Engineering*, pages 97–102, 2014. (Cited on pages 53 and 64.)

[75] M. Kintz. A semantic dashboard description language for a process-oriented dashboard design methodology. In *Proceedings of the 2nd International Workshop on Model-based Interactive Ubiquitous Systems, MODIQUITOUS*, pages 31–36, 2012. (Cited on pages 70, 86, 92, and 114.)

[76] M. Kintz, M. Kochanowski, and F. Koetter. Creating user-specific business process monitoring dashboards with a model-driven approach. In *Proceedings of the 5th International Conference on Model-Driven Engineering and Software Development, MODELSWARD*, pages 353–361, 2017. (Cited on pages 70, 86, 92, and 114.)

[77] B. Kitchenham. Procedures for performing systematic reviews. Technical Report TR/SE-0401, Keele, UK, Keele University, 2004. (Cited on page 6.)

[78] F. Koetter and M. Kochanowski. Goal-oriented model-driven business process monitoring using progoalml. In *Proceedings of the 15th International Conference on Business Information Systems, BIS*, pages 72–83, 2012. (Cited on pages 86, 92, and 114.)

[79] A. Kriebel. Visualvocabulary. Technical report, Tableau, 2018. Last accessed May 2021. (Cited on pages 69, 70, 71, 72, 73, 74, 75, 77, and 91.)

[80] R. A. Krueger. *Focus groups: A practical guide for applied research*. Sage publications, fifth edition, 2014. (Cited on page 6.)

[81] B. Kucukaltan, Z. Irani, and E. Aktas. A decision support model for identification and prioritization of key performance indicators in the logistics industry. *Computers in Human Behavior*, 65:346–358, 2016. (Cited on page 64.)

[82] A. Lavalle, A. Maté, and J. Trujillo. Requirements-driven visualizations for big data analytics: A model-driven approach. In *Conceptual Modeling - 38th International Conference, ER 2019, Salvador, Brazil, November 4-7, 2019, Proceedings*, pages 78–92, 2019. (Cited on pages 92 and 115.)

[83] A. Lavalle, A. Maté, J. Trujillo, and S. Rizzi. Visualization requirements for business intelligence analytics: A goal-based, iterative framework. In *27th IEEE International Requirements Engineering Conference, RE 2019, Jeju Island, Korea (South), September 23-27, 2019*, pages 109–119, 2019. (Cited on page 115.)

[84] A. Lavalle, A. Maté, J. Trujillo, M. A. Teruel, and S. Rizzi. A methodology to automatically translate user requirements into visualizations: Experimental validation. *Information and Software Technology*, 136:106592, 2021. (Cited on pages 92 and 115.)

[85] S. Leewis, K. Smit, and M. Zoet. Putting decision mining into context: A literature study. In *Digital Business Transformation*, pages 31–46. Springer, 2020. (Cited on page 145.)

[86] C. Lennerholt, J. van Laere, and E. Söderström. User-related challenges of self-service business intelligence. *Information Systems Management*, pages 1–15, 2020. (Cited on page 116.)

[87] J. Li, H. J. Wang, and X. Bai. An intelligent approach to data extraction and task identification for process mining. *Information Systems Frontiers*, 17(6):1195–1208, 2015. (Cited on page 28.)

[88] L. Liu and M. T. Özsu, editors. *Encyclopedia of Database Systems*. Springer, 2009. (Cited on page 36.)

[89] A. V. Looy and A. Shafagatova. Business process performance measurement: a structured literature review of indicators, measures and metrics. *SpringerPlus*, 5(1):1–24, 2016. (Cited on page 95.)

[90] X. Lu, S. A. Tabatabaei, M. Hoogendoorn, and H. A. Reijers. Trace clustering on very large event data in healthcare using frequent sequence patterns. In *Business Process Management - 17th International Conference, BPM, Vienna, Austria. Proceedings*, pages 198–215, 2019. (Cited on page 119.)

[91] J. Lyons and S. Evergreen. Qualitative chart chooser 3.0. `https://stephanieevergreen.com/qualitative-chart-chooser-3/`, 2017. Last accessed May 2021. (Cited on pages 71, 72, and 91.)

[92] E. McDaniel and S. McDaniel. *The Accidental Analyst: Show Your Data Who's Boss*. Freakalytics, 2012. (Cited on pages 69, 70, 72, 73, 74, 75, 76, and 77.)

[93] E. McDaniel and S. McDaniel. *Rapid Graphs with Tableau Software 8: The Original Guide for the Accidental Analyst*. CreateSpace Independent Publishing Platform, 2013. (Cited on pages 69, 70, 71, 72, 73, 74, 75, 76, 77, and 91.)

[94] M. Meuser and U. Nagel. The expert interview and changes in knowledge production. In *Interviewing experts*, pages 17–42. Springer, 2009. (Cited on page 6.)

[95] D. L. Morgan. *Focus groups as qualitative research*, volume 16. Sage publications, 1996. (Cited on page 6.)

[96] H. Nguyen, M. Dumas, M. L. Rosa, and A. H. M. ter Hofstede. Multi-perspective comparison of business process variants based on event logs. In *Conceptual Modeling - 37th International Conference, ER, Xi'an, China. Proceedings*, pages 449–459, 2018. (Cited on page 136.)

[97] C. K. Nussbaumer. *Storytelling with Data: A Data Visualization Guide for Business Professionals*. Wiley, 2015. (Cited on pages 69, 70, 71, 72, 73, 74, 75, 76, 77, 91, and 95.)

[98] A. Olivé. *Conceptual modeling of information systems*. Springer Science & Business Media, 2007. (Cited on page 6.)

[99] D. L. Padmaja and V. B. Vardhan. Comparative study of feature subset selection methods for dimensionality reduction on scientific data. In *IEEE 6th International Conference on Advanced Computing*, pages 31–34, 2016. (Cited on page 58.)

[100] T. Palpanas, P. Chowdhary, G. A. Mihaila, and F. Pinel. Integrated model-driven dashboard development. *Information Systems Frontiers*, 9(2-3):195–208, 2007. (Cited on pages 70, 86, 92, and 114.)

[101] D. Parmenter. *Key performance indicators: developing, implementing, and using winning KPIs*. Wiley, 2015. (Cited on pages 33, 53, and 94.)

[102] V. L. Parsons. *Stratified Sampling*, pages 1–11. Wiley, 2017. (Cited on page 58.)

[103] A. Partington, M. T. Wynn, S. Suriadi, C. Ouyang, and J. Karnon. Process mining for clinical processes: A comparative analysis of four australian hospitals. *ACM Transactions on Management Information Systems*, 5(4):19:1–19:18, 2015. (Cited on pages 3 and 136.)

[104] C. Pedrinaci and J. Domingue. Ontology-based metrics computation for business process analysis. In *Proceedings of the 4th International Workshop on Semantic Business Process Management*, pages 43–50, 2009. (Cited on pages 33 and 47.)

[105] C. Pedrinaci, D. Lambert, B. Wetzstein, T. van Lessen, L. Cekov, and M. Dimitrov. SENTINEL: a semantic business process monitoring tool. In *Proceedings of the First International Workshop on Ontology-supported Business Intelligence, OBI 2008*, pages 1–12, 2008. (Cited on pages 33 and 48.)

[106] M. Pelletier and P. Boily. Dashboards and data visualization, with examples. Technical report, Data Action Lab, 2019. Last accessed May 2021. (Cited on pages 72, 91, 92, and 95.)

[107] J. Peral, A. Maté, and M. Marco. Application of data mining techniques to identify relevant key performance indicators. *Computer Standards & Interfaces*, 54:76–85, 2017. (Cited on pages 60 and 64.)

[108] C. Pinna, M. Demartini, F. Tonelli, and S. Terzi. How soft drink supply chains drive sustainability: Key performance indicators (kpis) identification. In *51st CIRP Conference on Manufacturing Systems*, pages 862–867, 2018. (Cited on pages 60 and 64.)

[109] V. Popova and A. Sharpanskykh. Modeling organizational performance indicators. *Information Systems*, 35(4), 2010. (Cited on pages 33 and 47.)

[110] S. Pospiech, R. Mertens, S. Mielke, M. Städler, and P. Söhlke. Creating event logs from heterogeneous, unstructured business data. In *Multidimensional Views on Enterprise Information Systems, Proceedings of ERP Future 2014, Dornbirn, Austria, November 17-18, 2014*, pages 85–93, 2014. (Cited on page 28.)

[111] C. Ranjan, S. Ebrahimi, and K. Paynabar. Sequence graph transform (SGT): A feature extraction function for sequence data mining. *CoRR*, abs/1608.03533, 2016. (Cited on pages 119, 121, and 124.)

[112] C. Ranjan, S. Ebrahimi, and K. Paynabar. Sequence graph transform (sgt): A feature embedding function for sequence data mining (extended version). Technical report, Georgia Institute of Technology, 2016. Last accessed May 2021. (Cited on pages 119, 121, and 122.)

[113] J. Recker. *Scientific Research in Information Systems - A Beginner's Guide*. Springer, 2013. (Cited on pages 5 and 6.)

[114] A. Revina and Ü. Aksu. Towards a business process complexity analysis framework based on textual data and event logs. In *17th International Conference on Wirtschaftsinformatik (WI 2022), February 21-23, 2022, Nürnberg, Germany, Proceedings*, 2022 in press. (Cited on page 10.)

[115] A. Revina, Ü. Aksu, and V. G. Meister. Method to address complexity in organizations based on a comprehensive overview. *Information*, 12(10):423, 2021. (Cited on page 10.)

[116] A. Rozinat and W. M. P. van der Aalst. Conformance checking of processes based on monitoring real behavior. *Information Systems*, 33(1):64–95, 2008. (Cited on page 127.)

[117] S. Saitta. What is a good classification accuracy in data mining? `http://www.dataminingblog.com/what-is-a-good-classification-accuracy-in-data-mining/`, Apr. 2010. Last accessed May 2021. (Cited on page 62.)

[118] A. Sarikaya, M. Correll, L. Bartram, M. Tory, and D. Fisher. What do we talk about when we talk about dashboards? *IEEE transactions on visualization and computer graphics*, 25(1): 682–692, 2018. (Cited on page 91.)

[119] M. Schmidt, J. Schwöbel, M. Lienkamp, et al. Developing key performance indicators for variant management of complex product families. *DS 91: Proceedings of NordDesign 2018, Linköping, Sweden, 14th-17th August 2018*, 2018. (Cited on page 64.)

[120] D. M. M. Schunselaar, H. M. W. Verbeek, W. M. P. van der Aalst, and H. A. Reijers. Petra: A tool for analysing a process family. In *Proceedings of the International Workshop on Petri Nets and Software Engineering, co-located with 35th International Conference on Application and Theory of Petri Nets and Concurrency (PetriNets 2014) and 14th International Conference on Application of Concurrency to System Design (ACSD 2014), Tunis, Tunisia, June 23-24, 2014.*, pages 269–288, 2014. (Cited on page 29.)

[121] D. M. M. Schunselaar, J. Gulden, H. van der Schuur, and H. A. Reijers. A systematic evaluation of enterprise modelling approaches on their applicability to automatically generate ERP software. In *18th IEEE Conference on Business Informatics, CBI 2016*, 2016. (Cited on pages 15, 16, and 23.)

[122] J. D. Smedt, F. Hasic, S. K. L. M. vanden Broucke, and J. Vanthienen. Towards a holistic discovery of decisions in process-aware information systems. In *Business Process Management - 15th International Conference, BPM 2017, Barcelona, Spain, September 10-15, 2017, Proceedings*, pages 183–199, 2017. (Cited on page 145.)

[123] H. Snyder. Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104:333–339, 2019. (Cited on page 6.)

[124] M. Sokolova and G. Lapalme. A systematic analysis of performance measures for classification tasks. *Information processing & management*, 45(4):427–437, 2009. (Cited on page 57.)

[125] M. Song, C. W. Günther, and W. M. P. van der Aalst. Trace clustering in process mining. In *Business Process Management Workshops, BPM International Workshops, Milano, Italy. Revised Papers*, pages 109–120, 2008. (Cited on pages 119 and 128.)

[126] R. E. Stake. *The art of case study research*. Sage publications, 1995. (Cited on page 5.)

[127] D. W. Stewart and P. N. Shamdasani. *Focus groups: Theory and practice*. Sage publications, third edition, 2014. (Cited on page 6.)

[128] E. W. Steyerberg, A. J. Vickers, N. R. Cook, T. Gerds, M. Gonen, N. Obuchowski, M. J. Pencina, and M. W. Kattan. Assessing the performance of prediction models: a framework for some traditional and novel measures. *International Society for Environmental Epidemiology*, 21(1): 128, 2010. (Cited on page 62.)

[129] R. Tang, W. M. Shaw Jr., and J. L. Vevea. Towards the identification of the optimal number of relevance categories. *Journal of the American Society for Information Science*, 50(3):254–264, 1999. (Cited on page 55.)

[130] F. Taymouri, M. L. Rosa, M. Dumas, and F. M. Maggi. Business process variant analysis: Survey and classification. *CoRR*, abs/1911.07582, 2019. (Cited on pages 119, 125, and 136.)

[131] I. Teinemaa, M. Dumas, F. M. Maggi, and C. D. Francescomarino. Predictive business process monitoring with structured and unstructured data. In *Business Process Management - 14th International Conference, BPM 2016, Rio de Janeiro, Brazil, September 18-22, 2016. Proceedings*, pages 401–417, 2016. (Cited on page 145.)

[132] A. C. Telea. *Data visualization: principles and practice*. CRC Press, second edition, 2014. (Cited on page 91.)

[133] R. Tregear. *Reimagining Management*. Blurb, San Francisco, 2017. (Cited on pages 1, 91, and 94.)

[134] E. R. Tufte. *The visual display of quantitative information*. Graphics Press, 1992. (Cited on pages 69, 70, 71, 72, 73, 74, 75, 76, 77, 91, and 95.)

[135] J. van de Wiel, S. Jansen, M. Spruit, R. Mijwaart, and J. Gruppen. The automated generation of user specific dashboards in a multi-tenant erp product. Master's thesis, Utrecht University, 2017. (Cited on pages 92 and 115.)

[136] W. M. P. van der Aalst. Configurable services in the cloud: Supporting variability while enabling cross-organizational process mining. In *On the Move to Meaningful Internet Systems: OTM 2010 - Confederated International Conferences: CoopIS, IS, DOA and ODBASE, Hersonissos, Crete, Greece, 2010, Proceedings, Part I*, pages 8–25, 2010. (Cited on pages 3, 15, and 28.)

[137] W. M. P. van der Aalst. *Process Mining - Discovery, Conformance and Enhancement of Business Processes*. Springer, 2011. (Cited on page 29.)

[138] W. M. P. van der Aalst. Intra-and inter-organizational process mining: Discovering processes within and between organizations. In *IFIP Working Conference on The Practice of Enterprise Modeling*, pages 1–11, 2011. (Cited on page 3.)

[139] W. M. P. van der Aalst. *Process Mining - Data Science in Action, Second Edition*. Springer, 2016. (Cited on pages 2 and 124.)

[140] W. M. P. van der Aalst, A. K. A. de Medeiros, and T. A. J. M. M. Weijters. Process equivalence: Comparing two process models based on observed behavior. In *Business Process Management, 4th International Conference, BPM 2006, Vienna, Austria, September 5-7, 2006, Proceedings*, pages 129–144, 2006. (Cited on page 29.)

[141] W. M. P. van der Aalst, B. F. van Dongen, C. W. Günther, A. Rozinat, H. M. W. Verbeek, and T. Weijters. ProM: The process mining toolkit. In *Proceedings of the Business Process Management Demonstration Track (BPMDemos 2009)*, 2009. (Cited on page 37.)

[142] W. M. P. van der Aalst, A. Adriansyah, A. K. A. De Medeiros, F. Arcieri, T. Baier, T. Blickle, J. C. Bose, P. van Den Brand, R. Brandtjen, J. Buijs, et al. Process mining manifesto. In *Business Process Management Workshops - BPM 2011 International Workshops, Clermont-Ferrand, France, August 29, 2011, Revised Selected Papers, Part I*, pages 169–194, 2011. (Cited on page 3.)

[143] W. M. P. van der Aalst, A. Adriansyah, and B. F. van Dongen. Replaying history on process models for conformance checking and performance analysis. *WIREs Data Mining and Knowledge Discovery*, 2(2):182–192, 2012. (Cited on page 127.)

[144] H. van der Schuur, E. van de Ven, R. de Jong, D. M. M. Schunselaar, H. A. Reijers, M. Overeem, M. de Graaf, S. Jansen, and S. Brinkkemper. Next: Generating tailored erp applications from ontological enterprise models. In *The Practice of Enterprise Modeling Conference, PoEM 2017*, pages 283–298, 2017. (Cited on pages 35 and 37.)

[145] B. F. van Dongen, A. K. A. de Medeiros, H. M. W. Verbeek, A. J. M. M. Weijters, and W. M. P. van der Aalst. The prom framework: A new era in process mining tool support. In *Applications and Theory of Petri Nets, 26th International Conference, ICATPN, Miami, USA. Proceedings*, pages 444–454, 2005. (Cited on pages 125 and 126.)

[146] A. van Horenbeek and L. Pintelon. Development of a maintenance performance measurement framework—using the analytic network process (anp) for maintenance performance indicator selection. *Omega*, 42(1):33–46, 2014. (Cited on pages 53 and 64.)

[147] V. Vashisht and P. Dharia. Integrating chatbot application with qlik sense business intelligence (bi) tool using natural language processing (nlp). In *Micro-Electronics and Telecommunication Engineering*, pages 683–692. Springer, 2020. (Cited on pages 92 and 114.)

[148] A. Vázquez-Ingelmo, F. J. García-Peñalvo, and R. Therón. Application of domain engineering to generate customized information dashboards. In *Proceedings of the 5th International Conference on Learning and Collaboration Technologies, LCT, Held as Part of the 20th International Conference on Human Computer Interaction, HCI*, pages 518–529, 2018. (Cited on pages 70, 86, and 114.)

[149] A. Vázquez-Ingelmo, F. J. García-Peñalvo, R. Therón, and M. Á. Conde. Representing data visualization goals and tasks through meta-modeling to tailor information dashboards. *Applied Sciences*, 10(7):2306, 2020. (Cited on page 115.)

[150] A. Vázquez-Ingelmo, F. J. García-Peñalvo, R. Theron, D. A. Filvà, and D. F. Escudero. Connecting domain-specific features to source code: Towards the automatization of dashboard generation. *Cluster Computing*, 23(3):1803–1816, 2020. (Cited on pages 92 and 114.)

[151] G. M. Veiga and D. R. Ferreira. Understanding spaghetti models with sequence clustering for prom. In *Business Process Management Workshops, BPM International Workshops, Ulm, Germany. Revised Papers*, pages 92–103, 2009. (Cited on pages 119 and 128.)

[152] H. M. W. Verbeek, J. C. A. M. Buijs, B. F. van Dongen, and W. M. P. van der Aalst. Xes, xesame, and prom 6. In *Information Systems Evolution - CAiSE Forum 2010, Hammamet, Tunisia, 2010, Selected Extended Papers*, pages 60–75, 2010. (Cited on page 28.)

[153] I. Verenich, M. Dumas, M. L. Rosa, and H. Nguyen. Predicting process performance: A white-box approach based on process models. *Journal of Software: Evolution and Process*, 31 (6), 2019. (Cited on page 145.)

[154] Y. Wand and R. Weber. Research commentary: information systems and conceptual modeling—a research agenda. *Information systems research*, 13(4):363–376, 2002. (Cited on page 6.)

[155] J. D. Weerdt, S. K. L. M. vanden Broucke, J. Vanthienen, and B. Baesens. Leveraging process discovery with trace clustering and text mining for intelligent analysis of incident management processes. In *Proceedings of the IEEE Congress on Evolutionary Computation, CEC, Brisbane, Australia*, pages 1–8, 2012. (Cited on page 136.)

[156] J. D. Weerdt, S. K. L. M. vanden Broucke, J. Vanthienen, and B. Baesens. Active trace clustering for improved process discovery. *IEEE Transactions on Knowledge and Data Engineering*, 25 (12):2708–2720, 2013. (Cited on pages 119, 124, 127, and 128.)

[157] T. A. J. M. M. Weijters and J. T. S. Ribeiro. Flexible heuristics miner (FHM). In *Proceedings of the IEEE Symposium on Computational Intelligence and Data Mining, CIDM, Paris, France*, pages 310–317, 2011. (Cited on page 127.)

[158] S. Wexler, J. Shaffer, and A. Cotgreave. *The big book of dashboards: visualizing your data using real-world business scenarios*. Wiley, 2017. (Cited on pages 1, 69, 70, 71, 72, 73, 74, 75, 76, 77, 91, and 95.)

[159] M. T. Wynn, E. Poppe, J. Xu, A. H. M. ter Hofstede, R. Brown, A. Pini, and W. M. P. van der Aalst. Processprofiler3d: A visualisation framework for log-based process performance comparison. *Decision Support Systems*, 100:93–108, 2017. (Cited on page 136.)

[160] O. M. Yigitbasioglu and O. Velcu. A review of dashboards in performance management: Implications for design and research. *International Journal of Accounting Information Systems*, 13(1):41–59, 2012. (Cited on page 91.)

[161] R. K. Yin. *Case study research: Design and methods*. Sage publications, fourth edition, 2013. (Cited on pages 6 and 35.)

[162] G. Zelazny. *Say It with Charts Workbook*. McGraw-Hill, 2004. (Cited on pages 71, 72, and 91.)

[163] Z.-H. Zhou. *Ensemble methods: foundations and algorithms*. Chapman and Hall/CRC, 2012. (Cited on page 58.)

Curriculum Vitae

Ünal Aksu (1985) was born in Artvin, Turkey. From 2004 to 2008, he studied Computer Engineering at Ege University. In the years that followed, he worked as a software developer and subsequently as a software quality assurance consultant. In parallel to his work, he continued with the master of Software Management at Middle East Technical University (METU), which resulted in a Master of Science degree in 2013. He started his PhD research in the AMUSE project, which was a collaboration between Vrije Universiteit Amsterdam, Afas Software, and Utrecht University. On February 2, 2022, he defended his PhD thesis entitled *A Cross-Organizational Process Mining Framework* at Utrecht University.

The research and educational activities of Ünal Aksu focus on the areas of process mining, business intelligence, business process management, and data science. In addition, he provides lectures and workshops in these areas.

SIKS Dissertations

2011 01 Botond Cseke (RUN), Variational Algorithms for Bayesian Inference in Latent Gaussian Models
02 Nick Tinnemeier (UU), Organizing Agent Organizations. Syntax and Operational Semantics of an Organization-Oriented Programming Language
03 Jan Martijn van der Werf (TUE), Compositional Design and Verification of Component-Based Information Systems
04 Hado van Hasselt (UU), Insights in Reinforcement Learning; Formal analysis and empirical evaluation of temporal-difference
05 Bas van der Raadt (VU), Enterprise Architecture Coming of Age - Increasing the Performance of an Emerging Discipline.
06 Yiwen Wang (TUE), Semantically-Enhanced Recommendations in Cultural Heritage
07 Yujia Cao (UT), Multimodal Information Presentation for High Load Human Computer Interaction
08 Nieske Vergunst (UU), BDI-based Generation of Robust Task-Oriented Dialogues
09 Tim de Jong (OU), Contextualised Mobile Media for Learning
10 Bart Bogaert (UvT), Cloud Content Contention
11 Dhaval Vyas (UT), Designing for Awareness: An Experience-focused HCI Perspective
12 Carmen Bratosin (TUE), Grid Architecture for Distributed Process Mining
13 Xiaoyu Mao (UvT), Airport under Control. Multiagent Scheduling for Airport Ground Handling
14 Milan Lovric (EUR), Behavioral Finance and Agent-Based Artificial Markets
15 Marijn Koolen (UvA), The Meaning of Structure: the Value of Link Evidence for Information Retrieval
16 Maarten Schadd (UM), Selective Search in Games of Different Complexity
17 Jiyin He (UVA), Exploring Topic Structure: Coherence, Diversity and Relatedness
18 Mark Ponsen (UM), Strategic Decision-Making in complex games
19 Ellen Rusman (OU), The Mind's Eye on Personal Profiles
20 Qing Gu (VU), Guiding service-oriented software engineering - A view-based approach
21 Linda Terlouw (TUD), Modularization and Specification of Service-Oriented Systems
22 Junte Zhang (UVA), System Evaluation of Archival Description and Access
23 Wouter Weerkamp (UVA), Finding People and their Utterances in Social Media
24 Herwin van Welbergen (UT), Behavior Generation for Interpersonal Coordination with Virtual Humans On Specifying, Scheduling and Realizing Multimodal Virtual Human Behavior
25 Syed Waqar ul Qounain Jaffry (VU), Analysis and Validation of Models for Trust Dynamics
26 Matthijs Aart Pontier (VU), Virtual Agents for Human Communication - Emotion Regulation and Involvement-Distance Trade-Offs in Embodied Conversational Agents and Robots
27 Aniel Bhulai (VU), Dynamic website optimization through autonomous management of design patterns
28 Rianne Kaptein (UVA), Effective Focused Retrieval by Exploiting Query Context and Document Structure
29 Faisal Kamiran (TUE), Discrimination-aware Classification
30 Egon van den Broek (UT), Affective Signal Processing (ASP): Unraveling the mystery of emotions
31 Ludo Waltman (EUR), Computational and Game-Theoretic Approaches for Modeling Bounded Rationality
32 Nees-Jan van Eck (EUR), Methodological Advances in Bibliometric Mapping of Science
33 Tom van der Weide (UU), Arguing to Motivate Decisions
34 Paolo Turrini (UU), Strategic Reasoning in Interdependence: Logical and Game-theoretical Investigations
35 Maaike Harbers (UU), Explaining Agent Behavior in Virtual Training
36 Erik van der Spek (UU), Experiments in serious game design: a cognitive approach
37 Adriana Burlutiu (RUN), Machine Learning for Pairwise Data, Applications for Preference Learning and Supervised Network Inference
38 Nyree Lemmens (UM), Bee-inspired Distributed Optimization
39 Joost Westra (UU), Organizing Adaptation using Agents in Serious Games
40 Viktor Clerc (VU), Architectural Knowledge Management in Global Software Development
41 Luan Ibraimi (UT), Cryptographically Enforced Distributed Data Access Control
42 Michal Sindlar (UU), Explaining Behavior through Mental State Attribution
43 Henk van der Schuur (UU), Process Improvement through Software Operation Knowledge
44 Boris Reuderink (UT), Robust Brain-Computer Interfaces
45 Herman Stehouwer (UvT), Statistical Language Models for Alternative Sequence Selection

46 Beibei Hu (TUD), Towards Contextualized Information Delivery: A Rule-based Architecture for the Domain of Mobile Police Work
47 Azizi Bin Ab Aziz (VU), Exploring Computational Models for Intelligent Support of Persons with Depression
48 Mark Ter Maat (UT), Response Selection and Turn-taking for a Sensitive Artificial Listening Agent
49 Andreea Niculescu (UT), Conversational interfaces for task-oriented spoken dialogues: design aspects influencing interaction quality

---

2012 01 Terry Kakeeto (UvT), Relationship Marketing for SMEs in Uganda
02 Muhammad Umair (VU), Adaptivity, emotion, and Rationality in Human and Ambient Agent Models
03 Adam Vanya (VU), Supporting Architecture Evolution by Mining Software Repositories
04 Jurriaan Souer (UU), Development of Content Management System-based Web Applications
05 Marijn Plomp (UU), Maturing Interorganisational Information Systems
06 Wolfgang Reinhardt (OU), Awareness Support for Knowledge Workers in Research Networks
07 Rianne van Lambalgen (VU), When the Going Gets Tough: Exploring Agent-based Models of Human Performance under Demanding Conditions
08 Gerben de Vries (UVA), Kernel Methods for Vessel Trajectories
09 Ricardo Neisse (UT), Trust and Privacy Management Support for Context-Aware Service Platforms
10 David Smits (TUE), Towards a Generic Distributed Adaptive Hypermedia Environment
11 J.C.B. Rantham Prabhakara (TUE), Process Mining in the Large: Preprocessing, Discovery, and Diagnostics
12 Kees van der Sluijs (TUE), Model Driven Design and Data Integration in Semantic Web Information Systems
13 Suleman Shahid (UvT), Fun and Face: Exploring non-verbal expressions of emotion during playful interactions
14 Evgeny Knutov (TUE), Generic Adaptation Framework for Unifying Adaptive Web-based Systems
15 Natalie van der Wal (VU), Social Agents. Agent-Based Modelling of Integrated Internal and Social Dynamics of Cognitive and Affective Processes.
16 Fiemke Both (VU), Helping people by understanding them - Ambient Agents supporting task execution and depression treatment
17 Amal Elgammal (UvT), Towards a Comprehensive Framework for Business Process Compliance
18 Eltjo Poort (VU), Improving Solution Architecting Practices
19 Helen Schonenberg (TUE), What's Next? Operational Support for Business Process Execution
20 Ali Bahramisharif (RUN), Covert Visual Spatial Attention, a Robust Paradigm for Brain-Computer Interfacing
21 Roberto Cornacchia (TUD), Querying Sparse Matrices for Information Retrieval
22 Thijs Vis (UvT), Intelligence, politie en veiligheidsdienst: verenigbare grootheden?
23 Christian Muehl (UT), Toward Affective Brain-Computer Interfaces: Exploring the Neurophysiology of Affect during Human Media Interaction
24 Laurens van der Werff (UT), Evaluation of Noisy Transcripts for Spoken Document Retrieval
25 Silja Eckartz (UT), Managing the Business Case Development in Inter-Organizational IT Projects: A Methodology and its Application
26 Emile de Maat (UVA), Making Sense of Legal Text
27 Hayrettin Gurkok (UT), Mind the Sheep! User Experience Evaluation & Brain-Computer Interface Games
28 Nancy Pascall (UvT), Engendering Technology Empowering Women
29 Almer Tigelaar (UT), Peer-to-Peer Information Retrieval
30 Alina Pommeranz (TUD), Designing Human-Centered Systems for Reflective Decision Making
31 Emily Bagarukayo (RUN), A Learning by Construction Approach for Higher Order Cognitive Skills Improvement, Building Capacity and Infrastructure
32 Wietske Visser (TUD), Qualitative multi-criteria preference representation and reasoning
33 Rory Sie (OUN), Coalitions in Cooperation Networks (COCOON)
34 Pavol Jancura (RUN), Evolutionary analysis in PPI networks and applications
35 Evert Haasdijk (VU), Never Too Old To Learn – On-line Evolution of Controllers in Swarm- and Modular Robotics
36 Denis Ssebugwawo (RUN), Analysis and Evaluation of Collaborative Modeling Processes
37 Agnes Nakakawa (RUN), A Collaboration Process for Enterprise Architecture Creation
38 Selmar Smit (VU), Parameter Tuning and Scientific Testing in Evolutionary Algorithms
39 Hassan Fatemi (UT), Risk-aware design of value and coordination networks
40 Agus Gunawan (UvT), Information Access for SMEs in Indonesia
41 Sebastian Kelle (OU), Game Design Patterns for Learning
42 Dominique Verpoorten (OU), Reflection Amplifiers in self-regulated Learning
43 Withdrawn
44 Anna Tordai (VU), On Combining Alignment Techniques
45 Benedikt Kratz (UvT), A Model and Language for Business-aware Transactions
46 Simon Carter (UVA), Exploration and Exploitation of Multilingual Data for Statistical Machine Translation
47 Manos Tsagkias (UVA), Mining Social Media: Tracking Content and Predicting Behavior
48 Jorn Bakker (TUE), Handling Abrupt Changes in Evolving Time-series Data
49 Michael Kaisers (UM), Learning against Learning - Evolutionary dynamics of reinforcement learning algorithms in strategic interactions
50 Steven van Kervel (TUD), Ontologogy driven Enterprise Information Systems Engineering
51 Jeroen de Jong (TUD), Heuristics in Dynamic Sceduling; a practical framework with a case study in elevator dispatching

---

2013 01 Viorel Milea (EUR), News Analytics for Financial Decision Support
02 Erietta Liarou (CWI), MonetDB/DataCell: Leveraging the Column-store Database Technology for Efficient and Scalable Stream Processing
03 Szymon Klarman (VU), Reasoning with Contexts in Description Logics

04 Chetan Yadati (TUD), Coordinating autonomous planning and scheduling
05 Dulce Pumareja (UT), Groupware Requirements Evolutions Patterns
06 Romulo Goncalves (CWI), The Data Cyclotron: Juggling Data and Queries for a Data Warehouse Audience
07 Giel van Lankveld (UvT), Quantifying Individual Player Differences
08 Robbert-Jan Merk (VU), Making enemies: cognitive modeling for opponent agents in fighter pilot simulators
09 Fabio Gori (RUN), Metagenomic Data Analysis: Computational Methods and Applications
10 Jeewanie Jayasinghe Arachchige (UvT), A Unified Modeling Framework for Service Design.
11 Evangelos Pournaras (TUD), Multi-level Reconfigurable Self-organization in Overlay Services
12 Marian Razavian (VU), Knowledge-driven Migration to Services
13 Mohammad Safiri (UT), Service Tailoring: User-centric creation of integrated IT-based homecare services to support independent living of elderly
14 Jafar Tanha (UVA), Ensemble Approaches to Semi-Supervised Learning Learning
15 Daniel Hennes (UM), Multiagent Learning - Dynamic Games and Applications
16 Eric Kok (UU), Exploring the practical benefits of argumentation in multi-agent deliberation
17 Koen Kok (VU), The PowerMatcher: Smart Coordination for the Smart Electricity Grid
18 Jeroen Janssens (UvT), Outlier Selection and One-Class Classification
19 Renze Steenhuizen (TUD), Coordinated Multi-Agent Planning and Scheduling
20 Katja Hofmann (UvA), Fast and Reliable Online Learning to Rank for Information Retrieval
21 Sander Wubben (UvT), Text-to-text generation by monolingual machine translation
22 Tom Claassen (RUN), Causal Discovery and Logic
23 Patricio de Alencar Silva (UvT), Value Activity Monitoring
24 Haitham Bou Ammar (UM), Automated Transfer in Reinforcement Learning
25 Agnieszka Anna Latoszek-Berendsen (UM), Intention-based Decision Support. A new way of representing and implementing clinical guidelines in a Decision Support System
26 Alireza Zarghami (UT), Architectural Support for Dynamic Homecare Service Provisioning
27 Mohammad Huq (UT), Inference-based Framework Managing Data Provenance
28 Frans van der Sluis (UT), When Complexity becomes Interesting: An Inquiry into the Information eXperience
29 Iwan de Kok (UT), Listening Heads
30 Joyce Nakatumba (TUE), Resource-Aware Business Process Management: Analysis and Support
31 Dinh Khoa Nguyen (UvT), Blueprint Model and Language for Engineering Cloud Applications
32 Kamakshi Rajagopal (OUN), Networking For Learning; The role of Networking in a Lifelong Learner's Professional Development
33 Qi Gao (TUD), User Modeling and Personalization in the Microblogging Sphere
34 Kien Tjin-Kam-Jet (UT), Distributed Deep Web Search
35 Abdallah El Ali (UvA), Minimal Mobile Human Computer Interaction
36 Than Lam Hoang (TUe), Pattern Mining in Data Streams
37 Dirk Börner (OUN), Ambient Learning Displays
38 Eelco den Heijer (VU), Autonomous Evolutionary Art
39 Joop de Jong (TUD), A Method for Enterprise Ontology based Design of Enterprise Information Systems
40 Pim Nijssen (UM), Monte-Carlo Tree Search for Multi-Player Games
41 Jochem Liem (UVA), Supporting the Conceptual Modelling of Dynamic Systems: A Knowledge Engineering Perspective on Qualitative Reasoning
42 Léon Planken (TUD), Algorithms for Simple Temporal Reasoning
43 Marc Bron (UVA), Exploration and Contextualization through Interaction and Concepts

---

2014 01 Nicola Barile (UU), Studies in Learning Monotone Models from Data
02 Fiona Tuliyano (RUN), Combining System Dynamics with a Domain Modeling Method
03 Sergio Raul Duarte Torres (UT), Information Retrieval for Children: Search Behavior and Solutions
04 Hanna Jochmann-Mannak (UT), Websites for children: search strategies and interface design - Three studies on children's search performance and evaluation
05 Jurriaan van Reijsen (UU), Knowledge Perspectives on Advancing Dynamic Capability
06 Damian Tamburri (VU), Supporting Networked Software Development
07 Arya Adriansyah (TUE), Aligning Observed and Modeled Behavior
08 Samur Araujo (TUD), Data Integration over Distributed and Heterogeneous Data Endpoints
09 Philip Jackson (UvT), Toward Human-Level Artificial Intelligence: Representation and Computation of Meaning in Natural Language
10 Ivan Salvador Razo Zapata (VU), Service Value Networks
11 Janneke van der Zwaan (TUD), An Empathic Virtual Buddy for Social Support
12 Willem van Willigen (VU), Look Ma, No Hands: Aspects of Autonomous Vehicle Control
13 Arlette van Wissen (VU), Agent-Based Support for Behavior Change: Models and Applications in Health and Safety Domains
14 Yangyang Shi (TUD), Language Models With Meta-information
15 Natalya Mogles (VU), Agent-Based Analysis and Support of Human Functioning in Complex Socio-Technical Systems: Applications in Safety and Healthcare
16 Krystyna Milian (VU), Supporting trial recruitment and design by automatically interpreting eligibility criteria
17 Kathrin Dentler (VU), Computing healthcare quality indicators automatically: Secondary Use of Patient Data and Semantic Interoperability
18 Mattijs Ghijsen (UVA), Methods and Models for the Design and Study of Dynamic Agent Organizations
19 Vinicius Ramos (TUE), Adaptive Hypermedia Courses: Qualitative and Quantitative Evaluation and Tool Support

20 Mena Habib (UT), Named Entity Extraction and Disambiguation for Informal Text: The Missing Link
21 Kassidy Clark (TUD), Negotiation and Monitoring in Open Environments
22 Marieke Peeters (UU), Personalized Educational Games - Developing agent-supported scenario-based training
23 Eleftherios Sidirourgos (UvA/CWI), Space Efficient Indexes for the Big Data Era
24 Davide Ceolin (VU), Trusting Semi-structured Web Data
25 Martijn Lappenschaar (RUN), New network models for the analysis of disease interaction
26 Tim Baarslag (TUD), What to Bid and When to Stop
27 Rui Jorge Almeida (EUR), Conditional Density Models Integrating Fuzzy and Probabilistic Representations of Uncertainty
28 Anna Chmielowiec (VU), Decentralized k-Clique Matching
29 Jaap Kabbedijk (UU), Variability in Multi-Tenant Enterprise Software
30 Peter de Cock (UvT), Anticipating Criminal Behaviour
31 Leo van Moergestel (UU), Agent Technology in Agile Multiparallel Manufacturing and Product Support
32 Naser Ayat (UvA), On Entity Resolution in Probabilistic Data
33 Tesfa Tegegne (RUN), Service Discovery in eHealth
34 Christina Manteli (VU), The Effect of Governance in Global Software Development: Analyzing Transactive Memory Systems.
35 Joost van Ooijen (UU), Cognitive Agents in Virtual Worlds: A Middleware Design Approach
36 Joos Buijs (TUE), Flexible Evolutionary Algorithms for Mining Structured Process Models
37 Maral Dadvar (UT), Experts and Machines United Against Cyberbullying
38 Danny Plass-Oude Bos (UT), Making brain-computer interfaces better: improving usability through post-processing.
39 Jasmina Maric (UvT), Web Communities, Immigration, and Social Capital
40 Walter Omona (RUN), A Framework for Knowledge Management Using ICT in Higher Education
41 Frederic Hogenboom (EUR), Automated Detection of Financial Events in News Text
42 Carsten Eijckhof (CWI/TUD), Contextual Multidimensional Relevance Models
43 Kevin Vlaanderen (UU), Supporting Process Improvement using Method Increments
44 Paulien Meesters (UvT), Intelligent Blauw. Met als ondertitel: Intelligence-gestuurde politiezorg in gebiedsgebonden eenheden.
45 Birgit Schmitz (OUN), Mobile Games for Learning: A Pattern-Based Approach
46 Ke Tao (TUD), Social Web Data Analytics: Relevance, Redundancy, Diversity
47 Shangsong Liang (UVA), Fusion and Diversification in Information Retrieval

2015 01 Niels Netten (UvA), Machine Learning for Relevance of Information in Crisis Response
02 Faiza Bukhsh (UvT), Smart auditing: Innovative Compliance Checking in Customs Controls
03 Twan van Laarhoven (RUN), Machine learning for network data
04 Howard Spoelstra (OUN), Collaborations in Open Learning Environments
05 Christoph Bösch (UT), Cryptographically Enforced Search Pattern Hiding
06 Farideh Heidari (TUD), Business Process Quality Computation - Computing Non-Functional Requirements to Improve Business Processes
07 Maria-Hendrike Peetz (UvA), Time-Aware Online Reputation Analysis
08 Jie Jiang (TUD), Organizational Compliance: An agent-based model for designing and evaluating organizational interactions
09 Randy Klaassen (UT), HCI Perspectives on Behavior Change Support Systems
10 Henry Hermans (OUN), OpenU: design of an integrated system to support lifelong learning
11 Yongming Luo (TUE), Designing algorithms for big graph datasets: A study of computing bisimulation and joins
12 Julie M. Birkholz (VU), Modi Operandi of Social Network Dynamics: The Effect of Context on Scientific Collaboration Networks
13 Giuseppe Procaccianti (VU), Energy-Efficient Software
14 Bart van Straalen (UT), A cognitive approach to modeling bad news conversations
15 Klaas Andries de Graaf (VU), Ontology-based Software Architecture Documentation
16 Changyun Wei (UT), Cognitive Coordination for Cooperative Multi-Robot Teamwork
17 André van Cleeff (UT), Physical and Digital Security Mechanisms: Properties, Combinations and Trade-offs
18 Holger Pirk (CWI), Waste Not, Want Not! - Managing Relational Data in Asymmetric Memories
19 Bernardo Tabuenca (OUN), Ubiquitous Technology for Lifelong Learners
20 Lois Vanhée (UU), Using Culture and Values to Support Flexible Coordination
21 Sibren Fetter (OUN), Using Peer-Support to Expand and Stabilize Online Learning
22 Zhemin Zhu (UT), Co-occurrence Rate Networks
23 Luit Gazendam (VU), Cataloguer Support in Cultural Heritage
24 Richard Berendsen (UVA), Finding People, Papers, and Posts: Vertical Search Algorithms and Evaluation
25 Steven Woudenberg (UU), Bayesian Tools for Early Disease Detection
26 Alexander Hogenboom (EUR), Sentiment Analysis of Text Guided by Semantics and Structure
27 Sándor Héman (CWI), Updating compressed colomn stores
28 Janet Bagorogoza (TiU), Knowledge Management and High Performance; The Uganda Financial Institutions Model for HPO
29 Hendrik Baier (UM), Monte-Carlo Tree Search Enhancements for One-Player and Two-Player Domains
30 Kiavash Bahreini (OU), Real-time Multimodal Emotion Recognition in E-Learning
31 Yakup Koç (TUD), On the robustness of Power Grids
32 Jerome Gard (UL), Corporate Venture Management in SMEs
33 Frederik Schadd (TUD), Ontology Mapping with Auxiliary Resources

34 Victor de Graaf (UT), Gesocial Recommender Systems
35 Jungxao Xu (TUD), Affective Body Language of Humanoid Robots: Perception and Effects in Human Robot Interaction

2016 01 Syed Saiden Abbas (RUN), Recognition of Shapes by Humans and Machines
02 Michiel Christiaan Meulendijk (UU), Optimizing medication reviews through decision support: prescribing a better pill to swallow
03 Maya Sappelli (RUN), Knowledge Work in Context: User Centered Knowledge Worker Support
04 Laurens Rietveld (VU), Publishing and Consuming Linked Data
05 Evgeny Sherkhonov (UVA), Expanded Acyclic Queries: Containment and an Application in Explaining Missing Answers
06 Michel Wilson (TUD), Robust scheduling in an uncertain environment
07 Jeroen de Man (VU), Measuring and modeling negative emotions for virtual training
08 Matje van de Camp (TiU), A Link to the Past: Constructing Historical Social Networks from Unstructured Data
09 Archana Nottamkandath (VU), Trusting Crowdsourced Information on Cultural Artefacts
10 George Karafotias (VUA), Parameter Control for Evolutionary Algorithms
11 Anne Schuth (UVA), Search Engines that Learn from Their Users
12 Max Knobbout (UU), Logics for Modelling and Verifying Normative Multi-Agent Systems
13 Nana Baah Gyan (VU), The Web, Speech Technologies and Rural Development in West Africa - An ICT4D Approach
14 Ravi Khadka (UU), Revisiting Legacy Software System Modernization
15 Steffen Michels (RUN), Hybrid Probabilistic Logics - Theoretical Aspects, Algorithms and Experiments
16 Guangliang Li (UVA), Socially Intelligent Autonomous Agents that Learn from Human Reward
17 Berend Weel (VU), Towards Embodied Evolution of Robot Organisms
18 Albert Meroño Peñuela (VU), Refining Statistical Data on the Web
19 Julia Efremova (Tu/e), Mining Social Structures from Genealogical Data
20 Daan Odijk (UVA), Context & Semantics in News & Web Search
21 Alejandro Moreno Célleri (UT), From Traditional to Interactive Playspaces: Automatic Analysis of Player Behavior in the Interactive Tag Playground
22 Grace Lewis (VU), Software Architecture Strategies for Cyber-Foraging Systems
23 Fei Cai (UVA), Query Auto Completion in Information Retrieval
24 Brend Wanders (UT), Repurposing and Probabilistic Integration of Data; An Iterative and data model independent approach
25 Julia Kiseleva (TU/e), Using Contextual Information to Understand Searching and Browsing Behavior
26 Dilhan Thilakarathne (VU), In or Out of Control: Exploring Computational Models to Study the Role of Human Awareness and Control in Behavioural Choices, with Applications in Aviation and Energy Management Domains
27 Wen Li (TUD), Understanding Geo-spatial Information on Social Media
28 Mingxin Zhang (TUD), Large-scale Agent-based Social Simulation - A study on epidemic prediction and control
29 Nicolas Höning (TUD), Peak reduction in decentralised electricity systems - Markets and prices for flexible planning
30 Ruud Mattheij (UvT), The Eyes Have It
31 Mohammad Khelghati (UT), Deep web content monitoring
32 Eelco Vriezekolk (UT), Assessing Telecommunication Service Availability Risks for Crisis Organisations
33 Peter Bloem (UVA), Single Sample Statistics, exercises in learning from just one example
34 Dennis Schunselaar (TUE), Configurable Process Trees: Elicitation, Analysis, and Enactment
35 Zhaochun Ren (UVA), Monitoring Social Media: Summarization, Classification and Recommendation
36 Daphne Karreman (UT), Beyond R2D2: The design of nonverbal interaction behavior optimized for robot-specific morphologies
37 Giovanni Sileno (UvA), Aligning Law and Action - a conceptual and computational inquiry
38 Andrea Minuto (UT), Materials that Matter - Smart Materials meet Art & Interaction Design
39 Merijn Bruijnes (UT), Believable Suspect Agents; Response and Interpersonal Style Selection for an Artificial Suspect
40 Christian Detweiler (TUD), Accounting for Values in Design
41 Thomas King (TUD), Governing Governance: A Formal Framework for Analysing Institutional Design and Enactment Governance
42 Spyros Martzoukos (UVA), Combinatorial and Compositional Aspects of Bilingual Aligned Corpora
43 Saskia Koldijk (RUN), Context-Aware Support for Stress Self-Management: From Theory to Practice
44 Thibault Sellam (UVA), Automatic Assistants for Database Exploration
45 Bram van de Laar (UT), Experiencing Brain-Computer Interface Control
46 Jorge Gallego Perez (UT), Robots to Make you Happy
47 Christina Weber (UL), Real-time foresight - Preparedness for dynamic innovation networks
48 Tanja Buttler (TUD), Collecting Lessons Learned
49 Gleb Polevoy (TUD), Participation and Interaction in Projects. A Game-Theoretic Analysis
50 Yan Wang (UVT), The Bridge of Dreams: Towards a Method for Operational Performance Alignment in IT-enabled Service Supply Chains

2017 01 Jan-Jaap Oerlemans (UL), Investigating Cybercrime
02 Sjoerd Timmer (UU), Designing and Understanding Forensic Bayesian Networks using Argumentation
03 Daniël Harold Telgen (UU), Grid Manufacturing; A Cyber-Physical Approach with Autonomous Products and Reconfigurable Manufacturing Machines
04 Mrunal Gawade (CWI), Multi-core Parallelism in a Column-store

05 Mahdieh Shadi (UVA), Collaboration Behavior
06 Damir Vandic (EUR), Intelligent Information Systems for Web Product Search
07 Roel Bertens (UU), Insight in Information: from Abstract to Anomaly
08 Rob Konijn (VU) , Detecting Interesting Differences:Data Mining in Health Insurance Data using Outlier Detection and Subgroup Discovery
09 Dong Nguyen (UT), Text as Social and Cultural Data: A Computational Perspective on Variation in Text
10 Robby van Delden (UT), (Steering) Interactive Play Behavior
11 Florian Kunneman (RUN), Modelling patterns of time and emotion in Twitter #anticipointment
12 Sander Leemans (TUE), Robust Process Mining with Guarantees
13 Gijs Huisman (UT), Social Touch Technology - Extending the reach of social touch through haptic technology
14 Shoshannah Tekofsky (UvT), You Are Who You Play You Are: Modelling Player Traits from Video Game Behavior
15 Peter Berck (RUN), Memory-Based Text Correction
16 Aleksandr Chuklin (UVA), Understanding and Modeling Users of Modern Search Engines
17 Daniel Dimov (UL), Crowdsourced Online Dispute Resolution
18 Ridho Reinanda (UVA), Entity Associations for Search
19 Jeroen Vuurens (UT), Proximity of Terms, Texts and Semantic Vectors in Information Retrieval
20 Mohammadbashir Sedighi (TUD), Fostering Engagement in Knowledge Sharing: The Role of Perceived Benefits, Costs and Visibility
21 Jeroen Linssen (UT), Meta Matters in Interactive Storytelling and Serious Gaming (A Play on Worlds)
22 Sara Magliacane (VU), Logics for causal inference under uncertainty
23 David Graus (UVA), Entities of Interest — Discovery in Digital Traces
24 Chang Wang (TUD), Use of Affordances for Efficient Robot Learning
25 Veruska Zamborlini (VU), Knowledge Representation for Clinical Guidelines, with applications to Multimorbidity Analysis and Literature Search
26 Merel Jung (UT), Socially intelligent robots that understand and respond to human touch
27 Michiel Joosse (UT), Investigating Positioning and Gaze Behaviors of Social Robots: People's Preferences, Perceptions and Behaviors
28 John Klein (VU), Architecture Practices for Complex Contexts
29 Adel Alhuraibi (UvT), From IT-BusinessStrategic Alignment to Performance: A Moderated Mediation Model of Social Innovation, and Enterprise Governance of IT"
30 Wilma Latuny (UvT), The Power of Facial Expressions
31 Ben Ruijl (UL), Advances in computational methods for QFT calculations
32 Thaer Samar (RUN), Access to and Retrievability of Content in Web Archives
33 Brigit van Loggem (OU), Towards a Design Rationale for Software Documentation: A Model of Computer-Mediated Activity
34 Maren Scheffel (OU), The Evaluation Framework for Learning Analytics
35 Martine de Vos (VU), Interpreting natural science spreadsheets
36 Yuanhao Guo (UL), Shape Analysis for Phenotype Characterisation from High-throughput Imaging
37 Alejandro Montes Garcia (TUE), WiBAF: A Within Browser Adaptation Framework that Enables Control over Privacy
38 Alex Kayal (TUD), Normative Social Applications
39 Sara Ahmadi (RUN), Exploiting properties of the human auditory system and compressive sensing methods to increase noise robustness in ASR
40 Altaf Hussain Abro (VUA), Steer your Mind: Computational Exploration of Human Control in Relation to Emotions, Desires and Social Support For applications in human-aware support systems
41 Adnan Manzoor (VUA), Minding a Healthy Lifestyle: An Exploration of Mental Processes and a Smart Environment to Provide Support for a Healthy Lifestyle
42 Elena Sokolova (RUN), Causal discovery from mixed and missing data with applications on ADHD datasets
43 Maaike de Boer (RUN), Semantic Mapping in Video Retrieval
44 Garm Lucassen (UU), Understanding User Stories - Computational Linguistics in Agile Requirements Engineering
45 Bas Testerink (UU), Decentralized Runtime Norm Enforcement
46 Jan Schneider (OU), Sensor-based Learning Support
47 Jie Yang (TUD), Crowd Knowledge Creation Acceleration
48 Angel Suarez (OU), Collaborative inquiry-based learning

2018 01 Han van der Aa (VUA), Comparing and Aligning Process Representations
02 Felix Mannhardt (TUE), Multi-perspective Process Mining
03 Steven Bosems (UT), Causal Models For Well-Being: Knowledge Modeling, Model-Driven Development of Context-Aware Applications, and Behavior Prediction
04 Jordan Janeiro (TUD), Flexible Coordination Support for Diagnosis Teams in Data-Centric Engineering Tasks
05 Hugo Huurdeman (UVA), Supporting the Complex Dynamics of the Information Seeking Process
06 Dan Ionita (UT), Model-Driven Information Security Risk Assessment of Socio-Technical Systems
07 Jieting Luo (UU), A formal account of opportunism in multi-agent systems
08 Rick Smetsers (RUN), Advances in Model Learning for Software Systems
09 Xu Xie (TUD), Data Assimilation in Discrete Event Simulations
10 Julienka Mollee (VUA), Moving forward: supporting physical activity behavior change through intelligent technology
11 Mahdi Sargolzaei (UVA), Enabling Framework for Service-oriented Collaborative Networks
12 Xixi Lu (TUE), Using behavioral context in process mining
13 Seyed Amin Tabatabaei (VUA), Computing a Sustainable Future

14 Bart Joosten (UVT), Detecting Social Signals with Spatiotemporal Gabor Filters
15 Naser Davarzani (UM), Biomarker discovery in heart failure
16 Jaebok Kim (UT), Automatic recognition of engagement and emotion in a group of children
17 Jianpeng Zhang (TUE), On Graph Sample Clustering
18 Henriette Nakad (UL), De Notaris en Private Rechtspraak
19 Minh Duc Pham (VUA), Emergent relational schemas for RDF
20 Manxia Liu (RUN), Time and Bayesian Networks
21 Aad Slootmaker (OUN), EMERGO: a generic platform for authoring and playing scenario-based serious games
22 Eric Fernandes de Mello Araujo (VUA), Contagious: Modeling the Spread of Behaviours, Perceptions and Emotions in Social Networks
23 Kim Schouten (EUR), Semantics-driven Aspect-Based Sentiment Analysis
24 Jered Vroon (UT), Responsive Social Positioning Behaviour for Semi-Autonomous Telepresence Robots
25 Riste Gligorov (VUA), Serious Games in Audio-Visual Collections
26 Roelof Anne Jelle de Vries (UT),Theory-Based and Tailor-Made: Motivational Messages for Behavior Change Technology
27 Maikel Leemans (TUE), Hierarchical Process Mining for Scalable Software Analysis
28 Christian Willemse (UT), Social Touch Technologies: How they feel and how they make you feel
29 Yu Gu (UVT), Emotion Recognition from Mandarin Speech
30 Wouter Beek, The "K" in "semantic web" stands for "knowledge": scaling semantics to the web

---

2019 01 Rob van Eijk (UL),Web privacy measurement in real-time bidding systems. A graph-based approach to RTB system classification
02 Emmanuelle Beauxis Aussalet (CWI, UU), Statistics and Visualizations for Assessing Class Size Uncertainty
03 Eduardo Gonzalez Lopez de Murillas (TUE), Process Mining on Databases: Extracting Event Data from Real Life Data Sources
04 Ridho Rahmadi (RUN), Finding stable causal structures from clinical data
05 Sebastiaan van Zelst (TUE), Process Mining with Streaming Data
06 Chris Dijkshoorn (VU), Nichesourcing for Improving Access to Linked Cultural Heritage Datasets
07 Soude Fazeli (TUD), Recommender Systems in Social Learning Platforms
08 Frits de Nijs (TUD), Resource-constrained Multi-agent Markov Decision Processes
09 Fahimeh Alizadeh Moghaddam (UVA), Self-adaptation for energy efficiency in software systems
10 Qing Chuan Ye (EUR), Multi-objective Optimization Methods for Allocation and Prediction
11 Yue Zhao (TUD), Learning Analytics Technology to Understand Learner Behavioral Engagement in MOOCs
12 Jacqueline Heinerman (VU), Better Together
13 Guanliang Chen (TUD), MOOC Analytics: Learner Modeling and Content Generation
14 Daniel Davis (TUD), Large-Scale Learning Analytics: Modeling Learner Behavior & Improving Learning Outcomes in Massive Open Online Courses
15 Erwin Walraven (TUD), Planning under Uncertainty in Constrained and Partially Observable Environments
16 Guangming Li (TUE), Process Mining based on Object-Centric Behavioral Constraint (OCBC) Models
17 Ali Hurriyetoglu (RUN),Extracting actionable information from microtexts
18 Gerard Wagenaar (UU), Artefacts in Agile Team Communication
19 Vincent Koeman (TUD), Tools for Developing Cognitive Agents
20 Chide Groenouwe (UU), Fostering technically augmented human collective intelligence
21 Cong Liu (TUE), Software Data Analytics: Architectural Model Discovery and Design Pattern Detection
22 Martin van den Berg (VU),Improving IT Decisions with Enterprise Architecture
23 Qin Liu (TUD), Intelligent Control Systems: Learning, Interpreting, Verification
24 Anca Dumitrache (VU), Truth in Disagreement - Crowdsourcing Labeled Data for Natural Language Processing
25 Emiel van Miltenburg (VU), Pragmatic factors in (automatic) image description
26 Prince Singh (UT), An Integration Platform for Synchromodal Transport
27 Alessandra Antonaci (OUN), The Gamification Design Process applied to (Massive) Open Online Courses
28 Esther Kuindersma (UL), Cleared for take-off: Game-based learning to prepare airline pilots for critical situations
29 Daniel Formolo (VU), Using virtual agents for simulation and training of social skills in safety-critical circumstances
30 Vahid Yazdanpanah (UT), Multiagent Industrial Symbiosis Systems
31 Milan Jelisavcic (VU), Alive and Kicking: Baby Steps in Robotics
32 Chiara Sironi (UM), Monte-Carlo Tree Search for Artificial General Intelligence in Games
33 Anil Yaman (TUE), Evolution of Biologically Inspired Learning in Artificial Neural Networks
34 Negar Ahmadi (TUE), EEG Microstate and Functional Brain Network Features for Classification of Epilepsy and PNES
35 Lisa Facey-Shaw (OUN), Gamification with digital badges in learning programming
36 Kevin Ackermans (OUN), Designing Video-Enhanced Rubrics to Master Complex Skills
37 Jian Fang (TUD), Database Acceleration on FPGAs
38 Akos Kadar (OUN), Learning visually grounded and multilingual representations

---

2020 01 Armon Toubman (UL), Calculated Moves: Generating Air Combat Behaviour
02 Marcos de Paula Bueno (UL), Unraveling Temporal Processes using Probabilistic Graphical Models
03 Mostafa Deghani (UvA), Learning with Imperfect Supervision for Language Understanding
04 Maarten van Gompel (RUN), Context as Linguistic Bridges
05 Yulong Pei (TUE), On local and global structure mining
06 Preethu Rose Anish (UT), Stimulation Architectural Thinking during Requirements Elicitation - An Approach and Tool Support

07 Wim van der Vegt (OUN), Towards a software architecture for reusable game components
08 Ali Mirsoleimani (UL),Structured Parallel Programming for Monte Carlo Tree Search
09 Myriam Traub (UU), Measuring Tool Bias and Improving Data Quality for Digital Humanities Research
10 Alifah Syamsiyah (TUE), In-database Preprocessing for Process Mining
11 Sepideh Mesbah (TUD), Semantic-Enhanced Training Data AugmentationMethods for Long-Tail Entity Recognition Models
12 Ward van Breda (VU), Predictive Modeling in E-Mental Health: Exploring Applicability in Personalised Depression Treatment
13 Marco Virgolin (CWI), Design and Application of Gene-pool Optimal Mixing Evolutionary Algorithms for Genetic Programming
14 Mark Raasveldt (CWI/UL), Integrating Analytics with Relational Databases
15 Konstantinos Georgiadis (OUN), Smart CAT: Machine Learning for Configurable Assessments in Serious Games
16 Ilona Wilmont (RUN), Cognitive Aspects of Conceptual Modelling
17 Daniele Di Mitri (OUN), The Multimodal Tutor: Adaptive Feedback from Multimodal Experiences
18 Georgios Methenitis (TUD), Agent Interactions & Mechanisms in Markets with Uncertainties: Electricity Markets in Renewable Energy Systems
19 Guido van Capelleveen (UT), Industrial Symbiosis Recommender Systems
20 Albert Hankel (VU), Embedding Green ICT Maturity in Organisations
21 Karine da Silva Miras de Araujo (VU), Where is the robot?: Life as it could be
22 Maryam Masoud Khamis (RUN), Understanding complex systems implementation through a modeling approach: the case of e-government in Zanzibar
23 Rianne Conijn (UT), The Keys to Writing: A writing analytics approach to studying writing processes using keystroke logging
24 Lenin da Nobrega Medeiros (VUA/RUN), How are you feeling, human? Towards emotionally supportive chatbots
25 Xin Du (TUE), The Uncertainty in Exceptional Model Mining
26 Krzysztof Leszek Sadowski (UU), GAMBIT: Genetic Algorithm for Model-Based mixed-Integer opTimization
27 Ekaterina Muravyeva (TUD), Personal data and informed consent in an educational context
28 Bibeg Limbu (TUD), Multimodal interaction for deliberate practice: Training complex skills with augmented reality
29 Ioan Gabriel Bucur (RUN), Being Bayesian about Causal Inference
30 Bob Zadok Blok (UL), Creatief, Creatieve, Creatiefst
31 Gongjin Lan (VU), Learning better – From Baby to Better
32 Jason Rhuggenaath (TUE), Revenue management in online markets: pricing and online advertising
33 Rick Gilsing (TUE), Supporting service-dominant business model evaluation in the context of business model innovation
34 Anna Bon (MU), Intervention or Collaboration? Redesigning Information and Communication Technologies for Development
35 Siamak Farshidi (UU), Multi-Criteria Decision-Making in Software Production

---

2021 01 Francisco Xavier Dos Santos Fonseca (TUD),Location-based Games for Social Interaction in Public Space
02 Rijk Mercuur (TUD), Simulating Human Routines:Integrating Social Practice Theory in Agent-Based Models
03 Seyyed Hadi Hashemi (UVA), Modeling Users Interacting with Smart Devices
04 Ioana Jivet (OU), The Dashboard That Loved Me: Designing adaptive learning analytics for self-regulated learning
05 Davide Dell'Anna (UU), Data-Driven Supervision of Autonomous Systems
06 Daniel Davison (UT), "Hey robot, what do you think?" How children learn with a social robot
07 Armel Lefebvre (UU), Research data management for open science
08 Nardie Fanchamps (OU), The Influence of Sense-Reason-Act Programming on Computational Thinking
09 Cristina Zaga (UT), The Design of Robothings. Non-Anthropomorphic and Non-Verbal Robots to Promote Children's Collaboration Through Play
10 Quinten Meertens (UvA), Misclassification Bias in Statistical Learning
11 Anne van Rossum (UL), Nonparametric Bayesian Methods in Robotic Vision
12 Lei Pi (UL), External Knowledge Absorption in Chinese SMEs
13 Bob R. Schadenberg (UT), Robots for Autistic Children: Understanding and Facilitating Predictability for Engagement in Learning
14 Negin Samaeemofrad (UL), Business Incubators: The Impact of Their Support
15 Onat Ege Adali (TU/e), Transformation of Value Propositions into Resource Re-Configurations through the Business Services Paradigm
16 Esam A. H. Ghaleb (UM), BIMODAL EMOTION RECOGNITION FROM AUDIO-VISUAL CUES
17 Dario Dotti (UM), Human Behavior Understanding from motion and bodily cues using deep neural networks
18 Remi Wieten (UU), Bridging the Gap Between Informal Sense-Making Tools and Formal Systems - Facilitating the Construction of Bayesian Networks and Argumentation Frameworks
19 Roberto Verdecchia (VU), Architectural Technical Debt: Identification and Management
20 Masoud Mansoury (TU/e), Understanding and Mitigating Multi-Sided Exposure Bias in Recommender Systems
21 Pedro Thiago Timbó Holanda (CWI), Progressive Indexes
22 Sihang Qiu (TUD), Conversational Crowdsourcing
23 Hugo Manuel Proença (LIACS), Robust rules for prediction and description
24 Kaijie Zhu (TUE), On Efficient Temporal Subgraph Query Processing
25 Eoin Martino Grua (VUA), The Future of E-Health is Mobile: Combining AI and Self-Adaptation to Create Adaptive E-Health Mobile Applications

	26	Benno Kruit (CWI & VUA), Reading the Grid: Extending Knowledge Bases from Human-readable Tables
	27	Jelte van Waterschoot (UT), Personalized and Personal Conversations: Designing Agents Who Want to Connect With You
	28	Christoph Selig (UL), Understanding the Heterogeneity of Corporate Entrepreneurship Programs
2022	1	Judith van Stegeren (UT), Flavor text generation for role-playing video games
	2	Paulo da Costa (TU/e), Data-driven Prognostics and Logistics Optimisation: A Deep Learning Journey
	3	Ali el Hassouni (VUA), A Model A Day Keeps The Doctor Away: Reinforcement Learning For Personalized Healthcare
	4	Ünal Aksu (UU), A Cross-Organizational Process Mining Framework