THE FIT BETWEEN BUSINESS PROCESSES AND PROCESS MINING RELATED ACTIVITIES

A PROCESS MININING SUCCESS MODEL

MARCO JUTTEN
Business Information Technology
Faculty of Electrical Engineering,
Mathematics and Computer Science

July 3, 2015
Enschede

UNIVERSITY OF TWENTE.

Deloitte.

THE FIT BETWEEN BUSINESS PROCESSES AND PROCESS MINING RELATED ACTIVITIES

Enschede, 03-07-2015

AUTHOR

Marco Jutten

Study Program *Business Information Technology*
Faculty of Electrical Engineering, Mathematics and Computer Science

Student No. *0162752*

E-Mail *m.g.jutten@student.utwente.nl*

GRADUATION COMMITTEE

Maria Iacob, PhD

Department *Industrial Engineering and Business Information Systems*

E-Mail *m.e.iacob@utwente.nl*

UNIVERSITY OF TWENTE.

Martin van Sinderen, PhD

Department *Information Systems*

E-Mail *m.j.vansinderen@utwente.nl*

UNIVERSITY OF TWENTE.

Tijn van der Heijden, MSc

Department *Enterprise Architecture*

E-Mail *tvanderheijden@deloitte.nl*

Deloitte.

Edith Boschman, Drs

Department *Enterprise Architecture*

E-Mail *eboschman@deloitte.nl*

Deloitte.

Preface

This master thesis concludes my master study 'Business & IT' at the University of Twente and also signifies the end of my time as a student. Eight years ago I made the decision to go for the small University based on instinct. A decision that brought me five exciting years in Enschede which I enjoyed to the fullest.

The five years of smooth and joyful studying were followed by two rough years which were just as much part of life. The words of Henry Ford kept me on my feet from day to day.

"People who think they can or think they can't are both usually right"

Life is not just a process which is caught in sequential phases and managed by merely planning and contemplating all options. Many people were there to support me who I will not name here with the chance of missing someone out.

The employees of NS helped me get back on track during my internship for which I am really grateful. After the internship the graduation project was the only thing left on my study path. A project of which I was wary from the start. At Deloitte, Tijn van der Heijden and Edith Boschman created a perfect environment to start my graduation with. During the bumpy start and the acceleration towards the end they were always there for support and guidance, which deserves a special notion. The same accounts for Maria Iacob and Marten van Sinderen who showed flexibility in the time needed.

I have experienced graduating to be a process which is more or less the same for every student and challenges everybody in a similar way but there is no single roadmap which leads to guaranteed success. It is rather a process of best practices which one accumulates during the graduation period. Therefore a special thanks goes out to the fellow graduate interns and freshly graduated employees at the EA service line of Deloitte. I was able to take advantage from their best practices and the given advices.

The research really prospered from the interviews with experts in the field of Process Mining. Not only were they willing to invest their time but also share their interesting view on the topic. The discussions show that theory and practice are rarely perfectly unified but nonetheless depend on each other.

Finally I would like thank my family, my friends and especially my girlfriend Femia. She gives me all the reason to think I can, which makes the question whether I can or can't irrelevant.

Marco Jutten

Management summary

Process Mining is a collection of techniques to analyze the information stored in event data produced by information systems. Where traditional process models are a static abstraction of reality and built rather on opinions and limited observations than real life event data, they tend to miss operational details. On the other hand data analytics on Business Processes are very dependent on delivering preformatted reports with a defined set of Key Performance Indicators. Process Mining is able to produce an objective and dynamic view of the Business Process from different angles.

While the relatively young research discipline Process Mining has produced quite some literature on a wide range of techniques and their applicability in case studies, there is only little knowledge of the success factors of Process Mining in organizations. This research contributes to the knowledge of Process Mining Success by introducing a Success Model which describes a different focus for Process Mining related activities for different Business Processes.

The Success Model is based on the well-established Task-Technology Fit model of Zigurs and Buckland. A semi-structured literature review resulted in four categories of Process Mining related activities (Preprocessing, Discovery Analysis, Organizational Mining and Performance Analysis) and three Business Process characteristics (Variety, Analyzability and Interdependence).

By analyzing 11 Process Mining case studies on the Business Process characteristics and the added value of Process Mining related activities, three Fit scenarios are defined (Ad Hoc, Routine and Standardized). The Fit scenarios describe the added value of Process Mining related activities depending on the Business Process characteristics.

The model has been operationalized with indicators from the literature on Process Mining and Business Process characteristics. To validate the model, semi-structured interviews were conducted with 11 experts who applied Process Mining on operational Business Processes. The transcripts of the interviews were coded with the operationalized model.

The results support the model except for the Organizational Mining related activities as shown in Figure 1. These activities were hardly mentioned during the interviews. As a result there is no support to either discard or include Organizational Mining in the model. This might be explained due to the fact that Organizational Mining has less exposure in practice than in literature.

Since the Success Model is able to give insight into the focus of Process Mining related activities based on Business Process characteristics, the model contributes to understanding Process Mining success.

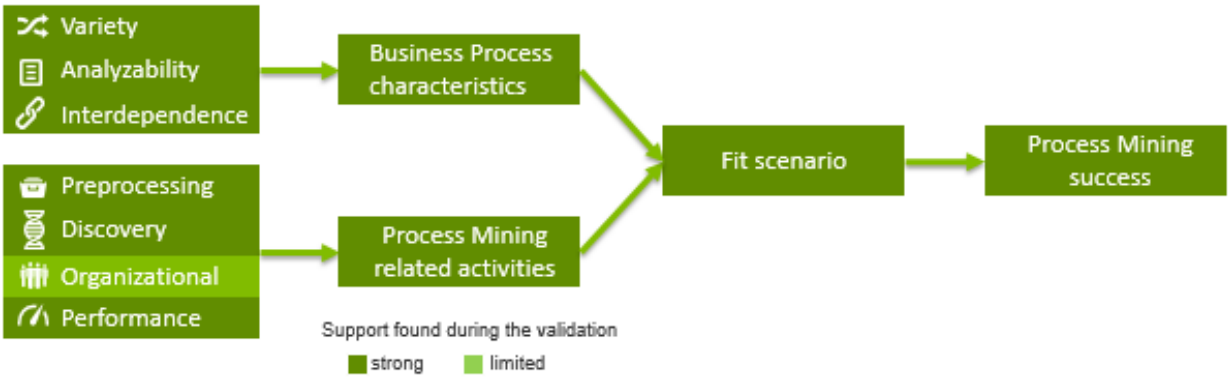

Figure 1 Process Mining Success Model

1 Table of contents

1. Introduction

The first chapter starts with the background and an introduction of the problem. Followed by the research questions and the research design to answer the research questions. The last part of this chapter outlines the structure of the rest of the document.

1.1 Background

Process Mining is a relative young research discipline which combines the practices of process modelling and event data analyses to discover, monitor and improve Business Processes (W.M.P. van der Aalst, 2011). Where current analyses of Business Processes start with inferring models from discussions and brown paper sessions, Process Mining uses event data from systems which support Business Processes like Enterprise Resource Planning (ERP)-systems. With the ability to combine both event data and Business Processes, a Process Miner can drill down in the Business Process to identify the exact source of the problem.

The research area of Process Mining has produced quite some techniques (W.M.P. van der Aalst, 2011) which are being applied on a wide area of different Business Processes varying from the testing of the wafer production process (Rozinat, Mans, Song, & van der Aalst, 2009) to visualizing healthcare processes (R. Mans, Schonenberg, Song, Aalst, & Bakker, 2011). Yet Process Mining is not a widely accepted practice for analyzing Business Processes. Figure 1 shows the L * life-cycle model depicting the several stages which are part of Process Mining (W.M.P. van der Aalst, 2011). A lot of research has gone into creating control-flow models and process models but the research lacks knowledge of success factors which contribute to the adoption of Process Mining.

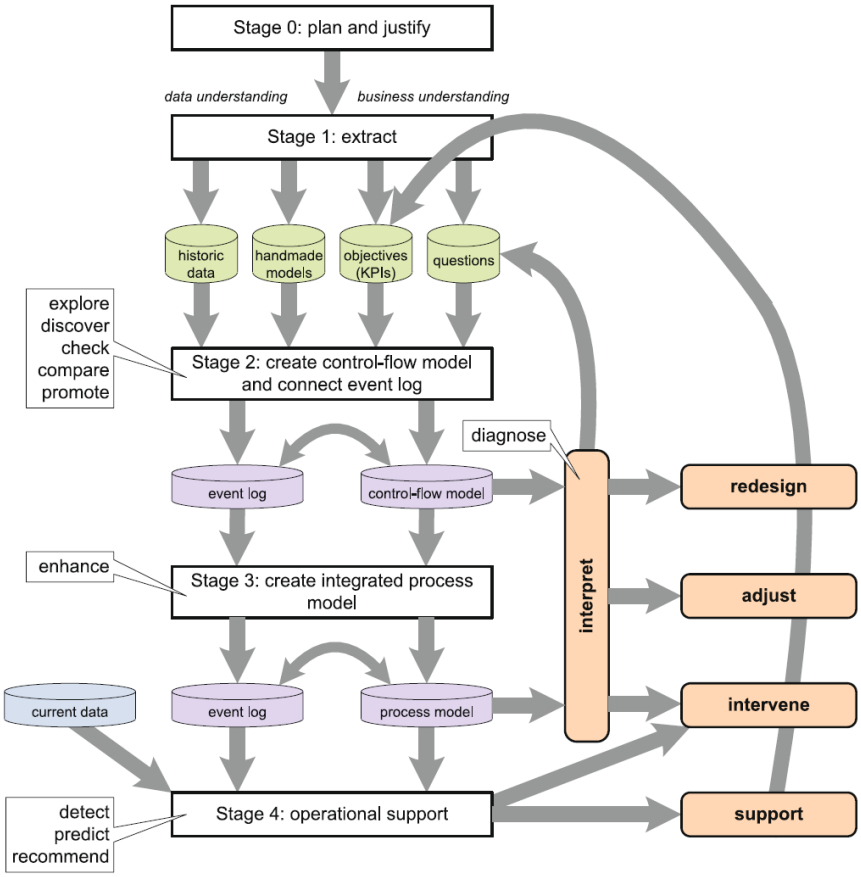

Figure 2 the L * life-cycle model (van der Aalst, 2011)

Mans et al. (2014) also recognize this discrepancy in literature and initiated an explorative research combining success factors from related fields such as data mining, information systems and process modeling. The focus of the research of Mans et al. (2014) was to explain the success of a Process Mining project. Although the research of Mans et al. (2014) contributes to a first overview of factors influencing Process Mining project success, a lot of factors are project and process change related and are applicable to any other project which includes process change. Therefore this research focuses on the analytical capabilities of Process Mining and how they contribute to analyzing Business Processes.

Problem statement

Process Mining is often used as a tool or technique to analyze Business Processes with the purpose of improving or redesigning the processes. The research of Mans et al. (2014) focused on the success factors of the Process Mining projects in a holistic manner which resulted in several high level success factors. To better understand the success factors which are more Process Mining specific this research is scoped on the analytical phase of such an improvement or redesign project.

Van der Aalst (2011) states that the kind of analysis which can be applied on Business Processes depends on the characteristics of a process. When Business Processes have a 'lasagna' structure and stakeholders have a reasonable understanding of the flow of work, all Process Mining techniques can be applied (W.M.P. van der Aalst, 2011). On the other hand when Business Processes have a less clear structure they tend to require more experience, intuition and vague qualitative information. Therefore this research looks into the Fit between the Business Process characteristics and Process Mining techniques.

Solution direction

This research uses the model of Zigurs & Buckland (1998) to explain Process Mining success according to the Fit between Process Mining related activities and Business Process characteristics.

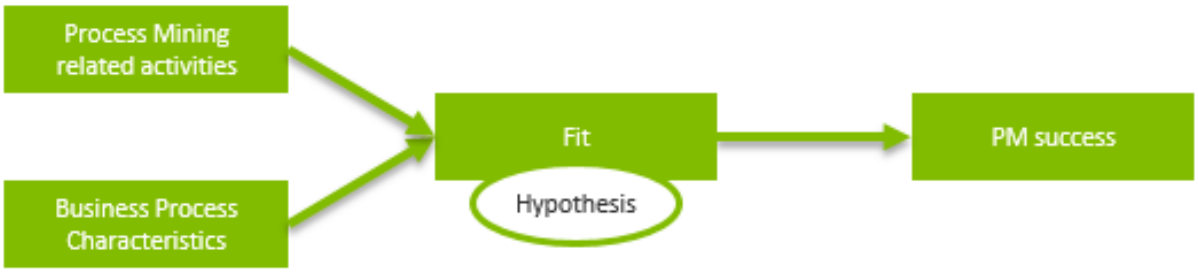

Figure 3 The adapted model and hypothesis based on the Fit model of Zigurs & Buckland (1998)

Figure 2 shows the conceptual model adapted from Zigurs & Buckland (1998). The green blocks represent the constructs of the research model. The hypothesis underpinning this research is that according to different Business Process characteristics, a Process Mining project should focus on different Process Mining related activities to reach Process Mining success.

1.2 Research questions

The main question to address the problem and research the hypothesis is:

"How can the Process Mining success be explained with the Business Process-Process Mining related activities Fit?"

The main question is divided into four sub questions (SQs) which together answer the main question. An overview of the research questions related to the research model is given in Figure 4.

"SQ 1: Which Process Mining related activities are used in practice?"

Although much research has been done into producing new tools, methods and algorithms to add extra functionality to Process Mining tools, not all functionalities are commonly being applied (Ronny S. Mans et al., 2014). To be able grasp the added value of using the Process Mining related activities this research only focuses on related activities which are applied and reported in case study research.

"SQ 2: "What are Business Process characteristics which are relevant for applying Process Mining?"

BPM research has reported many Business Process characteristics which influence the BPM success. Since Process Mining has an overlap with BPM (Goedertier, De Weerdt, Martens, Vanthienen, & Baesens, 2011), this research area will be the focus to find Business Process characteristics.

"SQ 3: How can Process Mining success be measured?"

The hypothesis is that according to the Business Process characteristics there is a specific Fit with Process Mining related activities. To express this Fit it is important to measure the Fit and therefore measure the Process Mining success.

"SQ 4: Which Process Mining related activities-Business Process characteristics Fit leads to Process Mining success?"

The last sub question combines the Business Process characteristics and the Process Mining related activities in different Fit scenarios and expresses how a specific Fit will lead to Process Mining success.

Figure 4 The sub questions (SQs) of this research

1.3 Research design

The nature of this research is explanatory, since an effort is done to explain how the Fit between Process Mining Related activities and Business Process characteristics contribute to Process Mining success. Due to the limited amount of practical applications of Process Mining, this research takes a qualitative approach. Figure 5 shows the steps which guide this research and the relevant research methodology per step.

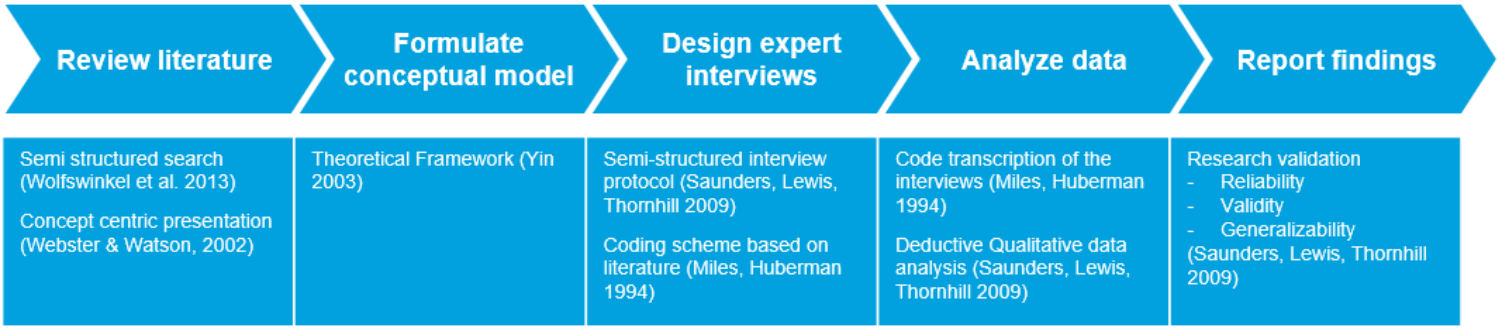

Figure 5 research approach and corresponding research methodology

The research starts with a semi-structured literature review to accumulate the knowledge in literature of Process Mining related activities (SQ 1), Business Process characteristics (SQ2) and Process Mining success (SQ3). Since literature only reports parts of how Process Mining related activities applied on Business Processes leads to Process Mining success, the case studies will be accumulated to specify the Fit (SQ4) and to formulate a conceptual model.

To validate the conceptual model it be subjected to expert interviews. In order to rigorously test the conceptual model the interviews are semi-structured and therefore conducted with a protocol. During the data collection the conceptual model and the interview questions cannot be altered, otherwise the interview questions might be subject to bias (Saunders, Lewis, & Thornhill, 2009).

To be able to analyze the data from the interviews and explain the conceptual model, a coding scheme will be made based on literature (Miles & Huberman, 1994). If the coding of interview transcriptions supports the conceptual model, it represents an explanation for the Process Mining success (Saunders et al., 2009). If there is a mismatch an alternative explanation has to be found (Saunders et al., 2009).

The last step is to report the findings. To rigorously conduct the interviews and qualitative analysis, a research has to guarantee the reliability, validity and generalizability (Saunders et al., 2009).
Since interviews are rarely repeatable, because they are part of a specific context (time and situation), they impose a threat on the reliability of the research. By structuring the interview with a predefined protocol and transcribing the conversation the reliability of the research is maintained.

The validity represents the extent to which the measurement model measures the concepts of the research. Due to a different point of view, questions might be interpreted differently by the interviewer and the interviewee. To establish the validity of this research the analytical framework (coding scheme) for analyzing the data is based on the literature review. The coding shows a clear relation between the answers given in the interviews and the concepts mentioned in the literature. This chain of evidence enables readers to trace back conclusions to the actual research questions (Saunders et al., 2009).

To establish the generalizability it is important to clearly describe the context of the situation and define the boundaries which apply to the model, so other researcher can check whether the model is applicable to their situation.

1.4 Document structure

This thesis is structured as follows: Chapter 2 presents the literature review with theoretical background for the sub research questions. Chapter 3 describes the operationalization of the conceptual model and the Fit scenarios. Chapter 4 shows how the expert interviews are designed, how the data is collected and the data is analyzed. Chapter 5 elaborates on the coding of the interviews and the results of testing the model in practice. The final Chapter presents conclusions and implications for both practice and theory.

2 Literature review

This chapter describes the approach on how the relevant literature was selected and synthesized, followed by a summary of the literature on the context of Process Mining, Process Mining related activities, relevant characteristics of Business Processes and finally Process Mining success.

2.1 Approach

This research uses the aspects of a case study approach and utilizes the methodology by Yin (2009). Yin recommends to start a case study with a literature study to develop a conceptual model for the case study. The difference between a case study and expert interviews is that the latter uses only the view of one person on a case. A case study research typically uses several interviews with different people on the same case. The advantage of using expert interviews is that it allows the researcher to gather more data and look at the differences between interviews which better suits this research.

The conceptual modal is used to take a deductive approach in this research. This provides several advantages, as it ties the research into the existing body of knowledge, helps research get started and directs the analysis of the collected data (Saunders et al., 2009). The literature is searched with the approach described by Wolfswinkel et al. (2011). The approach consists of the steps Define, Search, Select, Analyze and Present. The last step Present is structured based on the concept-centric presentation of (Webster & Watson, 2002).

Define, search and select

Since the research area of Process Mining is relatively young, there is only one article writing on the Process Mining success (Ronny S. Mans et al., 2014). Therefore first the literature searched for more general information system success models which are further discussed in section 0.

Further an initial search was done on Process Mining and related research areas:

- Process Mining and success
- Business Process Management and success
- Process Modeling and success
- Business Process Analysis and success

The literature was searched on Google scholar and Scopus, was from 2005 and later, was with the first 30 results sorted on relevance and citation count.

Analyze and Present

The Process Mining related literature was selected on the description of a comprehensive application of Process Mining and/or important success factors/measures with Process Mining in an organizational context. This resulted into a list of 11 extensive case studies which can be found in Appendix A.

The related research areas only delivered either high level factors or very detailed factors which the author deemed not interesting enough for composing a Process Mining success model. Therefore the Process Mining case studies were open coded in the search for Business Process characteristics, Process Mining related activities and Process Mining success measures.

Based on the results for Business Process characteristics additional literature was searched for based on the keywords:

- Business Process complexity
- Business Process standardization

2.2 Context

This section introduces the related research fields and definitions used during this research.

2.2.1 Business Processes

Work systems started with supporting relative easy work in parts of an organization and caught increasing interest because of the abilities to perform more work with higher quality. In the seventies and eighties work systems had the sole purpose of storing, retrieving and presenting information (De Weerdt, Schupp, Vanderloock, & Baesens, 2013). With the increasing adoption of information technology in organizations, information systems began to play a more important role in supporting the organization and its systems. Information systems than became a typical kind of work systems that uses information technology to capture, transmit store, retrieve, manipulate or display information (Ronny S. Mans et al., 2014).

With the increase of the functionality of Information Systems they also became more complex and harder to optimize. Since Business Processes became the fundamental unit of analysis, Business Process Redesign was advocated in the beginning of the 1990s to radically redesign Business Processes with the power of information technology (Ronny S. Mans et al., 2014). With this management technique organizations tried to understand inefficiencies and how routines were actually executed (De Weerdt et al., 2013). Since Business Process Redesign not only involves the technical challenge to redesign the Information System but also has a large socio-cultural challenge, it is not a trivial effort (Reijers & Liman Mansar, 2005).

The challenges in both technical research and socio-cultural research have led to a large body of research into Business Processes. The large amount of definitions given to Business Processes make it hard to distinguish the clear concept and it is therefore not possible to give a single definition which includes all aspects of Business Processes (Vergidis, Tiwari, & Majeed, 2008). Bandara, Gable, & Rosemann (2005) argue that looking at an organization as a compilation of Business Processes, is a way to deconstruct organizational complexity. This view implies that by building an organization on Business Processes an organization is able to simplify itself. A more formal and descriptive definition is given by Mathias Weske (2007) who defines Business Processes as:

> ”a set of activities that are performed in coordination in an organizational and technical environment. These activities jointly realize a business goal. Each business process is enacted by a single organization, but it may interact with business processes performed by other organizations.”

This definition recognizes that in order to realize business goals, a set of activities have to be performed in a coordinated matter, which are influenced by both the organizational and technical environment.

2.2.2 Business Process Management

As Business Processes can be seen as the core of an organization it is important to be able to manage Business Processes in the sense of quality, costs and time.

A lot of research is now accumulated under the term Business Process Management which include concepts, methods and techniques to support the design, administration, configuration, enactment and analysis of Business Processes (Weske, 2007).

Due to the ever changing environment of organizations at an increasing pace, organizations are constantly occupied with adapting their Business Processes. A well-known way to keep adapting Business Processes is called the Business Process Management Lifecycle as depicted in Figure 6.

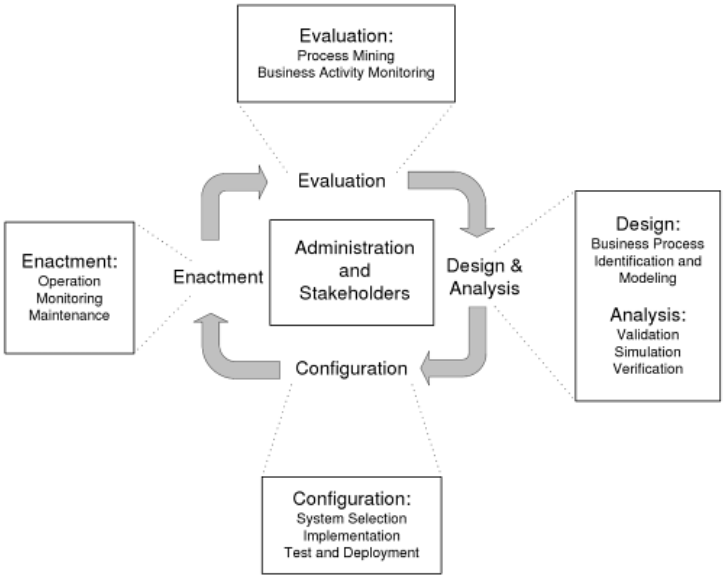

Figure 6 Business Process Management Lifecycle (Weske 2007)

This lifecycle is one of many versions but in general they all acknowledge the phases and the cyclic nature of Business Processes. Also Weske (2007) states that the dependencies in Figure 6 do not imply a strict temporal order.

The lifecycle is generally entered at the Design and Analysis phase in which surveys and workshops are organized to design business process models which are an abstraction of the reality so different stakeholders can communicate efficiently. It is important during the phase to verify whether the formalized description in the model reflects the desired and real behavior of a Business Process (Weske, 2007).

The following phase includes implementing the model which does not necessarily have to be supported by a Business Process Management System. The systems which support the Business Processes need to be configured according to the organizational environment which is an important step as many organizations deal with legacy systems across different functional departments.

The enactment phase is concerned with the real time execution of Business Processes. Systems which manages the Business Process actively control the execution of Business Process instances as defined in the Business Process Model. During the enactment of Business Processes valuable execution data is gathered in the form of a log file.

Once Business Processes are up and running the evaluation phase is concerned with searching for opportunities to improve the business process models and their implementations. The logs gathered from the systems that enact the Business Process are valuable in the sense that they contain information about the quality of the business process models and adequacy of the execution environment.

2.2.3 Process Mining

Process Mining comprises a collection of techniques to analyze the information stored in event logs, where the analysis focuses on the discovery, monitoring and improvement of processes (De Weerdt et al., 2013). It is important to notice that the definition of Process Mining is broader than just the application of algorithms on event data. While the mining algorithms are probably the techniques where Process Mining is known from, the research has accumulated other techniques such as clustering event data and Organizational Mining. Yet the goal of Process Mining remains to improve operational processes (W.M.P. van der Aalst, 2011)**.**

Assumptions

It is important to notice that Process Mining is relying heavily on the data quality. A Business Process that does not record any business steps is not analyzable with Process Mining techniques for obvious reasons. Therefore an event log must at least contain a case (a distinct resource going to the Business Process, an event which is related to one case (an action performed on the resource), a timestamp is connected to the event to order the events chronologically (the moment the an event started) (W.M.P. van der Aalst, 2011). Often an event log also registers a resource or attribute such as the activity specifics, costs or the actor performing the event.

Goedertier et. al (2011) mention four important assumption which often remain implicit in research:

- There is a one-to-one mapping between a system event and a business event
- It is possible to identify meaningful process instances in an event log
- The events in the log are generated by exactly one underlying process
- The processes take place in a structured fashion

Limitations

Even if all assumptions of Process Mining are correct it can be very challenging to get an understandable business process model out of event data. Many challenges are related to the quality of the event log (Goedertier et al., 2011; W.M.P. van der Aalst, 2011).

Incomplete logs: Data produced by the Business Process can be scattered and logged in different systems. Therefore producing an event log which includes the right level of abstraction and can be merged into a meaningful collection of knowledge is often a challenge. Also mining algorithms are very dependent on finding sequential activities and therefore the timestamps with events are of the uttermost importance. Although many information systems do record timestamps it might occur that the level of detail is too low and all events occurred on one day. Than a mining algorithm does not have the ability to find the correct order of the events.

Noise: Human-centric processes are prone to exceptions and logging errors. Process models which include this behavior become overwhelming and do not represent the frequent behavior of the Business Process.

Unsupervised learning: Event logs do not contain situations which did not happen but are possible according to the process model. Therefore it does not show all possibilities of the structure of a Business Process.

Scoping: Although the assumption of Process Mining is often made that one event log contains events which are only related to one Process in practice it is hard to distinguish the clear border of the process. Especially cross functional boundary processes in ERP systems gather data throughout the organization. It is important to only select data which increases insight into the Business Process.

- *Representational bias:* Modelling languages vary from free format (flowchart) to highly structured and strict languages (BPMN). Most Process Mining settings use a procedural language to describe end-to-end processes which are less subject to interpretation but can fail to capture the rich human behavior.

Advantages

The earlier mentioned limitations and assumptions do not withhold Process Mining from delivering important insights into Business Processes. Because Process Models are an abstraction of reality and build rather on opinions and limited observations than real life data, they tend to miss operational details (De Weerdt et al., 2013). On the other hand data analytics on Business Processes are very dependent on delivering static preformatted reports and thus are very dependent on a right set of Key Performance Indicators (W.M.P. van der Aalst, 2011).

Major advantages of Process Mining are that it is able to produce an objective view of the Business Process being executed in a relatively short time and answer specific questions about the Process. van der Aalst et al. (2007) mined the data from a Workflow Management System (WfMS) of an invoicing process. They were able to produce the underlying Process Model, give insight to the variation in the Process and quantify the results of the variation. Mans et. al. (2009) were able to structure the billing Process of hospital, give insight into how their Business Process was operating and what actors are involved in the Process. Although their findings correlate with the flowchart which was present, the automatically generated model required less effort.

Tools

Several tools are available which combine event log data and business process models to visualize the knowledge in data. This research does not focus on delineating all tools and their functionality but rather looks at the related activities which are being applied in practice. Therefore this research tries to be tool independent. Still it is necessary to distinguish what is assumed to be a Process Mining tool. The author defines a Process Mining tool as “having the capabilities to deduce a Business Process Model from any event log which contains cases, events and timestamps and visualize the result”.

Several tools which are known for these capabilities are ProM, Disco and Aris PPM (W.M.P. van der Aalst, 2011). Disco is a proprietary tool by Fluxicon which is able to handle large event logs with an algorithm based on fuzzy mining and allows for seamless abstraction and generalization based on the cartography metaphor (W.M.P. van der Aalst, 2011). Aris PPM is another proprietary tool which is able to extract knowledge from event data and produce it into performance information. Aris also provides the ability to mine a social network and show the connectedness of employees (Wil M. P. van der Aalst, 2009).

The last tool ProM is very popular in research because it is open source and supports the ability to build plugins to add extra functionality. The academic nature of ProM causes it to be less user-friendly than Disco and Aris PPM but the open source approach resulted in a plethora of functionality and therefore unprecedented (W.M.P. van der Aalst, 2011).

Since ProM is very popular in research it is important to notice that most of the literature, on which this research is based, used ProM functionality.

2.3 Task-Technology Fit

Several information system success models have been developed over the past few decades. The variant which is deemed best Fit by the author is based on Goodhue & Thompson (1995). Their research states that a good Fit between the information system and the task at hand leads to better performance of the individual. This situation is quite comparable to a situation where a process miner has the choice between several Process Mining related activities at hand to analyze a Business Process with the goal to enhance the analytical capabilities of the individual. Therefore this research focuses on finding the Fit between Process Mining related activities and Business Process characteristics to explain Process Mining success.

The three most cited Task-Technology Fit models are from Goodhue & Thompson (1995), Dishaw & Strong (1999) and Zigurs & Buckland (1998).

The initial model of Goodhue & Thompson (1995) contains five concepts as shown in Figure 7. The characteristic of the Task and Technology are correlated with the Fit which is correlated with both the Utilization and the Performance impacts. The model is empirically tested with strong support and therefore considered the basis of Task-Technology Fit.

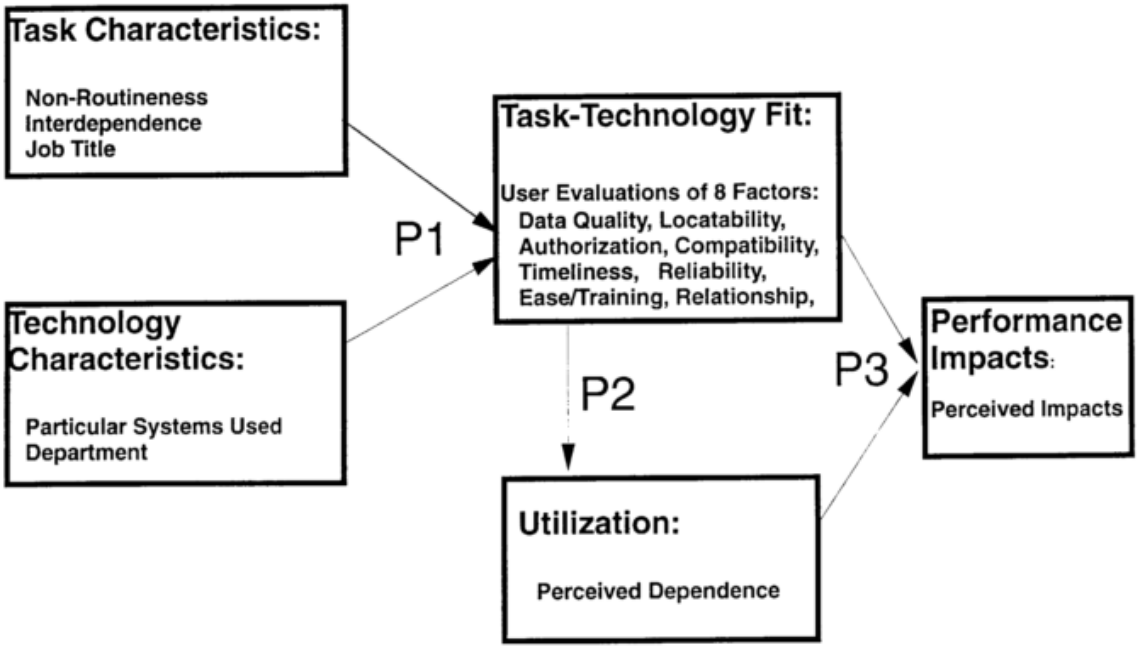

Figure 7 Task-Technology Fit model - Goodhue & Thompson

Dishaw & Strong (1999) combined the Technology Acceptance Model and the Task-Technology Fit model and were to have a bigger explanatory power (Figure 8). But combining both models makes it also inherently more complex to understand. Further the Technology Acceptance part of the model adds many soft factors related to the Use of the tool. In the case that employees are forced to use a specific information these concepts are relevant. But in the case of Process Mining most people use it on an explorative basis to understand whether the tool is useful for their work. Therefore the model is considered too complex for this research to measure the Fit between Business Processes and Process Mining related activities.

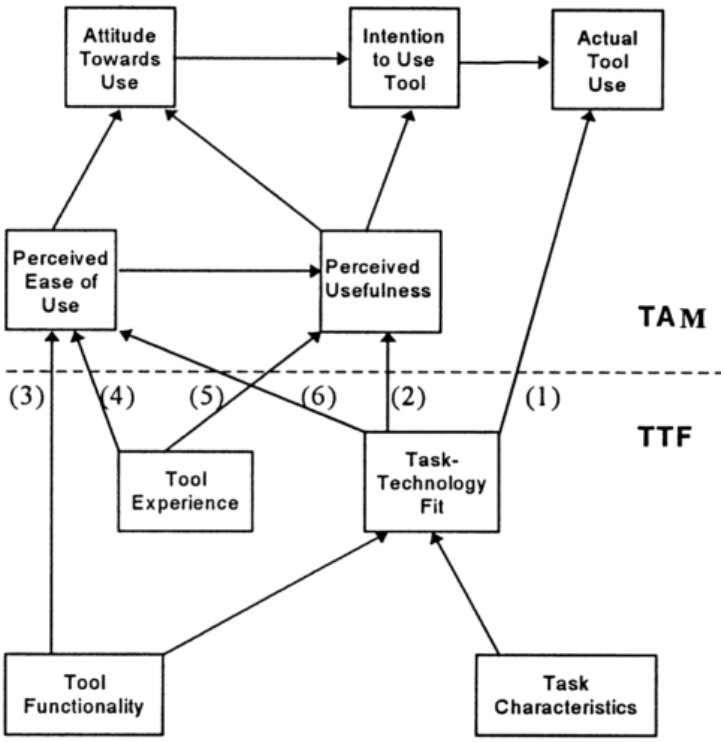

Figure 8 Task-Technology Fit and Technology Acceptance Model - Dishaw & Strong

The Task-Technology Fit model was used by Zigurs & Buckland (1998) to classify the tasks which are performed in groups and the support which was given by Group Decision Support Systems (GDSS). After this classification the amount of tasks possible were too high and were reduced to five common tasks performed by GDSS; simple, problem, decision, judgement and fuzzy task.

Based on the classifications Zigurs & Buckland (1998) made a Fit profile of a task and GDSS functionality category for which results in the best group performance.

This matches with the goal of this research to find the Fit between Business Processes and Process Mining related activities. Also the model is based on task complexity which shows resemblance with Business Process complexity.

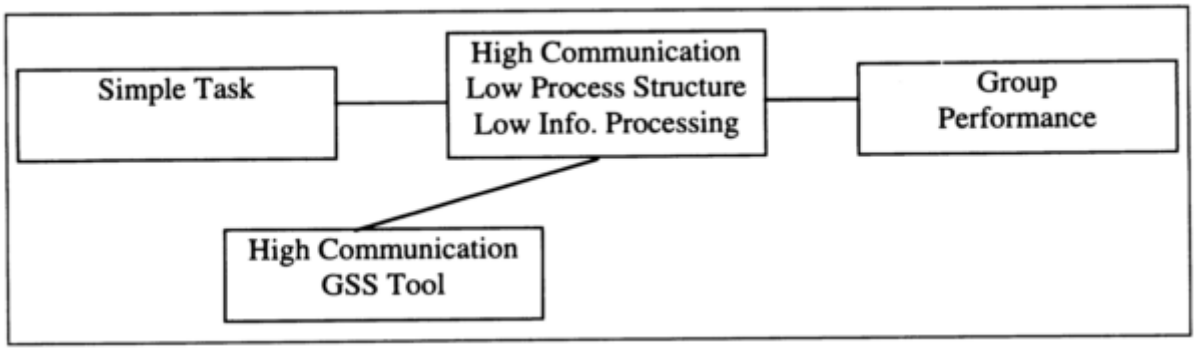

Figure 9 Task-Technology Fit - Zigurs & Buckland

2.4 Process Mining related activities

This chapter starts with defining a classification of Process Mining related activities according to the steps in Process Mining methodologies. The remainder of the chapter will elaborate on the Process Mining related activities associated with the classification.

Process Mining methodologies

Many techniques from Process Modelling and Data Mining have been applied in a Process Mining context and thus became part of the Process Mining research area (Ronny S. Mans et al., 2014). To be able to make a distinction in Process Mining related activities the author chose to look at the methodologies which have been developed for Process Mining. Several Process Mining methodologies exist in literature which are all building on each other and are applied in different contexts. This research focuses on the methodologies which have been applied in practice, describe the context of the application of Process Mining techniques and clearly report the findings of the Process Mining project.

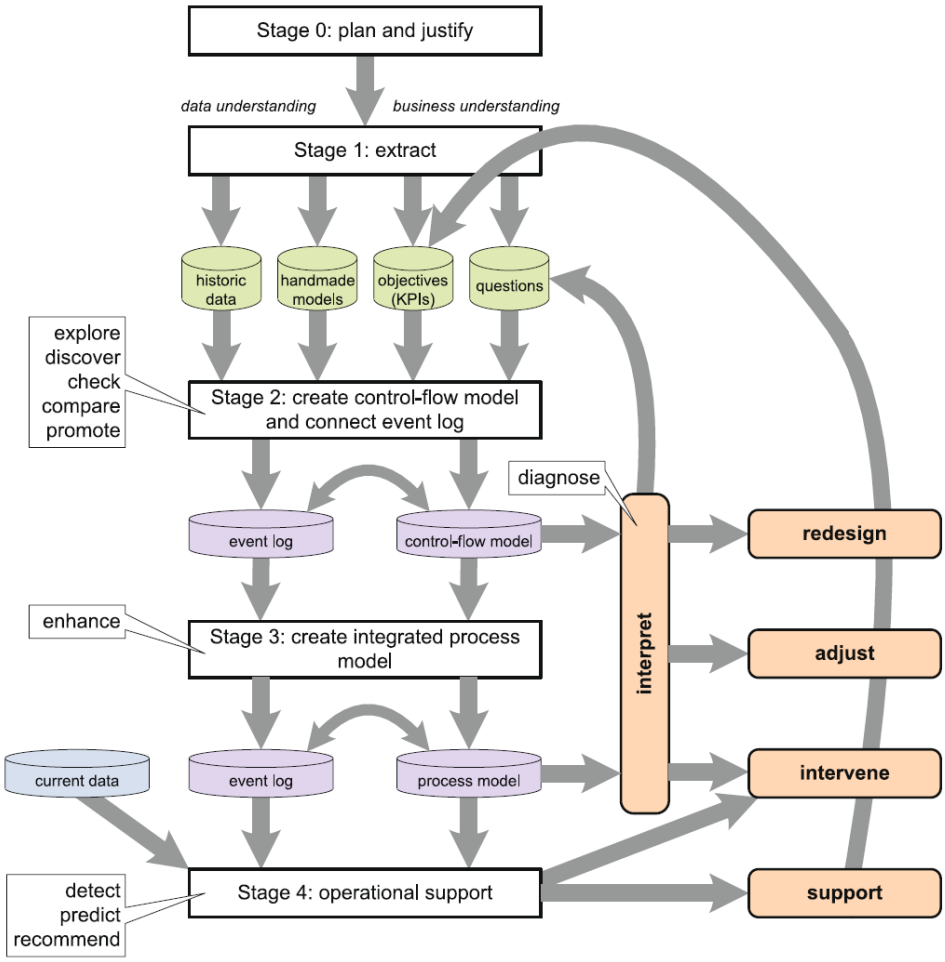

Figure 10 Process Mining L * cycle (Aalst 2011)

The L * cycle model (Figure 10) combines ten Process Mining techniques described by van der Aalst in five stages (W.M.P. van der Aalst, 2011). The model briefly describes the stages and how the ten techniques are separated over the stages. The first two stages (Plan and justify, Extract) are concerned with scoping the project and extracting the data from information systems. While it is only briefly

described, many applications of Process Mining require a great deal of effort to extract and combine data from information systems (Bose, Mans, & Van Der Aalst, 2013). Stage 2 consists out of halve of the techniques described by van der Aalst which is not surprising since Process Mining is widely known for the ability to construct a Process Model from an event log. The third stage is concerned with enhancing the discovered model with data other than regular Process Models contain. Often this is data like who initiated events and what resources were used or produced. The last stage introduces the ability to support operational processes while they are being enacted. The right part of the model shows the diagnostic technique and several steps (Redesign, adjust, intervene and support) in which Process Mining can deliver extra value.

Although the model is complete in the sense that it summarizes all Process Mining techniques, it lacks detail and how it is applied in practice. The model seems to be based mostly on the experience of the author and lacks the description of practical applications. Especially stage four 'operational support' is hardly described in literature and therefore seems to be mostly future work. Therefore this methodology is deemed not useful for this research.

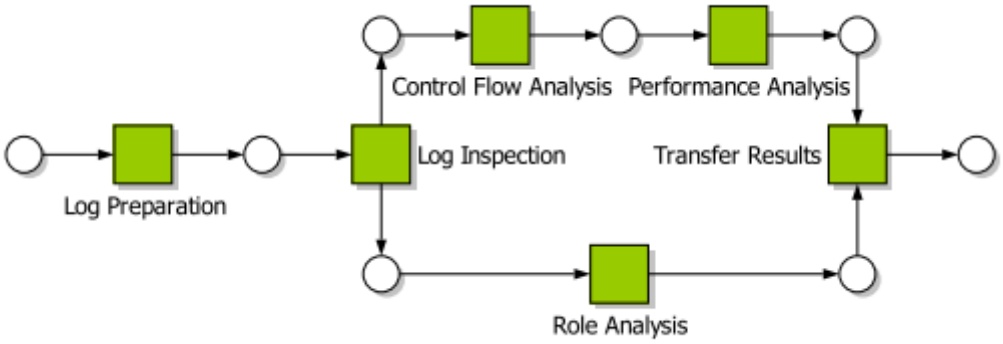

Figure 11 Business Process Diagnostics (Bozkaya et. al. 2009)

The Business Process Diagnostic methodology for Process Mining is designed to give quick results (Bozkaya, Gabriels, & Werf, 2009). The author of the model emphasizes the importance of delivering quick and understandable results to show the value of Process Mining. The method is applied in a case study format on a document managements system at a Dutch government. The goal of the case study was to gain insights into the document issuing process and how it could be further optimized. In just 50 man hours and without any prior knowledge of the industry the Process Miner was able to get results which did impress the stakeholders.

The author does not specify whether the model is only applicable to specific Business Processes which suggest that it is applicable to all Processes. Because of the clear description of the steps and the results delivered, the model is useful for this research and described in more detail in the following sections.

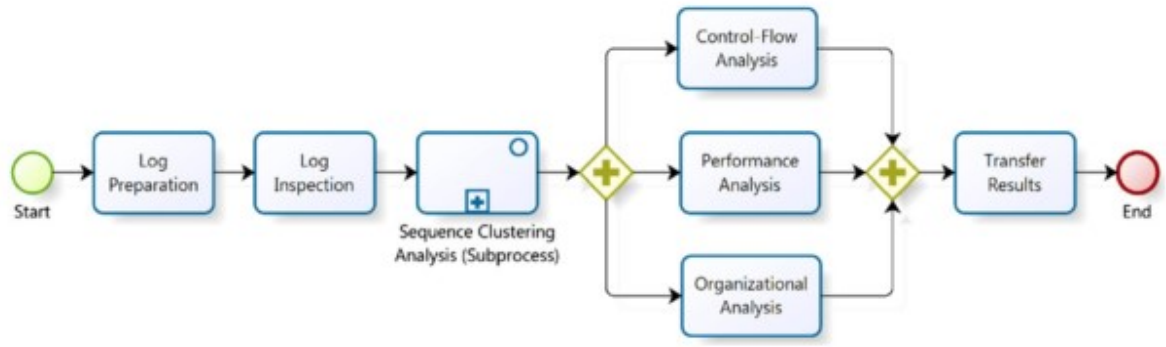

Figure 12 Process Mining for healthcare (Rebugé 2012)

The Process Diagnostic methodology of Bozkaya et al. (2009) was adapted by Rebuge & Ferreira (2012) specific for a healthcare environment (Figure 12). The adaptations were done to make the model more resilient for the complex and ad hoc nature of medical processes. The sequence clustering analyses step is introduced before the other analysis step which focuses on reducing the deviations in the event log. Next to the extra step it also good to notice that the Process Diagnostic methodology sees discovering the control-flow as a necessary step before being able to do a performance analysis while the healthcare focused methodology recognizes that these steps occur concurrently. The methodology is applied on an emergency care process in a Portuguese hospital which is support by a centralized hospital information system. A special Process Mining tool is built to extract the data from the system, apply Process Mining techniques and give insight into process improvement opportunities.

Because of the complex nature of the Business Processes in which this methodology was applied and the extensive description of the results of techniques used, this methodology is used in the next sections to describe the Process Mining related activities.

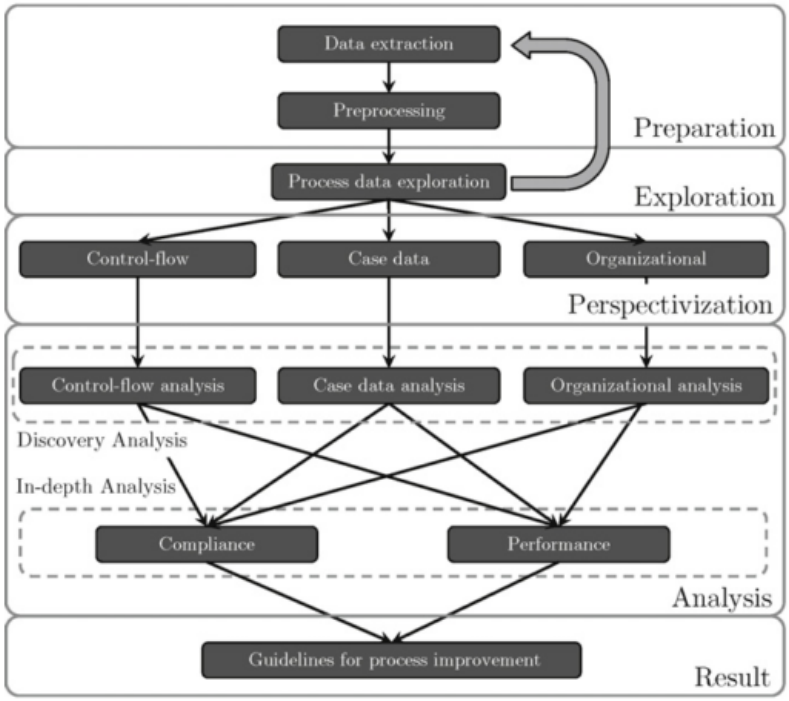

Figure 13 Process Mining Methodology Framework (Weerdt et. al. 2012)

The Process Mining Methodology Framework (Figure 13) is in line with both the Process Diagnostic and the healthcare specific methodology (De Weerdt et al., 2013). It also designed to be able to handle complex Business Processes which, according to the author, are often seen in service industries opposed to production industries. Like the other methodologies it acknowledges the importance of preprocessing but emphasizes the different perspectives and the difference between discovery and in-depth analysis. The methodology is tested in a case study at a large Belgian insurance company to improve the document management process supported by a document management system. Although the stakeholders of the process had several statistics on a regular basis, they did lack real knowledge on how the Business Process was executed in real life.

Using different perspectives is also reported valuable by van der Aalst but it is often seen as part of a Process Mining tool. In this methodology the author suggest a different approach by separating event logs according to the different purpose they serve to handle complex processes. This approach is extensively described and gives a different view on how to use Process Mining and is therefore used in this research.

Categorizing Process Mining related activities according to methodologies

The methodologies describe the steps needed to deliver Process Mining results. Since they are adapted for specific situation they also indicate how Process Mining related activities can be categorized. Table 1 shows an overview of the three Process Mining methodologies described in this section and their steps (the methodology by van der Aalst was not included because of a lack of practical reports). The categorization is based on the main differences in the methodologies.

Process Mining related activities	Preprocessing	Process Discovery	Organizational mining	Performance analysis
(Bozkaya et al., 2009)	- Log preparation - Log inspection	- Control flow	- Role analysis	- Performance - transfer results
(Rebuge & Ferreira, 2012)	- Log preparation - Log inspection - Sequence clustering	- Control flow	- Organizational	- Performance - transfer results
(De Weerdt et al., 2013)	- Preparation - Exploration - Perspectivization	- Discovery analysis - Case data - Control flow	- Discovery analysis - Organizational	- In-depth analysis - Performance - Compliance

Table 1 Process Mining related activities categorized according to the methodologies

The *preprocessing* related activities to create a Process Mining ready event log are recognized by all three methodologies and take an iterative approach. The difference between the methodologies are ranging from the amount of effort which is needed to make the event log Process Mining prove.

Process *discovery* is a main part of the methodologies concerned with discovering the Process Model based on event data. Where one methodology is able to discover a Process with limited effort, others require to adjust the scope of the project and re-enter the preprocessing phase. The *organizational* analysis is seen as a technique which can mine an event log without needing a Process Model.

The *Performance* analysis focuses on quantifying the found differences and analyzing the throughput time. While one methodology focuses on answering very specific questions the other describes it briefly and already produces interesting findings by just producing a Process Model.

Conformance is a well-known Process Mining technique to compare the discovered model to existing models. Since the methodologies do not describe this as a separate step in the methodology but all

mention the technique to be useful, it is deemed to be an integral part of Process Mining which is useful in all Process Mining projects. Conformance will be elaborated on in the sections 'Process Discovery' and 'Performance analysis'.

The next sections will describe the Process Mining related activities according to the classification in Table 1.

2.4.1 Preprocessing an event log

Before the process model can be produced an event log has to be produced. Rarely this is a trivial process. Consequently the author decided to separate the functionality of Process Discovery related to building an event log and to deducing the Process Model based on a given event log. In practice the functionality to preprocess an event log and to discover a Process Model is used in an iterative way. The reason to separate these types is because some applications report extensive effort into building an event log (R.S. Mans et al., 2009) and some applications require no effort at all (Măruşter & Beest, 2009) which implicates contextual influences on Process Mining effort.

Preprocessing raises questions such as 'what is the specific case we are analyzing', 'what are the activities and events we take into account' and 'how do we find the correct timestamp for a Process' (Bozkaya et al., 2009). A first glance of the statistics of the event log provides the miner with an impression on the number of cases, number of events, distribution of number of cases per number of events and the number of different sequences (Rebuge & Ferreira, 2012).

The building of an event log is mostly about gathering all available data, in the right level of detail, in a format which can be processed by the tool (De Weerdt et al., 2013). Most of the work like gathering the data and combining it into one log requires data specialists and process specialists. Once a log is combined often data needs to be omitted to get the right level of detail. Functionality which Process Mining tools provide are clustering and filtering.

Filtering is used to remove cases which are not finished or are logged in a wrong manner. Bozkaya et al. (2009) did filter the log to remove cases which are irrelevant and incomplete and do not add any value to be included in the Process Model. It is important to notice that this has to be done in accordance with a data and process specialist to interpret the meaning of an event and the consequences of excluding it.

Clustering can be based on several metrics but often boils down to separating cases based on the frequency of occurrence to create event logs which are more homogeneous (De Weerdt et al., 2013). This is done with the hospital case because the large amount of variety in the patients being treated causes for an unreadable log (R.S. Mans et al., 2009). The same clustering is also used to separate the emergency care flow into seven homogenous groups (Rebuge & Ferreira, 2012).

An example of an event log which does not require any preprocessing is that of the government fine collecting case were the information was produced by a workflow management system (Măruşter & Beest, 2009).

2.4.2 Process discovery

The Process Mining related activities of Process Discovery focuses on constructing a Process Model solely on events in system logs (W.M.P. van der Aalst, 2011) to reproduce the observed behavior (Rozinat, de Jong, Günther, & van der Aalst, 2009). The model gives insight into the complexity of the Process and many case studies report their first model to be unreadable because of the 'Spaghetti' structure. The goal of Process Discovery is to find a model that correctly summarizes the behavior in the event log, striking the right balance between generality (allowing enough behavior) and specificity

(not allowing too much behavior) (Goedertier et al., 2011). This model often shows the first deviations with reality which gives interesting insights.

To get the right generality and specificity two terms are introduced; noise and incompleteness (W.M.P. van der Aalst, 2011). Noise**,** is rare and infrequent behavior rather than logging errors. Although the less data problems a log has the better Process Mining works, it is impossible to get an error free event log. Therefore, once an event log is created and produces a 'spaghetti' model, often more filtering is applied to clean the event log. Incompleteness refers to the situation when the log contains too few events to be able to discover some of the underlying control-flow structure. While noise suffers from too much data, incompleteness is a problem related to little data to generalize behavior (W.M.P. van der Aalst, 2011).

Related activities in Process Discovery can be viewed in three perspectives; the process perspective, the organizational perspective and the case perspective. The process perspective focuses on the control flow (the right ordering of events) (W. M P van der Aalst et al., 2007). Many different algorithms have been produced to deduce the control flow from the event log but this research limits to the most divergent algorithms.

- *α-algorithm,* produces a place/transition net but is unable to mine certain constructs such as loops, duplicate and invisible tasks. It has also limited support for dealing with incompleteness and noise in event logs (W.M.P. van der Aalst, 2011). The algorithm can be seen as one of the first algorithms which is now superseded by other algorithms.

- *Heuristics miner*, constructs a model based on the frequencies of tasks. It is unable to deal with non-free choice and duplicate tasks but it is robust to noise in logs (W.M.P. van der Aalst, 2011). The model produced with the heuristics miner can be easily adjusted by increasing the threshold to show processes or not and therefore has an interactive way to deal with noise. Noise is a very common problem which makes this algorithm very popular in research. An example is given in Figure 14 Heuristic miner Maruster et al (2009).

- The *fuzzy miner* is best suited for mining less structured processes (spaghetti). It is able to abstract from details, although its design causes it to lack support for mining specific splits and joins in a process. The fuzzy miner can also deal well with noise (W. M P Van Der Aalst & Günther, 2007). Like the heuristics miner it is able to aggregate infrequent behavior based on a threshold which makes it intuitive to analyze Business Processes.

- *Genetic algorithm*, the genetic algorithm constructs a process model according to an approach that is similar to the process of evolution in biological systems. It is able to deal with all constructs, apart from duplicate tasks. It is also robust to noisy logs. One drawback is the long computational time required (W.M.P. van der Aalst, 2011). The long computational time and the less intuitive interface make it a less popular tool used in literature.

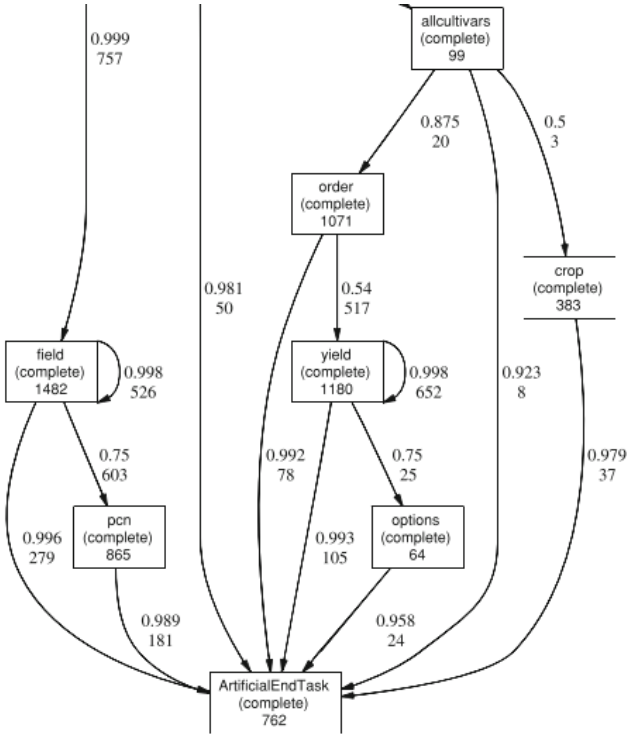

Figure 14 Heuristic miner Maruster et al (2009)

The case perspective focuses on cases which can be characterized by their path in the process or by the originators working on a case (W. M P van der Aalst et al., 2007). Like the organizational perspective, the case perspective does not depend on finding a relevant Business Process Model. The organizational perspective is presented in a different category because this view shows different information while the author deems the case view more as another view on the same kind of information as in a Process Model. Figure 15 Dotted chart, van der Aalst (2011) shows an example of a case perspective where every row represents a case, every dot represents an event and the color of the dot represents the task associated with the activity.

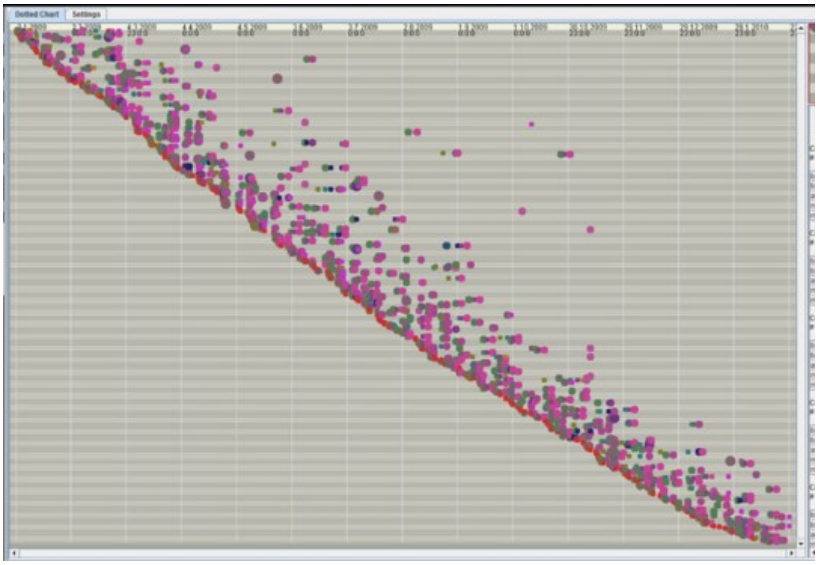

Figure 15 Dotted chart, van der Aalst (2011)

An example of the dotted chart (and case perspective) is the invoice process at a Dutch municipality where it was used to show activities that are executed in batches. When a column shows multiple dots

right under each other, this is an indication of a batch which leads to bottlenecks (W. M P van der Aalst et al., 2007).

It is important to notice that the techniques presented above do overlap in functionality (the case and control flow perspective can both show bottlenecks). The goal of Process Discovery is not to find the ultimate Process Model but rather present different views on reality. Whether a Process Model is suitable or not, ultimately depends on the questions one would like to answer (W.M.P. van der Aalst, 2011).

Conformance Checking is used to check whether the modeled behavior matches the observed behavior (Rozinat, de Jong, et al., 2009). It is important to notice that Conformance Checking requires an a priori model to check if the reality as recorded in the data of the information system, conforms the model and vice versa (W.M.P. van der Aalst, 2011).

Since most of the applications mentioned in research filter or split event logs to decrease the variety of traces in an event log, it is important to know how well the produced Business Model represents the data in the event log. The measure is also used often as a measure to define the quality of the mined log. van der Aalst (2011) calls this measure Fitness which is a number between 0 and 1 to describe the percentage of cases in the event log fit with the mined model. A fitness of 1 would mean that the Process Model can explain all behavior in the event log.

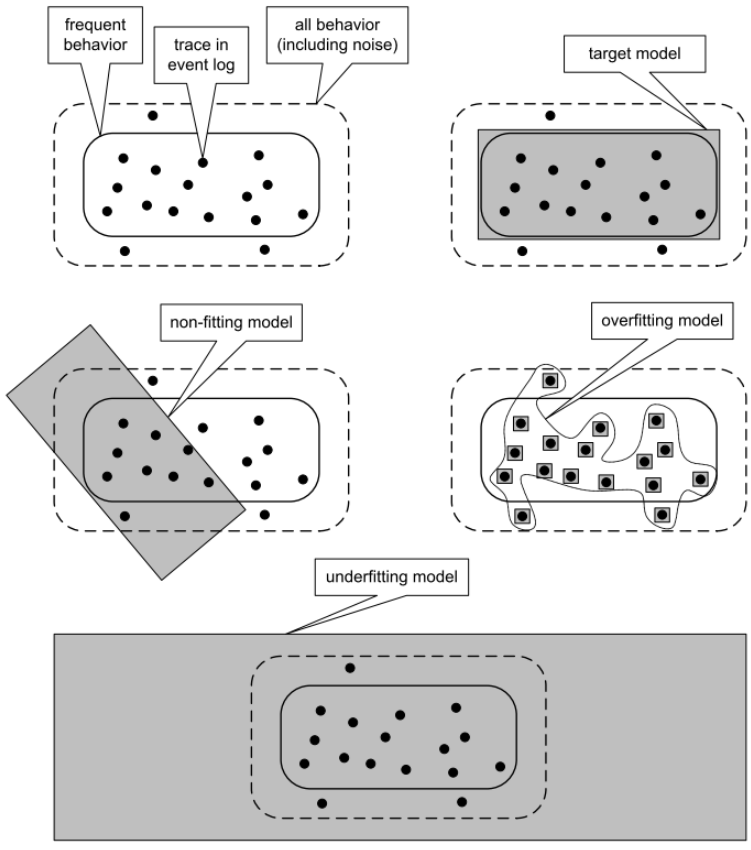

Figure 16 Challenges of process discovery (Aalst 2011)

As explained earlier the filtering and separation of event logs can be done to increase fitness. This is often done to tackle problems depicted in Figure 16 Challenges of process discovery (Aalst 2011). The goal is not always to reach a fitness of 1 because this can make a Process Model unreadable. A Process Model which is close to 0 cannot explain any behavior in the event log and is therefore useless. Although the fitness depends on the situation, a rule of thumb number is 0,8 (W.M.P. van der Aalst,

2011). When the number is below 0.8 this might limit the usefulness of Process Mining techniques which require a Business Process Model. Hence the case and organizational perspectives mentioned earlier do not depend on a comprehensible Process Model.

The heuristics miner mentioned earlier uses a threshold to include traces and depicts the amount of cases that follow a path with a fitness number. The higher the threshold, the more the miner generalizes behavior and the lower the fitness number. This approach is seen often in Process Mining applications because the effect of fitness is immediately visualized in the Process Model (Măruşter & Beest, 2009). When the fitness is considered too low, the Process Model is able to visualize the outlier cases and these can be filtered out of the event log to increase the fitness. Again it is important to do this in dialogue with data and process specialists because it might influence the credibility of the Process Model.

2.4.3 Organizational analysis

The organizational perspective focuses on the originator field. It analyses the event log and shows which performers are involved and how they are related (Figure 17 Social network analysis Bozkaya et al. (2009)). The goal is to either structure the organization by classifying people in terms of roles and organizational units or to show relations between individual performers (Song & van der Aalst, 2008). Organizational mining does not produce a Business Process Model and is therefore useful to give an insight when no Process Model can be made. When many actors are involved into the Business Process it increases the coordination load.

Mining a social network shows the handover networks of people working together which is based on different actors executing different events of a Process or handling specific resources in a Process. From the event data of the invoice system at a Dutch municipality a handover model was made which shows how a resource of the Business Process goes to different actors (Song & van der Aalst, 2008). This model shows that some resources remain a long time with one person which was a bottleneck in the Process.

Another application looked at a hospital where it is important to properly refer a patient to a different department. The model showed cases where this rule was broken (Rebuge & Ferreira, 2012). This is often referred to conformance checking.

Both Rebuge et. al. (2012) and De Weerdt et al. (2013) analyzed a Process which crosses many organizational boundaries and requires coordination and report interesting findings of the organizational mining while Bozkaya et al. (2009) argues that organizational mining requires business experience to interpret wanted and unwanted behavior to find interesting deviations.

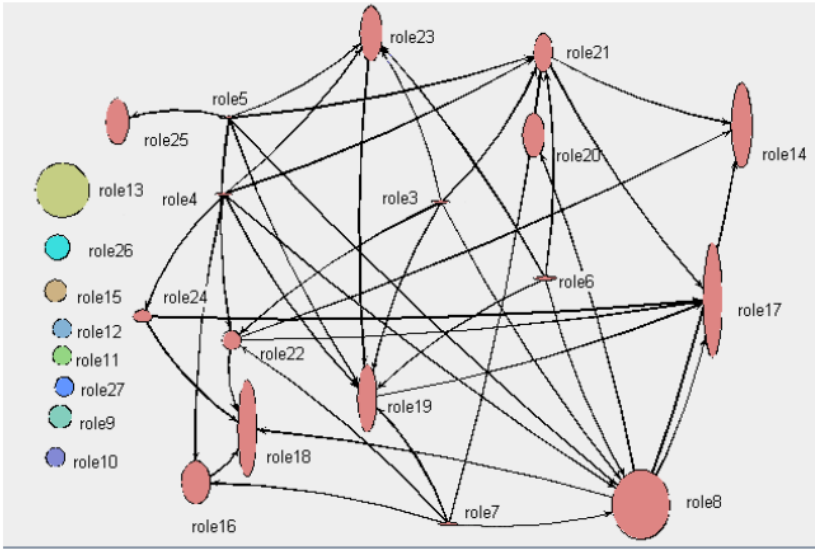

Figure 17 Social network analysis Bozkaya et al. (2009)

2.4.4 Performance analysis

Extending a process model by projecting information extracted from the log onto some initial model is referred to as Extension or Performance Analysis (Rozinat, de Jong, et al., 2009). Often event logs contain more information than just case related data, such as the performer of the task and which resource was used during the task.

Performance analysis is mostly about creating an integral overview of the Business Process ('Slicing') and zoom in on specific cases and instances ('Dicing') (W.M.P. van der Aalst, 2011). As mentioned in the previous section, this does require a Process Model which has a high fit with the connected event log.

Functionality of visualizing a Business Process is also seen in Process Discovery but in this stage the other resources which are available in the event log are added. In most case studies this resource is the time perspective.

Slicing refers to visualize the model with the cartography metaphor (W.M.P. van der Aalst, 2011); *Aggregate* Business Process activity names when zooming out (a bigger map shows only the capitals instead of smaller villages), *Emphasize* the most occurring traces of a process to show deviations (Main roads have different colors) and *Customize* the information in the map according to the need of the end user (There is no one single map for the world. Maps are specialized on a defined local context, have a specific level of detail and a dedicated purpose) (W. M P Van Der Aalst & Günther, 2007).

Dicing allows users to drill down and find the root cause of a deviation. Because the model is directly connected to the event log, all relevant data is available to answer specific questions. Most questions are related to the KPIs Time, Costs and Quality. Examples of KPIs related to time are lead time, service time, waiting time, and synchronization time. KPIs related to quality may refer to compliance, customer satisfaction, number of defects, etc. (W.M.P. van der Aalst, 2011).

The findings from visualizing Business Processes based on throughput time are often bottleneck related (W. M P van der Aalst et al., 2007), actors who save up small work to process it in batches (Măruşter & Beest, 2009) and rework of wrongly interpreted work (Goedertier et al., 2011)

By dicing and drilling down on the specific cases these findings can be quantified and their effects on the whole chain. The testing process of ASML is focused on decreasing throughput time. The

visualization showed the deviations and by drilling down the root cause of the problem was found (Rozinat, de Jong, et al., 2009). The specific Key Performance Indicators based on the throughput time of invoices can be answered by combining the event data and the model (Măruşter & Beest, 2009).

2.4.5 Summary

In this chapter the related activities of Process Mining found in case study literature are reviewed according to the classification made based on Process Mining methodology. There is no one single categorization which will have crisp boundaries to classify each related activity but the table below gives a general overview of which kind of Process Mining related activities can be distinguished.

Preprocessing	Discovery analysis	Organizational mining	Performance analysis
- Clustering - Filtering	- Algorithms - Fitness	- Handover network - Role analysis	- Slicing - Dicing - Quantifying results

Table 2 Classification of Process Mining related activities

2.5 Characteristics of Business Processes

This chapter starts with defining the structure which is chosen to present the Business Process characteristics. The remainder of the chapter elaborates on the characteristics.

Characteristics associated with the information processing view

Process Mining is applied on a very broad area of Business Processes varying from hospital care processes to machine testing processes. This suggests that Process Mining is not limited to a specific set of processes. Although all kind of Business Processes can be analyzed with Process Mining, case studies report that one Business Process is easier to analyze than the other. For example Rozinat et. al. (2009) mine the wafer scanner testing process and obtain a spaghetti like model which they do not blame on the limitations of Process Mining, but on the inherent complexity of the Business Process itself. Therefore it is important to understand what characteristics of Business Processes influence the effectiveness of Process mining.

Business Processes can be seen as a way to deal with organizational complexity (Bandara et al., 2005). This complexity has a strong negative effect on Business Process standardization (Schäfermeyer, Rosenkranz, & Holten, 2012). But the question remains what causes this complexity. Mani et. al. (2010) explain this complexity with the amount of information that has to be processed by the organization and therefore its processes and points out three dimensions; analyzability, variety and interdependence.

These three dimensions are deemed by the author to be suitable to explain the differences in Process Mining success and are therefore used in this research as constructs. The literature which is related to these constructs is discussed in the next sections. An overview is given in the table below.

	Analyzability	Variety	Interdependence
(Zigurs & Buckland, 1998)	- Outcome multiplicity - Solution scheme outcome multiplicity	- solution scheme multiplicity	- Interdependence
(Paul Lillrank, 2003)	- Assessment - Conversion rules - Logic	- Repetitiveness - Acceptance criteria	
(Mani et al., 2010)	- Analyzability	- Variety	- Interdependence
(Daft, Lengel, Science, & May, 1986)	- Task analyzability	- Task variety	- Interdependence
(Weske, 2007)			- Intra-organizational

Table 3 Literature on Business Process complexity

2.5.1 Analyzability

Mani et al (2010) refer to an analyzable Business Process when the outcomes are well understood, actors in the Process have a clear objective and a standard solution to resolve problems. When a Process is harder to analyze it is difficult to establish rules, procedures and predetermined responses to potential problems during process execution and management. Analyzability also increases the amount of information which has to be processed by the actors of the Process.

A clear description is given by Daft et al. (1986) “When the conversion process is analyzable, employees typically follow an objective, computation procedure to resolve problems. When work is not analyzable, participants have difficulty developing exact procedures, and hence rely on judgment and experience rather than on rules or computational routines”.

Outcome multiplicity means that there is more than one desired outcome of a task which increases the information load and diversity. Each outcome requires a separate information processing stream and is essentially a criterion against which a potential solution is evaluated.

Solution scheme outcome multiplicity is defined as the extent to which there is uncertainty about whether a given solution scheme will lead to a desired outcome. Examples are when the scope of the problem is too large, little historical information is available or when the outcomes are difficult to measure.

Lillrank (2003) also describes Business Processes in input, output and conversion rules which are strongly related to the degree of structure. These are combined with acceptance criteria and repetition to describe Processes as Standard, Routine and Nonroutine (Figure 18 Process classification (Lillrank 2003)). Acceptance criteria and repetition are considered part of the concept variety and will be discussed in the next section.

The *assessment* of the input for the Process refers to the knowledge needed to determine whether the input is acceptable or not. An analyzable Process has a predefined test or classification. A less analyzable Process requires human interpretation to decide to assess the impact of the input. *Conversion rules* refer to the kind of knowledge needed to execute the activities in the Process ranging from predefined algorithms to heuristics build on experience. *Logic* is related to the assessment that has to take place to determine the best approach. When the actor in the Process has to make an assessment based on a classification, the logic has a fuzzy nature while an acceptance test is stricter.

	Standard	Routine	Nonroutine
Acceptance criteria	Single variety	Bounded variety set	Open input set
Assessment	Acceptance test	Classification	Interpretation
Conversion rules	Switch, algorithm	Algorithm, grammar, habit	Heuristics
Repetition	Identical	Similar but not identical	Non-repetitive
Logic	Binary	Fuzzy	Interpretative

Figure 18 Process classification (Lillrank 2003)

2.5.2 Variety

Process variety is defined as the frequency of occurrence of process events that deviate from mean values of stability or uniformity of inputs/outputs, requiring different work processes than is the norm for completion of process objectives. High levels of process variety result in greater levels of process exceptions and deviations in the outsourced task environment and, hence, greater levels of information processing (Mani et al., 2010).

Daft et al. (1986) refer to variety in tasks as “the frequency of unexpected and novel events that occur in the conversion process. High variety means that participants typically cannot predict problems or activities in advance”.

Acceptance criteria determine the variety of resources a Business Process can handle (P. Lillrank, 2003). The more variety a process accepts, the more diverse the Process options will be and therefore will lead to less *repetitiveness*.

Repetitiveness is the amount of times a Process follows the same path.

Solution scheme multiplicity is defined as the amount of possible actions and their orders to reach a target. Often these solution schemes require the decision maker in the process to consider the consequences of their choices for a solution which increases the information load (Zigurs & Buckland, 1998).

2.5.3 Interdependence

Interdependent Business Processes cross organizational or functional boundaries and might be separately supported by different information systems, actors and managers. Process interdependence also necessitates variety in coordination efforts (Mani et al., 2010).

Interdependence in a Business Process means that choices made by actors or departments at the beginning can force adaptations to other stakeholders at the end of the production chain (Daft et al., 1986). High interdependence increases the information load of the actors in the Process to be able to cope with the adaptations and reach the goal of the Process. Low interdependence makes departments experience greater autonomy, stability and certainty with respect to coordination.

Conflicting interdependence may exist among solution schemes where adopting one scheme conflicts with adopting another possible solution scheme. In this case, the adoption of any one scheme substantially alters the situation such that the decision makers cannot simply change their minds, undo that adoption, and return to essentially the same conditions presented in the original task to make a new decision (Zigurs & Buckland, 1998)**.**

Intra-organizational processes versus process *choreographies*. Weske (2007) states that each Business Process is performed by a single organization and when a process does not interact with other parties it is classified as intra-organizational. When Business Processes interact with Business Processes from other organizations they form process choreographies. An intra-organizational process is primarily focused on streamlining internal processes by eliminating activities that do not add value and are often supported by traditional workflow management systems. Process choreographies on the other hand deal with communication aspects and legal matters. Also, choreographies are more challenging to support with technology because of the interoperability between the processes.

2.6 Process Mining success

To be able to measure Process Mining success it is important to define success measures. Will van der Aalst et al. (2003) address the importance of objectivity: “Modelling an existing process is influenced by perceptions, e.g., models are often normative in the sense that they state what ‘should’ be done rather than describing the actual process. As a result models tend to be rather subjective.” Also, Process Modelling is a tedious Process requiring a lot of effort (Bandara, Indulska, Chong, & Sadiq, 2007).

Others mention receiving a better understanding of who is involved in the Business Process, what is the main flow and how deviations affect the efficiency of the Business Process (R.S. Mans et al., 2009; Song & van der Aalst, 2008; W. M P van der Aalst et al., 2007). Goedertier et. al. (2011) introduce the terms justifiability and comprehensibility. Justifiability is the extent to which the induced model is in line with the existing domain knowledge. Comprehensibility is the extent to which and induced model is comprehensible to end-users.

The three measures which have been validated with Process Mining case studies by Mans et al (2014) are model quality, process impacts and project efficiency.

With the knowledge from the case studies, this research extends the three measures in the following manner:

- Model quality, the extent to which all desirable properties of a model created from Process Mining satisfy the needs of the model users
    - The model is justifiable in the sense that it is in line with the existing domain knowledge
    - The model is comprehensible to all end-users
- Process impacts, the overall effect of Process Mining on performance of a process
    - Process owners have a better understanding of:
        - who is involved in the process
        - what is the main flow
        - how deviations affect the efficiency of the process
- Project efficiency, the ratio of obtained outcomes over invested resources
    - Gaining similar results requires less effort with Process Mining or the results would require more investments than the expected returns with other techniques.

3 Conceptual model

The previous chapter describes the Process Mining related activities, the Business Process characteristics and the Process Mining success measures found in literature. This chapter describes how the Fit between Process Mining related activities and Business Process characteristics leads to Process Mining success with the Fit concept of Zigurs & Buckland (1998).

3.1 Fit

A high level Fit description is given based on the degree of structure of the Business Processes by van der Aalst. When a process model is mined and it is possible to structure the model with a fitness higher than 0.8 all Process Mining techniques can be applied (W.M.P. van der Aalst, 2011).

On the other hand when a Business Process cannot be structured than not all Process Mining related activities can be applied. Although this gives a general idea on the Fit this research looks for an explanation of this phenomenon in the characteristics of a Business Process. The structure which van der Aalst refers to is the structure seen in a mined process model. Some mined models show a lack of structure when first mined but after clustering and filtering specific events all Process Mining related activities can be applied (Rebuge & Ferreira, 2012). Therefore this Fit lacks precision.

Since there is no comprehensive description of Process Mining Fit in literature this chapter bases the Fit on the case studies found in literature. The case studies show three different compositions of the eight possible combinations (hence the three indicators for the Business Process can be either high or low and allow for eight different combinations).

The processes in the case studies all show high Interdependence which is not surprising since many practitioners and researchers consider Process Mining to be the technique to look at the process from end-to-end over functional boundaries. The last Fit which was not found in the case studies is; high Variety, high Analyzability and high Interdependence. This can be explained due to the fact that it is highly unlikely that a process with high Variety would have a formal description for each deviation and therefore a high Analyzability.

The three remaining Fits are explained in more detail in the following sections and follow the general model of Task-Technology Fit in Figure 19.

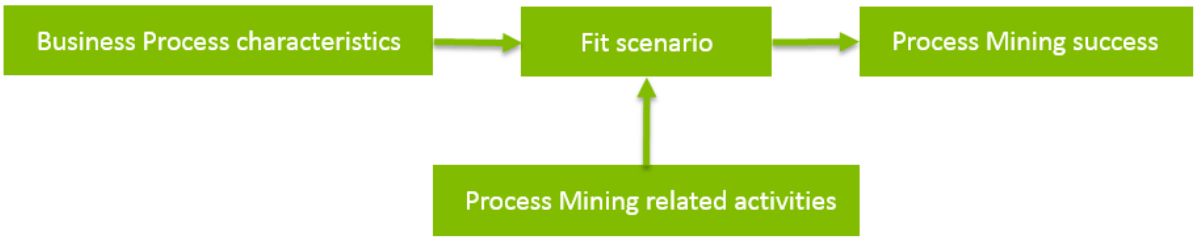

Figure 19 General Task-Technology Fit

The Fit in this is research prescribes the focus for added value of Process Mining related activities based on Business Process characteristics to get Process Mining success.

3.2 'Challenge' Fit

Figure 20 Challenge Fit

A Process Mining scenario where the Business Process shows 'Challenge' characteristics faces problems like a high Variety and therefore it is hard to construct a process model. These scenarios typically require a lot of external knowledge to make sense of the data. Process Mining offers technology to visualize clusters of Business Process instances that follow the same route. Preprocessing the event log and filtering out the infrequent cases reduces the Variety which can give insight in the main flow of the Business Process. Organizational Mining can deliver value to give insight into how actors in the Business Process collaborate. When a homogeneous group of cases can be found which are still of interest to the stakeholders the situation might be upgraded to the 'Best' Fit.

Proposition 1. In a 'Challenge' Fit scenario the project gets the added value of Process Mining related activities from Preprocessing and Organizational Mining to get Process Mining success

An example of a 'Challenge' Fit is the healthcare process of the emergency department (Rebuge & Ferreira, 2012). The event log contains 627 cases with many variations. The management prohibited physicians from forwarding patients to other physicians but they did not have any way to measure this phenomenon. The social network miner was able to show the cases being forwarded to other physicians. Important is to notice that it was possible to mine a process because the process miner did put a lot of effort in separating the event log into 7 clusters. In this case each cluster can be seen as a new process for which the variety has been reduced.

The model quality is justifiable and comprehensible in the sense that physicians did recognize the behavior in the model. The process impacts and project are high because management now has a way to control the unwanted behavior which they did not have before.

3.3 'Best' Fit

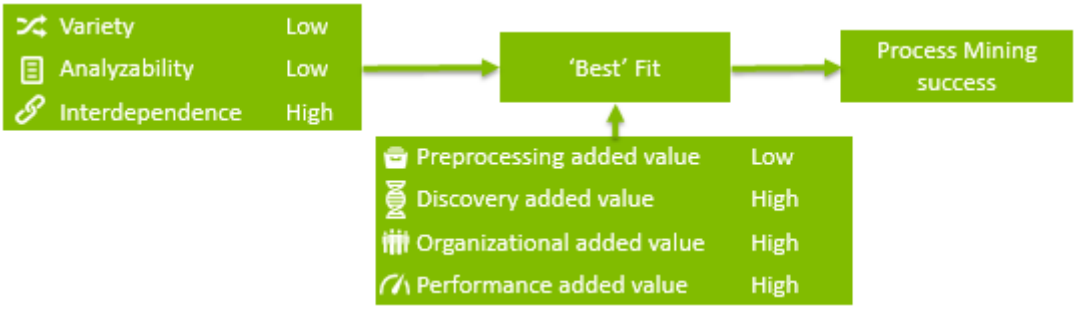

Figure 21 Best Fit

The second Fit scenario to get insight into the Business Process is a 'Best' scenario were the Business Process shows a high level of repetition but the process owners lack the insight into how the Business Process is executed in practice. Some effort is needed to gather all the separated data and produce one event log. By filtering out incomplete cases and reducing the Variety with help of, for example, the heuristics miner a process model can be produced with a fitness of 0.8 and higher. The preprocessing in this case does deliver less added value but is considered a necessary step for the rest of the project. Depending on the resources available in the log Process Mining can deliver insights in throughput time, handover metrics and the deviations from the main flow.

Proposition 2. In a 'Best' Fit scenario the project gets the added value of Process Mining from Discovery Analysis, Organizational Mining and Performance Analysis to get Process Mining success

A 'Best' Fit scenario is the document management system of a life insurance company (De Weerdt et al., 2013). The log contains 44,880 cases and although management has statistics on the performance they have no clear idea on what happens in the process. The Variety is low since there is a bounded set of variety input and a limited amount of steps that can be followed. The interdependence is high since a lot of different brokers are involved. With the help of filtering the event log and discover a model based on the heuristics miner, the final model showed a fitness of 0.77. The model enhanced with throughput time indicated deviation from the standard process model. This deviation showed that documents were misclassified. The ability to directly quantify the deviation led to a business case for buying an automatic document scanner to prevent misclassification.

In terms of model quality the model was justifiable and comprehensible for the stakeholder although they did not have a background in modelling. The actors in the process had no real idea about what was happening in real life although they received several statistics on the performance which indicates high process impacts and Project efficiency.

3.4 'Specific' Fit

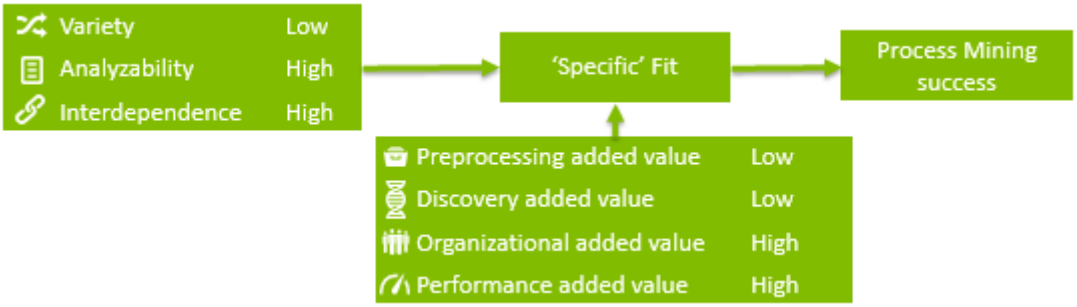

Figure 22 Specific Fit

When a Business Process shows low Variety and the actors in the process have a good understanding of the Process and how they should manage the Process it requires a 'Specific' Fit. It will be relatively straightforward to extract the data of the system and produce a process model with high fitness. Although the mined process model will have a high correlation with practice the discovery of the model will often be of less interest since the process owners were already aware of how the process is enacted. Because the model often is very accurately Process Mining functionality can enhance the model with extra information such as the throughput time, answer specific question about KPIs and allows for creating an end-to-end oversight of the process. In this case it often helps that the process owners have specific questions or metrics which they want to analyze.

Proposition 3. In a 'Specific' scenario the project gets the added value of Process Mining from Organizational Mining and Performance Analysis to get Process Mining success

The last scenario is ‘Specific’. The internal fraud detection on the procurement process recorded in SAP is an example of ‘Specific’ scenario (Jans, Van Der Werf, Lybaert, & Vanhoof, 2011). The event log contains 10,000 cases. The procurement process is very well described and does not allow for much variation. The process is interdependent as it crosses multiple organizational boundaries and work is handed over in the process. The process must follow the four eye principle and that is the specific task which the process miner needs to check. With the fuzzy miner a fitness of 0.9 was reached. With this model the miner was able to check for very specific process characteristics. Although the process model was already known with the users the dicing (specific checks) were able to control the information.

The model was justifiable and comprehensible because of the high fitness. Although the process owners already had a good insight in the process, the checks could be executed easier and ad hoc then before and therefore the project was also efficient.

3.5 Measurement model

To be able to measure the propositions the concepts in the model have to be operationalized. The concepts are Business Process characteristics related (variety, analyzability and interdependence), Process Mining activity related (Preprocessing, Discovery, Organizational and Performance) and Process Mining success related.

The Business Process characteristics related concepts are adapted from the literature to fit in a Process Mining setting. The Process Mining activity related concepts are operationalized by analyzing the case study literature according to the definitions given in the literature review. The Process Mining success measures are adapted from (Ronny S. Mans et al., 2014) and several measures were added from the case study literature.

Concept	Operationalization	Author	Question
Variety	Actors in the process often have to deviate from the process description	(Daft et al., 1986; P. Lillrank, 2003; Mani et al., 2010)	12 a/b/c
	The process does not impose restrictions on the input of the process	(P. Lillrank, 2003)	
	The process is hardly executed in the same way	(P. Lillrank, 2003)	
	There are many ways to attain the goal of the process	(Zigurs & Buckland, 1998)	
Analyzability	Process administrators follow an objective, computational procedure to resolve problems	(Daft et al., 1986; Mani et al., 2010)	12 d/e
	The actors in the process have a clear description how to execute the process	(P. Lillrank, 2003)	
	The actors in the process have good knowledge what the outcomes of an event can be	(Zigurs & Buckland, 1998)	
	The actors in the process know when the goals of the process are met	(Zigurs & Buckland, 1998)	
Inter-dependence	Decisions at the beginning of the process reduce the options later on in the process	(Zigurs & Buckland, 1998)	12 f/g/h
	The process has more than one process owner with disjunctive objectives	(Mani et al., 2010)	
	The process is not modular in the sense that individual steps can be analyzed, modified and enhanced independent of other process parts	(Mani et al., 2010)	
Preprocessing	Specific events were removed from the event log	(Rozinat, de Jong, et al., 2009)	2/4/5/13/14
	The event log was separated into smaller event logs with cases with similar properties	(R.S. Mans et al., 2009)	
	Incomplete cases were removed	(W.M.P. van der Aalst, 2011)	
Discovery	The Process Mining tool was able to visualize the Process Model based on the event log	(W.M.P. van der Aalst, 2011)	2/4/5/13/14
	The mined model was able to visualize the main behavior and the deviations	(Măruşter & Beest, 2009)	
	The events in the model were aggregated to emphasize the most important information	(R.S. Mans et al., 2009)	
	The mined model was able to explain 80% or more behavior of the event log	(W.M.P. van der Aalst, 2011)	
Organizational mining	A handover model was made based on different actors handling the same case	(Song & van der Aalst, 2008)	2/4/5/13/14

	A social network analysis was made based on actors executing the same events	(Song & van der Aalst, 2008)	
Performance	The consequences of the found deviations from the main flow were quantified	(W. M P van der Aalst et al., 2007)	2/4/5/13/14
	The mined model can be benchmarked to show performance differences	(Măruşter & Beest, 2009)	
	The enhanced model was able to show the performance of specific cases	(R.S. Mans et al., 2009)	
	The enhanced model was able to visualize bottlenecks	(Măruşter & Beest, 2009)	
Process Mining success	The extent to which all desirable properties of a model created from Process Mining satisfy the needs of the model users	(Ronny S. Mans et al., 2014)	15/16/17/18
	The model is justifiable in the sense that it is in line with the existing domain knowledge	(Goedertier et al., 2011)	
	The model is comprehensible to all end-users	(Goedertier et al., 2011)	
	Stakeholders are better aware of the consequences of their actions	(W. M P van der Aalst et al., 2007)	
	Stakeholders have a better understanding of the actors involved in the process	(Song & van der Aalst, 2008)	
	The stakeholders of the process have a better understanding of the differences between practice and the description of the process	(Song & van der Aalst, 2008)	
	The mined model shows the same information as handmade models but require less effort	(R.S. Mans et al., 2009)	
	The information retrieved through Process Mining was able to answer all the questions of the stakeholders	(Rozinat, de Jong, et al., 2009)	

Table 4 Operationalization of concepts

4 Data collection and analysis

This chapter elaborates on the methodology used to test the propositions of the previous chapter. The first section will explain how the data will be gathered and how to ensure the rigor. The next section will explain how the gathered data is analyzed.

4.1 Expert interviews

To validate the propositions of the Fit in the previous chapter, interviews with Process Mining experts are conducted. Interviews are most advantageous when the questions are either complex or open-ended (Saunders et al., 2009). Because the concepts are heavily context dependent and the interviewee might have a different naming for the concepts it is important to have the possibility to probe for answers (Saunders et al., 2009). It is especially interesting to retrieve the decisions and the context in which the expert has made the decisions.

Although it is stimulated to have an a priori analytical framework or conceptual model to analyze the results (Yin, 2009) it is important to be aware of the bias it might introduce. The interaction between the interviewer and interviewee might impact the data that is collected and therefore the conceptual model will only be discussed at the end of the interview.

The interview is guided by a protocol and done in a semi-structured manner. The protocol ensures that the interviews collect the data in a same way and reduce the impact of bias (Bhattacherjee, 2012). The interview is semi-structured for when discussions arise which can be valuable for the research or the interviewee gives a different view on the matter.

The interviews will be conducted with Process Mining experts since they both have the experience with Process Mining related activities and the Business Process characteristics that define the context of the Process Mining project. Therefore they are the most suitable persons to assess the Fit. To be able to distinguish the Fit in different contexts the interviews have to cover the three Fit scenarios described in the conceptual model.

To be able to assess the impact of other potential factors that influence the Process Mining success, the interview protocol will start with getting an overview of the context and the Process Mining experience.

4.2 Semi-structured interview protocol

A semi-structured interview allows for questions to be changed in order and to probe for answers but a protocol is needed to ensure that the data is collected in a structured manner (Saunders et al., 2009).

The protocol for the interview can be found in appendix B. The interview is divided into three main parts; context, practice and model related questions.

The *context* related questions take success factors into account of Mans et al. (2014) to establish whether these factors influence the results. The rest of the context related questions are based on the knowledge of the Process Miner of Process Mining related activities, the view on what Business Process characteristics are important and how Process Mining success should be measured. This gives an initial view on what the interview participant finds relevant for Process Mining. Since the interview participant might be unaware of Process Mining related activities or is limited to using specific functionality on a specific case, it is important to establish the base knowledge.

To further structure the questions and probe for the decisions made Saunders et al. (2009) advice to take a specific case and let the interview participant chronologically walk through the decisions made. Therefore the next group of questions are *case* related. First the interviewer is asked to elaborate on

the context of the project and the Business Process which is analyzed. Followed by questions about how Process Mining related activities was applied in several steps and how the related activities contributed to the results of the project.

After the case related questions the model will be more explicitly tested to ensure the model related information is collected. These questions are only asked at the end of the interview on purpose to limit the bias. The *model* related questions first aim to find out whether the interview participant would have taken a different Process Mining approach when the Business Process had different characteristics. The last questions explicitly state the model and ask whether the interview participant agrees with the propositions.

For all interviews, the following processing procedure will be followed (Bandara et al., 2005):
1. The interview is recorded
2. The relevant parts of the interview will be transcribed
3. The notes will be presented to the interviewee for review
4. Remarks are processed into final interview results
5. Recordings are erased once the case studies were completed

4.3 Coding

The conceptual model in chapter 3 is used to take a deductive approach in this research. This provides several advantages, as it ties the research into the existing body of knowledge, helps your research get started and directs the analysis of the collected data (Saunders et al., 2009; Yin, 2009).

The approach is referred to as pattern matching where the conceptual model (based on existing theory) is used to explain the findings from the collected data. If the pattern of the data matches the conceptual model it is an explanation of the phenomenon (Saunders et al., 2009). In this case the propositions based on the Fit model will be tested in order to answer the main question. If the data shows one or more outcomes than the model can explain, an alternative explanation will have to be presented (Yin, 2009).

The data in the interviews are coded or like Miles & Huberman (1994) state: “Codes are tags or labels for assigning units of meaning to the descriptive or inferential information compiled during a study”. This is done with the Atlas program.

4.4 Validity

Yin (2009) suggests different tactics that should be applied in case study research, to ensure the four common measures of data quality in empirical research are achieved. Although this research used expert interviews the validity measures are still deemed applicable.

Construct validity, is about establishing correct operational measures for the concepts being studied. This research established a chain of evidence to show where the operationalization of the concepts originate from. First the main topics on Process Mining where drawn from Process Mining case study literature and Process Mining experts within Deloitte. The related concepts are based on Process Mining methodologies and Business Process literature. The concepts are operationalized with corresponding literature in a measurement model shown in the table in section 3.5. The same table shows how the operationalization relate to the question protocol in Appendix B. The interviews are transcribed and coded based on the measurement model and the coding process described in the previous section. This chain of evidence allows other researcher to measure the same concepts.

Internal validity, examines to what degree the observed change in a dependent variable is caused by a change in the independent variable and not by changes in the environment or other factors (Bhattacherjee, 2012). Pattern matching is used to ensure the internal validity. If the pattern of your data matches that which has been predicted through the conceptual model you will have found an explanation, where possible threats to validity of your conclusions can be discounted (Saunders et al., 2009). This research describes three propositions (Fit scenarios) based on the conceptual model. The propositions have been operationalized in testable hypothesis. The hypothesis match the data from the interviews and therefore the model forms an explanation for the phenomenon. Also the interview protocol includes contextual elements from the Process Mining project success model of Mans et al. (2014). The results are included in the summaries of the interviews in section 5.1.

External validity, establishes the generalizability of findings. The model is based on literature which uses different Process Mining tools in different organizations. The same accounts for the variety of the interviews. They include different organizations in different industries and used different Process Mining tools which increases the generalizability of the findings.

Reliability, is about ensuring that a later researcher following the same procedures and conducting the same case studies would arrive at the same findings and conclusions (Bhattacherjee, 2012). An interview protocol is set up with a clear conceptual model for analyzing the results. This allows other researchers to conduct a same research which theoretically should uncover the same results.

5 Results

The interviews, to validate the conceptual model, are presented in this chapter. First an explanation is given how the interviews are analyzed and interpreted with the coding scheme. Then each interview is discussed separately followed by a section which discusses the accumulated results of all interviews. The last section includes a re-specification of the model.

The interviews are sorted according to the fits. An overview of the interviews, the organization at which the Process is analyzed and the Process which was subject of the project are presented in Figure 23.

Scenario	#	Organization	Process
Challenge	1	Hospital	Intestine cancer care
	2	DUO	Scholarship contact
	3	Hospital	Head Neck care chain
Best	4	Municipality	Invoicing
	5	Shipbuilder	Invoicing
	6	Hospital	Gall bladder removal
	7	Municipality	Invoicing
Specific	8	Provider	Invoicing
	9	Asset financing	Loan application call center
	10	Provider	Subscription activation
	11	Bank	IT service desk

Figure 23 Overview of the interviews

5.1 Coding process and individual interview results

The transcripts of the interviews are coded according with the measurement model in section 0 and the tool 'Atlas.ti'. This section describes how the coding is applied and translated into the results.

First a summary of the interview is given consisting of the experience of the interviewee, a description of the goal of the project, the Business Process which is analyzed, the Process Mining related activities applied and finally what the results of the project were. Several interviewees applied several projects. To be able to make a good comparison between the interviews, the interviewees were asked to choose one specific case.

After the summary, the coding results and the interpretation of interview results are given. The coding of the transcript is done with the measurement model from section 3.5. The measurement model describes indicators for the concepts in the model. The indicators for the Business Process characteristics describe whether the indicator of the process is high or low. An example of an indicator for Variety is 'The process is hardly executed in the same way'. When this code is applied to a section in the transcript of the interview, it gives an indication that the Business Process has a high Variety. To get a more accurate coding of the interview the indicator was separated into two codes, one which is applied in the case when a Business Process is 'hardly executed in the same way' and one when a Process is 'often executed in the same way.

The first code results into a reference to a high Variety and the second results in a reference to low Variety. This separation can be seen in Figure 25 under the Business Process Characteristics part.

Figure 24 is a visual representation of the 'Challenge' Fit from the conceptual model described in section 3.2. After coding the transcripts of the interviews the interview was classified into one of the scenarios. When there are more codings under high Variety opposed to low Variety then the Business Process is classified to have a high Variety. This is done for all three indicators of the Business Process.

A Business Process which is classified as a 'Challenge' can expect to have the most added value from Preprocessing and Organizational Mining to get Process Mining success as shown in Figure 24.

Figure 24 Model with a 'Challenge' fit

The model of a 'Challenge' Fit and its coding references are visualized as shown in Figure 25. The high and low codings for the indicators of the Business Process characteristics are shown in the left part, the according added value expected from the Process Mining related activities is described in the center and the right part shows the codings for the Process Mining success. Figure 25 shows a template for a 'Challenge' Fit. The expected codings according to the Fit classification are shown in the figures under the expected columns.

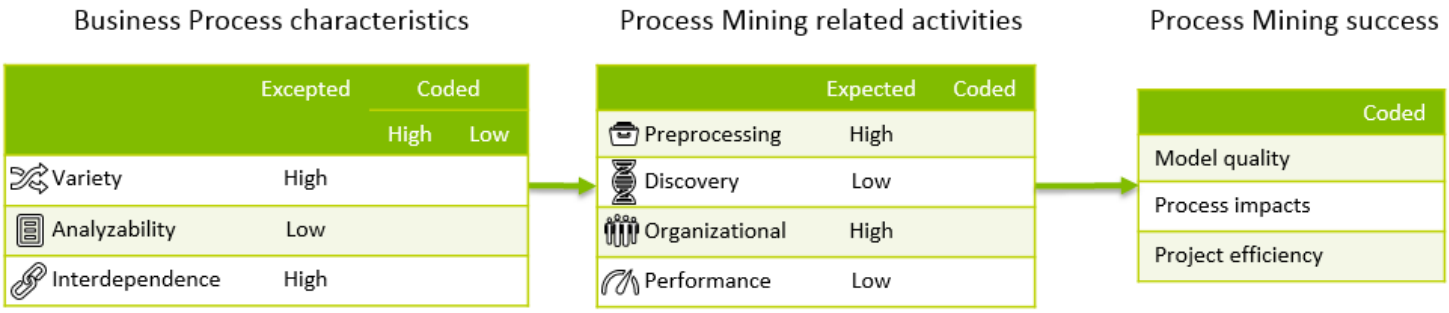

Figure 25 Coding template for a 'Challenge' Fit

The codings under 'Process Mining related activities' represent the references made by the interviewee to the added value of the type of Process Mining related activities. For example when the interviewee mentions that he or she got insight from visualizing the process model, it is coded under Discovery Analysis.

The codings under 'Process Mining success' are made when the interviewee mentions a result of the Process Mining project which is specified in the measurement model. The three indicators are based on the research of Mans et al. (2014). No specific hypothesis has been made on whether the different Fit scenarios would deliver differences in Process Mining success. The codings for the indicators are shown for the separate interviews but no hard conclusions can be drawn from the differences between the interviews due to the limitations of qualitative research.

5.1.1 Interview 1

Experience

The interviewee got his PhD in applying Process Mining in the healthcare sector. Through his education and his graduation he got in touch with Process Mining. He applied Process Mining on several processes in different hospitals. Although some processes in hospitals are quite comparable, a distinction can be made between low and high complex processes. The project which this interview focuses on, took place in a hospital on the patients diagnosed with intestine cancer. The project had a clear goal and wanted to get answers on questions like 'what are the most important paths, are we doing activities that should not take place and where are the bottlenecks'.

Process characteristics

The interviewee made a classification of healthcare processes which is also partly based on Lillrank (2003) as is this research. The classification extends the classification by dividing the healthcare into elective and non-elective. Processes which can be postponed and are elective can be ranked according to the Lillrank classification and the non-elective processes are divided into emergency and urgent. These classification are mainly based on the time the actors in the process have to plan and execute the process.

The process of diagnosing and treating intestine cancer was even complex at a fundamental level. The Gastroenterology discipline (a combination of the stomach, intestine and liver doctors) work together with the surgery discipline. When treating the patients with intestine cancer both disciplines are involved and they wanted to know whether they were executing double steps. An example is a double request for a blood research from both disciplines. This makes the process high interdependent.

The process starts with a patient which is seen by a doctor on the policlinic. Then several diagnostic steps are requested and then a treatment plan is made and the surgery is planned. Once the patient is operated the patient revisits the doctor on the policlinic and a further treatment plan is discussed. The process is completely dependent on the condition of the patient. Depending on this condition a specific plan is made for every patient which follows a specific path. When the patient has received a treatment (an operation, chemo- or radiotherapy) the doctor often requests for a diagnostic step. The further approach of the patient, and therefore also the process, is dependent on the outcomes of the diagnostic step and the condition of the patient. This makes the process highly variable.

The doctors have their medical protocols to rely on but the process is mainly based on their experience with patients. The throughput time of the process is regulated nationwide based service levels. These service levels give an indication of the waiting time of the patient before diagnosis and treatment, but it is hard to relate these targets to the specific actions the actors take to reach these goals. A general step is that after a treatment the doctor requests a blood research or a scan, but the results are mostly based on the interpretation of doctor rather than a formal description.

Process mining related activities

The interviewee started with a presentation to explain Process Mining and suggest which questions can be answered to raise awareness of the results of Process Mining. This also help to search for some initial questions which the stakeholders are eager to get answers on. These questions where already clear at the hospital.

To scope the patients which should be included, the interviewee decided to only include patients which were operated. To visualize the process data, the event log was filtered and events were

aggregated until a maximum of 20 events were present in the model. Otherwise the model would be incomprehensible for the stakeholders.

The aggregation of the events is done in dialogue with the doctors and partly based on experience of the interviewee. An example of aggregation is whether MRI or CT scans should be seen as separated events or should be aggregated to diagnostic research.

A specific question of the hospital was to get more insight into the cooperation between the different disciplines, but due to the high variety in patients this proved to be too big a challenge. The patients were supposed to enter the process at the policlinic of Gastroenterology and see a doctor. But in practice it also occurred that patient flowed into the process via another route.

This was also verified with the doctors who recognize the situation. Although these people do not follow the main stream, they were included into the calculations of the throughput time which gave a good estimation of the waiting time for patient which had to undergo surgery.

Process Mining success

The waiting time before operation appeared to be higher than expected. A lot of time was lost once a surgery was ordered and when it was finally planned. In several cases it took a week before the order was handled to be planned. This caused the throughput time to be higher than expected. The throughput time is a result of selecting the patient group which follows the main flow of the process, aggregating the events and calculating the throughput time based on this model with a Process Mining tool. This is a clear added value of Process Mining.

The clear answer is also a result of the clear question which the hospital had, so the interviewee was aware where to focus his effort on.

On the other hand the hospital asked for more insight into the collaboration between the different disciplines but this turned out to be difficult due to the high variety of the paths which the patients follow.

Scenario 'challenge'

Based on the amount of references (Figure 26 Interview 1 - Coding results) made and the description of the process characteristics, the case can be categorized as a 'challenge' scenario (High variety, Low analyzability, High interdependence).

The coding gives a good match with the 'challenge' scenario. The added value of Process Mining related activities comes from the preprocessing of the event log as expected. The interviewee selected a specific group of patients which at least were operated.

After this selection several steps were aggregated until finally a comprehensible model was found which was able to show the main steps and the deviations on a high level. For these specific patients it was possible to visualize the general steps and to calculate the throughput time.

What is interesting is that the hospital requested to give insight into the collaboration of the disciplines. For this goal Organizational Mining is very useful. But it received no codings to add value. Instead the interviewee explicitly mentioned that there was no use in visualizing the collaboration due to the high variety of the patient group flowing in the process. Where a patient should start at a specific discipline, several patients also skipped this parts of the process. This can be seen as a result but there was no clear result on how the disciplines work together.

Business Process characteristics

	Excepted	Coded	
		High	Low
Variety	High	4	0
Analyzability	Low	0	7
Interdependence	High	2	0

Process Mining related activities

	Expected	Coded
Preprocessing	High	4
Discovery	Low	2
Organizational	High	0
Performance	Low	2

Process Mining success

	Coded
Model quality	1
Process impacts	2
Project efficiency	2

Figure 26 Interview 1 - Coding results

5.1.2 Interview 2

Experience

The interviewee started as a developer for the workflow management software of Pallas Athena. The company acquired an organization which developed Process Mining software which know is known as Perceptive software. The interviewee has extended knowledge of different Process Mining tools and applied them at several organizations.

The interview focused on a project he is currently doing at the call center of 'Dienst Uitvoering Onderwijs' (DUO) which is a department of the educational ministry. He is part of a SCRUM team to enhance the contact inbound and outbound contact with students. The goal is to reduce the amount of unplanned contact.

Process characteristics

The process can be seen as a large number of smaller processes where each smaller process is a channel of a specific group of students contacting or being contacted by the organization. An example of a group which is analyzed is the group of foreign students who experience problems with signing in with their DigiD account. Another example is categorizing the incoming mail. A lot of the emails are status requests which could have been avoided if the students were able to find the information on the website.

When you look at the process on a more abstract level it is less complicated. A student or an ex-student which has to repay his loan, can initiate contact or make an adjustment via post, phone or via the website portal. The process allows for the student to choose his own way to gain information or make an adjustment. The employees of the call center try to prevent phone calls by informing students via mail before the student calls. But in practice these mailings can contain too less information and lead to extra phone calls instead.

The preventive mailings start with an employee detecting a possible information gap with a group of students. Once the group and the information gap are identified a preventive mailing is composed and set up to be send. Unique with this approach is that the SCRUM team looks at process from the customer perspective, which often has only a few steps, opposed to the long processes inside the organization.

The freedom for the student is enforced by law. A student must be able to make adjustments via forms on the website. When a student has a question about these forms, it does occur that a call agent again refers a student back to the form without further information while this is not the desired path. This is because the employees do not have a clear description to execute the process. The actors do get a short education but due to changes this education tends to get outdated. Changes are mailed around but often do not stick. Therefore the actors in the process mostly rely on experience instead of formal descriptions.

The process is currently managed on durations of phone calls. It does occur that people phone again or send another mail. In that sense it leads to rework. The phone calls and mails are routed to the right people. Although the impact is low when a call or mail is wrongly classified it does again lead to more rework.

Process mining related activities

As a consultant he always starts with exploring the available data to get a sense of what is available and what is the quality. Often this goes parallel with mapping the problem which the customer is

experiencing. This is then followed by defining hypotheses which can be tested. Once you are able to make a process model or extend the existing model to verify the hypothesis you create insight for the customer.

At this specific Project, Process Mining is applied in a highly iterative way. Before the Process Mining tools are used, the interviewee and the SCRUM team formulate clear hypotheses which can be tested with analyzing event data. The interviewee stresses that Process Mining is a means but should not be the goal. When a simple chart in excel will give a clear answer than that should suffice.

When the hypotheses are stated then they explore which data would give insight into the hypothesis and gather the data. The hypothesis often concern a specific group of students such as foreigners. The data is than gathered specifically on this group and their activities. An event log is made and loaded into a Process Mining tool to visualize the complexity. Often the process model is too complex to give insight and therefore cases are removed from the event log and steps are aggregated.

When a homogenous group is found and the data can be visualized in a Process Mining tool, it helps understanding the process and validate the hypothesis. After this insight is created, the solution to the problem is often a creative process without a specific Process Mining functionality. Once a solution is implemented, the Process Mining tool can quantify the throughput time before and after the implementation and therefore quantify the results.

Process Mining success

The added value of Process Mining at DUO is the visualization of the changes in the process. The interviewee states that giving insight might not always lead to a change in the process. When the stakeholders are guided through a circle of identifying the problem, stating a solution and measuring the results, the added value of Process Mining is often bigger.

In one case the team tried to reduce the amount of replies on a preventive mailing by giving more detail in the first mail. The results were both measured before and after but were not as effective as expected. With Process Mining they were able to quantify the results.

Analyzing the complete process is too complex and therefore the process is separated into groups of students with similar problems or characteristics. These groups are then analyzed separately and the hypothesis are measured with Process Mining or in some cases with MS Excel.

Scenario 'Challenge'

Based on the amount of references (Figure 27 Interview 2 - Coding results) made and the description of the process characteristics, the case can be categorized as a 'challenge' scenario (High variety, Low analyzability, High interdependence). The codings in Figure 27 show a good match with the 'challenge' scenario. The hypothesis of the model is that the Preprocessing phase delivers added value which is also the case in the results.

The interviewee and the SCRUM team use a highly iterative approach and pre-select the group which they want to analyze. Not all insights come from a Process Mining tool, sometimes it is easier to analyze the event data with Excel. The homogenous groups which can be visualized with Process Mining often also benefit from quantifying the deviations and lead to more concrete business cases. This also explains the difference between the expected and coded numbers with the Process Mining related activities in Figure 27.

The Organizational mining on the other hand was not coded in the interview while it was expected to add value in this scenario. The interviewee has quite some experience with Process Mining so it is likely that he has knowledge of the functionality.

Business Process characteristics

	Excepted	Coded	
		High	Low
Variety	High	3	1
Analyzability	Low	0	8
Interdependence	High	2	0

Process Mining related activities

	Expected	Coded
Preprocessing	High	4
Discovery	Low	2
Organizational	High	0
Performance	Low	2

Process Mining success

	Coded
Model quality	0
Process impacts	2
Project efficiency	2

Figure 27 Interview 2 - Coding results

5.1.3 Interview 3

Experience

The interviewee is employed in a Dutch hospital and has quite some experience in simulation models for surgery rooms and visualizing data networks. In his experience this is strongly related to Process Mining. Via a colleague he heard about Process Mining and started applying it for trial purposes.

The current problem with the management information on which the hospital is basing his knowledge, is mainly either financial data or qualitative data opposed to process information. When you ask five persons to draw the process they are operating in, then it is likely that you will receive five different models. The processes in general in the hospital are highly complex and therefore simplistic models are too far away from the reality to be of any use.

Therefore the interviewee applies Process Mining as a pilot to give the actors in the process insight in how complex their process is and learn them to think in a process way. The interviewee has applied Process Mining on to care chains. This interview is focused on the head-Neck chain which was initiated information driven. The interviewee often deals with process related questions by the doctors and in this case answered them with the help of a Process Mining tool. The goal was to give insight into the complexity of the chain, which is already analyzed for about ten years.

Process characteristics

In previous sessions actors in the chain were asked to draw a model of their process. Most of the time this is a linear model while the process contains many choices and loops. The actors in the process are not aware of the complexity of the process.

The process of the Head-Neck chain, like many care flows, crosses many medical disciplines and therefore departments of the hospital. The doctors and nurses of one discipline have a view on their part of the chain but are rarely aware of what specifically happens in other departments.

The patient is first assigned to a specific doctor who makes the care plan which determines the remainder of the process. This plan is consequently evaluated based on diagnostic results and can be heavily subject to change based on the characteristics of the patient. Consequently the process is hardly exactly executed in the same way.

In specific medical situations care practitioners have a medical protocol which they follow but these only cover certain paths in the process. The complete chain is operated on experience rather than process descriptions. Also on an administrative level, which currently allows for too much flexibility according to the interviewee.

Process mining related activities

The first step in the project was to analyze which diagnoses are correlated with the care chain and should be included into the analysis. Based on the knowledge of the interviewee he draws a first model which includes the high level steps and the related diagnoses. He discussed this first model with the care practitioners to check whether his initial perception of the process is correct and whether steps like diagnosis should be separated into MRI and CT or should be aggregated.

This initial model guides the search for event data which is logged by the different information systems. Including all data would make a too complex model which is incomprehensible but on the other hand you will need data from, for example, scheduling systems to be able to add specific information which is important for the process. An example is the multidisciplinary meeting where

several patients are discussed but only the meeting is present in the main system, not which patients are specifically discussed. Therefore the data is enhanced with the appointment system.

Then the data is preprocessed in excel and loaded into the Process Mining tool. Then you hope that you are able to answer the questions by applying filters and using the sliders to aggregate the behavior. When this is impossible you will have to go back and add extra data or remove cases.

The difficulty with analyzing the process is that it is data driven and there are no clear hypotheses or questions which can be answered with the Process Mining tool. There is not yet a clear goal. But visualizing the process does help in showing the complexity. By aggregating steps you are able to reduce the complexity. This complexity is often still present in the model even after preprocessing the data in Excel.

Process Mining success

The Process Mining tool helps building the bridge between the care professionals on the one hand and the logistic and process professionals on the other hand. When you are able to show an objective picture of the process based on facts rather than opinion, the stakeholders have more trust in the results. When you are trying to improve a process you both need the medical and process knowledge.

With the Head-Neck care chain it was possible to show the throughput time on a high level. The doctors also recognize the deviations and are able to explain the deviations from own knowledge, but it was not based on the results of the Process Mining analysis.

Since there was no definition of what diagnoses are associated with the care chain and what medical executions are applied it is also interesting to see what data is available and what patterns surface in the event log.

Scenario 'challenge'

Based on the amount of references (Figure 28 Interview 3 - Coding results) made and the description of the process characteristics, the case can be categorized as a 'challenge' scenario (High variety, Low analyzability, High interdependence). The coding clearly shows the characteristics of a challenge scenario. The process shows a high variety and is mostly executed based on experience. Although medical protocols are defined, they are more a guideline than a process description.

As expected with a challenge scenario, the most value comes from preprocessing the event log. Not in the sense of clustering but on preselecting which data is going to be included to visualize a meaningful process model which is comprehensible. Specific cases with patients that follow a very distinct path are left out to show the general flow. Also the process model only contains the main behavior on a high level. Still this gives insight into the complexity of the process. Where the care practitioners thought their process was a straight line, it actually was rather a collection of deviations. The specific set of patients that follow a main flow were visualized and also gave an insight into the general throughput time from the first consult until the multidisciplinary meeting and the treatment path.

Although the Organizational perspective was thought to deliver value, it was not mentioned. This can be explained due to the fact that the interviewee was not aware of this functionality and used the Process Mining tool as a pilot. Another remark can be made that the Process Mining success comes from the Process impacts and not from Model Quality or Project Efficiency. The model gave very important insight and was comprehensible by all users (which have no background in process modelling).

Business Process characteristics

	Excepted	Coded	
		High	Low
Variety	High	3	0
Analyzability	Low	1	6
Interdependence	High	3	0

Process Mining related activities

	Expected	Coded
Preprocessing	High	5
Discovery	Low	3
Organizational	High	0
Performance	Low	1

Process Mining success

	Coded
Model quality	1
Process impacts	6
Project efficiency	0

Figure 28 Interview 3 - Coding results

5.1.4 Interview 4

Experience

The interviewee got his Process Mining experience via his master operations management and logistics at the University of Eindhoven. At the consulting company where he is currently employed, he did a Process Mining project at a Dutch municipality and looked at the invoicing process of ten different departments. Both the accounts receivable and payable were part of the analyses. It is important to notice that the interviewee had no direct contact to the actors in the process but only via his manager. To give more insight into the specific process he provided extra documentation of the final results of the Process Mining project. The direct contact with the client was with the data experts and their superiors.

The goal of the project was to get insight into the throughput time and the cause of possible delays of the throughput time. Specifically the project also aimed to check the effectiveness of the 'four-eye' principle when an invoice had to be cleared above a specific value. The project was part of a larger BPM program. No analyses was performed on performance of any person specific due to privacy issues.

Process characteristics

Every invoice which the municipality receives and sends has a clear payment deadline which can differ per supplier or client. The processing of the invoices differs per department which led to an initial 10.000 process variants on a first glance at the data. People tend to deviate a lot more from the process than the managers at the municipality anticipated. This deviation is due to the different views of the actors on how to enact the process correctly. People in the process do not recognize the importance of sending an invoice and correctly logging the date of sending it.

To get the right data in the right format, the interviewee had several conversations with Oracle data experts of different departments on how they interpret the data. There was no clear description on how the interpret the data. The supervisors of the process on the other hand, did have a high level overview.

Process Mining related activities

The project was separated into three phases. First the extracting the data, then cleaning the data and then analyzing the data with a Process Mining tool.

The first step included specifying the required data. The data was extracted by the municipality and delivered in a pre-specified format. The issues with receiving the appropriate data for Process Mining differed per department. While data from some departments was directly usable for the Process Mining tool, other departments put a lot of effort in extracting the data and converting it to a usable format.

Next, the event log was analyzed and cleaned in cooperation with the data experts at the municipality. The event log had relatively few events; the logging of the activities seemed to be aggregated and therefore contained less detail about all the steps in the process. The data was also hard to interpret without knowledge of the Oracle system and the invoicing process. Therefore the interviewee had contact to both the data experts at the municipality as experts at the consulting company. The cleaning of the log was focused on removing duplicates, adding sequence numbers to activities and adding timestamps to events that had no timestamp.

The last step was considered 'real Process Mining'. After the data had been cleaned no further adjustments were done to the data. Several analyses were done to visualize the main flow with 80% of the cases and show the deviations. The interviewee created an individual enhanced process flow for each department with the amount of cases flowing through the model and the throughput time. The pre-defined KPIs were measured and quantified. These analyses and findings were presented to the stakeholders of the process in workshops.

After workshops were conducted with the actors and process owners, they were able to specify process improvements and formulate better checks on managing the process. The main finding was that the ten departments had large differences in the way they executed the process and therefore also differed in throughput time.

Process Mining success

The main results of the Process Mining project were insights into the unexpected deviations from the presumed process. The process owners did not believe that the process had this many variants. Also the different views on each department allowed for benchmarking the departments and formulate Improvement opportunities in decreasing the throughput time.

Scenario 'best'

Based on the amount of references (Figure 29 Interview 4 - Coding results) made and the description of the process characteristics, the case can be categorized as a 'best' scenario (Low variety, Low analyzability, High interdependence). The variety got both a high and low reference but was categorized low. A first glance at the data showed 10.000 process variants. In the transcript of the interview this was coded as High Variety. Several iterations of cleaning up the data showed that a part of the variety was caused by data quality issues such as unrecognizable data labels. Once these were filtered out it was possible to explain 80 percent of the cases with a simple model. Therefore this process is categorized as a Low Variety process. There were no codings for Preprocessing, because this cleaning is not considered an added value of Process Mining related activities.

With the 'best' scenario, added value can be expected from discovery, organizational and performance. The interviewee did also apply several iterations of preparing the event log, but this focuses on cleaning the data and adding timestamps. It did not focus on clustering the data to find a main flow. There were no references to Organizational Mining which can be explained due to the fact that the interviewee was not aware of this functionality. As the scenario describes, the added value of Process Mining was found in visualizing the main flow and its deviations and quantifying the differences. The model was also able to explain the questions of the stakeholders and quantify the KPIs.

Business Process characteristics

	Excepted	Coded	
		High	Low
Variety	Low	1	1
Analyzability	Low	1	5
Interdependence	High	3	0

Process Mining related activities

	Expected	Coded
Preprocessing	Low	0
Discovery	High	9
Organizational	High	0
Performance	High	4

Process Mining success

	Coded
Model quality	4
Process impacts	4
Project efficiency	0

Figure 29 Interview 4 - Coding results

5.1.5 Interview 5

Experience

The interviewee used Process Mining during his graduation project to research the applicability on fraud detection. He had a relevant education where he learned how to apply BPM and managing data. The Process Mining knowledge was mainly gathered from literature research. For the graduation project he used two real life data sets and applied Process Mining.

The interviewee analyzed two cases which were quite comparable, therefore this interview focused on one case where the procurement process of a large shipbuilder was analyzed for fraud detection. The analysis was done separately from the company. There was no contact between the shipbuilder and the interviewee until the results were presented.

Process characteristics

The procurement process is, from a fraud detection perspective, a very interesting process to analyze since it is very sensitive to fraud. Also the process is very rich in the sense that it contains a lot of information on which you can sort and create views. The procurement employees of the shipbuilder have a good understanding of the high level steps but it is unlikely that everyone would be able to draw the exact process. It also depends on the function, the experience accumulated during the years and the part of the process he or she is executing on how well the actor is able to reproduce the formal events in the process.

A procurer will be able to explain the steps but has limited view on what activities take place in the warehouse. Actors receiving the goods in the warehouse mostly have less knowledge of the complete process, but several employees have been employed for quite some time and have an idea of how the complete process works. The steps of the process are not literally captured in process descriptions, apart from the signing approvals, but most actors know the process from experience.

A special situation occurred when the extra required signature for approving an order was ignored on purpose. This is done to prevent further delays in the throughput time of the order because the consequences of this delay will cause a delay in a project which will cost even more than a possible fraud case.

Process mining related activities

The data set from the shipbuilder was handed to the interviewee for Process Mining.

After creating an event log it was possible to load the information directly into the Process Mining tool. The ERP system which supports the process was able to produce data which was directly readable for the Process Mining tool, but some events recorded timestamps only on a daily basis. Therefore some extra timestamps were added to get sequentially.

The interviewee applied several Process Mining tools which were all able to directly visualize the information. Based on his own business experience he was able to aggregate steps and find so called outliers. These deviations were then subjected to further analysis such as ‘which people are involved, what caused this deviation and what are its implications’. Based on the knowledge he accumulated during these first visualizations of these deviations he looked at the complete model again and zoomed in all the cases which were handled by the person or the group which caused the initial deviation. These cases were benchmarked against all other cases to check whether they were structurally deviating from the general process.

For deviations which stand out, the interviewee made a matrix on which persons executed which steps for these deviations. This was not only interesting for fraud but could also show which people were performing most of the cases and were also the bottleneck of the process.

In this case specific, the interviewee made several references to the usefulness of visualizing the process flow and visualizing deviations. Especially for fraud, the outliers (deviations) are considered most interesting and subject to further analysis.

Process Mining success

It is interesting to notice that the interviewee mentions that the flexibility of the program is both a good thing and a problem because there is no way to know for sure when you are done analyzing the process. He mentions that for BPM purposes to improve the process and find the bottlenecks, you are done after several iterations.

The main deliverable of the project was an oversight of quantified deviations that stand out without having a clear explanation. Mainly lists of suppliers and orders that take strikingly more time to be processed. The deviation does not directly mean fraud but it gives insight into how many times deviations happen and who is involved. These insights are clear for all stakeholders and they can relate them to the things which happen in reality. They give the process owner a concrete situation which should be analyzed into more detail, because it is difficult to base your results solely on Process Mining insights.

A specific deviation which happened several times was that some packages were received only five seconds after the purchase order was put into the system. When the process was strictly followed this would have been impossible because a delivery time of five seconds would have been unrealistic. This indicated that the process was not always strictly followed but can be explained by the fact that someone forget to put in the purchase order into the system.

Scenario 'best'

Based on the amount of references (Figure 30 Interview 5 - Coding results) made and the description of the process characteristics, the case can be categorized as a 'best' scenario (Low variety, Low analyzability, High interdependence). What was interesting is that the process had clear definitions of what the approval limits were for the amount of money associated with the order, but also parts which were based on experience. An example is the deviation from the amount of signatures required for an order above a certain payment limit. Based on the knowledge that the delay would cause more financial damage than the possible fraud damage some signatures were skipped. This knowledge is not captured in process descriptions but many actors are aware of the phenomenon.

The interviewee had several specific questions which were answered with the help of Process Mining, but the most important insights created were list of deviations that happen relatively often and who is involved. In the added value of Process Mining related activities it is noticeable that the interviewee was very pleased with being able to generate many different views based on different attributes of the process in short time. The interviewee had no deep knowledge of the process but was able to get insight in the deviations, find the actors and groups involved and present these findings to the stakeholders.

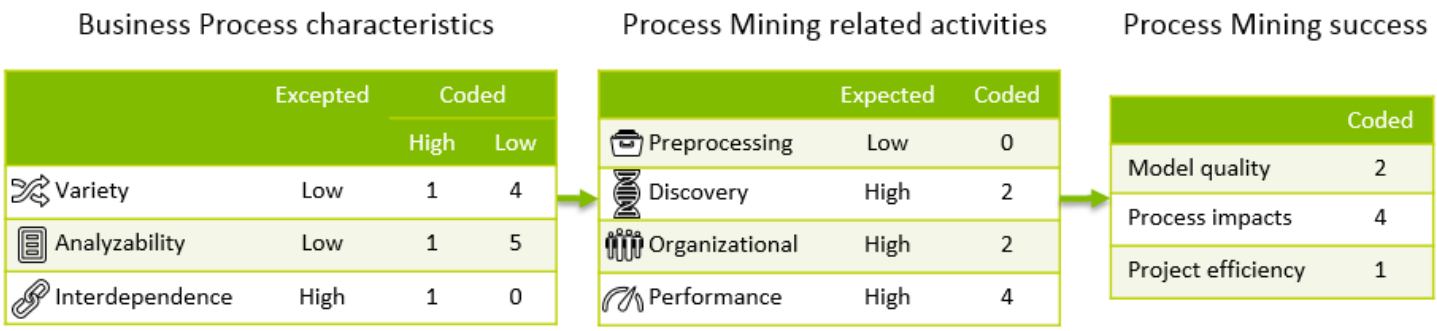

Figure 30 Interview 5 - Coding results

5.1.6 Interview 6

Experience

The interviewee is employed at a Dutch hospital and she is getting her degree in clinical informatics at the TU Eindhoven. During her education she got experience with analyzing processes and processing Information in a medical environment to reduce the gap between care professionals and IT professionals. During her education she received basic knowledge of Process Mining.

The Process Mining project at the hospital was her first experience with Process Mining. The project was executed in combination with a process expert from an external consultancy company and the interviewee was involved in all steps which were part of the Process Mining project. The occasion for applying Process Mining was because the hospital is moving to a new hospital information system to support the processes and they wanted to get more insight into these processes, since there were no accurate process descriptions available at that time. A main reason to choose Process Mining as a pilot over traditional sessions with brown papers, is to avoid spoiling time from care professionals such as doctors. The final goal of the project was to show a proof of concept and get a good description of the high level process.

Process characteristics

The process which was analyzed was the 'laparoscopic gall bladder removal'. This process was chosen with care because the interviewee wanted to analyze a process which had a clear beginning and end without any chance of reoccurrence. Hence, once a gall bladder is removed, it cannot be removed for a second time. A chronical disease often has patient flowing through the process multiple times and is therefore harder to analyze. Further the process is chosen because it crosses multiple disciplines and systems, which is interesting from an information system perspective.

Patients which undergo a procedure for removing their gall bladder often do not have a secondary medical problem. Therefore most of the patients undergo a similar procedure. Although general healthcare processes have a high variety, this process was specifically chosen based on the relatively low variety.

To get accurate process descriptions was one of the goals of the project and were not present at the hospital at that time. The process is mainly based on experience from the actors.

The process crosses multi disciplines. First the patient is seen by a doctor to assess the correct care path. The doctor also enters a score in the system to assess the readiness for the patient to undergo the operation. This is further on used by the preoperative screening at the anesthetist. In that sense the steps in the process are strongly dependent of the earlier made assessments. Although the doctor will have an idea what is done with the assessment code he enters into the system, but it is not likely that he oversees the consequences of the specific number. A higher number, for example, requires more steps at the anesthetist to assess the condition of the patient.

Process mining related activities

The first step was to critically select a process which shows comparable elements of importance for selecting an information system, but is not too complex to analyze. After the process was chosen, the interviewee and the project team defined which information was necessary to get a good view of the process. A first high level process was defined based on the national guidelines since no accurate descriptions were present at the hospital.

After defining the high level process, the interviewee consulted the data experts to verify whether the required data was present in one or more systems. The data came from three systems and were delivered in separate Excel files. Combining these files into one event log appeared to be challenging. The interviewee acted as an interpreter to translate data activities to process steps. Another example of an unforeseen problem was the definition of the duration of the operation. Several timestamps were recorded during the operation but there was no single definition. After the data problems were removed a first visualization of the process was produced.

A first look on the process showed that hardly any patient follows a comparable path. A group of events which caused a spaghetti process was the step of blood research. Each value for which the blood was checked was separately registered. These separate steps were aggregated to blood research, which quickly reduced the amount of deviation. Other data problems were found such as different abbreviations for the same medical procedure.

Once the data was cleaned from these specific situations, the Process Mining tool was able to show the main flow and the deviations. Although there were still quite some deviations, it already gave insight into the main flow. Steps which could not happen in a certain order did happen in reality. When these cases were analyzed specifically it appeared that one timestamp was registered on day level while the other contained a specific time. This caused the Process Mining tool to misjudge the order in which events happen.

The first findings were presented to the doctors which recognized the steps because national naming conventions were used for the steps in the process. With their knowledge they were surprised to see that patients were seen more than once after surgery which in practice did never occur according to the doctors. This appeared to be a wrong registration of the system. The interviewee did not notice this herself, but due to the medical knowledge and practical experience the doctors were able spot this deviation.

Process Mining success

The Process Mining project resulted in a process model based on the real life situation which shows the main steps from a logistic and IT process perspective. The model is comprehensible by doctors but also shows enough detail to be of value for IT specialists.

Important was that only little time of the doctor was necessary to make the process model. It did require extra time on the other hand from data analysts and the interviewee but this is experienced as an added value by the organization since the time of the doctor is scarce and should mainly be focused on treating patients.

The main deviations were found and their root cause. This also gave insight into the data requirements for both a new Process Mining pilot and the new hospital information system.

Also the doctors realized the added value of thinking in processes opposed to specific patients for managing the process. Where doctors are specifically trained to choose the best care path for a specific patient, for defining the most important steps of the process you need to think in steps and options.

Scenario ‘best’

Based on the amount of references (Figure 31 Interview 6 - Coding results) made and the description of the process characteristics, the case can be categorized as a ‘best’ scenario (Low variety, Low analyzability, High interdependence).

The codings match the expected scenario of best but it is important to notice that this assessment is based on the description given by the interviewee. She choose to abstract from several specific events such as the specific blood values which were researched. Each blood result was recorded as a specific event. For the purpose of finding the main flow it is not interesting to split these specific events in the process model. Even on a more abstract level the interviewee argued that the process could be considered 'specific' (or standardized as mentioned in the presented model). But since there were no formal process descriptions describing how to execute the process and the interviewee used the national descriptions to map the process steps on, the process is considered Low Analyzable.

The event log was filtered based on patients which did not undergo surgery but were recorded into the system. After analyzing the log the interviewee also found out that a few patients went to the gynecological department. This was not expected but gave insight into the fact that even a relatively simple process in theory can get complicated due to complex patients in reality. Therefore the Preprocessing phase did also deliver insights.

The Discovery phase clearly gave the most added value since all stakeholders were able to comprehend the model and aggregated on a level which showed the main steps and their deviations.

What does differ from the model, was that the added value of Organizational Mining and Performance Analysis were expected high but show no coding references at all. The tool which the interviewee and her team used did not have the functionality to show Organizational Mining which does explain why it did not deliver any added value.

The Performance analysis was not specifically mentioned during the interview. This can be explained due to the fact that the goal of the project was to find the main process description and a proof of concept for the approach. The goal was not to increase the throughput time but to get insight into the main flows.

Business Process characteristics

	Excepted	Coded	
		High	Low
Variety	Low	1	2
Analyzability	Low	0	5
Interdependence	High	3	0

Process Mining related activities

	Expected	Coded
Preprocessing	Low	4
Discovery	High	9
Organizational	High	0
Performance	High	0

Process Mining success

	Coded
Model quality	4
Process impacts	2
Project efficiency	1

Figure 31 Interview 6 - Coding results

5.1.7 Interview 7

Experience

The interviewee has a background in data analytics and is employed at a consultancy company. Via a colleague he got in touch with Process Mining and has been involved in several Process Mining projects. The interviewee has good knowledge and experience with process optimization with Lean and other similar methods. He considers the fact based information of Process Mining to be the biggest advantage.

The Process Mining projects took place at Dutch municipalities where he managed the projects. It must be noted that the interviewee did not do the analyses in the Process Mining tool himself but he was closely involved in the choosing the type of analyses, interpreting the results and presenting the results to the stakeholders.

The focus of this interview is on one of the municipality projects. The goal of the project was to get insight into the invoicing process and find the bottlenecks. The process was already under review by a bigger project to improve the process structure.

Process characteristics

The process concerned both the procurement as the billing part of the process. The process was analyzed over seven departments which all had similar steps although there are no formal process models. Only protocols requiring invoices to be checked by two persons.

The managers received reports on throughput time and saw that it was too high, but he was not able to find the source of the bottleneck. The interviewee finds it hard to describe the amount of variety in the process since there is no formal process description but calls the process relatively simple. The process is small and therefore manageable.

Process mining related activities

The data had to come from different tables in the SAP system which supports the process. The system administrators delivered the data in several spreadsheets which had to be converted into one event log. Also the data in the system was not logged as part of a process step. A specific step contained ‘controller’ and a timestamp. After a discussion with the stakeholders of the process this appeared to be a step where the controller approves an invoice.

It is good to notice that the interviewee did not see the first visualizations of the process but does remember that the first results showed a spaghetti like process. After several iterations by cleaning the data and renaming steps more than 70 percent of the cases followed a standard path through the process.

The analyses done on the cases consisted of comparing several departments and to find a correlation between cases from a specific supplier and the throughput time.

Process Mining success

The final results were able to answer the questions of the stakeholders of the process, but also raised a lot more questions like ‘why is this bill always paid late?’ This would require further analysis which was out of scope of this assignment.

The added value of Process Mining in this case was twofold. The stakeholders of the process thought their process was really straightforward and executed according to a standard procedure. But in reality the cases showed a lot of deviation which surprised the stakeholders.

Second, the differences in throughput time between departments are far bigger than expected. The departments are executing the same process with the same steps but differ a lot in throughput time. By zooming in on the timestamps of the one department it appeared that a controller was the bottleneck of the process.

Scenario 'best'

Based on the amount of references (Figure 32 Interview 7 - Coding results) made and the description of the process characteristics, the case can be categorized as a 'best' scenario (Low variety, Low analyzability, High interdependence).

The process in this interview was considered an invoicing process at one department of the municipality. The process was expected to be executed in a comparable way in the different departments although no formal process descriptions were present. Remarkable is that benchmarking the different departments showed surprising differences in throughput time although the steps were practically the same. What is interesting to see is that this process is not coded as a high interdependent process and therefore does actually not fit exactly in one of the three Fit scenarios. A possible explanation is that the process was considered relatively small and therefore requires less coordination efforts.

It is good to notice that the interviewee saw the results of the Process Mining tool and was aware of the analysis that were produced but he did not perform the analysis in the tool himself. He was impressed by the results of the tool and mainly the objectivity of the analysis which he could confront the stakeholders with. The results raised many questions which might have influenced the acceptation of the results. Now the findings are based on their own data and therefore can be considered objective which was a great added value in this project. The results of the Process Mining tool were able to show the main behavior, the differences between the different departments and the effect of these deviations on the throughput time.

Business Process characteristics

	Excepted	Coded	
		High	Low
Variety	Low	0	2
Analyzability	Low	2	4
Interdependence	High	0	1

Process Mining related activities

	Expected	Coded
Preprocessing	Low	0
Discovery	High	1
Organizational	High	0
Performance	High	4

Process Mining success

	Coded
Model quality	1
Process impacts	2
Project efficiency	1

Figure 32 Interview 7 - Coding results

5.1.8 Interview 8

Experience

The interviewee has an education in both mathematics and industrial engineering and management. She got her Process Mining experience at a consulting company which made an enhanced Process Mining tool which is able to automatically extract large amounts of data from ERP-systems and perform advanced Process Mining based analytics on the event data. Her analytical skills based on her education and Lean/Six Sigma experience were useful during her Process Mining projects.

Via the consulting company she performed several international Process Mining projects on large data sets. In comparison with the previous interviews, the amount of data analyzed is far more voluminous. The processes she analyzed were all invoicing related and specifically procurement to pay or order to cash. The case which was highlighted during the interview took place at a large telecom provider and she was part of a bigger team.

Process characteristics

The project at the telecom provider took longer than the other projects the interviewee participated in due to the maturity of the process. The process was centralized in the sense that the procurement process was the same for all the establishments of the company in different countries. 'When an order is done of a specific category in Germany and is above a certain payment limit then it must be rerouted through Luxemburg and receive two approvals'.

By defining clear checks on how to route an order through a process the company reduced the variety during the process. The interviewee mentioned that it seemed like they were analyzing ten different variants of one process, where each variant handled a specific type of order. All business units in the different countries were all connected to each other to handle the orders on a companywide level. All process variants and options were literally documented in formal documentation.

Process mining related activities

The data of the company was stored in an ERP system which could be extracted into the Process Mining tool. All projects start with a so called 'quick scan' where the general statistics of the event data are analyzed on a high level to get a high overview standard questions are answered like: 'How big is the data set? How many distinctive paths are present? What is the variation in throughput time? What kind of information is available in the event log?' This allows you to get feeling with the process you are analyzing.

After these first insights were presented, the analytical team soon realized that this approach was of less interest to the telecom provider since they already had a quite advanced description and insight into their general processes. The interviewee and her team used the feedback of the telecom provider to sharpen the questions they had and let them formulate specific insights which they were looking for.

The telecom provider than provided the formal policies for each process and asked the interviewee to check these descriptions against the reality recorded in the system. Additionally the company provided a specific list of questions about the process which they were interested in. These questions were than analyzed in the Process Mining tool by zooming in on the specific part of the process and with the enhanced model.

Process Mining success

The main added value of the project was that the specific process descriptions were predefined and therefore could be checked against the reality recorded in the system. The telecom provider required a visualized model enhanced with performance statistics on throughput time and whether the specific policies, such as the 'four eye principle', were followed. With the help of the Process Mining tool the interviewee was able to give these insights.

Scenario 'specific'

Based on the amount of references (Figure 33 Interview 8 - Coding results) made and the description of the process characteristics, the case can be categorized as a 'specific' scenario (Low variety, High analyzability, High interdependence). The high analyzability was mentioned often which is not surprising because the process was described very detailed. Although there are many variants of the processes they were all often executed in the same way since every variant was described into detail.

Because this process was highly standardized the interviewee also needed more detailed questions to show the added value of analyzing processes with Process Mining. On the other hand all specific questions were answered and it was conformed that the process descriptions match the reality.

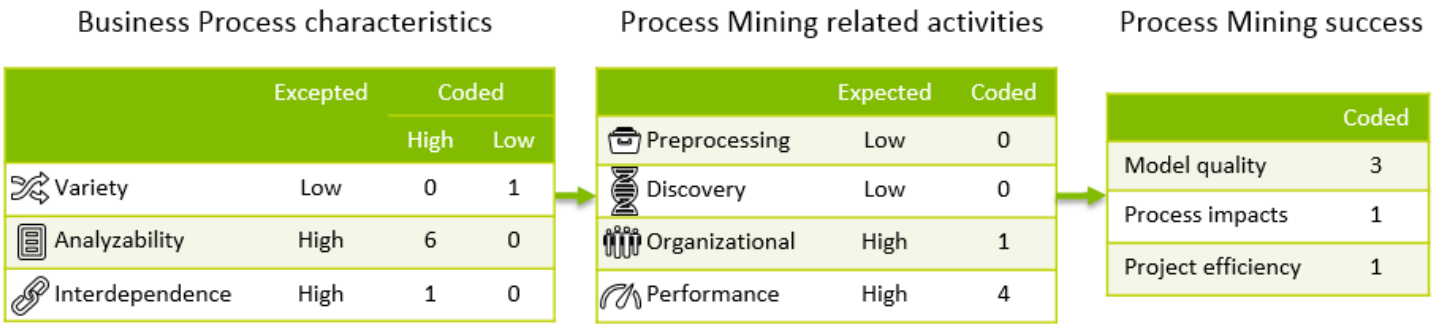

Figure 33 Interview 8 - Coding results

5.1.9 Interview 9

Experience

The interviewee studied Business Information Systems and got enthusiastic through the courses 'Business Modelling' and 'Business Improvement and Design' about Process Mining. His study gave him extensive knowledge on the subject of Process Mining which he applied during his graduation research.

During his graduation project he analyzed an application process at an asset financing company which provides revolving credits and personal loans. The process was part of a call center. The goal of the project was to find the bottlenecks and how to improve the process.

Process characteristics

The actors involved in the process are the customer, a call agent, an assessor, an acceptant and finally the administrative department which send out the proposals. The process was mainly done online but when a customer calls the call agent fills in the online form for the customer. The remaining process was identical.

The process does not allow for much deviation and is delimited. The roles of the actors in the process are defined and allow for little freedom which is also enforced by the information system supporting the process. When a problem occurs which a call agent could not handle he or she could ask one of the peers for guidance. This part is not specified but relies on experience.

Every actor in the process is aware of what other actors in the process are doing and therefore the knowledge of the complete process is considerably high. The call agents for example know where the assessor and the acceptant base their judgement on. This was partly due to the fact that all people were located in one department which is important for the success of the process. A good example is that the call agents inform clients to fill in the exact income since later on in the process, a payment check has to be send in. When the numbers do not exactly match the case is declined. This interdependence is known with the call agents and therefore prevents rework.

The process is managed based on specific KPIs such as the amount of proposal requests, how many were judged and how many led to a contract.

Process mining related activities

In theory the best way to execute Process Mining, according to the interviewee, is to predetermine how the event data of a process should be logged properly for Process Mining purposes and then gather the data. Mostly Process Mining is applied on historical available data but in his graduation project the interviewee was able to specify the required event data which was accumulated during three months.

The event data was directly loaded into the Process Mining tool. With the help of the fuzzy miner the interviewee was able to handle the noise in the system and aggregate to the main behavior. Most time then was spend on delivering management information to the stakeholders by selecting specific cases and analyzing the throughput time.

Process Mining success

The results of the project gave insight into the real execution of the process and answer questions like 'where are the bottlenecks? What are the unexpected things? What are the deviating cases and their root cause for deviation?' He was able to show the bottlenecks in the process and also found several

specific patterns on how to increase the success rate of the amount of contracts issued. Which was considered a great added value by the stakeholders. The most important thing, according to the interviewee, is a satisfied customer. The model should have added value and show exceptions or deviations which are interesting for the customer.

The general process was described on forehand but the interviewee was able to give a far more detailed description of the process, highlight the actors responsible for decisions and the underlying choices.

Scenario 'specific'

Based on the amount of references (Figure 34 Interview 9 - Coding results) made and the description of the process characteristics, the case can be categorized as a 'specific' scenario (Low variety, High analyzability, High interdependence). This process has clear boundaries and all actors involved have a clear defined role and therefore the process has a low Variety and a high Analyzability. The correct flow of the application of the process is dependent on the client providing the correct information. This is both enforced by the information system and the protocol the call agents follow. The codings in Figure 34 with Interdependence therefore show a balance in high and low.

The added value of Process Mining was on the one hand the detailed process map and on the other hand the quantified KPIs. The codings show that Performance related functionality added the most value. What is interesting to see is that Organizational Mining got no codings. It might be possible that the process model was already able to show how the actors in the process collaborating and therefore a handover matrix is less useful. But Organizational Mining also includes for example visualizing the organizational hierarchy which can be of added value in this case but was not mentioned.

Business Process characteristics

	Excepted	Coded	
		High	Low
Variety	Low	0	2
Analyzability	High	5	1
Interdependence	High	1	1

Process Mining related activities

	Expected	Coded
Preprocessing	Low	0
Discovery	Low	1
Organizational	High	0
Performance	High	4

Process Mining success

	Coded
Model quality	2
Process impacts	2
Project efficiency	1

Figure 34 Interview 9 - Coding results

5.1.10 Interview 10

Experience

The interviewee got his Process Mining experience through his education. He followed several courses and also graduated on the subject of Process Mining. He considers his Process Mining knowledge both practically and theoretically extensive. He has done several projects and is involved in all Process Mining steps.

The specific case chosen in this interview is done at T-Mobile on the mobile phone subscription activation process. The process was mainly executed by systems and the stakeholders were not sure on how the process was executed and where the bottlenecks were. Several subscriptions took significantly longer to be activated and they were interested in the root cause.

Process characteristics

The process starts when an already existing customer calls to extend his current subscription. This is often done when the subscription is close to an end or special reductions are given on extending the subscription.

The call agent enters the information into a CRM system. This system is connected to a provisioning system which sends the subscription details (such as the bundle information associated with the subscription) to a back-end system. This back-end system then activates or changes the subscription and sends an acknowledgement back to the provisioning system which updates the status in the CRM system.

Since the process is mainly automatically executed, it does not allow for much freedom and variety. The call agent has a description how to process the subscription into the system. The administrators of the system every two weeks analyzed the system for any subscriptions which were not processed. These were then processed by hand. Therefore the process can be considered interdepend since the wrong definition of the reduction on the subscription caused the back-end system to fail activating the subscription.

Process mining related activities

The first step included defining the questions which needed to be answered and collecting the data. The data originated from different systems and were combined into an event log with Access. The data was cleaned by relabeling steps, adding case IDs and correcting timestamps.

Once the data was transformed into one usable event log the interviewee used ProM to visualize the process and get a look at the main flow and its deviations. These deviations appeared to be the bottlenecks due to the fact that the discounts were wrongly processed by the system. Therefore these cases took longer on average. Normally a system administrator would fix these problems every two weeks.

By quantifying how often these cases happen and multiplying this with the costs of a customer calling, the total costs of these deviations were quantified.

Process Mining success

By visualizing the process in ProM the interviewee was able to find the deviations and where to look for the root cause. These deviations were quantified and gave the stakeholders insight both in how often this problem occurred and how to solve it.

Another advice was given on the amount of checks the system administrator did. It was considered acceptable for the system administrator to check the system every two weeks, but it could occur that a subscription was delayed with two weeks.

Scenario 'specific'

Based on the amount of references (Figure 35 Interview 10 - Coding results) made and the description of the process characteristics, the case can be categorized as a 'specific' scenario (Low variety, High analyzability, High interdependence).

The process was mainly automated which is interesting from a Process Mining perspective since it might be expected that a system does exactly what the documentation of the systems describe. In this case multiple systems were connected. The stakeholders of the process were aware that the throughput time of some cases was higher and could get stuck in the process but they were unable to find the root cause.

Interviewee was able to visualize the specific cases which had a higher throughput time compared to the main flow. This visualization led to finding the root cause of the problem.

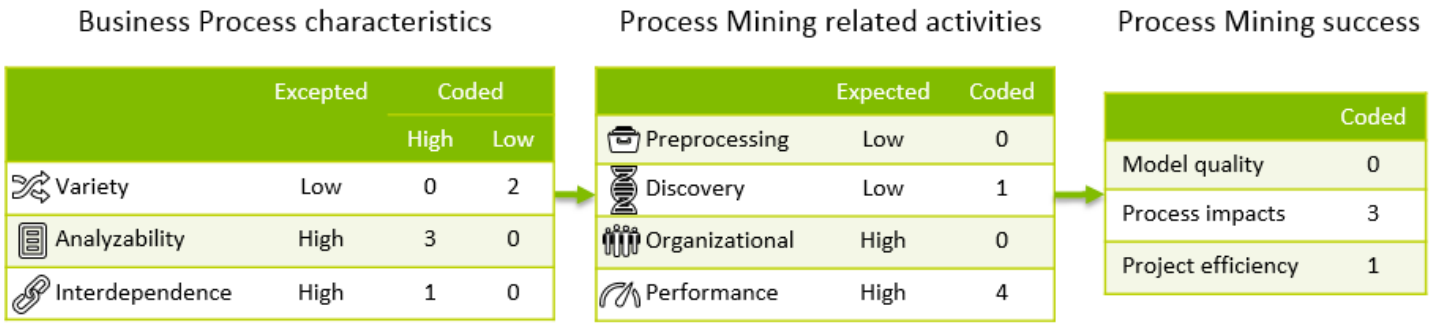

Figure 35 Interview 10 - Coding results

5.1.11 Interview 11

Experience

The interviewee graduated on process diagnosis with sequence diagrams before Process Mining existed. The dynamics and peculiarities of designing processes always fascinated him. Especially how discrepancies exist or grow between process descriptions and reality. After his graduation he started working in a different field but he regained his interest in the topic in 2008 when he read a paper how to apply Process Mining for IT auditing purposes. It had great resemblance with his graduation subject but instead of investing time in mapping sequence diagrams, the work was largely done automatically.

Since then he applied several Process Mining projects at the bank he is currently employed at. One of the projects is applied at the IT service desk with the goal to find the bottleneck and the potential savings. Due to the financial crises, the need for cost reduction was high and therefore the driver of the project. Also the proof of concept of Process Mining was an important part of the project.

The call center handles incoming phone calls about IT related problems. When a problem takes longer than 10 minutes to be resolved, the problem is labeled as an incident which is forwarded to special groups. The report of the call and the handling of the incident were analyzed with Process Mining.

Process characteristics

The process is supported by a workflow management system which enforces the actors to make a record of every call. The records of the incidents are directed to a problem solving group. It is important that the problems are forwarded to the correct team, otherwise in practice these problems take longer to be resolved. The wrong forwarding can happen due to the lack of oversight of the call agents about the current IT problems at hand.

An example of the lack of oversight is when a specific system goes offline and the IT support has already detected this issue but has had no time to inform the service desk. The service desk then receives many calls about the problem but is unable to inform the caller that the problem is already detected and being resolved. This causes for unnecessary long calls.

The system is designed to support all the communication between the service desk and the person reporting the problem, but in some occasions the employee of the service desk bypasses the system by directly sending emails to the problem reporter.

The incidents are also rated on the severity. Based on the severity of the problem there are specific protocols to handle the situation. These procedures and other work instructions are also taught during training sessions. The actors in the process know how to execute their own work but have only limited amount of knowledge of the complete process. The team leaders have a better understanding of the complete process and try to transfer their knowledge to the other actors.

The process is managed based on Service Level Agreements which has targets such as the satisfaction level and the throughput time. Such indicators are mainly based on output.

Process mining related activities

The interviewee had access to all data of the workflow system and based on the high level process descriptions they were able to find out which data is logged in the systems. To get a real grip on the exact logging behind a certain button or step, the interviewee needed the knowledge of the process and data manager. With help of the technical experts of the system extra attributes were added to the event log to enhance the model. Attributes such as whether the problem was reported via phone or mail.

A lot work has gone into defining steps in the process which can be marked as rework and therefore are waste in the process which has to be eliminated. The first process models showed a lot of bouncing behavior between actors. By sifting through the descriptions of the status updates the interviewee was able to really understand the meaning of the steps in the process. He also had help from process managers who both understood the process step as the recording step in the system.

When the data was again loaded into the Process Mining tool the steps in the model showed a good resemblance with reality. The interviewee was able to show the general steps and the average amount of steps needed to resolve a specific incident. The tool was able to visualize the bottlenecks which surprised the stakeholders since the bottlenecks were higher than expected.

Process Mining success

The project resulted in a solid business case which was based on facts extracted from the Process Mining tool. Although a lot of time was spent in analyzing the data and getting grip on the exact meaning of all steps and their logging, the added value of Process Mining was visualizing and quantifying the amount of deviations and their impact on the process.

Based on the enhanced process model with different attributes such as the severity of the problem, the interviewee was able to determine the baseline of how many steps a specific problem takes on average. When a case would take longer and include more steps than the average, the cases were further analyzed for the cause of the extra steps.

Scenario 'specific'

Based on the amount of references (Figure 36 Interview 11 - Coding results) made and the description of the process characteristics, the case can be categorized as a 'specific' scenario (Low variety, High analyzability, High interdependence).

The process has a good match with the specific scenario. The process is supported by a workflow system and process descriptions are present to guide the actors in the process. This limits the variety of the process and increases the analyzability. The call agent forwards the problem to a solving group. The Process Mining project evaluated these decisions and came to the conclusion that the call agents were able to make better decisions when they had better access to the current IT problems at hand. Once a problem was already identified the call agents are able to relate new incoming problems to problems which were already reported. This reduced the throughput time significantly.

The added value of Process Mining was the ability to show the deviations and quantify their results for a solid business case.

Business Process characteristics

	Excepted	Coded	
		High	Low
Variety	Low	1	2
Analyzability	High	8	1
Interdependence	High	6	0

Process Mining related activities

	Expected	Coded
Preprocessing	Low	0
Discovery	Low	1
Organizational	High	0
Performance	High	11

Process Mining success

	Coded
Model quality	3
Process impacts	5
Project efficiency	0

Figure 36 Interview 11 - Coding results

5.2 Combined results of all interviews

After all interviews are discussed separately in the previous sections, this section discusses the results of all interviews.

Figure 37 shows all coding results presented into one table. Based on these results there are three things which stand out.

First, the amount of coding references to Discovery and Performance in the Challenge Fit are higher than the model indicates. The Challenge Fit deals with cases where it is hard to define the process and which cases are part of the process and which are not. The model description in section 3.2 of the Challenge Fit describes that a Business Process with High Variety it might still be possible to find a homogenous group of cases. In the interviews all interviewees in the Challenge Fit were able to find a homogenous group. In essence this homogenous group can be further analyzed and can be considered a Best Fit.

This is also mentioned in the interviews with the hospital cases. They indicated that in detail their Business Processes are ultimately complex and are rarely executed in a same way, but when you start abstracting from the details and look at it from a more logistic perspective the Business Process becomes far simpler. Also, the patients itself make the Business Process complex and the Business Process itself from a logistic point of view is simpler and might even be considered a Specific Fit.

Second, Organizational Mining was hardly mentioned during the interviews. In literature Organizational Mining is described and applied in case studies and therefore it was included in the model. During the first part of the interview the interviewees were asked to elaborate on the Process Mining related activities which they are familiar with. Only two interviewees mentioned Organizational Mining as a Process Mining functionality. This might indicate that people are not aware of this functionality. Most of the interviewees do look at what actors are involved in cases but this is not considered Organizational Mining. Also the process model can have an overlap with the model of departments working together since several events in a Process are strongly correlated with a department. One specific instance in which Organizational Mining gave an interesting insight was the fraud analysis at the shipbuilder. This was because a specific check on the ‘four-eye principle’ was done which can be considered as an Organizational Mining related activity.

Finally, the model does show the different focus of added value of Process Mining per scenario. A Challenge scenario does focus more on Preprocessing and finding a homogenous group or a pattern in the event log. In these cases they were all able to find such a group and apply further analyses on this specific group. A Best scenario benefits both from visualizing the process model and enhancing the model with throughput time and other perspectives to analyze the Process. A Challenge scenario focuses more on the added value of throughput time and zooming on the specific deviations and the cause of their deviations. Some interviewees pointed out that it still might be useful to visualize the process model in a Challenge Fit because a Business Process almost always shows deviations. Even the Business Process has no deviations, stakeholders still appreciate the view of their Business Process being ‘in control’.

			Business Process characteristics						Process Mining related activities				Process Mining success			
			Variety		Analyzability		Interdependence		Preprocessing	Discovery	Organizational	Performance	Model quality	Process impacts	Project efficiency	TOTALS:
			High	Low	High	Low	High	Low								
Challenge	Hospital	Intestine cancer care	4	0	0	7	2	0	4	2	0	2	1	2	2	26
	DUO	Scholarship contact	3	1	0	8	2	0	4	2	0	2	0	2	2	26
	Hospital	Head Neck care chain	3	0	1	6	3	0	5	3	0	1	1	6	0	29
Best	Municipality	Invoicing	1	1	1	5	3	0	0	9	0	4	4	4	0	32
	Shipbuilder	Invoicing	1	4	1	5	1	0	0	2	2	4	2	4	1	27
	Hospital	Gall bladder removal	1	2	0	5	3	0	4	9	0	0	4	2	1	31
	Municipality	Invoicing	0	2	2	4	0	1	0	1	0	4	1	2	1	18
Specific	Provider	Invoicing	0	1	6	0	1	0	0	0	1	4	3	1	1	18
	Asset financing	Loan application call center	0	2	5	1	1	1	0	1	0	4	2	2	1	20
	Provider	Subscription activation	0	2	3	0	1	0	0	1	0	4	0	3	1	15
	Bank	IT service desk	1	2	8	1	6	0	0	1	0	11	3	5	0	38
	TOTALS:		14	17	27	42	23	2	17	31	3	40	21	33	10	280

Expected focus for added value

strong limited

Figure 37 Coding results of all interviews

5.3 Overview of the hypothesis

The results of the interviews are summarized in the previous section. These results show that the model is able to explain the difference in added value of Process Mining related activities for different Business Processes. This section gives the visual representations of the final model and the final Fit scenarios and there corresponding hypothesis. Figure 38 shows the initial research model.

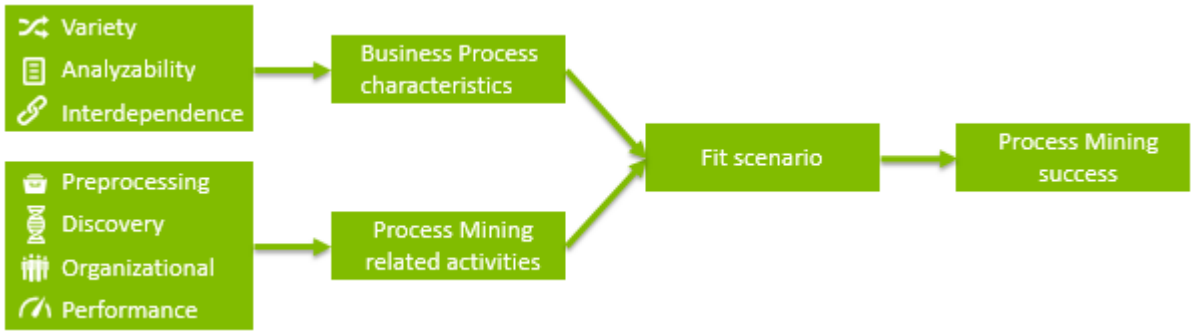

Figure 38 Initial Process Mining success model

Organizational Mining was included in the initial model as a Process Mining related activity which can deliver an added value. In the interviews in general Organizational Mining was hardly mentioned and therefore there is no support to either in- or exclude it from the Process Mining success model. The final model and the support is visualized in Figure 39.

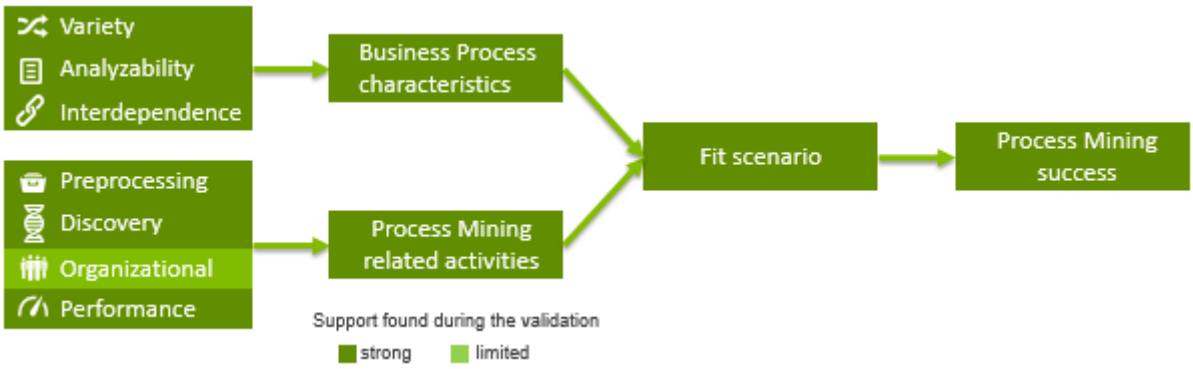

Figure 39 Final Process Mining success model

During the interviews the interviewees indicated that the naming of 'Challenge', 'Best' and 'Specific' Fit have a sentiment included. It could indicate that the 'Best' Fit is always the best situation to apply Process Mining. In practice this does not solely depend on the Business Process which is analyzed and therefore the labels of the fits are also changed. The labels are changed to match the Business Process which is analyzed. The labels are changed from 'Challenge', 'Best' and 'Specific' to 'Ad Hoc', 'Routine' and 'Standardized'.

In section 3 three propositions were defined concerning the relationships among the concepts. The concepts in the model were operationalized which makes the relationships measurable and therefore hypothesis. The results of the hypothesis are visualized in Figure 40 (Ad Hoc), Figure 41 (Routine) and Figure 42 (Standardized).

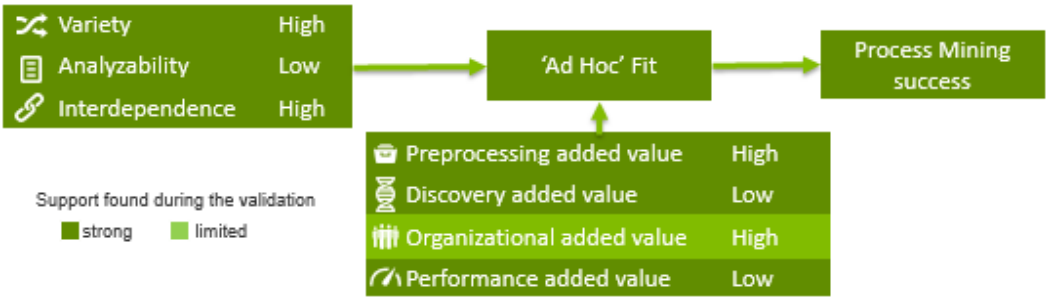

Figure 40 Ad Hoc Fit scenario

Hypothesis 1*: In an 'Ad Hoc' Fit scenario the project gets the added value of Process Mining related activities from Preprocessing and Organizational Mining to get Process Mining success*

There was no support found during the interviews for the added value of Organizational Mining. The remainder of the scenario is supported with the addition that an 'Ad Hoc' Fit can become a 'Routine' Fit when the project is able to reduce the Variety in the Business Process by finding a homogenous group of cases.

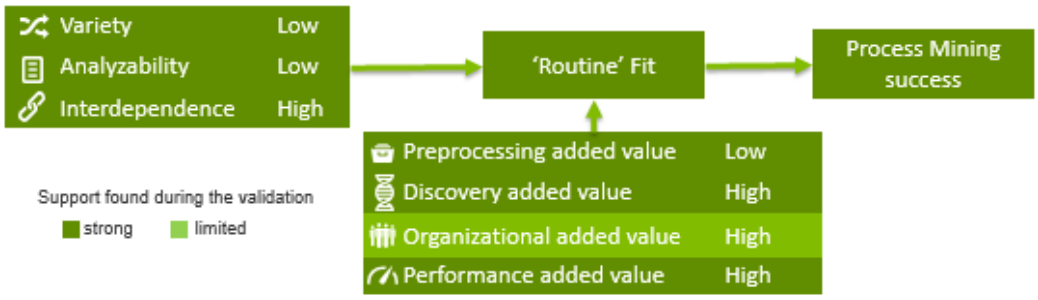

Figure 41 Routine Fit scenario

Hypothesis 2: *In a 'Routine' Fit scenario the project gets the added value of Process Mining related activities from Discovery Analysis, Organizational Mining and Performance Analysis to get Process Mining success*

There was no support found during the interviews for the added value of Organizational Mining. The remainder of the scenario is supported.

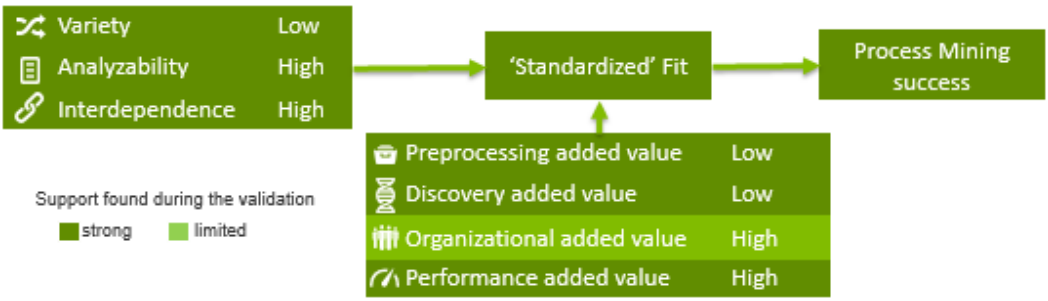

Figure 42 Standardized Fit scenario

Hypothesis 3: *In a 'Standardized' scenario the project gets the added value of Process Mining related activities from Organizational Mining and Performance Analysis to get Process Mining success*

There was no support found during the interviews for the added value of Organizational Mining. The remainder of the scenario is supported.

6 Conclusion

This final chapter summarizes the findings and the answers to the research questions, followed by the implications of this research both for practice and theory. Then the limitations of this research are stated with suggestions for further research.

6.1 Research questions

The main research question of this research is:

"How can the Process Mining success be explained with the Business Process-Process Mining related activities Fit?"

The main question is divided into four sub questions which together answer the main question.

Sub Question 1: *"Which Process Mining related activities are used in practice?"*

The literature was searched for Process Mining case studies which extensively describe the Process Mining related activities and the context of an appliance in practice. The literature describes three Process Mining methodologies. Based on the Process Mining case studies and the methodologies a grouping is made of four Process Mining related activities; Preprocessing the event log, Process Discovery, Organizational Mining and Performance Analysis.

In the interviews only a few references were made to Organizational Mining. This can be explained due to the fact that Organizational Mining has less exposure in practice than in literature. Therefore there is no support to either discard or include Organizational Mining in the model.

Sub Question 2: *"What are Business Process characteristics which are suitable for applying Process Mining?"*

The Process Mining case studies found in sub question one, were analyzed for Business Process characteristics. The literature was searched for characteristics which are often used in related fields such as BPM and process modeling. Three Business Process characteristics were found which are mentioned across the related research; Variety, Analyzability and Interdependence.

These characteristics were operationalized according to the related research and described in a measurement model. Zigurs & Buckland (1998) suggest to define the most relevant Fit scenarios in which the characteristics have a predefined value. All combinations of the differences in characteristics would lead to eight distinct scenarios. Therefore three scenarios were defined based on the occurrence of these scenarios in Process Mining case study literature:

- Ad Hoc: High Variety, Low Analyzability and High Interdependence
- Routine: Low Variety, Low Analyzability and High Interdependence
- Standardized: Low Variety, High Analyzability and High Interdependence

Sub Question 3: *"How can Process Mining success be measured?"*

The research of Mans et al. (2014) resulted in three success measures: Model Quality, Process Impacts and Project Efficiency. These measures were further specified with the success measures from the Process Mining case studies. The success measures were specifically mentioned by the interviewees as the success of their Process Mining project and therefore useful to indicate the Process Mining success.

Sub Question 4: *"Which Process Mining related activities-Business Process characteristics Fit leads to Process Mining success?"*

Based on the three scenarios which describe a combination of Business Process characteristics, the expected focus for added value of Process Mining related activities is described to reach Process Mining success. The three Fit scenarios were first based on the Process Mining case studies and then tested with interviews with Process Mining experts.

After the interviews, the collected data was analyzed with a coding scheme based on the measurement model. All three scenario Fits give insight which can lead to Process Mining success.

The answer to the main question is that by grouping the Business Processes based on the characteristics Variety, Analyzability and Interdependence three Fit scenarios Ad Hoc, Routine and Standardized give insight into the added value of Process Mining related activities to get Process Mining success.

6.2 Implications for practice

The Process Mining success model gives insight into the added value which Process Mining related activities can deliver based on characteristics of a Business Process.

Process Mining has become more popular recently but since the technology is already present for about 15 years in academic research, it is interesting to look why the adoption of Process Mining has not gone any faster. Many companies struggle with the complexity of their Business Processes and state that BPM is one of their biggest priorities (Schäfermeyer et al., 2012). Process Mining is typically a technique which can decompose the complexity of a Business Process.

One of the factors is that the appliance of Process Mining in practice requires investments in the sense of resources and knowledge, and therefore the organizations want insight into the added value which Process Mining can deliver (Ronny S. Mans et al., 2014). This research contributes to getting insights into the added value of Process Mining.

Based on the characteristics of the Business Process which an organization wants to analyze, they can use the model to look which Process Mining related activities are used in practice and what insights can be gained. A Business Process which can be considered Ad Hoc should focus on finding a homogenous group of cases and finding a main flow. By discarding the rest of the cases, the organization can reduce the Variety of the Business Process. It is important to notice that further analysis (such as throughput time) of this group is than subject to bias.

When a Business Process can be considered Routine then it is interesting to visualize the main flow and find the deviations. If the event log contains enough information such as timestamps and actors, then the effect of the deviations can be quantified and can results in an objective and to the point business case for improvements. The goal is to find the elephant trails where actors deviate from the main flow for some reason. The Process Mining tool actually gives insight into the decision points of the actors and the choices they make, which increases the Analyzability of the Business Process.

A standardized Business Process already has a low Variety and a high Analyzability due to clear process descriptions and the actors involved are process aware. Then a Process Mining project would profit from having a clear and specific question or hypothesis to answer. Visualizing the process model might give insight where to look for solution but the added value of Process Mining is to quantify the hypothesis.

6.3 Implications for theory

The field of Process Mining research is considered relatively young. Most of the research has gone into developing better algorithms and new techniques to find specific constructs which are known in process modelling but are hard to produce based on event logs. van der Aalst et al. (2007) started with applying Process Mining outside the academic environment to face the complexity of Business Processes in practice. Several Process Mining methodologies have been written to give practitioners grip in how to apply Process Mining and get to actual useful insights (Bozkaya et al., 2009; Măruşter & Beest, 2009; Rebuge & Ferreira, 2012).

Both the methodologies and the development of new Process Mining tools which are less focused on academic applications but more on industrial applications have caught attention of organizations. With the rising interest in practice it is interesting to look for the success factors and measures of applying Process Mining in organizations. Mans et al. (2014) looked at the success factors and measures of applying a Process Mining project. The factors and measures are based on studies from related fields such as Process Modelling and Data Mining.

This research focused on finding a Fit model between Business Process characteristics and Process Mining related activities to explain Process Mining success. For research it is interesting to know which factors influence Process Mining success. The Business Process characteristics are based on literature which looks at the factors influencing the success of standardizing and capturing the complexity of Business Processes (Mani et al., 2010). This is also often the goal for Process Mining projects and therefore these factors are also interesting for Process Mining research.

Also, this research is validated with a broad selection of interviews which looks at Process Mining projects analyzing Processes in different industries and different Process Mining tools. Therefore the implications of this research are not limited to a specific set of Processes and Process Mining tools.

6.4 Limitations and suggestions for further research

Although the model is generally considered interesting and usable by the interviewees, the model and the research has several limitations.

For this research the author made the decision to choose to either classify Business Process characteristics to be high or low, while in practice all characteristics should be put on a scale. The characteristics are tapped into existing literature which should allow other researchers to continue research and validate the results based on statistical research. This will also allow to measure the exact effect of the characteristics.

The interviews were also done with one person per project. This person might have a bias towards the project and therefore give a subjective view on the results of the project and the added value of Process Mining. With most of the interviews extra materials were available (such as reports of the results) and seen by the author of this research. No indication was found for the author to think the results were strongly biased. Yet statistical research which also includes all stakeholders of the project will rule out the chance of bias.

Next, Process Mining is considered a tool for analyzing Business Processes. Advanced techniques are able to continuously monitor Processes and even give real life predictions. Also Process Mining tools can help guiding a process improvement program to create support among the stakeholders and help as a change management instrument. But since in practice relatively few Process Mining projects have gone beyond the analysis and executed the improvement opportunities, this research was limited to the analysis and identifying the improvement opportunities. Further research into the effects of using

Process Mining during an improvement program opposed to brown paper sessions and other traditional analytics, might give interesting results.

In this research the interviews and their cases were not judged based on the amount of effort and resources invested in the Process Mining project. Further, interviewees mentioned that it was important to have a clear goal or hypothesis to start your Process Mining project. Almost every interviewee mentioned the importance of applying several iterations from cleaning the data, visualizing the event log and again clean the data based on the results. The effect of these factors on applying Process Mining have been mentioned in the summaries of the interviews but are not quantified. Further research is needed to analyze the effects of the resources invested in Process Mining and the need for having a clear hypothesis before starting the project.

Finally, one of the success factors which was mentioned with all interviews was the data quality and the knowledge of how the data can be interpreted into Process data. Eventually every interviewee was able to handle the poor data quality and find sufficient data expert knowledge, but this does affect the effort needed to get Process Mining results. This research did no go any further into the quality of the data on which the interviewees based their project. While poor data quality will require extra effort in cleaning the data, the results might be more valuable since the data was not available before. The effect of the quality of the data on Process Mining projects is therefore an interesting topic for further research.

7 References

Bandara, W., Gable, G. G., & Rosemann, M. (2005). Factors and measures of business process modelling: model building through a multiple case study. *European Journal of Information Systems*, *14*(4), 347–360. doi:10.1057/palgrave.ejis.3000546

Bandara, W., Indulska, M., Chong, S., & Sadiq, S. (2007). Major Issues in Business Process Management: An Expert Perspective. *ECIS 2007 - The 15th European Conference on Information Systems*, *2007*, 1240 – 1251. Retrieved from http://eprints.qut.edu.au/14345/

Bhattacherjee, A. (2012). Social science research: principles, methods, and practices. Creative Commons Attribution 3.0. Retrieved from http://scholarcommons.usf.edu/cgi/viewcontent.cgi?article=1002&context=oa_textbooks

Bose, R. P. J. C., Mans, R. S., & Van Der Aalst, W. (2013). Wanna improve process mining results? In *Proceedings of the 2013 IEEE Symposium on Computational Intelligence and Data Mining, CIDM 2013 - 2013 IEEE Symposium Series on Computational Intelligence, SSCI 2013* (pp. 127–134). doi:10.1109/CIDM.2013.6597227

Bozkaya, M., Gabriels, J., & Werf, J. (2009). Process Diagnostics: A Method Based on Process Mining. *2009 International Conference on Information, Process, and Knowledge Management*, (1). doi:10.1109/eKNOW.2009.29

Daft, R. L., Lengel, R. H., Science, S. M., & May, O. D. (1986). Organizational Information Requirements, Media Richness and Structural Design, *32*(5), 554–571.

De Weerdt, J., Schupp, A., Vanderloock, A., & Baesens, B. (2013). Process Mining for the multi-faceted analysis of business processes - A case study in a financial services organization. *Computers in Industry*, *64*(1), 57–67. doi:10.1016/j.compind.2012.09.010

Dishaw, M., & Strong, D. (1999). Extending the Technology Acceptance Model with Task- Technology Fit Constructs. *Information & Management*, *36*(1), 9–21. doi:10.1016/S0378-7206(98)00101-3

Goedertier, S., De Weerdt, J., Martens, D., Vanthienen, J., & Baesens, B. (2011). Process discovery in event logs: An application in the telecom industry. *Applied Soft Computing Journal*, *11*(2), 1697–1710. doi:10.1016/j.asoc.2010.04.025

Goodhue, D. L., & Thompson, R. L. (1995). Task-technology Fit and Individual Performance. *MIS Quarterly*, *19*(2), 213–236.

Jans, M., Van Der Werf, J. M., Lybaert, N., & Vanhoof, K. (2011). A business process mining application for internal transaction fraud mitigation. *Expert Systems with Applications*, *38*(10), 13351–13359. doi:10.1016/j.eswa.2011.04.159

Lillrank, P. (2003). The quality of information. *International Journal of Quality & Reliability Management*, *20*(6), 691–703. doi:10.1108/02656710310482131

Lillrank, P. (2003). The Quality of Standard, Routine and Nonroutine Processes. *Organization Studies*, *24*(2), 215–233. doi:10.1177/0170840603024002344

Mani, D., Barua, A., & Whinston, A. (2010). An Empirical Analysis Of The Impact Of Information Capabilities Design On Business Process Outsourcing Performance. *MIS Quarterly*, *34*(1), 39–62. doi:Article

Mans, R., Reijers, H., Berends, H., Bandara, W., & Rogier, P. (2014). Business process mining success. In *Proceedings of the 21st European Conference on Information Systems* (pp. 1–13). Association for Information Systems. Retrieved from http://eprints.qut.edu.au/67333/1/Mans_et_al._2013._Business_Process_Mining_Success.pdf

Mans, R. S., Reijers, H. A., Berends, H., Bandara, W., & Rogier, P. (2013). Business process mining success. In *Proceedings of the 21st European Conference on Information Systems* (pp. 1–13). Association for Information Systems. Retrieved from http://eprints.qut.edu.au/67333/1/Mans_et_al._2013._Business_Process_Mining_Success.pdf

Mans, R. S., Reijers, H. A., Berends, H., Bandara, W., Rogier, P., & Prince, R. (2014). Business process mining success. In *Proceedings of the 21st European Conference on Information Systems* (pp. 1–13). Association for Information Systems. Retrieved from http://eprints.qut.edu.au/67333/1/Mans_et_al._2013._Business_Process_Mining_Success.pdf

Mans, R. S., Schonenberg, M. H., Song, M., Van der Aalst, W. M. P., & Bakker, P. J. M. (2009). PROCESS MINING IN HEALTHCARE - A Case Study. In *Communications in Computer and Information Science, volume 25* (pp. 425–438). doi:10.1007/978-3-540-92219-3_32

Mans, R., & Schonenberg, H. (2008). Process mining techniques: an application to stroke care. *Studies in Health …*, *136*, 573–578. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18487792\nhttp://schonenberg.info/files/SHTI136-0573.pdf

Mans, R., Schonenberg, M. H., Song, M., Aalst, W. M. P. van der, & Bakker, P. J. M. (2011). Application of Process Mining in Healthcare – A Case Study in a Dutch Hospital. In *Communications in Computer and Information Science* (Vol. 127, pp. 224–237). doi:10.1007/978-3-642-18472-7

Măruşter, L., & Beest, N. R. T. P. (2009). Redesigning business processes: a methodology based on simulation and process mining techniques. *Knowledge and Information Systems*, *21*(3), 267–297. doi:10.1007/s10115-009-0224-0

Miles, M. ., & Huberman, A. . (1994). *Qualitative data analysis: An expanded sourcebook (2nd ed.)*. *Qualitative data analysis: An expanded sourcebook (2nd ed.).* Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1995-97407-000&site=ehost-live

Rebuge, Á., & Ferreira, D. R. (2012). Business process analysis in healthcare environments: A methodology based on process mining. *Information Systems*, *37*(2), 99–116. doi:10.1016/j.is.2011.01.003

Reijers, H. a., & Liman Mansar, S. (2005). Best practices in business process redesign: An overview and qualitative evaluation of successful redesign heuristics. *Omega*, *33*(4), 283–306. doi:10.1016/j.omega.2004.04.012

Rozinat, a., de Jong, I. S. M., Günther, C. W., & van der Aalst, W. M. P. (2009). Process mining applied to the test process of wafer scanners in ASML. *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, *39*(4), 474–479. doi:10.1109/TSMCC.2009.2014169

Rozinat, a., Mans, R. S., Song, M., & van der Aalst, W. M. P. (2009). Discovering simulation models. *Information Systems*, *34*, 305–327. doi:10.1016/j.is.2008.09.002

Saunders, M., Lewis, P., & Thornhill, A. (2009). *Research Methods for Business Students*. *Research methods for business students*.

Schäfermeyer, M., Rosenkranz, C., & Holten, R. (2012). The impact of business process complexity on business process standardization: An empirical study. *Business and Information Systems Engineering*, *4*(5), 261–270. doi:10.1007/s12599-012-0224-6

Song, M., & van der Aalst, W. M. P. (2008). Towards comprehensive support for organizational mining. *Decision Support Systems*, *46*(1), 300–317. doi:10.1016/j.dss.2008.07.002

Van der Aalst, W. M. P. (2009). Process-aware information systems: Lessons to be learned from process mining. *Transactions on Petri Nets and Other Models of Concurrency II*, 1–26. doi:10.1007/978-3-642-00899-3_1

Van der Aalst, W. M. P. (2011). Process Mining: Discovery, Conformance and Enhancement of Business Processes. In *Media* (Vol. 136, p. 352). doi:10.1007/978-3-642-19345-3

Van Der Aalst, W. M. P., & Günther, C. W. (2007). Finding Structure in Unstructured Processes: The Case for Process Mining. In *Proceedings - 7th International Conference on Application of Concurrency to System Design, ACSD 2007* (pp. 3–12). IEEE. doi:10.1109/ACSD.2007.50

Van der Aalst, W. M. P., Reijers, H. a., Weijters, a. J. M. M., van Dongen, B. F., Alves de Medeiros, a. K., Song, M., & Verbeek, H. M. W. (2007). Business process mining: An industrial application. *Information Systems*, *32*(5), 713–732. doi:10.1016/j.is.2006.05.003

Vergidis, K., Tiwari, A., & Majeed, B. (2008). Business Process Analysis and Optimization: Beyond Reengineering. *IEEE Transactions on Systems, Man, and Cybernetics, Part C*, *38*(1), 69–82. doi:10.1109/TSMCC.2007.905812

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review, *26*(2).

Weske, M. (2007). *Business Process Management: Concepts, Languages, Architectures*. Springer. Retrieved from http://books.google.nl/books?hl=nl&lr=&id=QMyu_B1KTZIC&oi=fnd&pg=PA2&ots=pZ8TqNXlB-&sig=_f8QxZGrd2Y18j2JnF8L2QRoVHk#v=onepage&q&f=false

Wolfswinkel, J. F., Furtmueller, E., & Wilderom, C. P. M. (2011). Using grounded theory as a method for rigorously reviewing literature. *European Journal of Information Systems*, *22*(1), 45–55. doi:10.1057/ejis.2011.51

Yin, R. K. (2009). *Case Study h Researc Design and Methods Fourth Edition*. *Applied Social Research Methods Seiries* (Vol. 5). Retrieved from http://cemusstudent.se/wp-content/uploads/2012/02/YIN_K_ROBERT-1.pdf\nISBN 978-1-412296099-1

Zigurs, & Buckland, B. (1998). Task / Technology Group Support Systems Defining task. *MIS Quarterly*, *22*(3), 313–334.

List of figures

Appendix A - Process Mining case studies

	Author	Scenario	Process	Event log	Fitness	Functionality
1	**(W. M P van der Aalst et al., 2007)**	best	Invoice WfMs	14,279 cases 147,579 event	0,9	- Filter noise and incompleteness - Heuristics miner - C-Net - Throughput time performance
2	**(Song & van der Aalst, 2008)**	specific	Invoice WfMs	570 cases 3,023 events		- Handover metrics
3	**(R. Mans & Schonenberg, 2008)**	challenge	Healthcare Intake, 4 hospitals	368 + 234 cases		- Combining event logs - Heuristics miner - Throughput time performance
4	**(R.S. Mans et al., 2009)**	challenge	Healthcare billing process	627 cases 376 events	Good	- Clustering - Heuristic miner - Fuzzy miner - Handover metrics
5	**(Bozkaya et al., 2009)**	best	Government Document Management	83,611 cases 276,333 events	0,81	- Filtering incomplete cases - Heuristics miner - Dotted chart - Handover metrics - Social network
6	**(Măruşter & Beest, 2009)**	specific	Gas capacity booking		0,99	- Data merged from separate systems - Filter out noise and incompleteness - Heuristic miner - Fuzzy miner - Throughput time performance - Bottleneck visualization
		specific	Government fine collecting WfMs	130,136 cases	0,97	- Heuristics miner - Fuzzy miner - Throughput time performance - Bottleneck visualization
		challenge	Decision Support System	500 cases	0,5	- Filtering outliers
7	**(Rozinat, de Jong, et al., 2009)**	challenge	ASML machine test process	24 cases 154,966 events		- Aggregating small steps into job steps - Heuristics miner
8	**(Goedertier et al., 2011)**	challenge	Second line customer service WfMs	17,812 cases	0,97	- Cluster to find a meaningful case - Filter noise and incompleteness - Genetic miner - Heuristics miner - AGNE miner - Throughput time performance - Bottleneck visualization
9	**(Jans et al., 2011)**	specific	Transaction fraud SAP system	10,000 cases 62,531 events	0,8	- Methodology used of Bozkaya 2009 - Filter cases to make an understandable model - Fuzzy mining - Handover metrics - Specific checks for fraud
10	**(Rebuge & Ferreira, 2012)**	challenge	Hospital emergency care WfMs	627 cases		- Cluster logs into 7 logs based on similar behavior - Heuristics miner - Throughput time performance - Dotted chart - Handover metrics - Specific compliance issue check
11	**(De Weerdt et al., 2013)**	best	Document management system for life insurance brokers	4,491 cases 44,880 events	77,5	- Separating event logs for different purposes - Heuristics miner - Handover metrics - Benchmarking - Specific compliance check - Throughput time performance

Appendix B - Interview structure

Introduction

- State the appreciation for the interview and get permission for an audio recording
- State that the interview will be transcribed and the findings will be presented for approval
- Give a short introduction into the research topic
- Give a short overview of the structure of the interview
- State that the questions will be asked in English but the answers can be given in Dutch

Part 1 – Context questions

Experience of the Process Miner related questions

1. How did you get acquainted with Process Mining?
2. How would you define your Process Mining knowledge?
    a. In terms of related education
    b. In terms of practical experience

Success factors

3. What would you consider to be success factors of Process Mining?

Process Mining functionality related questions

4. Can you describe the general steps to apply Process Mining?
    a. Do you use a specific methodology?
5. What kind of Process Mining functionality would you use with each step?'

Business Process characteristics

6. Can you describe characteristics of a Business Process that you consider relevant for Process Mining?
7. How do you think they influence your Process Mining approach?

Process Mining success

8. What would you consider to be the results of Process Mining?
9. How would you measure the results of Process Mining?

Part 2 – Practical specific questions

Case related questions

10. Can you describe specific situations in which you applied Process Mining?
    a. What was the reason for applying Process Mining?
11. How would you describe the context of the project?

Business Process characteristics

12. Can you describe the Process?
    a. What people are involved?
    b. To what degree are the people allowed to choose their own way to reach the goal of the process?

c. How often do people deviate from the process descriptions or rules and why?

d. When a problem occurs, do people follow a strict protocol or do they rely on experience or interpretation?
e. How would you describe the knowledge of the people of the complete process?

f. How do the Information system(s) support the process?
g. How is the process managed in the sense that the goals are met?
h. To what degree do decisions made in the beginning of the process influence decisions which have to be made further on in the process?

Process Mining functionality

13. Can you describe the steps taken?
14. Can you describe which functionality you have applied during the steps?

Fit

15. Why did you apply this functionality?
16. What were the results of the functionality applied? **(PM success)**
17. If you could do the project over again, would you apply any other functionality and why?

Process Mining success

18. What were the final results of the analysis of the process?

Part 3 – Testing the model

Depending on the Process specifics in the case before, the questions may vary

19. How would the project change in the sense of approach, Process Mining functionality and outcomes of the project when…
    a. The process would be *more/less* executed in the same way?
    b. The people in the process would have *more/less* knowledge of the Process?
    c. The different parts of the process would be *more/less* dependent of each other?

Fits

20. Do you think that a Process Mining project analyzing a process which is
    a. Rarely executed in the same way and it is not possible to find a homogenous process group which is trivial,
    b. people in the process mostly rely on experience and interpretation
    c. and all the parts of the process are highly dependent of each other

    Should focus mostly on
    d. clustering and filtering the event log and apply organizational mining

    To achieve Process Mining success?
21. Do you think that a Process Mining project analyzing a process which is
    a. Often executed in the same way,
    b. people in the process mostly rely on experience and interpretation
    c. and all the parts of the process are highly dependent of each other

    Will have less focus on clustering and filtering the event log, but will focus on
    a. visualizing the main process and its deviations
    b. organizational mining

c. quantifying findings, look for root causes of deviations and answer specific questions/KPIs about the process

To achieve Process Mining success?

22. Do you think that a Process Mining project analyzing a process which is
    a. Often executed in the same way,
    b. people in the process have good knowledge of the complete process and follow protocols
    c. but all the parts of the process are highly dependent of each other

    Will have less focus on clustering and filtering the event log and visualizing the main process and its deviations, but will focus on
    d. quantifying findings, look for root causes of deviations and answer specific questions/KPIs about the process

    To achieve Process Mining success?

Ending

- State appreciation
- State that the notations of the interview will be send for approval