FEDERAL UNIVERSITY OF THE STATE OF RIO DE JANEIRO
CENTER OF EXACT SCIENCES AND TECHNOLOGY
SCHOOL OF APPLIED INFORMATICS

Process Mining Applied to Process Redesign in a Domestic Navigation Business Company

Wesley da Silva Santos

Supervisor
Sean Wolfgand Matsui Siqueira

RIO DE JANEIRO, RJ – BRAZIL
FEBRUARY OF 2021

Computerized cataloging by the author

d 237

da Silva Santos, Wesley

Process Mining Applied to Process Redesign in a Domestic Navigation Business Company / Wesley da Silva Santos. -- Rio de Janeiro, 2021.

59

Orientador: Sean Wolfgand Matsui Siqueira.

Trabalho de Conclusão de Curso (Graduação) - Universidade Federal do Estado do Rio de Janeiro, Graduação em Sistemas de Informação, 2021.

1. Process Mining. 2. Process Redesign. 3. Case Study. 4. Domestic Navigation Business. I. Wolfgand Matsui Siqueira, Sean, orient. II. Título.

“*And life's gonna drop you down like the limbs of a tree*

It sways and it swings and it bends until it makes you see”

Madonna, Henry, J., Price, S., Jump, Confessions on a Dance Floor (2005).

Process Mining Applied to Process Redesign in a Domestic Navigation Business Company

Wesley da Silva Santos

Undergraduate Project presented at the School of Applied Informatics at the Federal University of the State of Rio de Janeiro (UNIRIO) to obtain the title of bachelor’s in Information Systems.

Approved by:

SEAN WOLFGAND MATSUI SIQUEIRA (UNIRIO)

BRUNA DIIRR GONÇALVES DA SILVA (UNIRIO)

HENRIQUE PRADO DE SÁ SOUSA (UNIRIO)

RIO DE JANEIRO, RJ – BRAZIL

FEBRUARY OF 2021

ACKNOWLEDGMENTS

When I decided to move on to Rio de Janeiro in 2014, I couldn’t imagine 10% of the things that would happen to me. Much has changed in me during these 6 years, and now it's so strange to look back and think that at that time of my life, I was a fresh high school graduated boy from a small rural town, and now I'm (almost) getting my bachelor’s degree in a federal university.

During my undergraduate course, I strangely imagined this section of my graduation work. What would be written, the people who would be mentioned, in short, all the details. This is one of my favorite parts of reading a work because I believe that here is recorded the most human part of the one who did it during the study, not only of the work developing but also of the undergraduate / master's / doctorate experience. Thus, I’d like to invite you to know a part of me and my transformations during my graduation period.

I would like first to thank my parents and my brother for all their support during this journey. It is difficult to measure how grateful I am to you. I know that sometimes I take ungrateful and exaggerated postures, and I would like to apologize for imploding these attitudes with you in times of pressure/despair. I am working on these issues internally and improving on this aspect every day!

I would like to thank my aunts Gilvaneide Vargas, Marilene Jeronimo, my uncle Moisés Cardoso and my cousins Moisés Junior, Luis Fernando, and Darley Santos, who welcomed me and gave me a home to live for most of my graduation period. Thank you for all your patience and friendship throughout this period. Without you, I would not have achieved what I have achieved so far!

To all my friends that supported me, from UNIRIO or that I met in life during this graduation period. I will not mention names because there were so many people, and you were all fundamental to my personal development. During my graduation period, unfortunately, I had to distance myself from many friends, and I apologize for the refusals of parties, fun, and meetings. I had to work hard and give up so much to accomplish all the achievements that I wanted/needed in graduation.

About these achievements, none of them I would have achieved if I had not counted on the infrastructure, faculty, and administrative staff of my college. I am extremely grateful to all public servants at UNIRIO, who work hard for the university to produce and teach with high quality. I am grateful to PRAE and CNPq for the scholarships that were essential for my stay in college while developing essential life skills.

In particular, I would like to thank some teachers who signed partnerships during my time at UNIRIO and provided me with out-of-the-curve learning in course monitoring projects in some subjects. They were: Bruna Diirr, Claudia Cappelli, Geiza Hamazaki, Gleison Santos, Mariza Almeida, and Paulo Sérgio. Thanks for everything.

During this period of graduation, I've been in touch with some professors that made possible for me to have a greater type of collaboration outside the classroom working on research projects. I would like to thank the teachers Fernanda Baião and Flávia Santoro for being my mentors for the most part of this journey. You were essential for my professional and academic development, and I have no words to describe how grateful I am for all the affection, friendship, and conversations!

Also, I would like to thank professor Sean Siqueira for tutoring me for the past two years and for accepting to be my advisor in this final project. Thank you for all your friendship, for the laughs, and, above all, for being patient and dedicated to knowledge transfer. I hope we have a lasting partnership collaborating on future projects!

Finally, I would like to thank everyone from Log-In Logística Intermodal who participated directly or indirectly in executing this work, particularly Vinicius Abreu and Leandro Barral. Professors Henrique and Bruna for agreeing to participate in my examining board and you who have read so far. I hope you enjoy reading this work as I enjoyed developing it.

SANTOS, Wesley da Silva. **Mineração de Processos Aplicada no Redesenho de Processos em uma Empresa de Cabotagem** . 2021. Trabalho de Conclusão de Curso (Graduação em Sistemas de Informação) – Escola de Informática Aplicada, Universidade Federal do Estado do Rio de Janeiro, Rio de Janeiro, 2021.

RESUMO

As técnicas de mineração de processos usam dados de eventos com muitos propósitos, como descobrir processos, verificar conformidade, analisar gargalos, comparar variantes de processos e sugerir melhorias. Devido ao estabelecimento de várias empresas de software que investem em produtos comerciais, a mineração de processos tem sido cada vez mais adotada e incorporada como parte das iniciativas de BPM por empresas em diferentes domínios. Assim, tornou-se relevante não apenas compreender, mas interpretar os resultados da utilização da mineração de processos nas práticas de gestão empresarial à luz do referencial teórico. O objetivo deste trabalho é apresentar e discutir os resultados de um estudo de caso de adoção da mineração de processos em uma empresa de navegação. O foco deste estudo foi a aplicação de técnicas de mineração de processos no redesenho do processo de “solicitação de pagamento de faturas por fornecedores de serviços”, e também, como consequência, propor seu redesenho e melhorias. Os resultados do estudo de caso foram positivos e permitiram o entendimento do processo rapidamente, possibilitando redesenhá-lo e sugerir melhorias. Nossas contribuições são: a discussão de um projeto real de mineração de processos como potencial fonte de aprendizado sobre a evolução de um negócio; e a descrição e análise de um caso de mineração de processos seguindo um método bem estruturado já proposto na literatura.

Palavras-chave: Mineração de Processos, Redesenho de Processos, Estudo de Caso, Cabotagem

SANTOS, Wesley da Silva. **Process Mining Applied to Process Redesign in a Domestic Navigation Business Company**. 2021. Undergraduate Project (Information Systems Bachelor) – School of Applied Informatics, Federal University of the State of Rio de Janeiro, Rio de Janeiro, 2021.

ABSTRACT

Process mining techniques use event data for many purposes, such as discovering processes, checking compliance, analyzing bottlenecks, comparing process variants, and suggesting improvements. Due to the establishment of several software companies investing in commercial products, process mining has been increasingly adopted and incorporated as part of BPM initiatives by enterprises in different domains. Thus, it became relevant not only to understand but also to interpret the results of process mining within business management practices in light of theory. This work aims to present and discuss results from a case study of the adoption of process mining in a navigation business company. This study's focus was applying process mining techniques in the redesign of the "invoice payment request by service suppliers" process, and moreover how, as a consequence, it was possible to propose its redesign and improvements. The results of the case study were positive and led to quick understood of the process, making it possible to be redesigned and suggest further improvements. Our contributions are twofold: the discussion of a real case process mining project as a potential source of learning about a business evolution; and the description and analysis of process mining case following a well-structured method already proposed in the literature.

Keywords: Process Mining, Process Redesign, Case Study, Domestic Navigation Business

Index

Image Index

1 Introduction

1.1 Motivation

Process models play an important role in the redesign and implementation of processes within Process-Aware Information Systems (PAIS) (van der Aalst, 2012). On the other hand, data produced by a PAIS is mainly used to support the enactment and monitoring as well as process analysis. Process mining techniques use event data for many purposes, such as discovering processes, conformance checking, analyzing bottlenecks, comparing process variants, and suggesting improvements (van der Aalst, 2016).

According to van der Aalst (2012), process mining is “both data-driven and process-centric”; therefore, many conformance and performance questions can be answered by applying a combination of event data and process models. Lately, due to the establishment of several software companies investing in commercial products, process mining has been more and more adopted and incorporated as part of BPM initiatives by enterprises in different domains (Dumas et al., 2019). Some of them have already been reported in the literature (Fernández-Llatas et al., 2015; Rebuge et al., 2012). However, despite the practical perspective, there is a lack of a sound theorizing about process mining (Langley et al., 2013).

1.2 Scenario and Problem Statement

Log-In is a Brazilian company that offers full door-to-door logistic solutions: planning, management, and operation for cargo movement using domestic navigation, complemented by roads. The company operates an integrated network that facilitates movement between ports and door to door transportation due to an extensive intermodal network, which allows for a geographical range in Brazil and Mercosur (The Southern Common Market formed by Argentina, Brazil, Paraguay, Uruguay, and Venezuela).

The intermodal transportation is characterized by using two or more transportation modals (maritime, road, air, and rail) to reduce logistics costs. Since containers can be easily transferred from one modal to another, the result is optimizing the cargo displacement and reducing risks involving accidents or losses. Each container is an independent cargo unit, with dimensions that follow an international standard. The measure adopted to indicate the capacity is known by the acronym TEU, which stands for Twenty feet Equivalent Unit.

Domestic Navigation is the navigation service between ports in the same country, in opposition to the long course navigation, which is executed between different countries' ports. It is considered a promising modal since Brazil has an extensive navigable coast, and the main cities, industrial locations, and big consuming centers are concentrated at the coastline or near cities. Compared to road and railroad transportation, in terms of cost, cargo capacity, and environmental impact, domestic navigation becomes a viable alternative to comprise the various sectors´ supplier chain. Log-In has invested more than 200 million dollars in Domestic Navigation, with a total capacity of 16,700 TEUs. It has more than 1,500 clients in Mercosur and operates with its own and chartered container ships.

The digitalization of the the process “invoice payment request by service suppliers” emerged as a demand from the Operational Cost Management area at Log-In. This process was considered poorly structured and with gaps in the interactions between the areas involved. Many emails were lost in the inbox of the process participants, and, as a result, updates on some requests to suppliers’ status were not sent. There was also a backlog of uncharged invoices due to the Enterprise Resource Planning (ERP) lack of integration. In 2017, the Financial department started a project to implement this process on a Business Process Management System (BPMS) aiming to digitalize the process. The areas involved were Procurement, Operational Cost Management, and Information Technology.

The initial scope was to create a portal through which the Log-In suppliers could start and follow the process. However, the process was not properly documented, and some typical implementation problems showed up: nobody knew the real end-to-end process; knowledge silos in each area; weak integration with other systems; different types of cases addressed the same way; and low efficiency. When the IT member of the team involved in the digitalization of this process left the company and demands for implementation of new services were requested, it was necessary to go over the process and associated subprocesses from the start. The challenge faced was understanding the process and how the data changed during its evolution along time to elicit the necessary requirements to implement the new services.

The application of process mining seemed to be an alternative to understanding the as-is situation and data/control behavior to conduct more objective interviews with users and validate hypotheses faster and in a greater amount. However, during this project execution, many unknown outcomes about the process were discovered, and in-depth discussions about them made it possible to suggest further improvements.

1.3 Objective

The goal of this work is to present and discuss the results from a case study of the adoption of process mining in the domestic navigation business company Log-in, in the light of the work of Grisold et al. (2020), who argues that process mining supports theorizing about changes in organizations and how its features can reveal insights about the dynamics of organizational routines. The present research focus was applying process mining techniques in the understanding and redesign of the process “invoice payment request by service suppliers”.

We aim to answer the three following research questions: **RQ1:** How can process mining be integrated into a company's IT landscape? **RQ2:** What is the business impact of process mining? **RQ3:** What strategic implications emerge from process mining within enterprises? The answers to the RQs allowed extracting propositions that support theorizing both in process mining as a research area in BPM and in the business domain in which it was applied.

The contributions are twofold: (i) the discussion of a real case process mining project as a potential source of learning about business evolution; and (ii) the description and analysis of a process mining case following a well-structured method proposed in the literature.

1.4 Structure

The present work is structured in chapters and, in addition to the first one, will be developed as follows:

- Chapter II: presents the theoretical background, discusses related work, and explains the methodological approach.
- Chapter III: describes the details of the case study.
- Chapter IV: discusses the results against the theory.
- Chapter V: presents some final remarks, limitations and future work.

2 Preliminaries

2.1 Theoretical Background

2.1.1 Business Process Management (BPM)

Dumas et al. (2013) state that Business Process Management (BPM) is a discipline that aims to study the way work "is performed in an organization to ensure consistent outcomes and to take advantage of improvement opportunities". BPM is a multidisciplinary field of knowledge that incorporates principles, methods, and tools to design, analyze and monitor business processes. These tools are useful to perceive handover of work, conformance checking, operational monitoring, improvements, etc. The discipline can be sum up as a continuous cycle comprising seven macro phases, as figure 1 depicts.

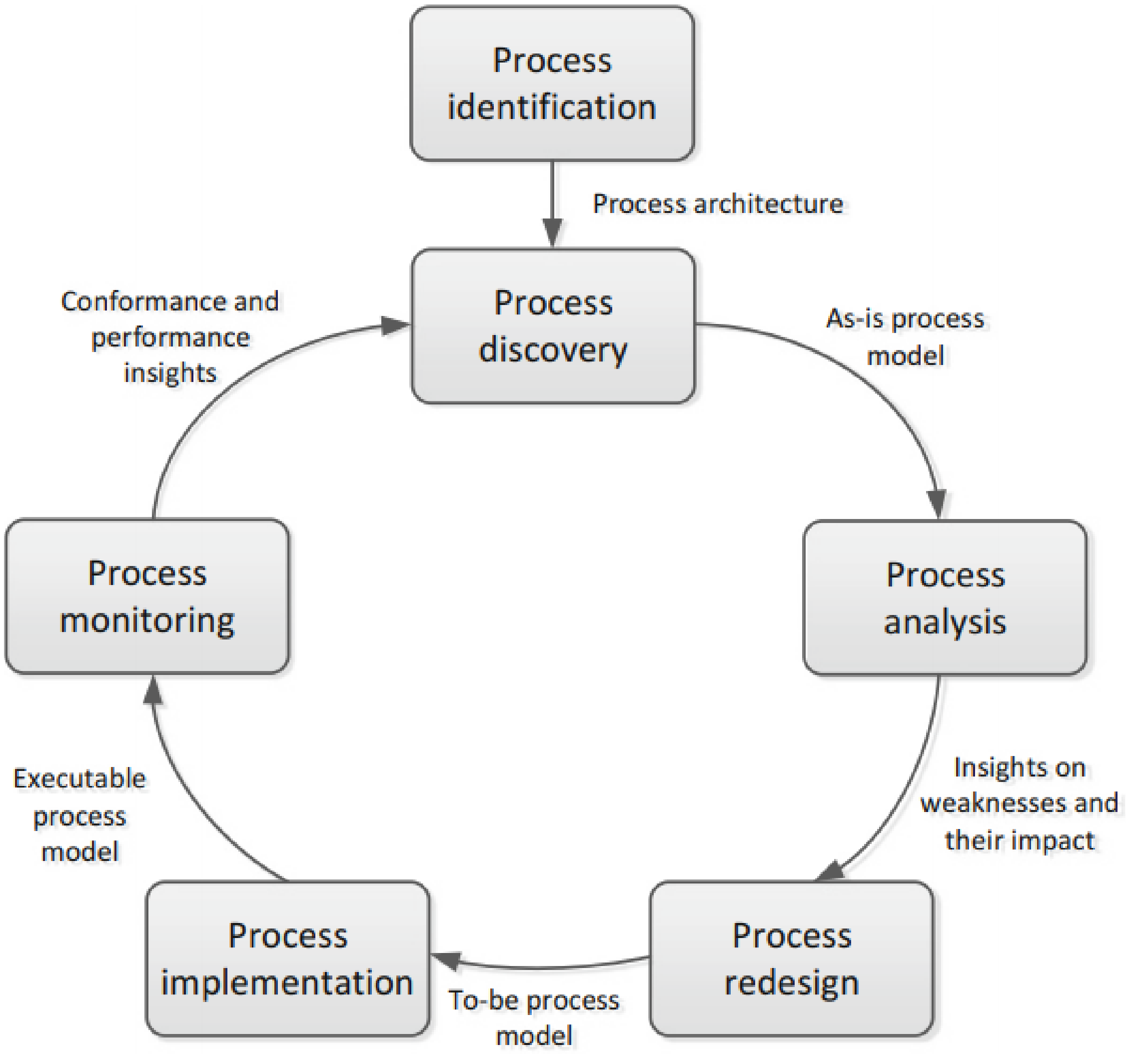

Figure 1: The BPM lifecycle (Dumas et al., 2013)

The phases can be summarized as follows:

- *Process identification*: phase where a faced business problem is selected, and then relevant processes related to the problem are identified. The main outcome of this phase is the process architecture, which is useful for improvement prioritization.
- *Process discovery*: this phase is dedicated to collect relevant artifacts about the processes selected in the process identification phase. Such artifacts can be process models and documentation. In case there's no artifacts about the processes it is common to follow other approaches to discoveries, being the most common the execution of interviews with process participants using a computational tool for process modeling.

- *Process analysis*: once the process discovery is over, the process analysis is done with the main goal to analyze the as-is situation of the selected processes, focusing in documenting the issues and performance measures. A possible outcome from this phase is a collection of issues.
- *Process redesign*: improvements are suggested based in the collection of issues discovered in the analysis phase. Such improvements are extracted using analysis techniques and the main output is a to-be process model.
- *Process implementation*: this phase is dedicated to implementing the changes proposed in the process redesign phase. The improvements may impact in many aspects of the process (i.e., enhancement of an IT system, addition of more participants in the scope of the process leading to an organizational structure change, etc.).
- *Process monitoring*: once the suggested changes are implemented, this phase focus in analyzing the data produced by the process to discover how the process is running. A possible output of this step is a collection of new issues, leading the cycle to the start phase.

These authors define business processes as a "collection of inter-related events, activities and decision points that involve several actors and objects, and that collectively lead to an outcome that is of value to at least one customer" (Dumas et al., 2013). Events are the occurrences that activate and finish a business process. They trigger the execution of a sequence of activities or decision points in a process. Decision points refer to points in the process when decisions are made that determine how a process is run. Figure 2 presents a conceptual model of business process ingredients and their relations.

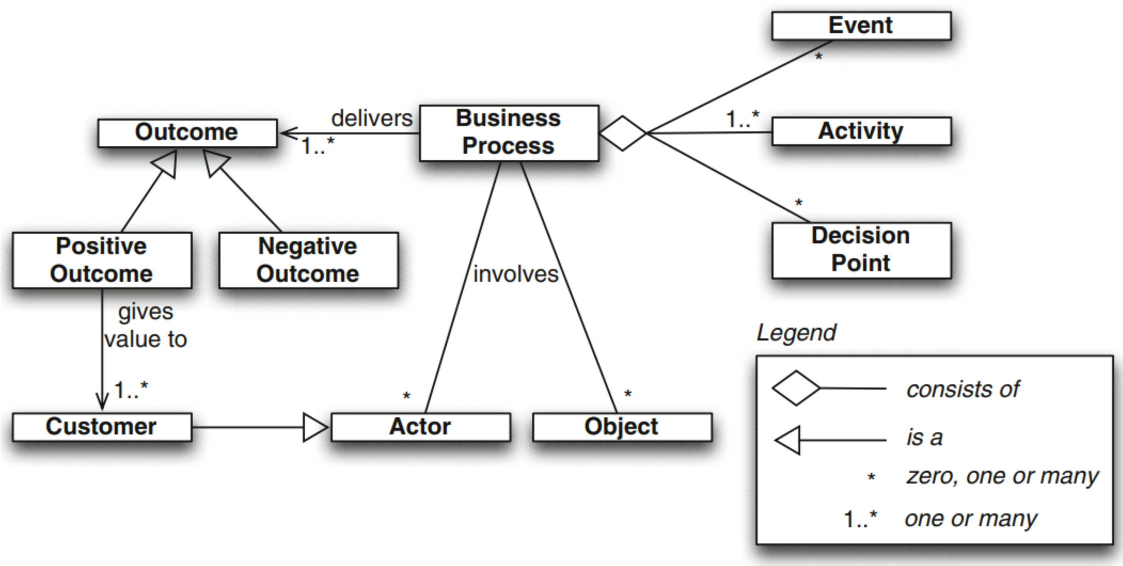

Figure 2: Ingredients of a business process (Dumas et al., 2013)

The main standard notation to represent business processes is the Business Process Modeling and Notation (BPMN). This notation represents concepts related to processes (i.e., actors, activities, message exchange, data objects, etc.). These concepts are useful to represent the knowledge of a process, decision points, handover of work between areas, etc. A simple model of an order-to-cash process is presented bellow in figure 3. In this figure, events, activities and sequence flows are represented by circles, rounded rectangles and arcs, respectively. However, BPMN has many more elements, such as gateways that are represented using diamond shapes.

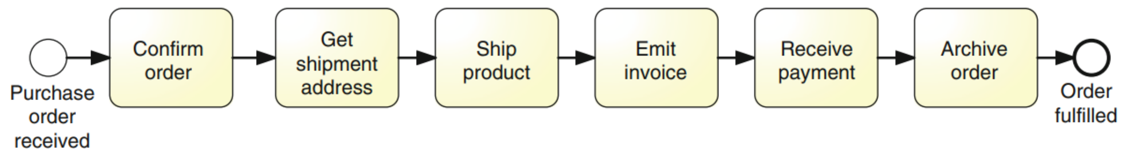

Figure 3: A model of a simple order-to-cash process

Processes often store and use data for many purposes. They can be stored in relational databases, non-relational databases, or common data repositories. The representation formats of data related to processes may vary since it can be stored like computer data files (i.e., XML, JSONs, TXTs, etc.) or physical objects formats (i.e., printed or drawn diagrams, documentation, etc.). Also, data produced by processes can be used by data mining algorithms to extract knowledge and insights about how a process is run.

2.1.2 Process Mining

According to van der Aalst (2016), Process Mining is a discipline within BPM and Data Mining disciplines that aims to discover, monitor, and improve processes supported by information systems by extracting knowledge from their event logs. Process mining includes, but is not limited to, process discovery, conformance checking, social/organizational mining, case prediction, and history-based recommendation.

The process mining algorithms combined with process logs are used to understand, analyze and improve processes through knowledge extraction from data. The event log is a data structure that contains case-level data from the process to be analyzed, and it is crucial for process mining to work. Such data can be related to actors, activities, activities' timestamp, information system through which the activity is executed, among many others.

The use of process mining as a computational method to support theorizing about change in organizations was argued by Grisold et al. (2020), who stated that process mining could uncover the ways work is being carried out within organizations and how changes take place over time. There are three reasons for this argument: first, process mining generates visual information about a routine; second, digital trace data might be "hidden" to observers or are not explicitly mentioned in interviews; and third, activities that might have been disguised could be discovered. Their claim was grounded on Organizational Change theories and Process Theorizing (Berente et al., 2019; Langley et al., 2013; Van de Ven et al., 2005; Volkoff et al., 2007).

According to Grisold et al. (2020), "process mining can support inductive and deductive theorizing by iterating between action traces from event logs and derived patterns of actions", and they highlight that some algorithms can be used for both. Additionally, the authors affirm that computational theorizing cannot replace human theorizing; thus, the insights from process mining alone are not enough to explain how or why changes occur, mainly because it does not reflect all actors' intentions, reasoning, or feelings. Consequently, it should be combined with interpretive approaches and, besides, the researchers would need to apply a pre-existing set of concepts, constructs, and their implied relationships as so-called a lexicon (Berente et al., 2019), which works as a basis to support interpreting data and relating it to scientific discourse.

2.1.2.1 eXtensible Event Strem (XES) Format

The standard format for representing process logs supported by most process mining applications is eXtensible Event Stream (XES[1]). The format is a grammar for the tag-based language that aims to provide designers of information systems with a unified and extensible methodology for capturing systems behaviors from event logs and event streams. XES is an XML-like format used for representing a digitalized process behavior. A sample of the used event log in this project is depicted bellow, presenting the main features of the XES file format:

```xml
<?xml version="1.0" encoding="UTF-8" ?>
<!-- This file has been generated with the OpenXES library. It conforms -->
<!-- to the XML serialization of the XES standard for log storage and -->
<!-- management. -->
<!-- XES standard version: 1.0 -->
<!-- OpenXES library version: 1.0RC7 -->
<!-- OpenXES is available from http://www.openxes.org/ -->
<log xes.version="1.0" xes.features="nested-attributes" openxes.version="1.0RC7">
    <extension name="Time" prefix="time" uri="http://www.xes-standard.org/time.xesext"/>
    <extension name="Lifecycle" prefix="lifecycle" uri="http://www.xes-standard.org/lifecycle.xesext"/>
    <extension name="Concept" prefix="concept" uri="http://www.xes-standard.org/concept.xesext"/>
    <classifier name="Event Name" keys="concept:name"/>
    <classifier name="(Event Name AND Lifecycle transition)" keys="concept:name lifecycle:transition"/>
    <string key="concept:name" value="Invoice payment process v3.csv (filtered on simple heuristics)"/>
    <trace>
        <string key="concept:name" value="SP2019-1958"/>
        <event>
            <string key="concept:name" value="Request invoice payment"/>
            <string key="END_RESOURCE" value="Sistema"/>
            <date key="time:timestamp" value="2019-04-08T13:53:00.000-03:00"/>
            <string key="START_RESOURCE" value="Marjorie Andrade"/>
            <string key="lifecycle:transition" value="start"/>

        </event>
        <event>
            <string key="concept:name" value="Request invoice payment"/>
            <string key="END_RESOURCE" value="Sistema"/>
            <date key="time:timestamp" value="2019-04-08T13:56:00.000-03:00"/>
            <string key="START_RESOURCE" value="Marjorie Andrade"/>
            <string key="INVOICE_AMOUNT" value="749,92"/>
            <string key="lifecycle:transition" value="complete"/>
        </event>
        ...
    </trace>
    ...
</log>
```

Figure 4: Fragment of a XES file

Figure 4 presents a sample of two events of the process, start and ending of an activity. Each event has its context attributes, such as resource, timestamp, activity name and lifecycle transition. These attributes can be used for many purposes, such as activity performance. The potential of process mining techniques depends on the context attributes of every activity.

[1] IEEE Standard for eXtensible Event Stream, available at: https://ieeexplore.ieee.org/document/7740858

2.1.2.2 Petri Nets

Petri nets are often used in the context of process mining. Various frameworks employ Petri nets as the internal representation used for process mining after processing the process log. Petri net is a graphic representation of workflows that allows concurrency modeling. It is possible to model a Petri net with a triplet N = (P, T, F), where P is a finite set of *places*, T is a finite set of *transitions* such that $P \cap T = \emptyset$, and $F \subseteq (P \times T) \cup (T \times P)$ is a set of directed arcs, called the flow relation. A marked Petri net is a pair (N,M), where N = (P,T,F) is a Petri net and where $M \in B(P)$ is a multi-set over P denoting the marking of the net. The set of all marked Petri nets is denoted N (van der Aalst, 2012). An example of a Coloured Petri Net, a subtype of Petri Net, is shown in the Conformance Checking application (Figure 18 and Figure 19).

2.1.3 Applications and Frameworks Used

With the popularization of Process Mining in the industry and academy, many frameworks and applications suites focusing on process mining solutions were developed. A comparison of the existing solutions could be found in 'Process Mining Software Comparison[2] website developed by the Chair of Digital Industrial Service Systems at the University of Erlangen-Nürnberg. In this work, we used the free open-source process mining suites: ProM[3], Apromore[4] , and bupaR[5]. The three solutions were combined to deliver satisfactory outputs to fulfill the goals of the project. The details about the used plug-ins are depicted in the case study section (section 3).

2.1.3.1 ProM

ProM is an open-source free process mining framework implemented in Java that supports a wide variety of process mining techniques in the plug-in formats to extract valuable knowledge from process logs (Rozinat et al., 2006). ProM uses a plug-in architecture allowing users to choose which techniques, algorithms, or visualization patterns they want. To exemplify, one can choose the alpha algorithm in the discovery phase while others choose for the heuristic miner. In this project, we used most of the

[2] Process Mining Software Comparison, available at: https://www.processmining-software.com/
[3] https://www.promtools.org/
[4] https://apromore.org/
[5] https://www.bupar.net/

plug-ins and techniques available in the framework, from process discovery to conformance checking.

2.1.3.2 Apromore

Apromore was an open-source process model repository that evolved into a complete process mining suite (Conforti et al., 2015), offering rich process mining capabilities grounded in state-of-the-art research to drive digital transformation and operational excellence. Apromore is a more mature process mining framework having two versions: open-source and commercialized. The commercialized features include premium features such as collaborative process mining. In this project, we used the open-source version to mine process models and plot statistics that are neither present in ProM or BupaR.

2.1.3.3 BupaR

bupaR is an open-source, integrated suite of R-packages for the application of a variety of process mining techniques (Janssenswillen et al., 2019). It consists of packages focusing on different perspectives of process mining related activities (i.e., data handling for creating and manipulating event logs, process visualizations in different perspectives; process dashboards, and exploratory and descriptive analysis). In this project, we used the process visualization and analysis features to do conformance checking.

2.2 Related Work: Process Mining in Industry

The application of process mining techniques has been described in the literature in different domains. Researchers start to present some conclusions that could be useful to foster further research and support new projects and findings. We discuss some of them in the following.

van der Aalst et al. (2007) confirmed the applicability of process mining in a case study of provincial offices of the Dutch National Public Works Department, responsible for the construction and maintenance of the road and water infrastructure. The authors applied techniques to analyze the handling of invoices sent by the sub-contractors and suppliers. Despite the outcomes that process mining could provide, one important consideration of this case was the relevance of engaging with the process owner and stakeholders to carry out a meaningful analysis.

De Weerd et al. (2013) agree on the need for real-life case studies discussing the conduct of process mining. The authors proposed a framework for a multi-faceted analysis of real-life events and evaluated it in a case study focused on the back-office process of a Belgian insurance company. Their results corroborated that process mining techniques can address organizational challenges by recommending improvements and promoting process awareness.

Jans et al. (2013) discussed the use of process mining by internal and external auditors to rethink how auditing should be carried out. In their paper, the authors remarked that only when auditors and researchers understand process mining benefits, this approach will become achievable in this domain.

Ruschel et al. (2020) applied process mining techniques to develop a probabilistic model in Bayesian Networks integrated into predictive models. The probabilistic model would simulate changes in the event occurrence probability, and several different scenarios can be visualized to support managers in scheduling maintenance activities. The results of this case study indicate that production losses can be reduced through optimally defined intervals between maintenance inspections.

There are also many health domain cases, such as described in Badakhshan et al., 2020, Battineni et al., 2019, Mans et al., 2008, Yoo et al., 2016. Mans et al. (2008) applied process mining techniques to discover typical paths followed by certain groups of patients in a hospital, in a real case of a gynecological oncology sector of a Dutch hospital. The authors analyzed the process from three different perspectives: control flow, organizational, and performance. The results obtained reinforced that process mining provided insights to support the improvement of existing care flows. Yoo et al. (2016) used process mining technology to analyze process shifts in a hospital environment due to diverse environmental changes to measure their effects in terms of consultation wait time, time spent per task, and outpatient care processes. The authors achieved some interesting conclusions like the total time spent in outpatient care did not increase substantially if compared to that before constructing a new building and that the consultation wait time had decreased. In Battineni et al. (2009), the idea was to obtain an unconventional event log of care flows to improve performance. The authors concluded that authorities should include process mining as a technique to support decision-making, whilst it is possible to use their proposal to organize and design hospital units.

Those papers provided examples of industry case studies that collected evidence of process mining applicability in real situations. We observe that in general: (a) they do not follow a systematic approach or methodology; so, it is difficult to replicate a similar case; (b) although the authors list limitations, the project outcomes are positive, so there is a lack of failure studies, which could not be necessarily true; and, (c) the complexity of the cases vary according to the domain, which does not ease reusing, e.g., research questions or the choice for specific algorithms to be applied. This case study explores the gaps of these previous studies.

2.3 Methodological Approach

We adopted the case study research method since we are dealing with observing a contemporary phenomenon, and we have no control over the events. We followed the steps suggested by Runeson and Höst (2009):

a. Case study design: objectives are defined, the case study is planned. This case study aimed to perform a process mining project to collect evidence of its applicability in practice and make propositions. Moreover, we also used interviews with process participants to confirm our findings and obtain more information about the process' evolution. Ten people directly involved in the process participated in the interviews: 1 Operational Costs Coordinator, 3 IT Analysts, 1 IT Assistant, 1 Business Consultant, 1 Supply Chain Analyst, 2 Operational Costs Analysts, 1 Operations Analyst. The project's execution followed a method proposed in the literature, and the details are provided in Section 3.
b. Preparation for data collection: procedures and protocols for data collection are defined. We identified the data sources to discover and analyze the process and elected the algorithms to be applied.
c. Collecting evidence: The data was collected and processed, generating different models that served as sources for analysis and interpretation. Moreover, we also

used interviews with process participants to confirm our findings and obtain more information about the evolution of the process.

d. Analysis of collected data. The information from the interviews and models/graphics generated were discussed, and propositions about the process mining project and the business domain were made.

Figures 5, 6, 7 and 8 represents the process flow of each macrophase of this case study with its main inputs and outputs.

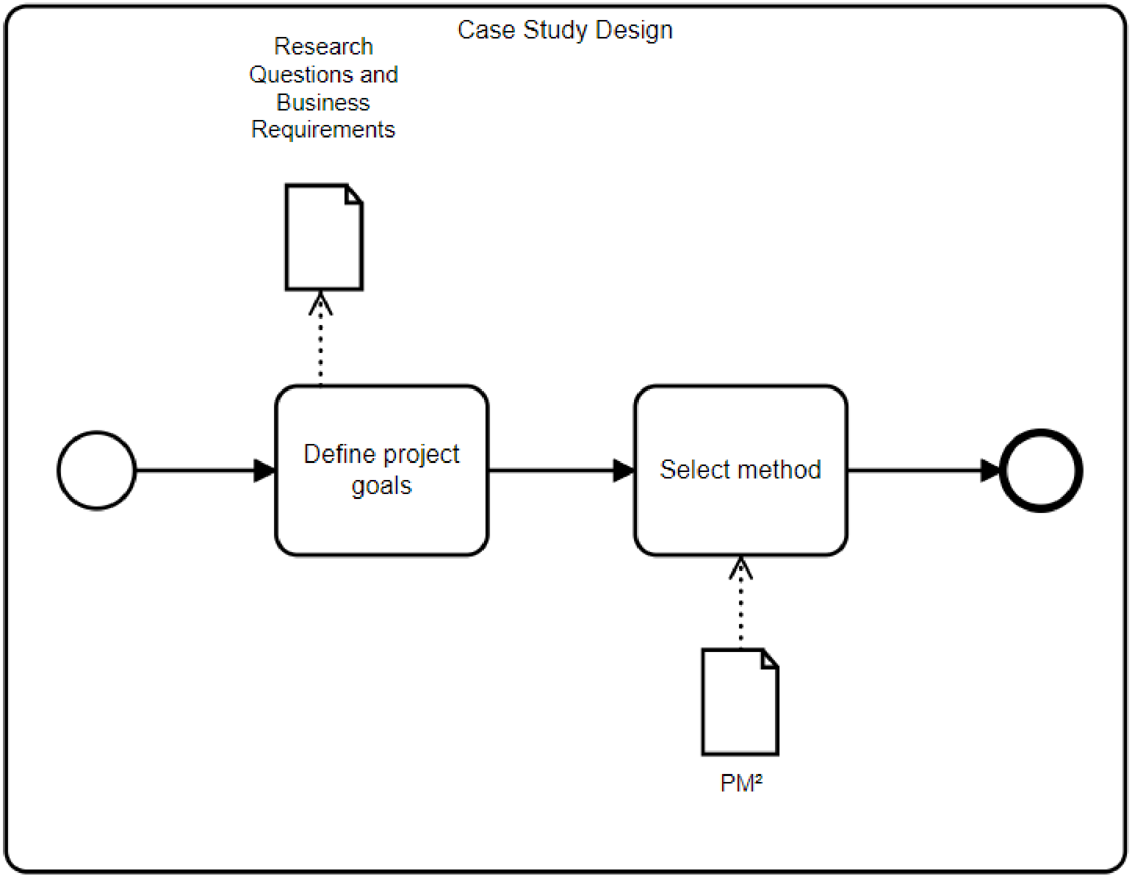

Figure 5: Case study Design Phase

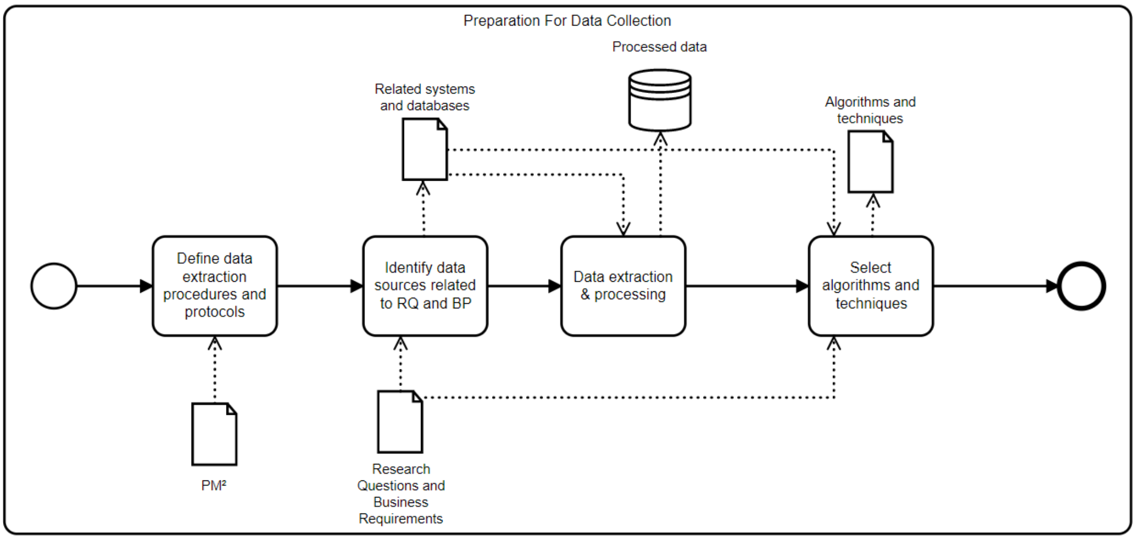

Figure 6: Preparation for Data Collection Phase

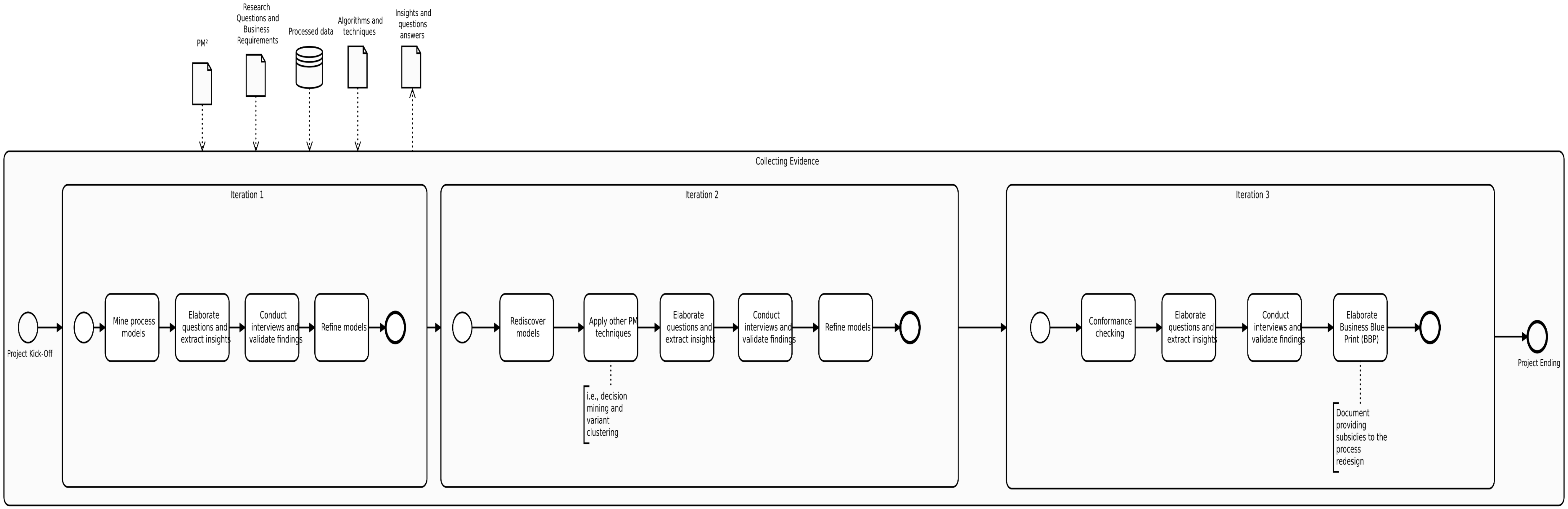

Figure 7: Collecting Evidence Phase

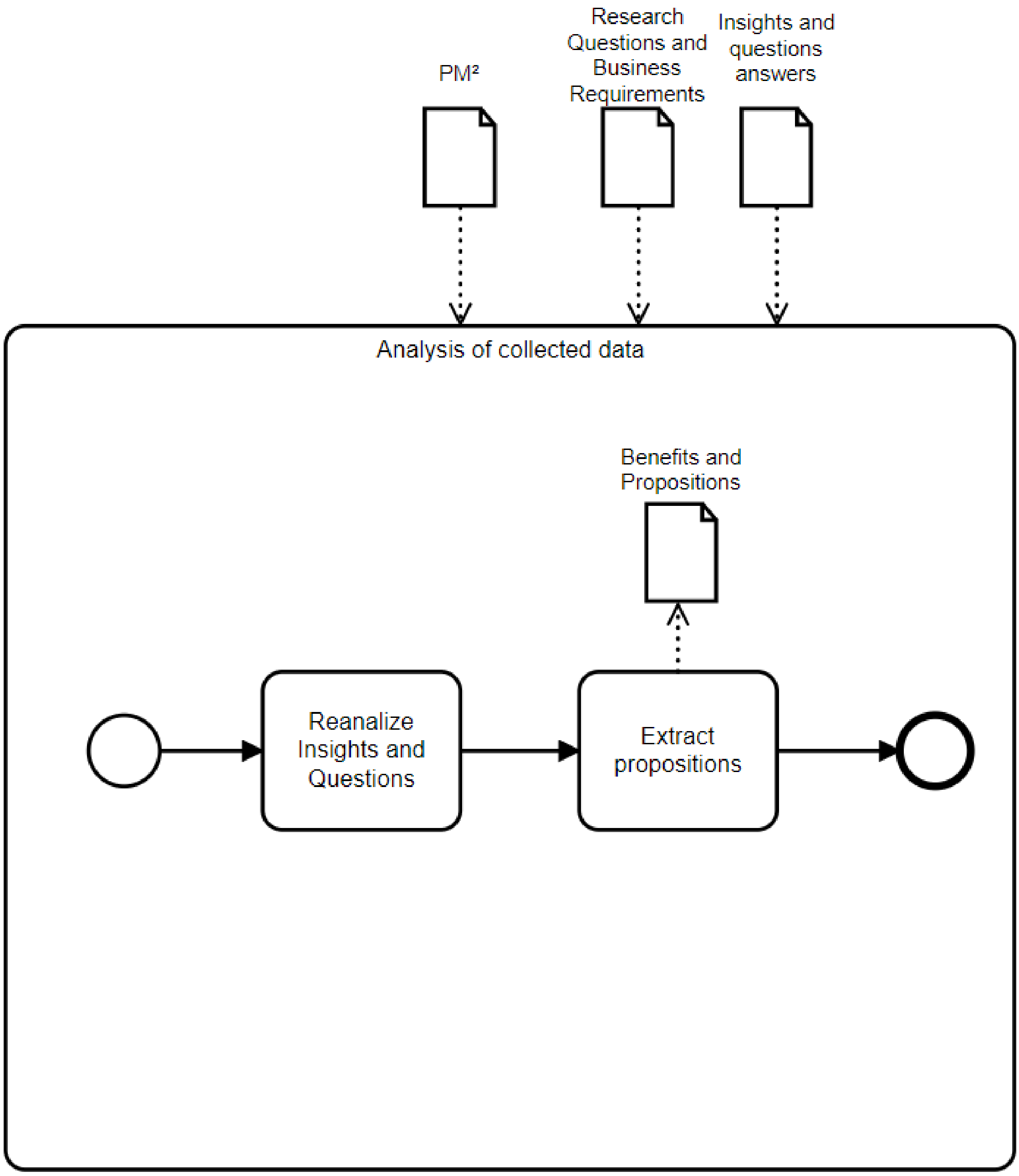

Figure 8: Analysis of Collected Data Phase

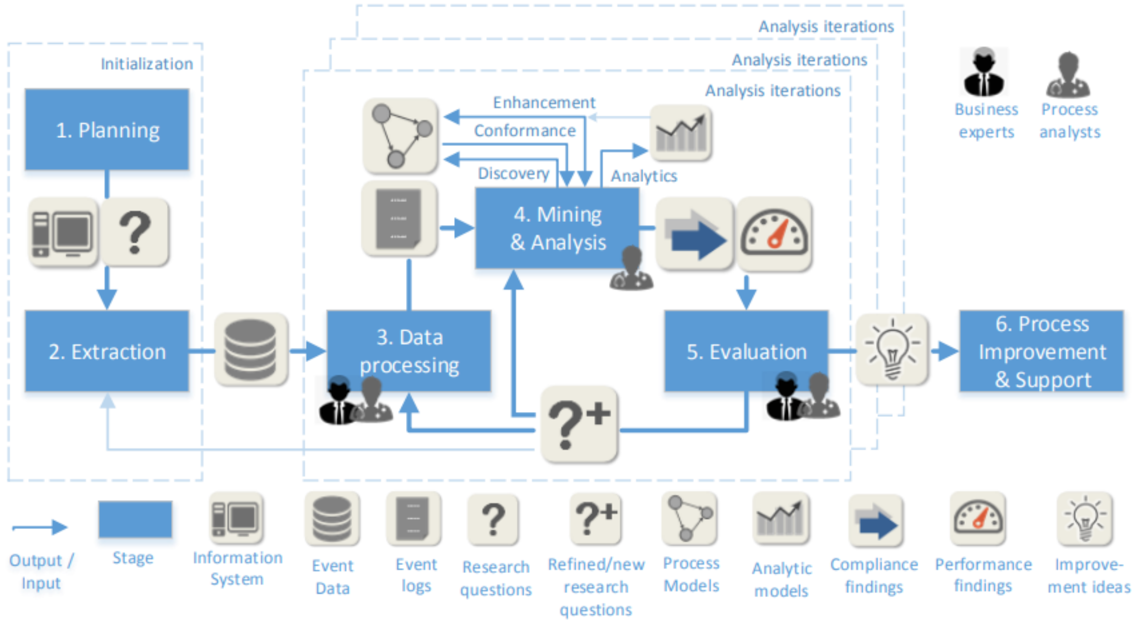

Figure 9: PM² methodology overview (van Eck et al., 2015)

The company followed the methodology PM² proposed by van Eck et al. (2015), which aims to guide organizations in process mining projects to improve performance or compliance to rules and regulations. The methodology comprises six stages as figure 9 depicts with several different input and output objects of the following types: goal-related objects, data objects, and models. The stages are: (1) planning; (2) extraction, (3) data processing, (4) mining & analysis; (5) evaluation; and (6) process improvement & support. In the stages (1) and (2), the initial research questions are defined, and event data are extracted. The stages (3), (4), and (5) can comprise one or more analysis iterations, possibly in parallel. Each iteration focuses on answering a specific research question by applying process mining and evaluating the discovered process models and other findings. If the conclusions are satisfactory, then they can be applied for improvement (6). The four goal-related objects are: (a) research questions derived from project goals, which are answered by (b) performance findings and (c) compliance findings, leading to (d) improvement ideas to achieve the goals.

3 The Log-In Case Study

3.1 Scenario description

This project started from a demand to implement two services in the process. Thus, the problem faced was understanding the process and how the data changed during its evolution along time to elicit the necessary requirements to implement the new services. In the next subsections, we describe and explain the conduction of the case step by step.

Planning

The goal of this step is to set up the project and to determine the research questions. Three activities take place: identifying research questions, selecting business processes, and composing a project team. Despite some problems related to the quality of data available to be collected, the selected process was "invoice payment request by service suppliers". In our case, the initial goal was to understand the process implemented some time ago because the as-is model did not correspond to the way it was actually running, and improve it by adding two new services.

The team was composed of 1 process analyst, 3 IT members with low or no expertise about the process, and 3 business experts. During the project execution, we noticed that the business experts only knew the process at a high level, but process mining allowed the team to find other people that understood the real process operation to answer detailed questions through resource mining.

Extraction

The current process model version in the production environment of the BPMS was considered the as-is reference model. The process model is shown in Figure 10.
After an initial evaluation, we observed that part of the process data was missing due to non-mandatory fields in the system, and many variants were addressed together. Besides, some activities' timestamp was not precise, which could introduce errors in the task order.

We found out that another attribute provided by the BPMS described the last time an activity was performed, so we decided to order the execution of activities from that attribute.

The process log was extracted from January to December 2019 (data from one year). The list of process attributes is as follows: case, process_version, type_of_activity, activity, timestamp_start, timestamp_end, process_owner, start_resource, end_resource, canceled, case_situation, service_type, data_atracacao, sailing, place, month, ship, harbor, subservice, terminal, voyage, storage, handling, repair, storage_value, handling_value, repair_value, aca, seller_aca, supplier, line, currency, organization, supervisor, invoice_amount, po_amount2, value_storage, value_storage_adjusted, value_handling_1, value_handling_adjusted, value_repair_1, value_repair_adjusted, amount_1, amount_2, date_apr_container, aca_1, invoice_apr_nmb, organization_1, status_erp and status_invoice.

We used ProM tools to perform the initial evaluation of this log. The log was composed of 4460 cases, 40234 events, and 16 event classes. Figure 11 and Figure 12 depict some details about the log: the evolution of the types of cases over time and the duration of different types of cases. Figure 11 shows an interesting behavior of the process: there are more cases at the end/beginning of the months. We discovered the participants responsible for each type of service provided in this process through resource mining. All data consumed by the process were also extracted. Only events from the BPMS execution were collected, neglecting the part of the process of launching the request for release of payment and approval, which is performed within the ERP.

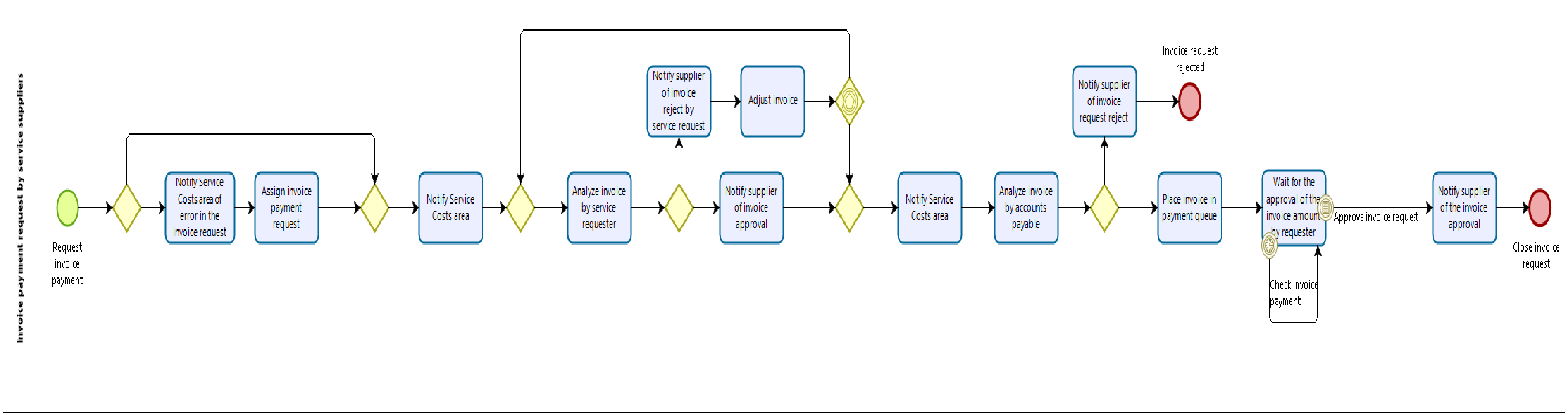

Figure 10: The "clean" AS-IS process model

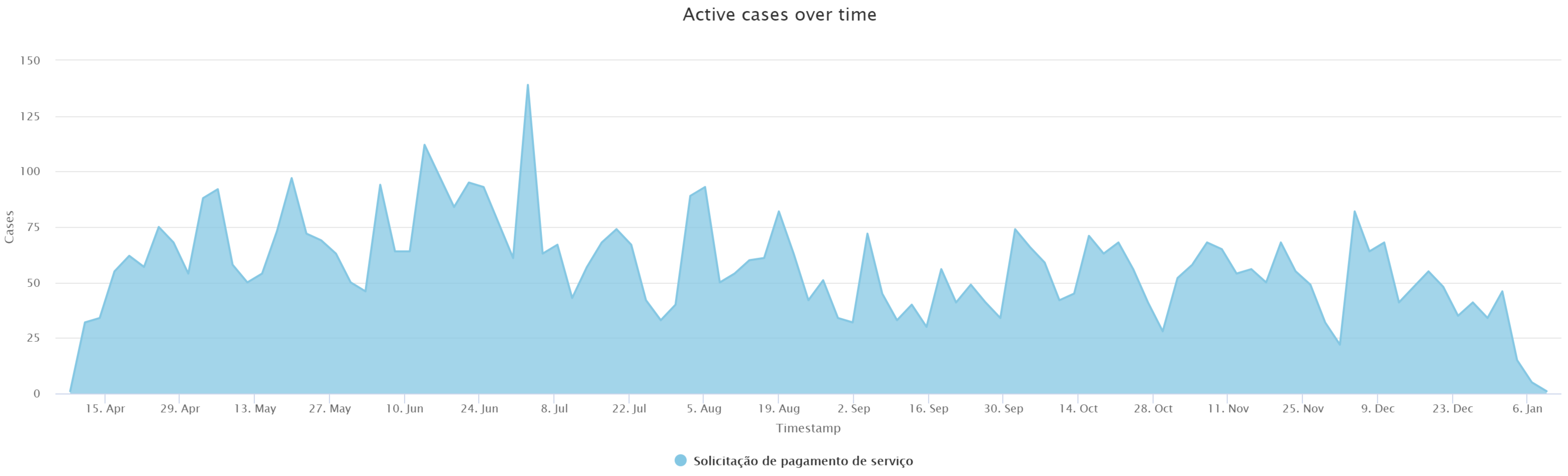

Figure 11: Log characteristics – the evolution of the types of cases over time

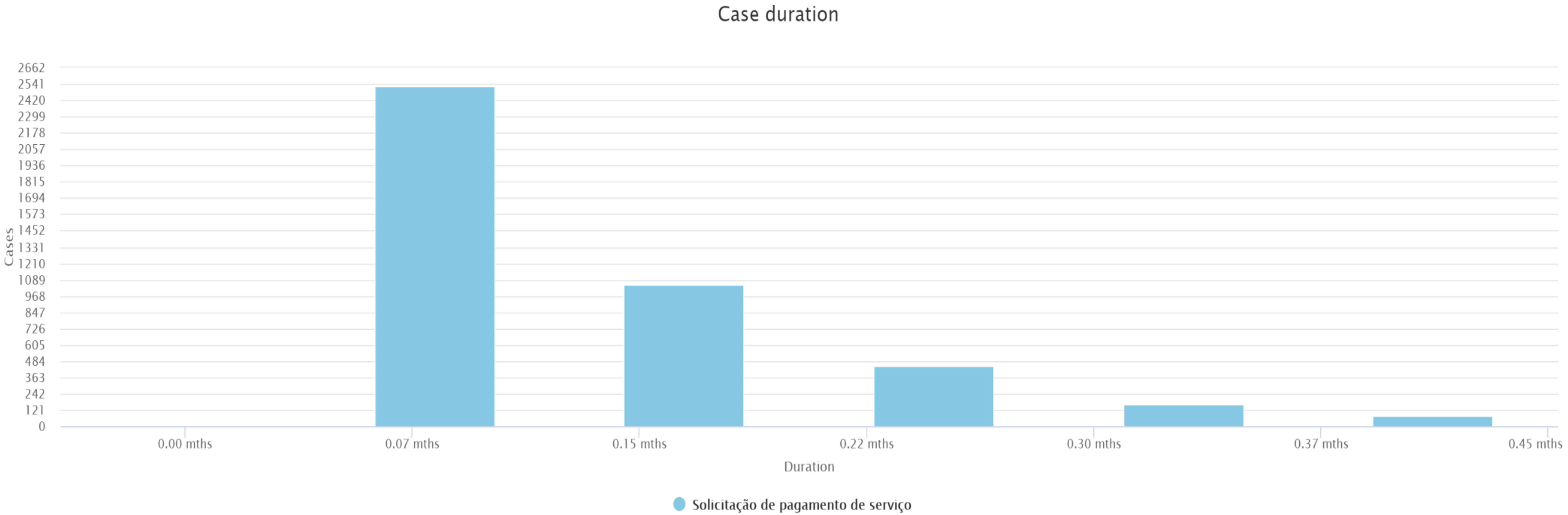

Figure 12: Log characteristics - duration of different types of cases

Data Processing

In this step, we applied filters to remove the open cases (slice and dice) from the process. The processing time attribute for each case was also added to the log. Events were renamed to improve the semantics and converted into activities because the ProM does not provide good support for converting from BPMN to Petri net, which is the format used by most plug-ins.

Mining & Analysis and Evaluation

After processing and filtering the log, the Apromore platform was used to discover a model that allows the quick examination of the process behavior within its dashboard functions. This stage was conducted in three iterations. The **Evaluation** step was performed concurrently with each iteration of the mining/analysis phase.

Iteration 1: We used the inductive visual miner plug-in in Apromore and ProM. Inductive Visual Miner aims to bridge a gap between commercial and academic tools by supporting the steps of process exploration and improving evaluation by a new notation and animation combining business process discovery and conformance checking through the discovery of a process model and then comparing with the event log allowing visualization of performance measures, queue lengths and animation (Leemans et al., 2014). The ID attribute of the process instance in the BPMS was used to identify the cases. We could visually detect the absence of activities and paths in the discovered model instead of the BPMS model. Thus, we raised some questions based on the model discovered, visually comparing it with the as-is reference model. For example, understand some details such as the lack of certain events in the model discovered with process mining and why there are more active cases during the end/beginning of the month.

At this point, it was not feasible to run conformance checking because we still did not have proper knowledge about the process activities and control flow, so we decided to carry out some meetings with process performers and started working on improving semantics for a better understanding of the flow. The most significant benefit reaped from process mining was knowing how to raise objective questions through the model and resources. The questions were asked concentrating on each type of service, which we could filter using the variant explorer plug-in. After the application of filters, it was possible to realize which paths one specific type of service took in the process and the average time for its resolution. As a result of these meetings, we created another reference model for future use on conformance checking. In this new model, we tried to represent the users' knowledge who performed the process and improved the semantic notation.

Moreover, during the interviews, we realized that many activities in the process take place outside the BPMS. When asked about this issue, the stakeholders of the process said that they decided to focus a priori on digitizing the basics of the process and on the most critical integration (with the ERP) in the approval part of the payment release request.

Iteration 2: After renaming the events, we applied the discovery algorithms again. The final discovered models are depicted in Figure 13 and Figure 14.

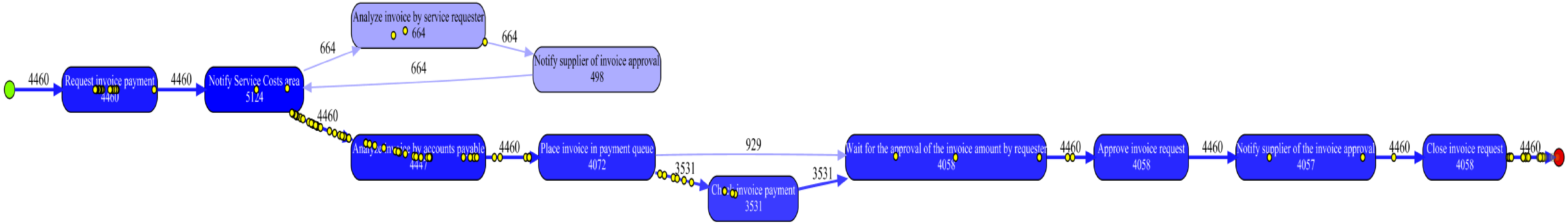

Figure 13: Process model extracted with ProM

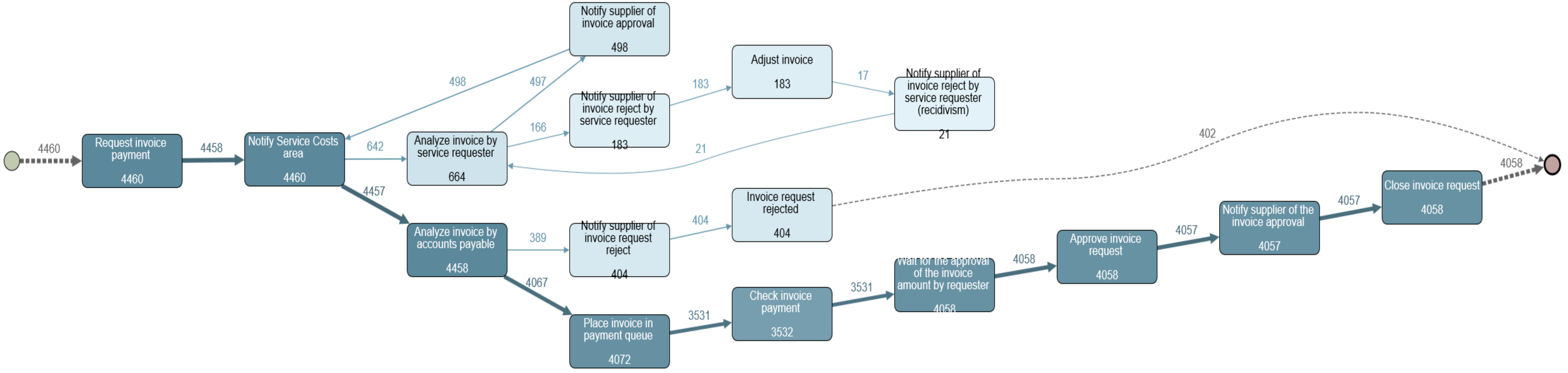

Figure 14: Process model extracted with Apromore

In ProM, the first analysis performed was the visualization of the dotted chart (Figure 15). We noticed that most cases start and end quickly, but many outliers should be analyzed in more detail. However, analyzing outliers regarding time indicators were out of scope of the project and the team thought that it could be counter-productive to engage efforts on this aspect. So, we included it in an analysis backlog. Another important insight gained during this initial analysis was that most cases happen during the working days, but a few requests are opened during the weekends, which is a surprise as some types of service are provided for the ship during weekends or in exchange for a stopover.

We also analyzed the types of services that lasted the longest. The ProM explore event log plug-in discovered 20 process variants (Figure 16). The first three variants show the “happy path” of the process, accounting for 84.7% of the log. The other variants indicate paths that contain rework (i.e., requests rejected, re-analyzed, and approved or disapproved). The primary outcome of this analysis was raising questions to the participants to help understand the causes that can result in rejected, re-analyzed, or approved instances. Furthermore, it was also important to precisely understand why there is a variant in which the process instance was rejected, but a payment request had been made in the ERP.

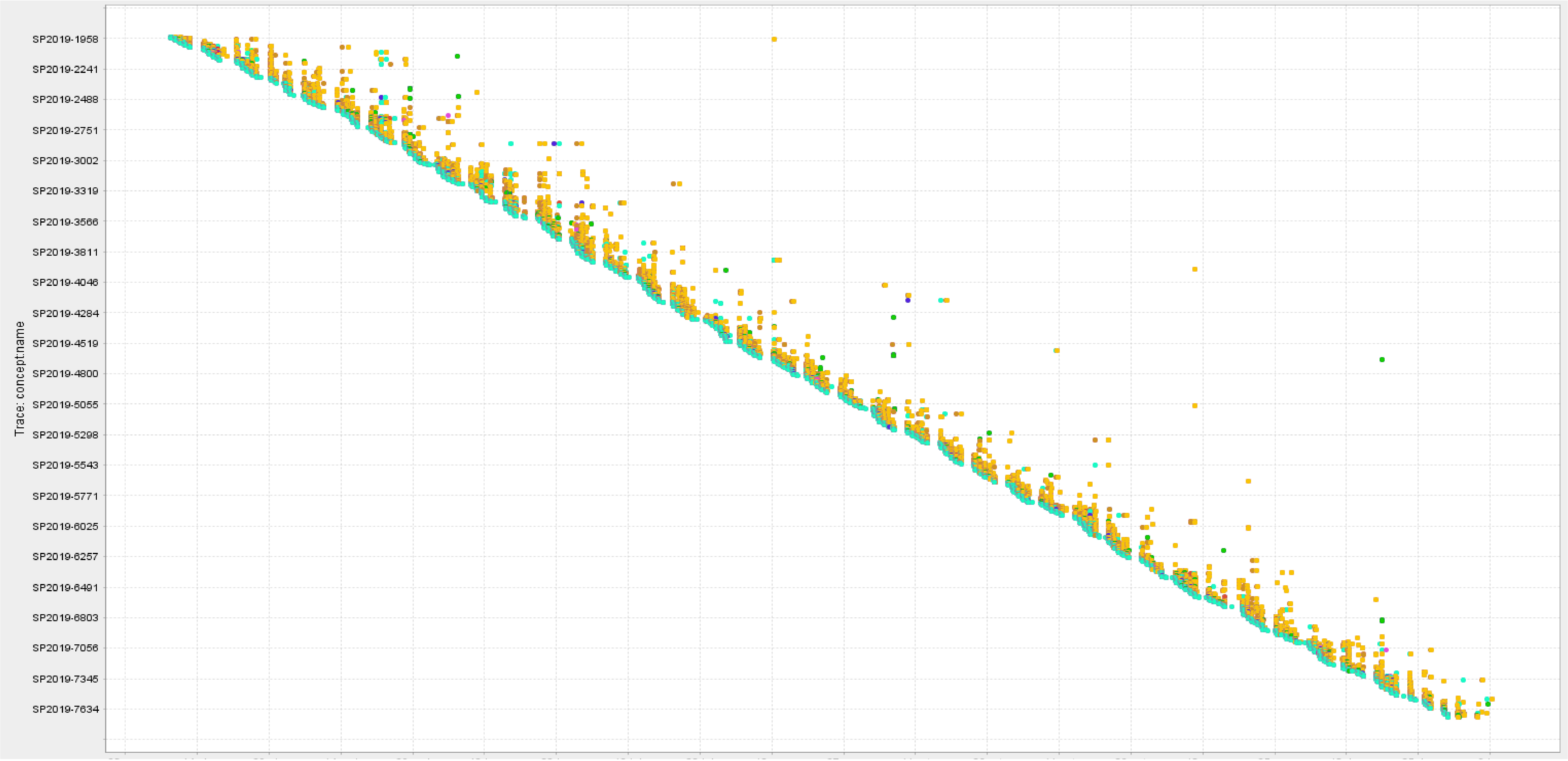

Figure 15: Dotted chart produced in ProM

Figure 16: Results of the event log explore plug-in in ProM

Another technique applied in the second iteration was the Discovery of Process Dataflow, using the Decision-Tree Miner plug-in in ProM. This plug-in falls into the category of decision mining algorithms. The goal of decision mining is to detect data dependencies that might affect the routing of a case. Machine learning techniques can be used to discover how data attributes influence the choices made in the process based on past process executions (Rozinat et al., 2006). The partial result is shown in Figure 17. The application of this technique allowed to validate some decisions that occur within the process (such as the container services following a particular flow with two approvals). Additionally, we could also realize that some suppliers fit into a group where all requests were initially rejected, and therefore, they had to correct the requests' errors.

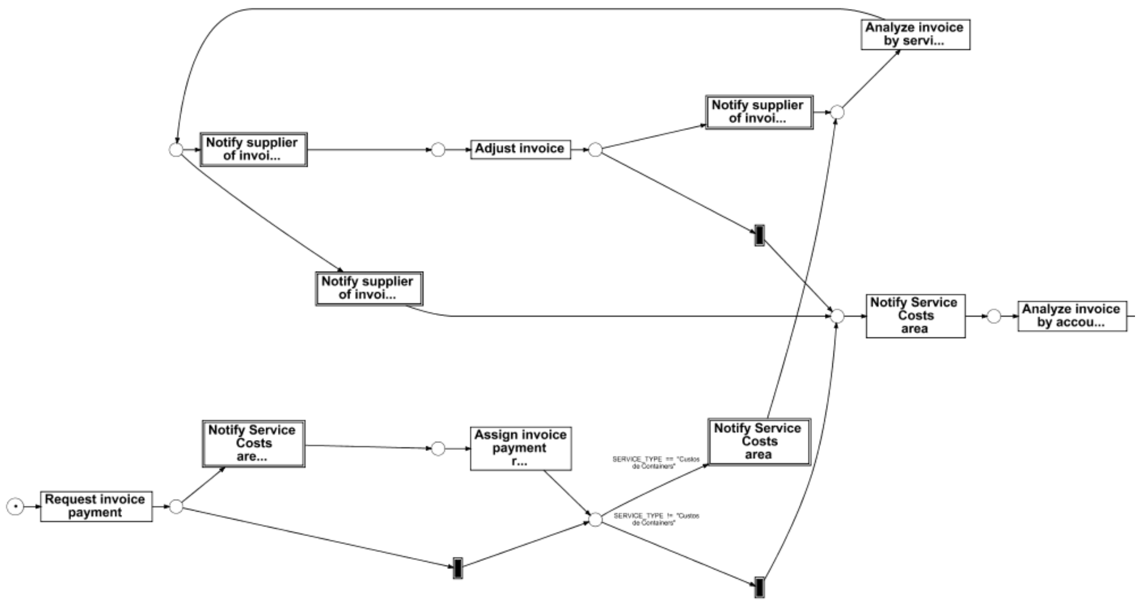

Figure 17: Extract of the Process Data-Flow model discovered

Iteration 3: We decided to apply conformance checking (Figure 18 and Figure 19) to analyze how the cases behaved against the reference as-is model (Figure 1). We used the plug-in "Replay a log on Petri net for conformance analysis" of ProM. The replay technique forces a test log to be run through the process model, even if the model and test log are not compatible, so the activities that are missing or left over during the execution are counted and inserted in the fitness calculation in order to obtain a compliance index (van der Aalst, 2012). This plug-in aims to compare a test log with a model so that a Petri net can be generated with various information about how a process is being performed when compared to its reference model, such as the fitness between the log and the model, quantification, and evidence of alternative paths to the model, points of resource overload and analysis of the individual performance of the modeled activities. We hypothesized that the fitness would be high because the BPMS structures the process, and the slider only showed infrequent paths when it came to represent 99% of the log data.

As we assumed, there were just a few non-conforming cases when aligned with the model. A more in-depth analysis showed the interaction of the former IT member and the BPMS administrator at some points of time in certain activities of the related cases. We could infer that the IT member and the BPMS administrator manually included the results of activities producing non-conforming behavior in the log.

Figure 20 depicts the analysis made with bupaR. This graphic shows the precedence relations between 2 activities of the process according to the log, i.e., the y-axis shows activity *a*, and the x-axis shows activity *b*, with activity *b* being a possible successor of *a*, if this has been registered in the log. The figure shows the absolute frequency of the occurrences of each activity that followed another one. It was applied as conformance checking because it synthesizes well all the behavior of the log. One example is the analysis of successors to the activity “Approve payment release” that is usually executed before the activity “Notify supplier of payment release”. We identified a case that skipped this activity directly to the final event “Payment request made successfully”.

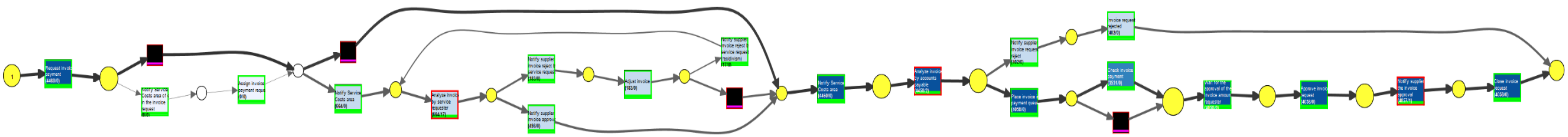

Figure 18: The model aligned with the log for conformance checking with log+model moves

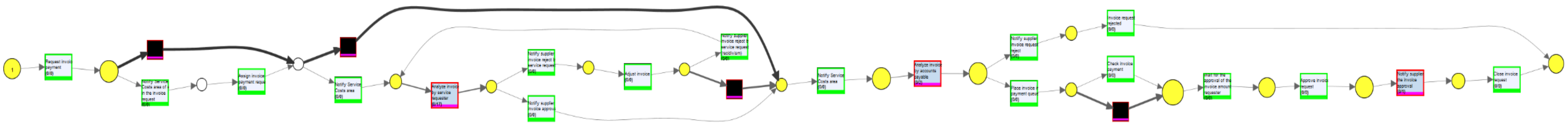

Figure 19: The model with model moves only

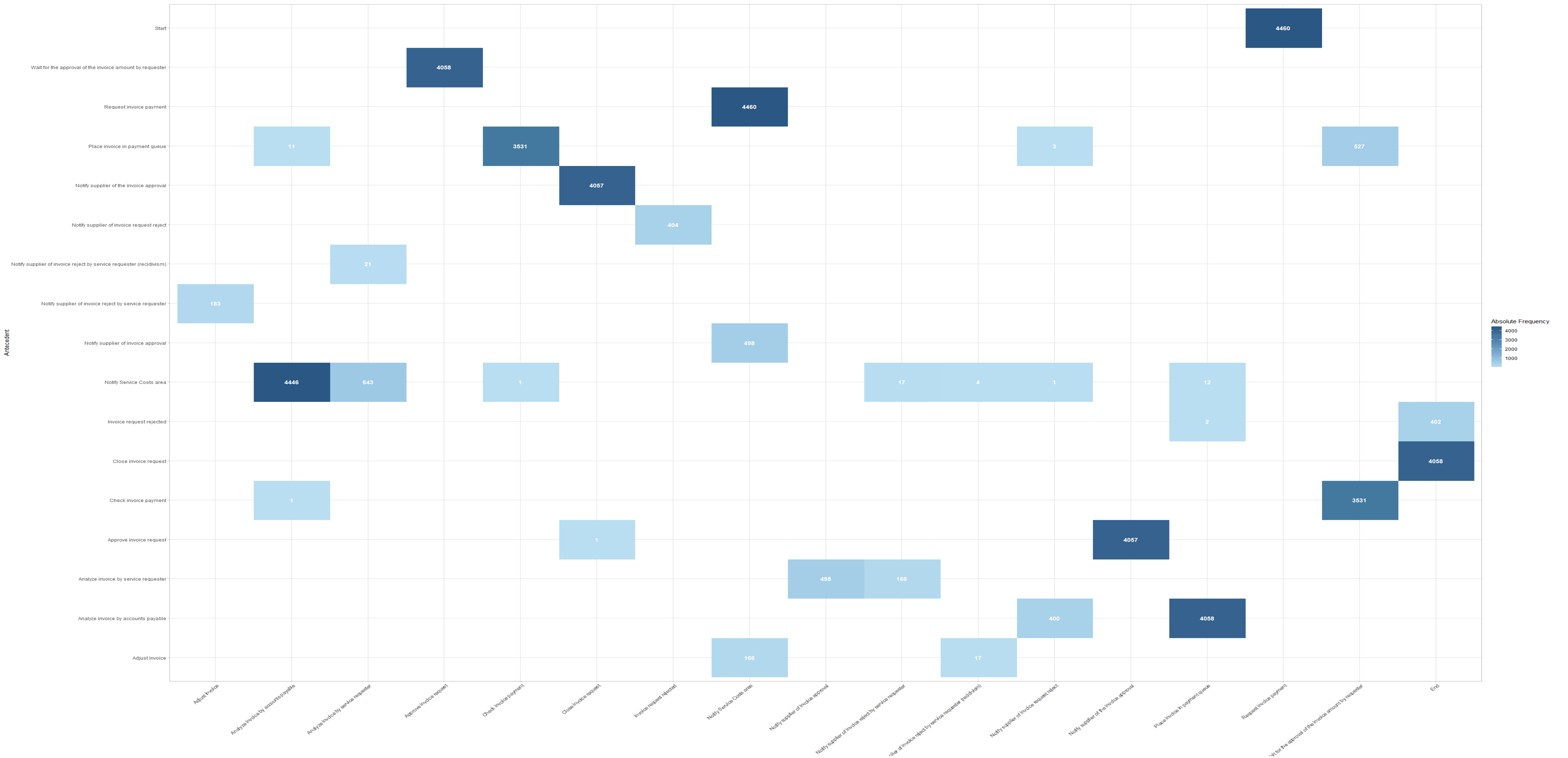

Figure 20: Precedence of activities matrix extracted from bupaR

There were some cases in which the events related to messages about the status of the payment release did not occur, and a more in-depth analysis together with the system's supply team has been carried out. We discovered that in fact the event did not happen even with the condition setup for it.

Another problem was in the activity aimed to analyze the costs of container services. At some point, the system should inform the users about the recurrence of disapproval of the request; but we observed that they were not receiving such notification. Thus, in their perception, as soon as a request is rejected, they should start a re-analysis, even before receiving the communication.

There was also a specific case where a timer event, which checks with the ERP if a payment request was approved occurred even before the request was made. In this specific case, the former IT member again altered the trigger of this event manually. Its execution was canceled, and the process proceeded normally. In summary, we realized that the process has been evolving over time, but the system was not correctly updated. All these problems were checked using filters with the conditions that generated the model move only.

Process Improvement & Support

After the analyzes, a Business Blue Print (BBP) suggesting improvements was prepared to provide subsidies to the redesign. In addition to the two new services that would be added to address the initial demand, some opportunities for improvement were obtained using process mining. Among them, the integration between the BPMS and the charging system for a better audit assessment of the launch of the Purchase Order (PO) in the ERP, and moreover, the integration with the systems related to the service request opening stage are noteworthy. This improvement would ensure that a payment request is automatically opened after the service is provided. Furthermore, there were many errors in requesting payment based on the supplier's input, causing the rework of opening them or correcting data.

We also found that the process treats macro activities related to several cases as one. This problem led to a very general process in which many activities specific to each service are carried out outside the BPMS. These specific activities would be important for obtaining more valuable insights from the process. In some types of services, the cost analysis activity is performed by comparing data from the process instance with data reported by other systems. For these cases, integration of the BPMS was suggested to carry out automatic and faster checking of these data. It is one of the most time-consuming activities in the system, with an average of 44 hours, almost two days.

4 Findings and Discussion

The goal of this research was to answer three main research questions. We discuss our findings related to each one of them as follows: **RQ1** (How can process mining be integrated into the IT landscape of a company?); **RQ2** (What is the business impact of process mining?); **RQ3** (What strategic implications emerge from process mining within enterprises?). In the end, we extract propositions about this process mining project in order to help to theorize about process changes in an organization.

RQ1: We found out that Log-In had more than one information system supporting this process (ERP, BPMS, and others). The process mining approach identified the lack of integration at some parts of the process, and based on the insights provided visually by its outcomes, we could infer how to solve the problem. Grisold et al. (2020) explain that some algorithms allow inductive and deductive theorizing. Thus, concerning the answer to **RQ1**, we can state the following propositions.

Proposition 1: Process mining (specifically inductive visual mining) is useful for the analysis and monitoring of IT resources integration,

Proposition 2: BPMS and ERP should integrate the different parts of the requesting PAYMENT from the SUPPLIER process, specifically the REQUEST and the CONSOLIDATION stages.

RQ2: The process analyzed in this case study is crucial to the domestic navigation business since the services provided by the suppliers could not be neglected. Some delays or re-works could mean significant losses. The adoption of process mining in this context impacted the identification of improvements and the reorganization of the whole macro-process. Thus, we summarize these findings as:

Proposition 3: Process mining (specifically variants identification) is useful for the re-organization of sub-processes.

Proposition 4: Different types of SERVICES should be addressed by different process models.

RQ3: We can infer some strategic implications of process mining usage in our case: people cannot prevent registering manual activities, and data collected from suppliers need to be checked and integrated within systems. In this sense, we state that:

Proposition 5: Process mining (specifically conformance checking) is useful for identifying activities performed out of the information systems.

Proposition 6: The NOTIFICATION of SUPPLIERS should be based on data shared within different types of SERVICES to guarantee the correctness of PAYMENT.

The company observed several benefits from this case. First, the most apparent benefits emerged in this case. The team could reach a fast and objective analysis of the process, avoiding long and tiring interviews or workshops that can take days until full understanding to analyze requirements for new features in a process. It was possible to start auditing the process from a more transparent and complete perspective, looking at the digitized parts of the process and the resources that perform the activities.

Moreover, we can also highlight the following gains: the discovery and monitoring of more complex Key Process Indicators (KPIs) and the possibility to perform applicability assessment of other automation technologies. Process mining enabled the monitoring of integrations between systems related to processes, allowing for the possibility of examining more complete KPIs. We can mention the creation of KPIs related to the time between the service provision and the opening of the request invoice payment by the supplier, the time to approve invoice request in the ERP, and the charging on the tax system.

In our case, some activities in the digital process that are performed manually could be integrated via BPMS REST API to perform data entry in other internal systems and even compare the invoice with another system to validate automated decisions. Process mining also made it possible to visualize activities where such integrations would not be a possible approach due to technological limitations of third parties (in our case, some suppliers); however, the orchestration of activities with a BPMS and the use of Robotic Process Automation (RPA) appears as a possibility to not only automate an activity but also keep it visible in logs for future analysis via process mining.

We found process discovery useful to track imperceptible activities for the process actors. Since many activities of a process may be automated (i.e., no input is taken by a user), users may not know that activity exists in the process. Thus, process discovery is useful to uncover and validate knowledge of a process from the Information Technology perspective.

Conformance checking can be used to explain hypotheses or propositions about a process deductively. Process discovery and variant explorer in conjunction with conformance checking was useful to theorize process goals inductively. In our case, conformance checking within interviews were useful to uncover endogenous changes in the process over time.

The process version control within the BPMS allowed us to track purposeful interventions in the process evolution. This BPMS feature supports exogenous punctuated changes as it allows to see past interventions in the evolution of the process (i.e., implementing new artifacts, integrations, actors, etc.). Moreover, Decision Mining can be used in conjunction with this feature to discover the reasons for an intervention in a process evolution. We can mention the integration of RPA in the process scope as productivity or rework issues in a process.

It is important to note that all the findings about the process and related systems came not only from the data but also from the interviews with participants. It confirms that the insights from process mining alone cannot explain everything about the context of a process or organizational changes, as Grisold et al. (2020) affirm. Moreover, since we did not have an *a priori* set of concepts in this domain, we tried to extract some relevant terms derived from the process models, which were used in the propositions formulated for Log-In (the words in uppercase).

We also have some considerations about the method adopted for the process mining project. Following a systematic approach was greatly beneficial in our case. The team could organize the work, and more than that, it was possible to demonstrate the value-added much easier step by step. However, in practice, besides the mining & analysis phase, which is planed through iterations, in which the results of each iteration stimulate the next one and are all over again reviewed, it would be important to provide feedback also to the other phases. For example, in our case, with the discovery of resources in the activities, we were able to compose the team better (phase 1) and gain more valuable insights from the process actors. Besides that, the evaluation step was performed together with the iterations in the mining & analysis phase; we felt it difficult to separate them. Finally, we missed a distinguished approach for structured vs. unstructured process, i.e., specific guidelines to deal with different types of findings and templates (document, spreadsheet, etc.) to represent the findings and suggestions implemented in the improvement stage. Apart from González et al. (2018), who followed the PM^2 method, but focusing on the three initial steps, and Johnson et al. (2019) that extended this method to cover specifically the health domain, to the best of our knowledge, there is not reported experiences with the use of a systematic approach for process mining projects.

5 Conclusions

The umbrella term “process science” is used by van der Aalst (2012) to refer to the broader discipline that combines knowledge from information technology and knowledge from management sciences to improve and run operational processes. This work presented the case study of a domestic navigation business company that applied process mining to learn about and improve its “invoice payment request by service suppliers” process. The prescriptions of Grisold et al. (2020) supported the analysis and provided insights about the findings of this case.

We argue that process mining made it explicit that the way the process was being conducted was obsolete, and some changes had already been made in practice. Besides that, the inspiration provided by the models obtained and data analysis served as a positive source to start a conversation with process participants. The result was a better and shared understanding of the process that made it possible to recommend more changes. This could be an adequate cycle to be adopted by the company.

This work contributed to a reflection on how to extract propositions from a process mining approach. As a contribution, we published part of our work as an 18th Int. Conference on Business Process Management (BPM 2020) Industry Track forum paper (Santos et al., 2020). The implication for the practice of BPM was to provide lessons on how enterprises actually adopt and use process mining software. The relevance of a proper method proved to be essential in our case and could be shared in other cases.

Most of the limitations and challenges of this project relies in the data quality of the systems related to the analyzed process as mentioned in the data extraction process. The data quality impacts in many aspects of the project, for example, in the scope of applicable algorithms and propositions extracting. In our case, we can mention the impossibility of carrying out social network analysis, which would be useful to extract propositions on how collaboration between areas occurs more precisely.

Future work includes the execution of other instances of this case not only in the same company but also in other companies. We also intend to create a framework with the results obtained, the propositions, and theoretical approach, besides building a lexicon for this domain based on our propositions.

References

Badakhshan P., Alibabaei A. Using Process Mining for Process Analysis Improvement in Pre-hospital Emergency. In: Baghdadi Y., Harfouche A., Musso M. (eds) ICT for an Inclusive World. Lecture Notes in Inf. Systems and Organisation, vol 35. Springer, Cham (2020).

Battineni, G., Chintalapudi, N., Karami, V., Nittari, G., Amenta, F., Tayebati, S.K. Process mining case study approach: Extraction of unconventional event logs to improve performance in Hospital Information Systems, Journal of Comp.Science, 17 No 4. 117-128 (2019).

Berente, N., Seidel, S., Safadi, H. Data-Driven Computationally-Intensive Theory Development. Information Systems Research, 30 (1): iii–viii (2019).

Conforti, R., Dumas, M., La Rosa, M., Maaradji, A., Nguyen, H., Ostovar, A., Raboczi, S. Analysis of Business Process Variants in Apromore. In Proceedings of the Demo Track of the 13th International Conference on Business Process Management (BPM'15) (2015).

De Leoni, M., van der Aalst, W.M.P & Dees, M. A general process mining framework for correlating, predicting and clustering dynamic behavior based on event logs. Information Systems, 56, 235-2 (2016).

De Weerdt, J., Schupp, A., Vanderloock, A., Baesens, B. 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 (2013).

Dumas, Marlon et al. Fundamentals of business process management. Heidelberg: Springer, 2013.

Fernández-Llatas C., Lizondo A., Monton E., Benedí J-M., Traver V. Process mining methodology for health process tracking using real-time indoor location systems. Sensors 15(12):29821-29840 (2015).

González López de Murillas, E., Hoogendoorn, G. E., & Reijers, H. A. Redo log process mining in real life: Data challenges & opportunities. In Business Process

Management Workshops - BPM 2017 International Workshops, Revised Papers (pp. 573-587). (Lecture Notes in Business Information Processing; Vol. 308). Springer/Verlag (2018).

Grisold T, Wurm B, Mendling J, vom Brocke J (2020) Using process mining to support theorizing about change in organizations. In: Proceedings HICSS 2020.

Jans, M., Alles, M., Vasarhelyi, M. The case for process mining in auditing: Sources of value added and areas of application. International Journal of Accounting Information Systems, 14 (1): 1-20 (2013).

Janssenswillen, G., Depaire, B., Swennen, M., Jans, M., & Vanhoof, K. bupaR: Enabling reproducible business process analysis. Knowledge-Based Systems, 163, 927-930 (2019).

Johnson O.A., Ba Dhafari T., Kurniati A., Fox F., Rojas E. (2019) The ClearPath Method for Care Pathway Process Mining and Simulation. In: Daniel F., Sheng Q., Motahari H. (eds) Business Process Management Workshops. BPM 2018. Lecture Notes in Business Information Processing, vol 342. Springer, Cham (2019).

Langley, A., Smallman, C., Tsoukas, H., Van de Ven, A.H. Process studies of change in organization and management: Unveiling temporality, activity, and flow. Academy of management journal, 56 (1): 1–13 (2013).

Leemans, S. J. J., Fahland, D., & Aalst, van der, W. M. P. Process and deviation exploration with Inductive visual Miner. In L. Limonad, & B. Weber (Eds.), BPM Demo Sessions 2014 (co-located with BPM 2014, Eindhoven, The Netherlands, September 20, 2014) (pp. 46-50). (CEUR Workshop Proceedings; Vol. 1295). CEUR-WS.org (2014).

Mans, R.S., Schonenberg, M.H., Song, M., van der Aalst, W.M.P., Bakker, P.J.M. Application of Process Mining in Healthcare – A Case Study in a Dutch Hospital. In: Fred A., Filipe J., Gamboa H. (eds) Biomedical Engineering Systems and Technologies. Communications in Computer and Information Science, vol 25. Springer, Berlin, Heidelberg (2008).

Rebuge A., Ferreira, D.R. Business process analysis in healthcare environments: a methodology based on process mining. Information Systems 37(2):99-116 (2012).

Rozinat A., van der Aalst W.M.P. Decision Mining in ProM. In: Dustdar S., Fiadeiro J.L., Sheth A.P. (eds) Business Process Management. BPM 2006. Lecture Notes in Computer Science, vol 4102. Springer, Berlin, Heidelber (2006).

Runeson, P., Höst, M.: Guidelines for conducting and reporting case study research in software engineering. Empir. Softw. Eng. 14 (2), 131–164 (2009).

Ruschel, E., Santos, E.A.P., Loures, E.d.R. Establishment of maintenance inspection intervals: an application of process mining techniques in manufacturing. Journal of Intelligent Manufacturing ,31, 53–72 (2020).

Santos, W. S., Santoro, F. M., Abreu, V. A., Barral, L. M. . Process Mining Applied to Process Redesign in a Domestic Navigation Business Company. In: BPM Industry Forum 2020 in conjunction with the Business Process Management Conference, 2020, Sevilla. Proceedings of the BPM Industry Forum 2020 in conjunction with the Business Process Management Conference, 2020. v. 1. p. 1.

Van de Ven, A.H., Poole, M.S. Alternative approaches for studying organizational change. Organization studies, 26 (9): 1377–1404 (2005).

van der Aalst, W.M.P. Process mining: data science in action. 2nd ed. Springer (2016).

van der Aalst, W.M.P., Adriansyah, A., van Dongen, B.: Replaying history on process models for conformance checking and performance analysis. WIREs Data Mining Knowl. Discov. 2(2), 182–192 (2012).

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. Business process mining: An industrial application, Information Systems, 32 (5): 713-732 (2007).

van Eck, M.L., Lu, X., Leemans, S.J.J., van der Aalst, W.M.P. PM2: A Process Mining Project Methodology. In: Zdravkovic J., Kirikova M., Johannesson P. (eds) Advanced Information Systems Engineering. CAiSE 2015, pp 297-313. Lecture Notes in Computer Science, vol 9097. Springer, Cham (2015).

Volkoff, O., Strong, D.M., Elmes, B. Technological embeddedness and organizational change. Organization science, 18 (5): 832–848 (2007).

Yoo, S., Cho, M., Kim, E., Kim, S., Sim, Y., Yoo, D., Hwang, H., Song, M., Assessment of hospital processes using a process mining technique: Outpatient process analysis at a tertiary hospital. International Journal of Medical Informatics, Volume 88, Pages 34-43 (2016).