Towards a method and a guiding tool for conducting process mining projects

Arthur Valle
Data Science Research Group - Waikato Institute of Technology (Wintec)
arthur.valle@wintec.ac.nz

Gilberto Aleces dos Santos, Cleiton dos Santos Garcia, Edson Emílio Scalabrin
Programa de Pós-Graduação em Informática - Pontifíca Universidade Católica do Paraná (PUCPR)
gilbertoaleces@gmail.com, cleitonsg@gmail.com, scalabrin@ppgia.pucpr.br

Eduardo Alves Portela dos Santos
Programa de Pós-Graduação em Engenharia de Produção - Pontifíca Universidade Católica do Paraná (PUCPR)
eduardo.portela@pucpr.br

ABSTRACT

Due to the increased use of information systems by organizations, information on the execution of processes is recorded. This enables using process mining as a tool for improving process performance. Process mining allows gaining insights regarding actual processes by extracting and processing data from existing systems. Many projects have been conducted for process discovery, conformance checking, etc. Despite of the existence of general methods for data analysis, there's a lack of specific methods to support process mining projects. Thus, completions of such projects are often dependent on expertise of the analysts. This paper presents a detailed method for conducting process mining projects and a tool for supporting its execution and retaining the outcomes of each step. A case is analysed for evaluating them. Organizations seeking process performance improvement can get benefit from a method that states how process mining techniques can be used in process mining projects.

Keywords: process mining, process mining methodology, process mining project.

1. INTRODUCTION

The amount of digital data being created globally is doubling every two years (Zwolenski and Weatherill, 2014). According to the Massachusetts Institute of Technology, only about 0.5 percent of that data are ever analysed. By 2020, Forrester predicts businesses that use data effectively will be collectively worth USD 1.2 trillion[1]. These companies may take advantage from applying process mining to discover, manage and improve business processes. Process mining is a relatively young research discipline considered as the bridge between data science and process modelling & analysis. It enables discovering, monitoring and improving real processes (i.e., not assumed processes) by extracting knowledge from event logs' elements (such as timestamps, case ID, activities, performers, etc) that are available in today's information systems (Van Der Aalst et al, 2011).

This technology has become available only recently, but it can be applied to any type of operational processes. Example applications include: analysing patient-treatment (Ghasemi and Amyot, 2016), improving insurance claiming (De Weerdt, Schupp, Vanderloock, and Baesens, 2013), understanding students' behaviour when attending eLearning courses (Aguirre, Parra, Alvarado, 2012), maximizing call center resolutions (Panpanich, Porouhan, and Premchaiswadi, 2015), analysing production line (Meincheim, Garcia, Nievola, and Scalabrin, 2017), among others. All these applications have in common that dynamic behaviour needs to be captured as processes.

Hence, it is no longer acceptable to just look at processes and data in isolation. In 2017, Gartner's process mining market estimation for new software product license and maintenance revenue was approaching 120 million[2]. This market is expected to easily triple or quadruple in size in the next few years.

A large follow-on market also exists for consulting and services in implementing these tools and the methods for using them. This opportunity for business process consulting services also represents a challenge for updating a large number of professionals about how to use process mining. It is estimated that consulting and service revenue significantly exceeds its software revenue, according to the Gartnert's market guide for process mining. Despite of a crescent proliferation of algorithms, tools and plug-ins, process mining is still incipient in terms of clear orientation on how to carry out process mining projects aiming to improve process performance aspects, such as lead-time. In this paper, we propose a detailed method that comprises stages, activities and tasks for guiding the conduction of process mining projects. We also present a guiding tool able to handle and record the outputs of each task in the method, supporting its execution.

The paper is organized as follows: Section 2 provides some background about existing process mining methodologies. Section 3 presents the research methodology. Section 4 introduces the proposed method. Then, section 5 presents a tool for guiding application of the method. A case study is presented

This quality assured paper appeared at the 10th annual conference of Computing and Information Technology Research and Education New Zealand (CITRENZ2019) and the 32nd Annual Conference of the National Advisory Committee on Computing Qualifications, Nelson, NZ, Oct. 9-11, during ITx 2019.

[1] https://go.forrester.com/press-newsroom/insights-driven-businesses-will-take-1-2-trillion-a-year-by-2020/

[2] https://www.gartner.com/doc/3870291/market-guide-process-mining

in section 6 and discussed in section 7. Section 8 describes the conclusions, limitations and further work.

2. RELATED WORK

According to van Eck, Lu, Leemans, and van der Aalst (2015), a process mining project is a way of applying process mining to achieve results such as improving process performance or checking the process compliance to rules and regulations. Van der Aalst (2016) recognizes three types of process mining projects:

- Data-driven. It is a process mining project based on the availability of event data. They have an exploratory character, so there is no defined question or goal: it is expected that valuable information will emerge from the analysis of data;
- Question-oriented. It is a process mining project that aims to answer specific questions, such as "why cases addressed by team A are faster than cases addressed by team B?" or "why are there more process deviations in cases performed by senior staff?".
- Goal-driven. A goal-driven process mining project aims to improve a process concerning to specific performance goals, such as cost savings or reducing lead-times.

Van der Aalst (2016) argues that several models describing the life cycle of a classic data mining or business intelligence project have already been proposed by academia and industry. For example, the CRoss-Industry Standard Process for Data Mining (CRISP-DM), a methodology with a life cycle consisting of six phases: a) understanding the business, b) understanding the data, c) preparing data, d) modeling, e) evaluation, and f) implementation. Similarly, the Sample, Explore, Modify, Model and Assess (SEMMA) approach consists of five phases: a) sampling, b) exploration, c) modification, d) modeling, and e) assessment.

According to van der Aalst (2016), both methodologies are very high level and provide little help. In fact, such methodologies are not suitable for process mining projects. Thus, van der Aalst proposes the L* life cycle model, comprising five stages: 0-Plan and justify, 1- Extract, 2-Create the control flow model and connect Events log, 3-Create Integrated Process Model, and 4-Operational support. The characteristics of each stage are as follows:

- Stage 0: *Plan and justify*. Similar to any project, a process mining project also needs to be carefully planned, which includes identifying the expected results of the project. In addition, one should identify the activities of the project, the resources allocation, the milestones and how project progress will be tracked continuously.
- Stage 1: *Extract*. Event data, models, goals and questions need to be extracted from systems, practitioners and management. If the project is goal-driven or question-oriented, such aspects are identified at this stage, through interactions with stakeholders (e.g, domain experts, end users, customers and managers).
- Stage 2: *Create the control-flow model and connect events log*. It aims to determine the actual control-flow model to be analysed. The process model is discovered by using process discovery techniques. However, if a good process model already exists, it is evaluated using compliance checking or it is evaluated in relation to the discovered model. Upon completion of stage 2, there is a control flow model connected to the event log, or the events in the log are mapped to activities in the model.
- Stage 3: *Create Integrated Process Model*. The model is enhanced by adding new perspectives in the control-flow model. For example, organizational perspective, case perspective, and time perspective. The outcome is an integrated process model that is used for various purposes. For instance, the model can be inspected directly for better understanding the process as it is or for identifying bottlenecks. It can also be used to answer selected questions and take the appropriate actions.
- Stage 4: *Operational support*. This stage relates to the aspects of detecting, predicting and recommending. For instance, it is possible to predict the remaining time for cases in progress. In addition, the result does not need to be interpreted only by process mining analysts and, instead, can be made available to end users. For example, a process deviation may result in automatically triggering an alarm in the shop floor of a factory.

Despite of the establishment of that L* life cycle for process mining projects, it has been considered not adequate, as it has deficiencies of being a general approach for all types of process mining projects (van der Heijden, 2012) (van Eck et al. 2015). Due to that, two methodologies have been proposed, which ended up having the same title: PMPM-Process Mining Project Methodology (van der Heijden, 2012) and PM^2-Process Mining Project Methodology (van Eck et al. 2015). The first one aims to be an appropriate methodology for process discovery, monitoring and improvement using process mining (van der Heijden, 2012). The second is designed to support process mining projects that aim to improve process performance or compliance with rules and regulations (van Eck et al. 2015). Both methodologies include six phases (or stages) and 18 activities. It is possible to note a strong similarity among them: only three activities are exclusive to one or other methodology.

A method based on building blocks has been proposed by Bolt, de Leoni, and van der Aalst, (2016). It considers four common analysis scenarios and six categories to organize the activities in building blocks according to capabilities orchestrated by the RapidProM tool (van der Aalst, Bolt, and van Zelst, 2017).

Other methods have been proposed targeting some business segments, such as healthcare or manufacturing. Sometimes this is needed to address segment's particularities. De Weerdt et al. (2013) proposes a multi-faceted method for financial services organisations. It is divided into four major phases: event log gathering; event log exploration; significant discoveries; and process improvement recommendations. For logistics, a method aiming to establish a logistics segment-oriented method for analysing material movements has been proposed by van Cruchten and Weigand (2018). Focused on health care process, a general framework proposal aimed to support the patient journey; identify practices for patient-centric process redesign, process conciliation, and management of decision support system (MDSS) has been proposed by Curry (2018).

A method for systematic support of knowledge-intensive business processes has been proposed by Mundbrod, Beuter, and Reichert (2015). It is based on involving more effective collaboration and coordination among employees. It has been validated using development projects for electrical and electronic components.

Another proposed method aimed to discover patterns of customer service request handling processes has been proposed in a methodology divided in four phases: business understanding, data collection & review, discovery and decision aid (Delias, Doumpos, and Matsatsinis, 2015).

An information system audit methodology, enabled by process mining, has been proposed by Zerbino, Aloini, Dulmin, and Mininno (2018). It is divided in stages: 0-Justification and planning; 1-Data extraction; 2- Control-flow model construction; 3-Model enrichment; and, 4-Conformance checking. Its validation was carried out in an export process in a Port Community System (PCS), to explore the main process deviations. Unlike sample-based audits, the proposed method focuses in all event logs, through quantitative approach and automatic tools.

Based on the above, we note a lack of segment-independent methods for process mining projects that are helpful, prescriptive and detailed (i.e covering task and subtask levels) while being generic enough to be used in process mining projects of any purpose.

In the next section, we present the research methodology that we selected for developing our method.

3. RESEARCH METHODOLOGY

In order to develop the (segment-independent) method and its supporting tool, we followed some steps inspired by a research approach called Design Science Research Methodology (DSRM). It focuses on the design and development of artifacts, such as systems, applications and methods and was developed by Peffers et al (2007). It has been selected because it is a research methodology that focusses on designing and developing artifacts, exactly as our objective. This research methodology comprises six activities: 1-Problem identification and motivation; 2-Define the objectives for a solution; 3-Design and development (of a solution); 4-Demonstration; 5-Evaluation and 6-Communication.

In our research, we performed these research activities considering the following aspects:

- In the first activity, in fact, the identification of the problem and the motivation were exposed in the sections 1 and 2 of this paper. Basically, the problem is the lack of a prescriptive method, detailing task and subtask levels, for conducting process mining projects. Additionally, there is a lack of a (software) tool that, along with actual process mining tools, supports the execution of process mining projects.
- The second activity covers the definition of the objectives for a solution to solve the exposed problem. In our case, the objective is to refine an integrated view – presented in section 4 - into a detailed method as well as to develop a (software) tool for supporting the execution of process mining projects. The outcome of this activity is the "requirement" for the detailed method, also presented in section 4.
- The third activity relates to the design and development of the method. Based on the identified requirements, the design is derived. After, the development of the method itself takes place. The same approach also applies for the supporting tool. Section 4 presents the first outcome of this activity: the design principles and the development of the method. Section 5 presents the second outcome: the tool that supports the execution of the method.
- For the fourth activity – demonstration – a case is used. It involves the execution of a real process mining project in a manufacturing company. The project has some specific objectives and for meeting them, the proposed method and tool are used. Section 6 presents the case study.
- In the subsequent activity – evaluation - the expectation is to evaluate the results of the case in term of how the method and its tool contribute for the conduction of the process mining project. Section 7-Discussion presents the evaluation.
- Finally, the activity called communication involves publishing the research results, which is made via this paper.

4. A METHOD FOR CONDUCTING PROCESS MINING PROJECTS

As commented in section 3, a high level, merged view of those two methods called PMPM (refer to section 2) was proposed by Valle, Santos, and Loures (2017) to guide the conduction of process mining projects. Such integrated view, comprising 14 activities distributed in 6 stages, is presented in Table 1.

Table 1: Integration of two methodologies for process mining projects

Stage	Activity
1-Scoping and Planning	1.1-Identify business processes and associated information systems, and gather basic knowledge 1.2-Determine goals and research questions 1.3-Determine the required team, data, techniques and tools.
2-Data understanding	2.1-Locate and explore required data in the system's logs 2.2-Verify the data in the system's log and select dataset in term of event context, timeframe and aspects
3-Data processing	3.1-Extract the set of required event data 3.2-Prepare the extracted dataset, by cleaning, constructing, merging, mapping, formatting and transforming the data 3.3-Familiarize and filter log
4-Process mining and analysis	4.1-Apply process mining techniques to answer (research) questions
5-Evaluation	5.1-Verify and validate process mining results 5.2-Accreditate process mining results 5.3-Present process mining results to the organization
6-Process improvement and support	6.1-Identify and implement improvements 6.2-Support operations

The first stage, *1-Scoping and Planning*, focuses on defining the scope (i.e business process, information systems, data types, etc) and the planning of the project, covering the definition of project goal(s); research question(s); project team composition; process mining tools, algorithms and techniques. The stage *2-Data understanding* covers locating, exploring and evaluating data in the information systems' logs. *3-Data processing* addresses the extraction, preparation, familiarization with and filtering the event log that is required for the application of the process mining techniques. The subsequent stage, *4-Process mining and analysis*, comprises the actual application of the previously identified process mining algorithms in order to answer the research question(s). Stage *5-Evaluation* aims to verify and validate the obtained process mining results as well as to present the results to the stakeholders. Verification here means technically assessing the outputs of the applied techniques to ensure correctness. Validation, in the other hand, means evaluating the degree to which the obtained outcomes represents the real process under analysis. Additionally, this stage covers the degree to which the outcomes meet the project goals and derived research question(s). The final stage *6-Process improvement and support* covers the identification and implementation of improvements and may also involve the provision of operational support through the use of process mining.

In fact, Valle et al (2017) have adapted such method to specifically extend a method called SCAMPI-Standard CMMI Appraisal Method for Process Improvement, which is used to carry out process appraisals based on CMMI-Capability Maturity Model Integration models. In this paper, in the other hand, we elaborated such extension in terms of tasks and subtasks enabling it to be used in process mining projects of any type: discovery, compliance checking or enhancement. For this, we defined the following "requirement" for the method: to be feasible, usable and useful, as characterised by Platts (1993). Taking this requirement into consideration, we applied the following design principles when refining the content of the detailed method:

- Completeness: the method should be comprehensive, i.e. to cover all the main aspects of a process mining project, right from its scoping and planning until the presentation of validated results and the support of operations, through process mining, if applicable.
- Simplicity: the method should be simple to follow, even by an analyst who is not so experienced in process mining.
- Flexibility: although with a high degree of prescriptive steps, including tasks and subtasks, the method should be flexible enough to address different types of project goals and/or questions.

Taking such principles into consideration, 24 tasks (and 21 subtasks) were derived from the pre-existing 14 activities. Another relevant aspect is that the method contains other guiding elements like examples of goals and questions, such as "which business rules conformance checking algorithms will be used?" and "what type of data will be needed for undertaking the project?". They aim to help accomplishing the project. Similarly, some checklists were designed and included to ensure that an activity (or task) has been entirely completed. Examples of checklist items are: "( ) verify process mining results" and "( ) validate process mining results".

The table 2 presents the tasks of each activity of the proposed method. The complete description of the method has several pages, but due to the length limitation of this paper, only an overview of the tasks is provided, below:

Table 2: Derived tasks of each activity in the method

Activity	Tasks
1.1	1.1.1-Identify business processes 1.1.2-Identify associated information systems 1.1.3-Gather basic knowledge
1.2	1.2.1-Define the goals for the process mining project 1.2.2-Based on the defined goals, derive relevant research questions
1.3	1.3.1-Identify the team 1.3.2-Identify the data that will be used in the process mining project 1.3.3-Identify the process mining techniques 1.3.4-Identify the tools required for process mining
2.1	2.1.1-Locate data reflecting the implementation 2.1.2-Explore the located data
2.2	2.2.1-Evaluate the quality of the event data 2.2.2-Select the data to be extracted
3.1	*<no task>*
3.2	*<no task>*
3.3	3.3.1-Load event log in identified Process Mining tool 3.3.2-Familiarize with the event log and process information that is contained in the event log 3.3.3-Filter an event log to consider only the interested data, by adding or removing information
4.1	4.1.1-Discover actual process from event log 4.1.2-Check conformance of event log with *de Jure* model 4.1.3-Compare conformance between *de Facto* model and *de Jure* model 4.1.4-Check conformance to business rules 4.1.5-Examine process mining results
5.1	5.1.1-Verify the process mining results 5.1.2-Validate the process mining results
5.2	5.2.1-Accreditate the process mining results
5.3	*<no task>*
6.1	*<no task>*
6.2	*<no task>*

Task 1.1.1 identifies the business process(es) that will be the object of the process mining project. Information systems related to such process(es) are identified in task 1.1.2. Contextual information is captured in task 1.1.3 called Gather basic knowledge. Task 1.2.1 defines the goal(s) for the process mining project while task 1.2.2 derives relevant research question(s) from the goal(s). The project team is identified in task 1.3.1 and the data to be used in task 1.3.2. Selection of process mining techniques and algorithms is done in task 1.3.3. Task 1.3.4 identifies the tool(s) required for process mining. Task 2.1.1 locates data reflecting the execution of process instances. Exploration of the located data to understand their purpose and structure is done in task 2.1.2. Task 2.2.1 evaluates data quality in terms of event context, timeframe and other pertinent aspects. Selection of data to be extracted occurs in task 2.2.2. In stage 3-Data processing, activity 3.1 creates the event log(s). The activity 3.2-Prepare the extracted dataset, by cleaning, constructing, merging, mapping, formatting and transforming the data has no tasks. Task 3.3.1 loads event log in the identified process mining tool. Familiarization with the event log occurs in task 3.3.2 while filtering of the event log happens in task 3.3.3. In the fourth stage, 4-Process mining and analysis, task 4.1.1 discovers the actual process from the event log. Conformance of event log with the de Jure model is checked in task 4.1.2. Task 4.1.3 compares conformance between de Facto model and de Jure model. Conformance checking to business rules is performed in task 4.1.4. Examination of process mining results, to gain detailed insight about the implementation of process instances within the organization is done in task 4.1.5. In stage 5.1-Evaluation, task 5.1.1 covers the verification of the process mining results while validation occurs in task 5.1.2. Accreditation of the process mining results happens in task 5.2.1. Activity 5.3-Present process mining results to the organization has no tasks. In the

final stage, 6-Process mining and support, there are no tasks in activities 6.1-Identify and implement improvements and 6.2-Support operations.

It is important to note that in addition to stages, activities, task and subtasks, the method also provides guidance on relevant elements present in the method, such as process mining techniques and algorithms. In this sense, the method brings examples for the following process mining techniques: conformance checking and business rules conformance checking. Additionally, for each technique, the method provides names of some existing algorithms. For instance, alpha miner, alpha++ miner, evolutionary tree miner, fuzzy miner, genetic miner, heuristic miner, inductive miner, multi-phase, organizational miner, role hierarchy miner, social network miner and transition system miner for process discovery. Similarly, conformance checker, graph matching analysis, differences analysis and footprint similarity for conformance checking.

5. A GUIDING TOOL FOR SUPPORTING THE METHOD

In order to meet the "requirement" for the method to be easy to use, we realised that a guiding tool, such as a software application, was needed to make that possible. When deciding how a tool would look like for supporting the execution of the method, we identified that the best way was to embed the guiding tool into a vendor software that actually implements the algorithms and plug-ins of process mining. So, we integrated the method into a process mining platform called UpFlux[3], which main features are process discovery and conformance checking, as presented in figure. 1.

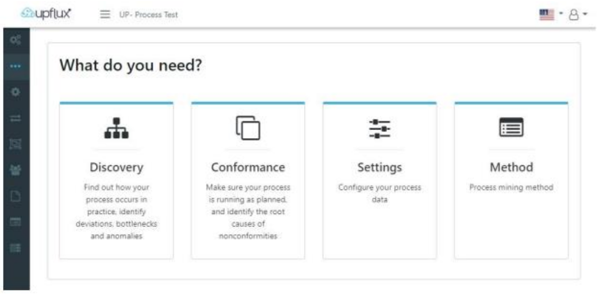

Figure 1. A process mining platform and its main features

The structure of the guiding tool can be divided in three parts: method maintenance, tailoring and execution. The first one focuses on maintaining (i.e creating, changing or removing activities in) the standard method. As presented in figure 2, our approach focuses on having a hierarchical structure where forms are built in the tool for supporting the execution of the stages, activities and tasks of the method when a process mining project is undertaken.

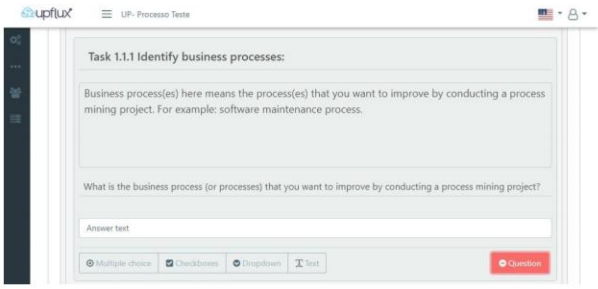

Figure 2. Guiding tool for maintaining the method

Secondly, there is a feature for customising, when applicable, the standard method for adding, changing or removing activities in order to create possible variations of the standard method to be later instantiated for each process mining project. This tailoring-feature allows different users to establish pre-defined variations of the standard method to be used for specific segments (e.g. healthcare), purposes (i.e. process discovery, conformance checking, enhancement) or types (e.g. question-oriented) of process mining projects.

As the third pillar, there is the "execution" instance that consists on following, step by step, the method version that has been instantiated for that specific project. In this feature, the tool records all the algorithms outputs as well as documents the analysis and its findings in order to be able to consolidate all this knowledge in a final project report (or presentation). Figure 3 presents a method task – 1.1.1-Identify business processes - ready to be undertaken. Note that this feature enables the process analyst to perform the method steps directly in the supporting tool. It also allows the analyst to write notes about analysis, filters, inconsistencies, strengths, weaknesses and other findings that, later on, will help analysing and deriving insights for process improvement.

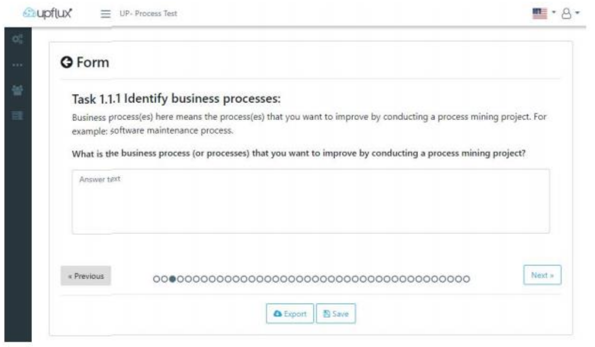

Figure 3. Guiding tool for executing the method

6. CASE STUDY

In order to evaluate our method (and the guiding tool), a case study was carried out. It concerns a process mining project of an industry setting.

6.1 Manufacturing production case

The case study examines the conduction of a process mining project in a large Brazilian manufacturing group that has over thirty thousand employees. The project goal, set in the 1-Scoping and planning stage, was to identify production performance issues in their paint factory. Questions such as "what is the most common production control flow?" and "which is the process variant with the worst lead-time?" were also defined in the first stage of the method, along with aspects such as related business process(es), associated information system(s), project team, data, and process mining tool and techniques.

In the second stage of the method, 2-Data Understanding, we located data in the system that support the execution of the process instances and we evaluated the quality of data. Some wrong event data were discarded, due to typing errors. Additional fields were mapped to the event log in order to compare departments, teams and shifts. After that, in the third stage called 3-Data Processing, relevant data were selected. The extraction was done using SQL expression resulting that

[3] http://upflux.net

database tables were transformed to event log. Later on, after loading the event log in the process mining platform, we familiarized with and properly filtered the event log.

In the subsequent stage, 4-Process Mining and Analysis, we discovered the actual process reflected by the event log. However, we did not check conformance or compare the discovered process model to the expected process model, as it was not a goal of this project. After, we applied the previously selected process mining techniques in order to answer the research questions.

In stage 5-Evaluation, it was the time to verify and validate the process mining results that were obtained as well as to present results to the stakeholders of the client organization.

In the final stage, 6-Process improvement and support, we identified improvements to be done. Implementation of such improvements were performed as per discretion of the client. The last task of this stage, where process mining may support the continuous operation, was not performed as it was dependent on decisions and further process changes by the client.

7. DISCUSSION

With the aim of evaluating the method and its execution in the selected case study, three criteria were considered: feasibility, usability and utility. We identified one question for each aspect in order to evaluate our method: a) Feasibility - Can the method be followed? b) Usability - Are method stages, activities, tasks and other instructions easy to use? c) Utility - Does the method provide a useful direction towards a solution for the problem that the method aims to address?

In terms of feasibility, the outcome is that the method can be followed, as it actually was for the case study. In terms of usability, the method and its steps, instructions and checklists, are simple to use, as no barriers or blocking points were faced. Finally, in terms of utility, the method actually provides guidance on what, how, when and where (to) apply process mining techniques and algorithms when undertaking process mining projects.

Concerning to the method content and structure, the overall perception from the process mining analyst leading the case, is that the method helps on increasing efficiency when running and managing a process mining project. "Following the method allows us to plan beforehand what to be done in each moment of the project, reducing anxiety, uncertainty and rework" said the analyst. However, the process mining analyst suggested to change the order of some tasks, such as 1.2.1-Define the goals for the process mining project, that could have been performed earlier. It was also recommended adding a step to identify and interview the process owner and project sponsor, before setting the project goal(s). In terms of example questions, it was suggested to consider the following ones:

- What are the process model variants? Are there any significant performance impacts among these variants?
- Which performance analysis can be applied? Are there relevant impacts caused by loops, or time-consuming activities, bottlenecks or waiting-time in some transitions?
- Are there any improvement opportunities related to the organizational structure? Do you recognise lack of available capacity in any role? When some department or role is allocated does occur some consequence such as delay or rework? Can you identify any overload of someone who is acting as a hub or as an activity distributor?
- What business process rules should be verified, such as separation of duties, limits of approval according roles, discount rules or free shipping rules?

Another feedback was how to handle conformance checking without a reference process model. As an answer, the most consistent process model could be taken as a basis, and then, to make some design adjustments and promote it as the reference model for conformance analysis.

Regarding the guiding tool and its dynamics, we divided our analysis in three aspects, as per the features in the software tool:

- In terms of the ability to create and maintain the standard method and its variation, there is a perception that, in many organizations the process improvement team will need to use - and sometimes to customise for specific situations - an organization's standard method rather than the generic standard method. Regarding that, our finding was that, this is expected, and it was one of the reasons for creating the feature that enables deriving variations of the standard method.
- In terms of the ability to customise a method variation for a particular project (i.e an instance), it demonstrated to be very useful in the case, although it was suggested to remove such steps from the method, making the method lighter to be followed when the process mining project does not aim to, for instance, check conformance.
- In terms of the execution of the method itself, in addition to a positive feedback, it was suggested to enhance the guiding tool, by automatically adding or hiding activities and tasks forms/elements in the method, according to the actual definition of some elements content, such as goals, questions, techniques or algorithms. For example, when conformance analysis is not pertinent, then conformance checking related method content (e.g tasks, questions) should not be displayed.

The last feedback from the case study is that the guiding tool needs to offer an alternative to skip or remove a set of activities when there is a native integration with process-aware system or an integration activated to a process mining platform used for continuous operational support.

8. CONCLUSION

The conclusion is that organizations seeking process improvement (or conformance) through process mining can now take advantage of an useful, usable and feasible step-wise method - and related guiding tool - that explicitly states which (how, where and when) process mining techniques and algorithms to use when carrying process mining projects out. As future work, in addition to addressing the feedback from the case study, we plan to extend the guiding tool in at least two aspects: 1) to enhance the integration of the execution of the method with the actual process miming algorithms and techniques used by the analyst. In other words, if there is a step in the method that requests, for instance, the use of the alpha miner algorithm, to enable the analyst to direct trigger such command in the tool, and 2) to control the execution of the process improvement proposals that are part of the report presented to the organization. In many organizations is very common to import these actions into a Kanban board tool for further agile project management, so we are evaluation on incorporating this feature into the tool.

Finally, a limitation of our work is that only one case is carried out, which may restrict our ability to judge how robust is the method for addressing other process mining types, such as conformance checking in periodic process audits. Aiming to address such limitation, more case studies are being planned in order to corroborate (or not) the findings presented here.

9. ACKNOWLEDGMENTS

The authors would like to thank the Waikato Institute of Technology and the Pontifícia Universidade Catolica do Paraná.

10. REFERENCES

Aguirre, S., Parra, C., & Alvarado, J. (2012, June). Combination of process mining and simulation techniques for business process redesign: a methodological approach. In International Symposium on Data-Driven Process Discovery and Analysis (pp. 24-43). Springer, Berlin, Heidelberg.

Bolt, A., de Leoni, M., & van der Aalst, W. M. (2016). Scientific workflows for process mining: building blocks, scenarios, and implementation. International Journal on Software Tools for Technology Transfer, 18(6), 607-628.

Curry, J. (2008). A Meta-methodology for the Design and Modeling of Patient Journeys.

Delias, P., Doumpos, M., & Matsatsinis, N. (2015). Business process analytics: a dedicated methodology through a case study. EURO Journal on Decision Processes, 3(3-4), 357-374.

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.

Ghasemi, M., & Amyot, D. (2016). Process mining in healthcare: a systematised literature review.

Meincheim, A., dos Santos Garcia, C., Nievola, J. C., & Scalabrin, E. E. (2017, November). Combining Process Mining with Trace Clustering: Manufacturing Shop Floor Process-An Applied Case. In 2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI) (pp. 498-505). IEEE.

Mundbrod, N., Beuter, F., & Reichert, M. (2015, September). Supporting knowledge-intensive processes through integrated task lifecycle support. In 2015 IEEE 19th International Enterprise Distributed Object Computing Conference (pp. 19-28). IEEE.

Muralidharan, K. (2015). Six Sigma for organizational excellence. Springer Proceedings of the Institution of Mechanical Engineers, 203(B1), 43-50.

Panpanich, P., Porouhan, P., & Premchaiswadi, W. (2015, November). Analysis of handover of work in call center using social network process mining technique. In 2015 13th International Conference on ICT and Knowledge Engineering (ICT & Knowledge Engineering 2015) (pp. 97-104). IEEE.

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. Journal of management information systems, 24(3), 45-77.

Platts, K. W. (1993). A process approach to researching manufacturing strategy. International Journal of Operations & Production Management, 13(8), 4-17.

Valle, A. M., Santos, E. A., & Loures, E. R. (2017). Applying process mining techniques in software process appraisals. Information and software technology, 87, 19-31.

van Cruchten, R. R., & Weigand, H. H. (2018, May). Process mining in logistics: The need for rule-based data abstraction. In 2018 12th International Conference on Research Challenges in Information Science (RCIS) (pp. 1-9). IEEE.

Van Der Aalst, W. (2016). Data science in action. In Process Mining (pp. 3-23). Springer, Berlin, Heidelberg.

Van Der Aalst, W., Adriansyah, A., De Medeiros, A. K. A., Arcieri, F., Baier, T., Blickle, T., ... & Burattin, A. (2011, August). Process mining manifesto. In International Conference on Business Process Management (pp. 169-194). Springer, Berlin, Heidelberg.

Van der Aalst, W. M., Bolt, A., & van Zelst, S. J. (2017). RapidProM: mine your processes and not just your data. arXiv preprint arXiv:1703.03740.

Van der Heijden, T. H. C. (2012). Process mining project methodology: Developing a general approach to apply process mining in practice. Master of Science in Operations Management and Logistics. Netherlands: TUE. School of Industrial Engineering.

Van Eck, M. L., Lu, X., Leemans, S. J., & van der Aalst, W. M. (2015, June). PM^2: A Process Mining Project Methodology. In International Conference on Advanced Information Systems Engineering (pp. 297-313). Springer, Cham.

Zerbino, P., Aloini, D., Dulmin, R., & Mininno, V. (2018). Process-mining-enabled audit of information systems: Methodology and an application. Expert Systems with Applications, 110, 80-92.

Zwolenski, M., & Weatherill, L. (2014). The digital universe: Rich data and the increasing value of the internet of things. Australian Journal of Telecommunications and the Digital Economy, 2(3), 47.