UNIVERSITY OF TARTU
Institute of Computer Science
Innovation and Technology Management Curriculum

Kateryna Kubrak

Visualizing Business Process Improvements

Master's Thesis (20 ECTS)

Supervisors: Fredrik Payman Milani
Alexander Udo Nolte

Tartu 2021

Visualizing Business Process Improvements

Abstract: Organizations have sets of business processes through which they produce value. To maintain high efficiency and better quality of services provided to their customers, organizations must constantly seek to improve their processes. While data-driven methods, such as process mining, can be used to analyze business processes and to identify improvement opportunities, process analysts take the specific decisions as to which changes to implement. To select which process changes to implement, the analysts need to understand the impact these changes would have on the process performance. Analysts often use process visualizations as part of their decision process to assess and determine which changes to pursue. Therefore, this thesis addresses the research question of how process improvement opportunities identified from event logs can be visualized to support the decision-making process of analysts. Thus, the contribution of this thesis is a set of guidelines to visualize process mining outputs for identifying improvement opportunities. This contribution is useful to those involved in the improvement of business processes. They benefit from this contribution as such visualizations of process changes support their decision-making and their communication with stakeholders.

Keywords: business processes, business process improvement opportunity, visualization

CERCS: P170 Computer science, numerical analysis, systems, control

Äriprotsesside parenduste visualiseerimine

Lühikokkuvõte: Organisatsioonidel on äriprotsesside kogumid, mille kaudu nad väärtust loovad. Nad peavad pidevalt otsima võimalusi oma protsesside parendamiseks, et tõsta klientidele pakutavate teenuste kvaliteeti ja säilitada kõrge efektiivsus. Kuigi äriprotsesside analüüsimiseks ja parendusvõimaluste tuvastamiseks saab kasutada andmepõhiseid meetodeid, näiteks protsessikaevet, on protsessianalüütikud need, kes võtavad vastu konkreetsed otsused selle osas, milliseid muudatusi ellu viia. Rakendatavate muudatuste valimiseks peavad analüütikud mõistma nende mõju protsesside toimimisele. Analüütikud kasutavad protsesside visualiseerimist sageli osana oma otsustusprotsessist, et hinnata ja määrata, milliseid muudatusi ellu viia. Seetõttu käsitletakse antud lõputöös uurimisküsimust, kuidas saab sündmuste logidest tuvastatud protsesside parendamise võimalusi visualiseerida, et toetada analüütikute otsustusprotsessi. Lõputöö panuseks on kogum juhiseid, mis aitab visualiseerida protsessikaeve väljundid parendusvõimaluste tuvastamiseks. Antud panus on kasulik neile, kes tegelevad äriprotsesside parendamisega, kuna selline protsessimuudatuste visualiseerimine aitab neid otsuste tegemisel ja huvirühmadega suhtlemisel.

Võtmesõnad: äriprotsessid, äriprotsesside parendamise võimalused, visualiseerimine

CERCS: P170 Arvutiteadus, arvanalüüs, süsteemid, kontroll

Contents

1 Introduction **6**

2 Background **9**
2.1 Business Process Management 9
2.2 Process Mining 11
2.3 Visualization 14

3 Related Work **16**

4 Methodology **19**
4.1 Requirements Elicitation Interviews 20
4.1.1 Data Collection 20
4.1.2 Data Analysis 26
4.2 Evaluation User Study 27
4.2.1 Data Collection 27
4.2.2 Data Analysis 29

5 Requirements **30**
5.1 Requirements Elicitation Interviews Findings 30
5.1.1 Themes 30
5.1.2 Findings 32
5.2 User Stories 39
5.3 Requirements Specification 40

6 Development **44**
6.1 Visualization Principles 44
6.2 Mockup Description 46

7 Evaluation **54**
7.1 User Study Findings 54
7.1.1 Themes 54
7.1.2 Findings 56
7.2 Additional Requirements 62
7.3 Additional Development 65

8 Discussion **68**
8.1 Limitations 69

9 Conclusion **71**

Acknowledgements **71**

References **72**

Appendix **77**
I. Interview Questions 77
II. User Stories 79
III. Requirements 82
IV. Licence 86

List of Figures

1 The BPM lifecycle [17] . . . 11
2 The Devil's Quadrangle [17] . . . 12
3 Three types of process mining [50] . . . 13
4 Research process . . . 19
5 Persona 1: Leo the Business Analyst . . . 21
6 Persona 2: Jane the Lead Process Analyst . . . 23
7 Persona 3: Lucy the Consultant . . . 24
8 Themes (Requirements elicitation interviews) . . . 30
9 Hierarchy of the mockup screens . . . 46
10 Screen 1: Overview - Summary . . . 47
11 Screen 2: Overview - Impact. Process Maps . . . 49
12 Screen 3: Overview - Impact. Process Performance . . . 50
13 Screen 4: Comparison (Separate) - Process Maps (Resources) . . . 51
14 Screen 5: Comparison (Separate) - Process Performance . . . 52
15 Screen 6: Comparison (Combined) - Process Maps (Resources) . . . 52
16 Screen 7: Comparison (Combined) - Process Performance . . . 53
17 Themes (User evaluation study) . . . 54
18 Screen 8: Overview - Summary (Additional Development) . . . 65
19 Screen 9: Comparison (Separate) - Prioritization (Additional Development) 66

List of Tables

1 Requirements elicitation interviews participants . . . 25
2 Evaluation user study participants . . . 28
3 Findings summary (Requirements elicitation interviews) . . . 38
4 Visualization principles summary . . . 45
5 Findings summary (Evaluation user study) . . . 62
6 Additional user stories (after evaluation) . . . 63
7 Additional requirements (after evaluation) . . . 64
8 User stories . . . 79
9 Requirements . . . 82

1 Introduction

Organizations have sets of business processes through which they produce value. Business processes are sets of coordinated activities performed in an organization to realize business goals [55], for example, an order-to-cash process which is initiated when the customer submits an order for a product or a service and is finished when the customer receives an order and makes a payment. Better designed and executed business processes lead to higher efficiency and better quality of services provided to the customers [17], for example, elimination of delays in delivery or reduction of costs related to the shipment for the order-to-cash process. As a result, this creates a competitive advantage for the organization, and so organizations must constantly improve their existing business processes. Thus, business process improvement is an essential step in managing business processes.

Business process management is defined as "a body of methods, techniques, and tools to identify, discover, analyze, redesign, execute, and monitor business processes to optimize their performance" [17]. These activities embody the business process management lifecycle. First, the organization identifies its processes and uses various criteria to assess which processes are of value. Those relevant to the targeted problem are then modeled as-is during the stage of process discovery, often with the help of diagrams, such as BPMN diagrams. Then, these processes are analyzed. The goal of process analysis is to gather insights about weaknesses in the process and their impact on the process. The outcome of the process analysis stage is the direct input for the next stage, process redesign. This stage aims to improve process performance through the performance dimensions, the most common of which are time, cost, quality, and flexibility. Thus, there can be multiple opportunities to redesign the process, and they should be properly analyzed and evaluated to select the right option [17]. Usually, all BPM lifecycle activities are carried out by a team that is engaged in a specific process improvement initiative [17]. Therefore, the team is responsible for process identification, modeling, analysis, and redesign. If the team completes the activities of identification and modeling manually, then the process improvement initiative is confidence-based, i.e., it describes the processes of the organization as they are presumed to be. This can evoke some limitations, such as the quality of the analysis and the scope of the analysis that the process improvement initiative can cover.

On the opposite, some stages of the BPM lifecycle, such as process discovery, can be completed using an evidence-based approach. It enables the process improvement opportunity to be based on the actual event data available in the organization. Event data is a set of records that is written in event logs that are generated by the information systems used by the companies [49]. The discovery of the business processes from the event logs is done through process mining. Process discovery, one of the process mining activities, allows for generating evidence-based process models, i.e., process models that

are built on the actual organization's data. Another way to use process mining is to do conformance checking, the comparison of the model, the modeled behavior, and the event log, the observed behavior, to identify deviations, violations, etc. [49]. Additionally, process mining can be used for performance analysis. For example, there are approaches to analyze process performance regarding its performance measures (time, resources, quality) to identify recurring working patterns, waiting times, resource performance, etc. [35]. Therefore, it is further possible to identify process improvement opportunities based on the retrieved information. A decision then can be made as to what to improve in the process. However, despite being supported by facts from event data, these activities still need to be carried out by specific decision-makers engaged in the process improvement project.

To make a well-informed decision about what changes to implement in the process, the decision-makers have to understand what impact these changes can have on the process performance. Sometimes, improving one of the performance dimensions of the process can decrease others [17]. For example, making a process quicker may result in decreasing its quality and vice-versa. Moreover, there might be several improvement opportunities for some processes, the combination of which may even make a higher impact on the process. Regardless, decision-makers need to understand the relationship between the process and the outcome [44]; therefore, they need to be able to predict the outcome while selecting the changes to the process. One well-established way to help the users comprehend data and support them in performing their tasks is visualization [47]. Thus, visualizing possible improvement opportunities and combinations of those could help the decision-makers to understand the relationship between the changes and their impact on the process. Current research in process mining visualizations focuses on a number of topics, such as visualizing process models information [47], visualizations for predictive process monitoring [33, 13], for process comparison [32, 38], and for change in process behavior [57, 25]. There also exist visualization frameworks, such as [45] that supports developers in designing process mining diagrams. However, these publications do not consider process improvement opportunities. This gap in visualization can limit decision-maker's ability to base the decisions concerning improvements in business processes on real-world data.

In light of this, we examine how to visualize business process improvement opportunities. This thesis addresses the following research question: **RQ.** *How can process improvement opportunities identified from the event log be visualized to support decision-making?* The contribution of this thesis is a set of guidelines to visualize process mining outputs for identifying improvement opportunities. This contribution will be useful to those involved in the improvement of business processes, for example, business analysts, process analysts, consultants, process owners, etc. They benefit from this contribution as such visualizations of process changes support their decision-making and their communication with stakeholders. To answer the research question, we followed the design

science approach. First, the requirements for the visualization were identified. Secondly, research on principles of information visualization was conducted. Next, the business process improvement visualization mockup was developed based on the requirements and literature on visualization. After that, user testing was conducted to evaluate the developed mockup and propose improvements. Lastly, a set of guidelines to visualize process mining outputs for identifying improvement opportunities was proposed based on the findings.

The rest of the thesis is structured as follows. Section 2 outlines the background where the concepts of business process management, process mining, and visualization are introduced. Section 3 gives an overview of the works related to this thesis and identifies the research gap. Section 4 outlines the research process. In Section 5, the requirements for the visualization mockup are given. Section 6 provides an overview of the mockup development process. In Section 7, the evaluation of the mockup with the users is described. Section 8 provides a discussion of the findings as a set of guidelines, and Section 9 concludes the thesis with a summary.

2 Background

This section introduces the key concepts and notions used in this thesis, such as business processes, business process management, process mining, and visualization.

2.1 Business Process Management

In a definition by Hammer and Champy [23], a business process is "a collection of activities that takes one or more kinds of input and creates an output that is of value to the customer". Dumas et al. [17] define it as "the chain of activities, events, and decisions which are performed by the organization to deliver value to its customers". An observation can be made that a business process describes any set of activities that an organization performs that at the end delivers value to the customer. Some examples of business processes include order-to-cash, a process that starts when a customer orders a process or service and ends when they receive and pay for it; issue-to-resolution, a process that starts when a customer submits a complaint and ends when the complaint is resolved, etc. [17]. The way business processes of an organization are designed and implemented has a direct impact on the customer. Thus, well-designed customer-oriented processes increase the quality of services that customers receive, and well-designed internal processes increase the efficiency with which they receive them [17]. Moreover, highly performing customer-oriented and internal business processes can help an organization to outperform another organization [17], so it is crucial for the organizations to constantly seek improvement opportunities for their processes to stay competitive.

Business process improvement is an essential part of managing business processes. As described by [17], "business process management (BPM) is a body of principles, methods, and tools to discover, analyze, redesign, implement and monitor business processes". These BPM activities follow the business process management lifecycle. First, business processes are identified in the process identification stage. The general purpose of this stage is to understand the process architecture of the organization, i.e., what processes the organization performs to achieve its business goals, and how they are interconnected. Having understood the process architecture, the organization can choose specific business processes that are investigated further. In the next stage, process discovery, business processes are modeled as-is, i.e., their current state is documented. Business processes can be modeled by the means of simple textual descriptions. However, the main purpose of process modeling is to facilitate communication between stakeholders [17], and diagrams serve that purpose better. The standard for process diagrams is the Business Process Model and Notation (BPMN) language. Once business processes are identified and discovered, they are analyzed. The goal of the process analysis stage is to identify and document the issues of the previously discovered business processes. To achieve that,

the business processes are quantitatively and qualitatively analyzed, and the issues are documented, quantified by performance measures, if possible, and prioritized based on their impact on the business processes. For example, combining qualitative methods as value-added analysis, waste analysis, and stakeholder analysis, and quantitative methods as flow analysis, queuing theory analysis, and process simulation allows for looking into the process from multiple perspectives [17]. The next step is to redesign the process to eliminate the issues identified in the previous stage during the stage of process redesign. The stage finishes with a to-be process design, i.e., a model of the improved process. In this stage, specific improvement opportunities are identified. According to [31], an improvement opportunity is "a pattern in the process execution that demonstrates weakness or has potential for improvement". It is essential to understand how different improvements would impact the process and its performance. Therefore, process analysis and process redesign stages are closely related as each redesign opportunity should be properly analyzed before it can be implemented [17]. Next, changes proposed in the previous phase are implemented in the process implementation stage. In this stage, changes required to go from the as-is model to the to-be model are performed. And finally, in the process monitoring stage, the new process is monitored and the data about the newly redesigned and implemented process is collected in order to analyze its performance. Process monitoring can help determine bottlenecks and errors, and also indicate the need to repeat the BPM lifecycle [17]. Figure 1 represents the BPM lifecycle.

Business process management can be useful for an organization for multiple reasons. As described above, well-designed internal and customer-oriented business processes can help an organization to outperform another organization. However, the processes cannot stay of the same high quality as the time progresses unless continuous improvement is applied to them. Moreover, if an organization does not seek improvement for its processes, its overall performance level decreases over time [2]. In order to evaluate their processes current performance, organizations can use process performance measures. Process performance measures are metrics usually derived from typical process performance dimensions of time, cost, quality, and flexibility [17]. Measures of time commonly encompass cycle time (the time from the start to the end of the process), those of cost concern any costs related to the process; quality can be measured as customer satisfaction, and flexibility measures the general ability of the process to reach changes [17]. It can be challenging to improve multiple performance measures at the same time. This challenge is well explained by The Devil's Quadrangle principle, that states that improving the process along one of the performance dimensions can decrease others [17]. As can be seen in Figure 2, moving along one of the directions can trigger the opposite one to move simultaneously with it. Thus, making a process quicker may result in its decreasing quality, and vice-versa.

There are multiple methods and techniques that facilitate business process improve-

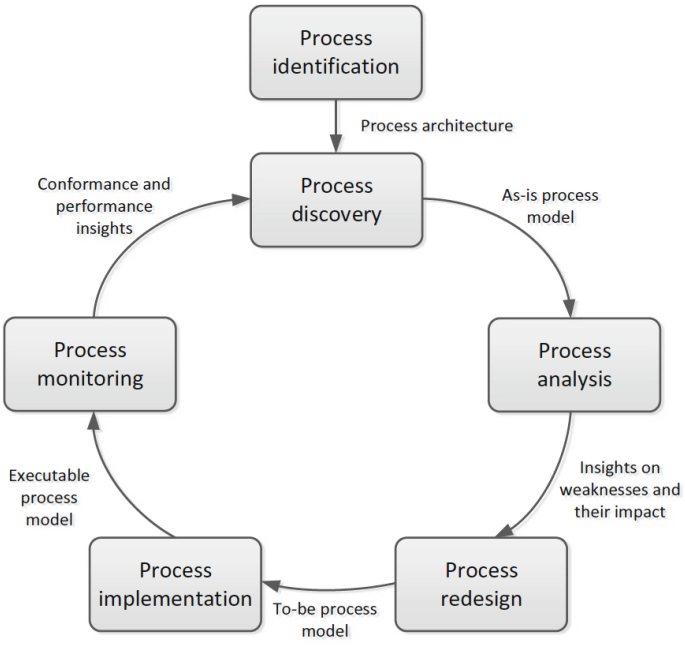

Figure 1. The BPM lifecycle [17]

ment during the different stages of the BPM lifecycle. For example, in the process analysis stage, such tools as the Pareto chart, Five whys analysis, and Cause-effect diagram can be used for the analysis of the specific drawbacks of the process [2]. Process redesign encompasses a number of transactional redesign methods (e.g., redesign heuristics, benchmarking, etc.) and transformational methods (e.g., business process reengineering, product-based design, etc.). However, before these methods for analysis and redesign can be applied, the process has to be modeled as-is. There are two approaches to process discovery: confidence-based and evidence-based. Process models can be modeled by organizations based on what the processes are presumed to be, i.e., these models are confidence-based. On the opposite, process models can be built on real data about organizations' activities and performance, i.e., such models are evidence-based. According to van der Aalst et al. [52], business process improvement can bring more to the organization's KPIs if it relies on the organization's event data. This data is the starting point of process mining.

2.2 Process Mining

According to the Process Mining Manifesto [50], "the idea of process mining is to discover, monitor and improve real processes (i.e., not assumed processes) by extracting knowledge from event logs readily available in today's (information) systems". Event

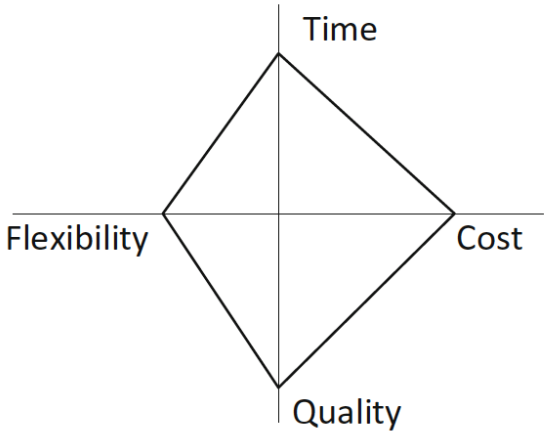

Figure 2. The Devil's Quadrangle [17]

logs contain data about events, which refer to an organization's activities, e.g. purchases made, orders placed, issues resolved. Data written in the event log can include the time of the activity, resources engaged in the activity, different attributes describing the activity, for example, what was the order placed for and how much money was paid for it. Therefore, process mining allows for analysis of these data to discover patterns, identify bottlenecks, explore frequencies, provide insights, etc. [49].

There are three main types of process mining: process discovery, conformance checking, and model enhancement (Figure 3). Process discovery aims at extracting process models from event logs [50]. Process discovery can be used when it is needed to discuss the problems among stakeholders (shared understanding of a process model helps reach consensus); to generate process improvement ideas (visual representation of the process and its problems reinforces re-engineering efforts) [49]. Conformance checking is used to check the alignment between the information extracted from the event logs and the process models existing in the organizations, such as business rules, procedural models, etc. Conformance checking can be used to audit and check compliance; to discover and analyze process deviations; to assess the quality of a process model [49]. The goal of model enhancement is to improve the existing process model with the information extracted from the event log [50]. For example, by looking into the resources performing different sets of activities, resource performance can be analyzed and assessed [49]. [35] introduces other examples of using process mining to analyze different aspects of process performance, for instance, measuring activities duration in hospital logs, analyzing waiting times in a make-to-order manufacturing firm. Aforementioned examples concern the performance dimension of time (activity duration and waiting time, respectively) but the authors additionally bring out examples for other performance

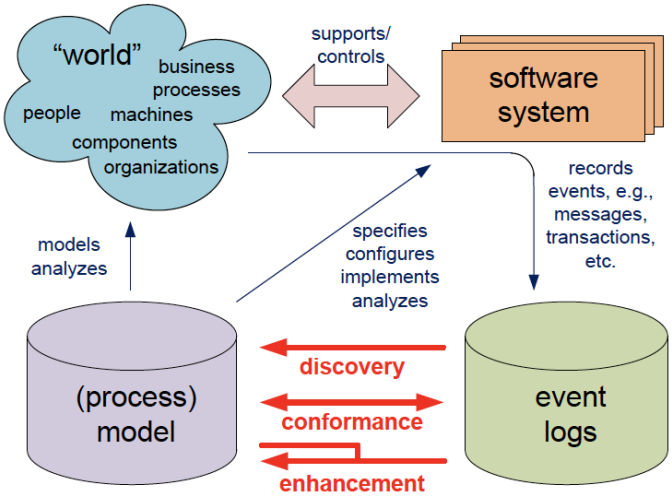

Figure 3. Three types of process mining [50]

measures (e.g., analyzing internal and external quality, resources performance, etc.) In summary, there exists evidence of using process mining to analyze process performance from the perspective of various performance measures.

Process mining provides an opportunity to identify the patterns and frequencies in the processes, therefore it can be used in most of the phases of the BPM lifecycle [52]. It is especially useful in the phase of process discovery, but it can also be applied in process implementation and process redesign. In these phases, insights learned during the analysis of the models derived from process mining can provide decision support [50]. Thus, a conclusion can be made that process mining can serve as a helper in the decision-making activities of the BPM lifecycle. According to the Process Mining Manifesto [50], "process mining provides an important bridge between data mining and business process modeling and analysis". The reasons for this derive from the following key characteristics of process mining: 1) it is process-oriented, unlike data mining which is data-oriented; 2) it is evidence-based, unlike traditional analysis approaches which are confidence-based (it generates information from existing data recorded in an organization's event logs) [49].

However, despite the constant growth of event data and therefore, growing abilities to generate more detailed real-world models from event logs, organizations still tend to lean more to easy-to-understand PowerPoint slides, management dashboards, etc. and make decisions based on those [49]. On the opposite, case studies that applied process mining to discover real processes in various organizations suggest that the insights derived from those come out as a surprise even to organizations themselves [49]. Thus, the link between business process improvement and process mining is crucial for finding process improvements based on evidence.

2.3 Visualization

Since each process improvement opportunity needs to be analyzed before it can be adopted, those deciding on implementing specific changes need to have a clear understanding of the potential impact of such opportunities on the process. One way to facilitate this is visualization as it can help the users comprehend data and support them in performing their tasks [47]. As stated by Chen et al. [11], "Graphic displays are often very effective at communicating information. They are also very not effective at communicating information"; so, what is important about visualization is not how much information it contains but rather how easy it is to understand that given information [48]. In the given context, it is crucial to find an understandable way to visualize potential changes to the business processes since a decision-maker needs to understand the possible outcomes of the change they are about to realize. Existing visualization principles and techniques provide guidance to selecting representations from a more general perspective, such as Shneiderman's Visual Information Seeking Mantra [43] or Munzner's Rules of Thumb [36], and on a more granular level, such as the Visual Comparison Model [20] for comparative visualization, or the Guide to Charts [58] on selecting those charts that best fit to the communicated information.

One of the most extensively utilized visualization principle is the Visual Information Seeking Mantra that is formulated as "Overview first, zoom and filter, then details-on-demand" [43]. This principle aims to describe how information should be represented so it is most effective for users [15]. It recommends to develop information visualizations in a series of "layers". First, an overview of the general context should be given where major components and their relationships are presented. Next, the user should be able to zoom and filter to reduce the complexity of the visualization and further organize the information. Lastly, details-on-demand means that the user can get additional details on the information of their interest by a certain simple action such as a mouse-over [15]. Moreover, Munzner [36] presents 8 rules of thumb that are advice and guidelines accumulated by the author based on the current state of knowledge and her own experience regarding which design choices are appropriate for what data and which task. Thus, according to the rule of "No Unjustified 3D", 3D visualizations should be used only to visualize inherently spatial data where the user's task is to understand a shape so a 3D visualization is required. Moreover, the rule "Get It Right in Black and White" suggests to ensure that the most important parts of the visualization are understandable even if they were presented in a black-and-white version. Another important rule of thumb is "Function First, Form Next" which implies that firstly, the visualization has to achieve its functionality goal, and after that, the shape can be perfected.

As to visualization on a more granular level, current research introduces several types of graphical elements and graphs that can be utilized for various purposes. For instance, [58] describes five basic charts (pie, bar, column, line, dot) and highlights that when choosing one, the kind of comparison for which they are chosen should be

considered. As such, if components are being compared, pie chart should be used as it aids in showing the structure of something, much like bar chart. However, bar charts are better suited for item comparison or representation of correlation in data. Additionally, line charts are suitable for time series representations, and clustered column charts should be used for comparison of sets of values for ranking. In addition to certain charts to use for visual comparison, there are also certain approaches to representing comparison as such. These approaches are juxtaposition, superposition and explicit encoding [20]. Juxtaposition is used to represent the compared objects separately. In such representation, it is the user who makes the connections between the objects. Superposition overlays the compared objects instead of representing them separately. This way, the user also makes the connections but such a representation helps since instead of trying to find the difference by imagining one object over another one, they are already overlaid. Explicit encoding calculates the differences between the compared objects and represents them using other visual encodings, so the user can clearly see the difference between the objects.

As highlighted by [1], for visualization to aid in decision-making, it has to consider the tasks that it is used for. As such, there are multiple case studies of using visualization particularly for process analysis, to understand which representations and elements to use for which purposes. Thus, in [46] the authors utilize the principles from [58] mentioned above. The area chart is used to visualize work in progress as it is a suitable representation to show quantitative values for different categories that have changed over time. Similarly, since column charts are suitable for visualizing the frequency and distribution for a range of quantitative groups, the authors utilize it for total cycle time distribution; a pie chart is used to represent the ratio between the average working time and the average waiting time. Thus, it is evident that selecting visual representations according to the task at hand facilitate decision-making and aid in process analysis [46].

In conclusion, there exist numerous visualizations guidelines and techniques that can also be adjusted for process mining purposes, such as which graphical elements to choose based on the communication purpose, and how to adjust visual comparison to the situation. One of such guidelines aims at helping to understand the task that the visualization is used for and select a proper graph based on that [58]. There exists evidence in the literature of using this framework for process mining visualizations, for example [46].

3 Related Work

In this section, an overview of related works is presented. The works discussed cover visualizations in the areas of process mining such as process discovery, process performance analysis, process comparison, predictive process monitoring, and change analysis. Additionally, existing visualization frameworks for process mining are described.

Visualization has been used in the topic of process mining for various purposes such as to visually present the outcomes of process discovery, visually facilitate process performance analysis, process comparison, and change analysis. The purpose of process discovery is to produce a model based on an event log that explains the most observed behavior [49]. To visualize such models, works on process discovery use various notations, such as BPMN [3, 4, 40], Petri nets [26, 54], causal nets [51]. Such discovered process models can be visually improved to reduce their complexity and increase understandability. For example, La Rosa et al. [30] describe a set of patterns that improve process model structure, reduce the model size, and increase the maintainability of the processes. Visualization is also used for process performance analysis, where the discovered process model is enhanced with data about the performance of this process with regard to its performance measures (time, resources, quality) to identify recurring working patterns, waiting times, resource performance, etc. [35]. For instance, [5] presents a set of visualization principles that allow to visually depict process execution dynamic on a process model. The example the authors propose visualizes activities' execution times and their frequency simultaneously on a process model. In [22], a rhythm-eye view, that uses a circular timeline to represent the running time of processes, is presented as a visual tool for process rhythm analysis to analyze time-related processes' behavior. [37] presents an approach to analyzing the dynamics of behavior of human resources over time and visualizing it by line charts. Thus, while existing work on process discovery and process performance visualization provides input to the research question of this thesis, such visualizations focus on showcasing the process model and enhance it with the process's performance. The focus of this thesis, however, is to build on such visualizations to visualize improvement opportunities.

In the area of process comparison, there also exist a number of works. Visualization has been used to illustrate differences between process variants. For instance, Pini et al. [38] propose techniques for multi-perspective process comparison that allow for visualizing process variants and their performances to make informed process improvement recommendations. [32] develops a number of visualization techniques to analyze resource and timing differences, and [14] presents an approach to calculate and visualize differences between process model variants in the context of multidimensional process mining. Visualization has also been used to compare process models. For example, [19] presents a difference graph visualization that uses color-coding and symbols to highlight the differences between the two process models. In [8], visualization is used to compare

event logs. The authors use different shades of color and thickness for nodes and edges in a transition system to highlight statistically significant differences between two event logs. Therefore, the aforementioned works showcase where the differences are, which can be a relevant input for this thesis, but they do not highlight specific improvement opportunities that can be derived from such comparisons.

Another area where visualizations are used is predictive process monitoring. [33] define predictive process monitoring as "a branch of process mining that aims at predicting at runtime and as early as possible the future development of ongoing cases of a process given their uncompleted traces". Therefore, organizations have the opportunity to take preventive measures in order to avoid violations, deviances, and delays in a process execution [16]. Examples of visualization for predictive process monitoring are [13] that extends a visual plug-in that uses map visualizing with risk-based metrics to support risk-informed decisions, and [53] that develops a technique to predict the remaining cycle time and visualizes the technique on a graph. Therefore, such visualizations provide insight for process management on a tactical level since they predict various characteristics of a process that has not yet come to an end. This thesis, however, focuses on visualizing improvement opportunities and thus, looks at a strategic level.

Visualizations are also used for the analysis of changes in process behavior, for example, in concept drift analysis. Concept drift is defined in [50] as "the situation in which the process is changing while being analyzed". Therefore, concept drift should be taken into account while analyzing the process to make a decision about the change [18]. [57] proposes a visual technique for drift detection in the form of an interactive system that uses maps, charts, graphs, and visualizes a number of statistics. [25] presents an approach to identify change in process behavior and, both programmatically and visually, compare clusterings of those to analyze the effect of changing behavior in the process. However, these techniques focus on visualizing changes that happen in the process due to various reasons while being analyzed. Our goal is to visualize where in the process there are improvement opportunities identified.

Additionally, there exist a number of various frameworks for visualization for process mining. For instance, [45] presents a framework for guiding developers of process mining techniques during the creation of process diagrams. The authors introduce a three-level hierarchy of the questions that can be answered while developing a process diagram that aims at easing the work of developers who often do not have experience in developing visualizations. Wynn et al. [56] developed a visualization framework for process performance comparison for multiple cohorts in a single visualization. The authors address the issue that visual comparisons of more than two process cohorts in one visualization are not well-developed, therefore, they create a new visualization using 3D space for global comparison. Therefore, these frameworks provide guidance on process discovery, process performance, and process change analysis but do not discuss identifying process improvement opportunities.

[45] furthermore provides an overview of commonly used methods for visualizing process mining outputs. According to the authors, node-link diagrams are the prevalent type of diagrams and are used in 24 out of the 48 studies analyzed. Thus, the aforementioned [5, 19, 8] include node-link diagrams in their visualizations. Other commonly used techniques are bar- and triangle charts, as exemplified in [38, 56], pie charts [5, 28], scatterplots [9, 6], line charts [37]. Therefore, while none of the works use the specified methods specifically to visualize improvement opportunities, they all are taken as input for this thesis.

The topic of improvement opportunities is also addressed in the commercial process mining solutions. For example, Celonis[1] offers a solution named Action Engine that is able to identify so-called "Signals" in the process and communicate them to the user. The Action Engine Signals are issues identified in the process and suggestions on the actions to take. They are communicated to the user via email or other communication channel, and are accessible in the form of a list in the Action Engine. However, such signals are minor issues in the process that do not relate to major improvement opportunities as such. Furthermore, they are presented to the user as text messages with description of the issue, while our goal is to find possibilities to visualize improvement opportunities in the process more elaborately.

In conclusion, there exist a number of works on visualizations in the field of process mining that can provide useful input to the research question of this thesis. However, existing research does not focus specifically on the visualization of process improvement opportunities. Therefore, this thesis contributes to the research with a set of guidelines on the visualization of improvement opportunities in a business process.

[1] https://www.celonis.com/

4 Methodology

This section details the research process and describes how the data required for answering the research question was collected and analyzed.

This research aims at understanding how to visualize business process improvement opportunities. The research question, formulated as RQ: *How can process improvement opportunities identified from the event log be visualized to support decision-making?*, relates to the broad topic of visualizations in process mining and covers a specific area of business process improvement opportunities. To address this research question, we followed the design-science research guidelines [21]. Design science fits our purposes since we aim to solve a problem (*how to visualize business process improvement opportunities*) by creating and evaluating an artifact (*a visualization mockup*). According to the guidelines, we first explore the problem relevance. We approach it through qualitative research involving practitioners from the industry. As a result, we elicit requirements for the artifact. We also base the development of the artifact on the visualization principles collected from the literature. Additionally, as suggested by the guidelines, we evaluate the artifact with its potential users through a user evaluation study. The research process is presented in Figure 4.

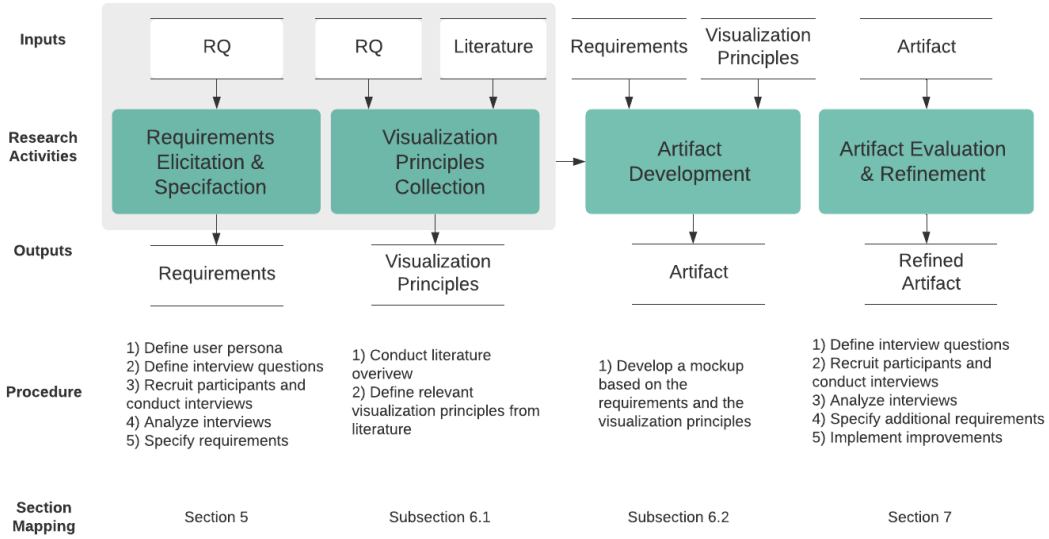

Figure 4. Research process

Firstly, requirements for the visualization were collected. To focus the visualization on its potential users, we created user persona, as they help consider user needs and goals when designing visualizations [34]. Then, we defined the interview questions that stem from the research question of the thesis. As a result, the interview questions cover

various aspects of working with improvement opportunities. Having the user persona and the interview questions ready, we recruited the interview participants from the industry and conducted the interviews. The goal was to deeply understand the needs of potential users, so interviews were selected as a form of user research [24]. Then, the interviews were analyzed using thematic analysis [10], and the requirements for the visualization were elicited from the interviews. The process of data collection and data analysis of the requirements elicitation stage is elaborated in subsection 4.1, and the requirements specifications are given in section 5. Secondly, we collected visualization principles from the literature on the visualization in general (briefly discussed in subsection 2.3) and visualization in process mining (briefly discussed in subsection 3). The visualization principles used are outlined in subsection 6.1. Thirdly, based on the elicited requirements and visualization principles, the artifact (a visualization mockup) was developed; the process is described in subsection 6.2. Lastly, we evaluated the developed artifact with the users in a user study. This method was chosen because it allows for identification of strengths and weaknesses of a visualization [29]. The design of the evaluation user study is outlined in subsection 4.2. The artifact was refined based on the findings from the evaluation user study (subsection 7).

4.1 Requirements Elicitation Interviews

This subsection describes how the data was collected (subsection 4.1.1) and how the collected data was analyzed (subsection 4.1.2) to elicit the requirements for the visualization mockup.

4.1.1 Data Collection

To elicit the requirements for the visualization of improvement opportunities identified from event logs, the perspective of those who will be using such visualization had to be considered. For the purpose of describing potential users, user personas were created (subsection 4.1.1.1). Based on the user persona, we recruited interview participants, and conducted individual online interviews with them (subsection 4.1.1.2).

4.1.1.1 User Persona

User personas allow for evaluating the visualization against the needs and goals of actual user personas [34]. For example, as pointed out by [47], some users may not have enough expertise to understand the layout of the process, which poses a challenge to make a visualization user-friendly and easy to understand. Therefore, introducing requirements from the perspective of user personas who have different levels of experience with process mining, different backgrounds, and different tasks to complete ensures the understandability of the visualization.

When creating user personas for the visualization, we diversified them against their experience (less experienced process analyst, more experienced process analyst) and their role (working internally in the company, working as a consultant). This way, the broad range of potential users could be considered when developing the visualization. The three developed user persona are presented in more detail further. For each user persona, the following points are covered: detailed description of the persona (their background, environment, and specific scenario that serves as the basis for identifying the persona's goals, needs, and pains), visualization of the persona.

Persona 1. Leo the Business Analyst (Figure 5)

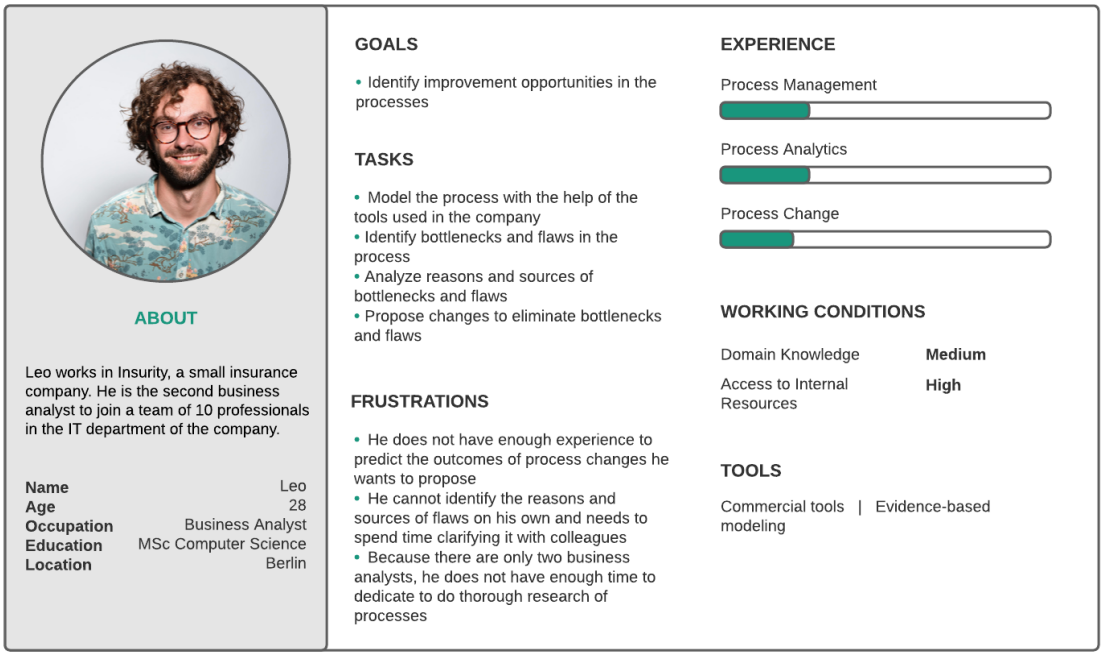

Figure 5. Persona 1: Leo the Business Analyst

Description. Leo is 28 years old and he works as a business analyst in Insurity, a small insurance company in Berlin. Leo is the second business analyst to join a team of 10 professionals in the IT department of the company.

Background: Leo has 5 years of experience in business analysis and project management. He holds a MSc degree in Computer Science. While studying, Leo had several courses on software development but has always been more interested in areas that did not include coding. Leo took the Business Analysis course as an elective and decided to explore the field in more detail after it. He took Business Process Management in the next semester and completed a few additional courses on Business Analysis online.

After graduation, he started working as a business analyst in a bank. He did not like the industry and switched jobs one year after. Before joining Insurity, he worked as a business analyst in another insurance company for three years. He has been employed by Insurity for half a year.

Environment: As one of the two business analysts in the company, Leo performs many tasks daily. They do not divide areas of responsibility with his colleague, therefore Leo constantly switches from one activity to another. Leo spent a few days last week preparing and conducting an internal workshop for their IT department on stakeholder management. This week, he is working on the business processes of the company. Next week, he will need to synthesize scattered documentation files produced by him and his colleague that have been piling up.

Scenario: Leo is working on identifying improvement opportunities for the company's claims handling processes. He is familiar with the main concepts and principles of business process management but spends a significant amount of time looking for theoretical and practical materials online. Leo uses a commercial tool for process mining. He has derived a few insights and identified bottlenecks from the process but he has little understanding of why things happen in the process in the way that they happen. He has to communicate with separate people in the company to analyze certain activities and events in the process. When he did so, going from person to person and asking questions took him a few days and in the end, it caused more confusion than clarifications. He is struggling to connect the flaws in the process with input that he received from his coworkers. He would complete his task more quickly if he had a tool that could help him identify opportunities to eliminate bottlenecks based on data.

Persona 2. Jane the Lead Process Analyst (Figure 6)

Description. Jane is 39 years old and she works as a lead process analyst in Softlabs, a medium-sized software development company in London. She works with 2 other process analysts and serves as a leader of their team.

Background: Jane demonstrates an extensive history of working with business processes since she has been engaged in jobs related to business process management for nearly 15 years. She obtained a MSc degree in Business Informatics and started her professional career as a business analyst in a small financial services firm. Shortly after, she decided to focus on process analysis and started working as a process analyst in Softlabs. Overall, she has been in this company for 12 years and took on the responsibilities of a team lead two years ago.

Environment: As a lead process analyst, Jane overlooks a few concurrent improvement projects. In her routine, she is responsible for producing documentation on process improvement opportunities and presenting it to stakeholders. It is her job to communicate the proposed changes to the stakeholders to ensure that they comprehend the impact of the changes on the business. The team uses an open-source tool for process mining and

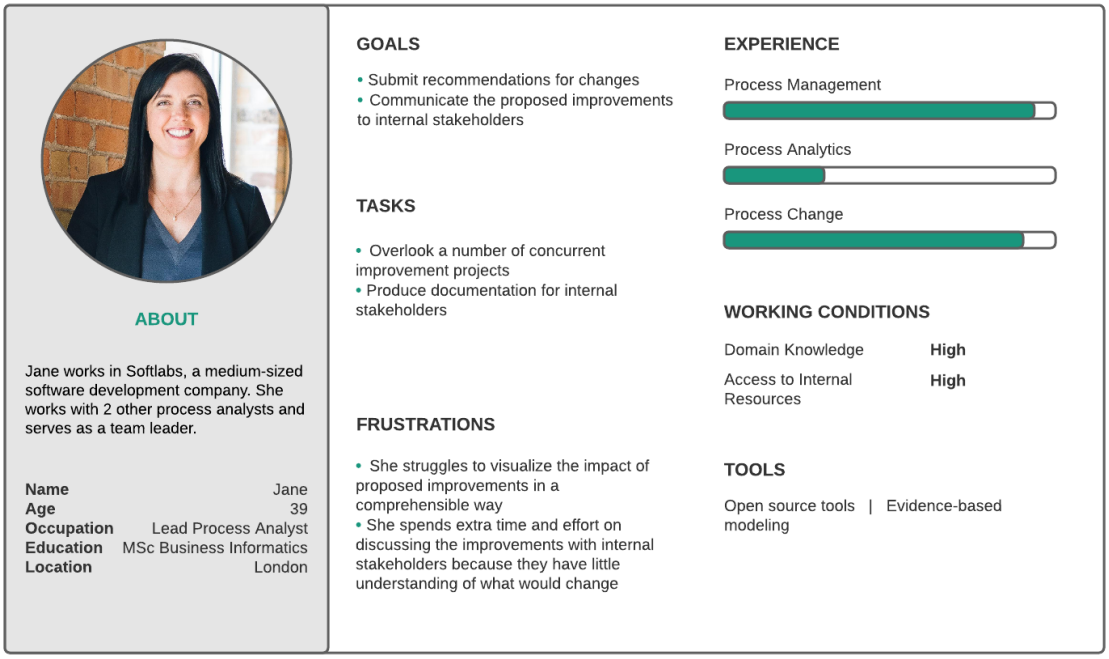

Figure 6. Persona 2: Jane the Lead Process Analyst

conducts analysis and redesign manually.

Scenario: Jane has a meeting with stakeholders tomorrow where she has to present the result of a month of work of her team. The problem the team has worked on was to investigate the quality assurance processes of the company and identify and propose changes to them. Some time ago, the team identified that the QA process was the slowest in the product development cycle, and had to investigate it. The team has identified a set of changes to the process which concern both changes to the resources and the time. Jane is trying to compose a comprehensive report to share with stakeholders. In her report, she would like to include concrete facts and figures to support the process changes her team proposes. However, she struggles to find a way to include such information, as she knows from her experience that raw numbers in the shape of diagrams and tables raise more questions from stakeholders and create confusion. She would be relieved if there was a way to visually describe this set of changes.

Persona 3. Lucy the Consultant (Figure 7)

Description. Lucy is 32 years old and she works as a consultant specializing in change management in Berata, a large consultancy firm in Stockholm. She works together with 15 other consultants.

Background: Lucy had known she wanted to be a consultant since she started her university degree. She holds a BSc in Management with a minor in change management.

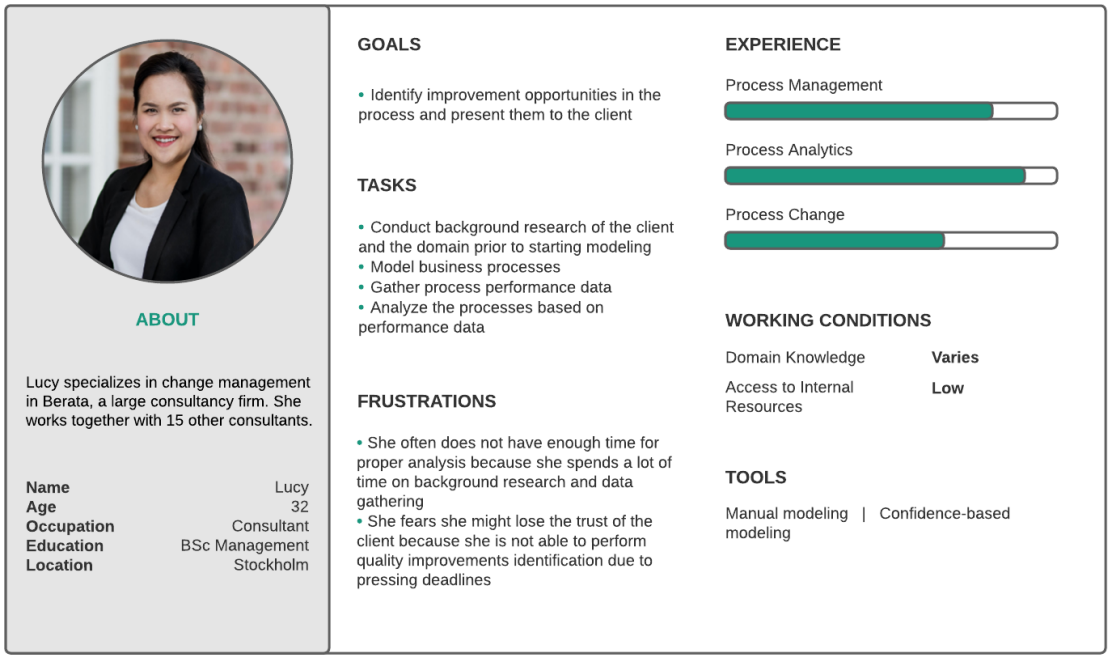

Figure 7. Persona 3: Lucy the Consultant

During her studies, she completed a few consultancy internships and started working full-time immediately after graduation. She never obtained a master's degree but she keeps learning and improving herself through various certification courses, trainings, and workshops. Up to now, she has 10 years of professional experience.

Environment: In her team, Lucy specializes in business process management. Her usual tasks require modeling business processes, gathering process performance data, and identifying improvement opportunities based on her analysis. Lucy performs modeling and data gathering manually. Together with her team, they compose and communicate the change management plan to the stakeholders. When Lucy is unsure about some of her decisions, she seeks advice from her colleagues who also have extensive experience in business process management.

Scenario: Lucy is working on a complex case posed by a client who is a healthcare company. The goal of the case is to improve existing business processes. Modeling the business processes of such a company requires a profound understanding of the context so Lucy spends a significant amount of time on interviewing, background research, and observations. Lucy also needs to spend additional time gathering performance data. Since there is a pressing deadline for the case, Lucy has to be efficient. She rushed data gathering due to lack of time so she missed a few important points. As a result, a list of proposed changes is not relevant to the case. Lucy would prefer it if she could have a tool that would support evidence-based process modeling and redesign. This way, she

would be more concerned with coining the most relevant change propositions rather than with modeling the initial process right.

4.1.1.2 Interviews

As the vision about the potential users of the visualization was obtained, we conducted user research. The user research goal was to obtain input on what information user personas need to see when working with improvement opportunities and how they interact with visualized information. As such, interviews were chosen as the method of user research since the aim was to deeply understand the expert knowledge of the participants [24].

In total, we conducted six interviews with professionals representing different user persona. Data saturation, i.e., no new information being provided by additional interviews [42], was reached after the fifth interview. However, one more interview was conducted to ensure the data was enough. Table 1 describes which domain each interviewee corresponds to, which user persona they represent, and how many years of experience with process mining they have. Participants that come from various domains and work on different process improvement projects were recruited for a broader range of perspectives. During the recruitment, we also ensured that the interviewees correspond to the previously described user persona and thus represent different levels of experience. An equal number of internal process analysts and consultants were selected.

Table 1. Requirements elicitation interviews participants

Code	Domain	Experience (years)	User Persona
I-01	Electrical engineering	2+	Leo
I-02	Data science	4+	Lucy
I-03	Auditing	2+	Lucy
I-04	Insurance services	1+	Leo
I-05	Process mining	5+	Jane
I-06	Public services	1+	Leo

The interviews were held online (Zoom, Microsoft Teams) individually with each of the participants. The duration of the interviews was between 29 and 46 minutes. During the interviews, multiple participants (I-01, I-03, I-04, I-06) shared the materials (screenshots, images shown directly in the software used, videos) of their projects in addition to orally describing them. The materials shared by the participants were also utilized during data analysis.

As the user research goal was to understand the information that the process analysts need while working with process improvements, the interviews were conducted in a semi-structured manner [41]. As pointed by [24], semi-structured interviews are used for

diving deeply into the topic and thoroughly understanding the answers provided. Each of the interviewees was asked to think about a research process improvement project that they were part of so that all the interview questions would be focused on one specific project. During the interview, we covered several areas that target the research question of this thesis. The areas are: general context of the project (what processes process analysts analyze), improvement opportunity (which process improvement opportunities they work with, how they prioritize them), analysis process (how they identify the opportunities, how and to whom they communicate them), and visualization (which visualizations they use for that). Firstly, we sought to understand the context of the process improvement project the interviewee was introducing: *"What was the case for which the process improvement project was initiated?", "Why was this case of interest?"*. Secondly, there was a set of questions regarding the improvement opportunity as such: *"What was the specific improvement opportunity identified?", "How was the improvement opportunity identified?"* These questions provide insight into which opportunities the interviewees work with in their projects and based on that, which of them can be used as examples in the visualization. Moreover, we also asked the participants about how they proritize the improvement opportunities: *"Were there any alternatives to the selected improvement opportunity? How was it decided which one to select? Who made this decision?"* This question area relates to the decision-making mentioned in the research question. Thirdly, to have an overview of the current state, we researched which visualization methods process analysts already use to identify their improvement opportunities: *"Were any visualizations used to help decide on the improvement opportunity? What was important to see visually to compare them?"* Additionally, we sought to understand how process analysts work with communicating identified improvement opportunities to the clients or business users. The motivation behind such questions was to assess whether the visualizations used for communication purposes are different (e.g., adjusted, simplified) from those utilized for process analysis. Examples of questions related to that are: *"Who were the results presented to?", "How were the results presented (images, dashboards, etc.)?"* Thus, these questions explore several areas that in turn allow for addressing the research question. The complete list of interview questions is available in Appendix I.

4.1.2 Data Analysis

All the interviews were transcribed with the Otter.ai[2] tool. After initially transcribing the recordings, they were then reviewed and edited manually since not all of the words were captured by the tool correctly.

The approach to analyzing the interviews was similar to thematic analysis [10]. First, it was essential to get familiarized with the data by reading it through and noting down ideas for codes and patterns. Next, we assigned initial codes to the data. As described

[2]https://otter.ai/

by [10], codes "identify a feature of the data that appears interesting to the analyst." We focused the initial codes on the questions areas described before. Some examples of the initial codes are: *"improvement opportunity"* for when the interviewees spoke about what improvement opportunities were identified in the process and how they were identified, *"improvement prioritization"* for when the means of prioritizing improvement opportunities were mentioned in the transcriptions, *"throughput times heatmap"* and *"constructing process maps"* for mentions of specific visualizations. As the next step, generated codes were sorted into themes. For this, the affinity diagram method was used [39]. It allows for quick organizing of data by transferring the data (interview excerpts with codes from the previous step) on sticky notes and arranging the sticky notes by creating groups of notes. Some codes appeared to be of a wider nature at this stage and thus became themes, and the interview excerpts inside the themes received new codes. For example, *"improvement opportunity"* was transformed into a theme named *"improvement opportunity description"* that now incorporated codes such as *"reworks"*, *"bottlenecks around waiting times"*. As this step was complete, the created themes were revised. At this point, some themes were collapsed together, while other themes had to be detailed more. Next, we defined and named the themes, thus clarifying the complete picture of the data. The themes and the interview findings are described in more detail in subsection 5.1.

4.2 Evaluation User Study

This subsection describes how the data was collected (subsection 4.2.1) and how the collected data was analyzed (subsection 4.2.2) to evaluate the developed artifact with the users.

4.2.1 Data Collection

To evaluate strengths and weaknesses of different visualization techniques and to test whether a visualization technique is useful in practice, user studies can be used [29]. The given user study was conducted in order to evaluate the artifact developed in section 6 from the perspective of industry practitioners, and in particular, to identify what the visualization is lacking and how it could be improved.

In total, eight practitioners participated in the evaluation user study (Table 2). We had the same requirements to recruiting the participants as in the interview round of the requirements elicitation (cf. subsection 4.1.1). The table has an additional column "Repeated" to mark the participants that had repeated participation in the interviews, i.e., they participated in both requirements elicitation and user evaluation interview stages. Having several interviewees participate twice allowed for comparing whether there was any difference in evaluation between the users who had already had some idea about the visualization and those who were completely new to it. Additionally, this

way we could assess whether the visualization needs were elicited adequately in the first round. Repeating interviewees have the same code as they had in the first round. New participants received codes with numbers following the previous batch (from I-07 onward). Data saturation [42] was reached after the sixth interview, but two more were conducted to ensure the evaluation was done from multiple perspectives. The interviews were held online (Zoom, Skype) individually with each of the participants. The duration of the interviews was between 32 and 49 minutes.

Table 2. Evaluation user study participants

Code	Domain	Experience (years)	User Persona	Repeated
I-01	Electrical engineering	2+	Leo	*
I-02	Data science	4+	Lucy	*
I-03	Auditing	2+	Lucy	*
I-07	Banking services	0.6+	Leo	
I-08	Banking services	2+	Leo	
I-09	Consultancy services	2+	Lucy	
I-10	Public services	4+	Lucy	
I-11	Robotics	4+	Jane	

We had a mixed approach to the evaluation user study, both observing the participants while they were interacting with the mockup, and interviewing them. At the beginning of the interview, a demonstration of the visualization was given to the interviewee. For the demonstration, the following scenario was used: *"We are analyzing a claim-to-resolution process in an insurance company. The process mining tool we are using has identified four improvement opportunities in the process: Order queues, Unnecessary job handovers, Activity rework, and Low resource capacity. Our task is to assess the identified improvement opportunities and choose which one to proceed with"*. During the demonstration, all of the developed screens (cf. section 6) were shown to the participant, based on the introductory scenario. It allowed for introducing all of the aspects that the visualization was built for, e.g., an overview of all the improvement opportunities together, an individual opportunity in more detail, prioritizing improvements, etc. After that, the participant received access to the mockup to use it themselves. During this stage, the participants were asked three questions that relate to the research question areas described in subsection 4.1.1.2. Firstly, we sought to understand which elements of the visualization aid the participants in identifying and analyzing an improvement opportunity: *"How do you proceed with assessing an individual improvement opportunity? What is important to see to assess it?"* (improvement opportunity). Secondly, we asked the participants to prioritize the identified improvement opportunities to explore which visualizations they use for that: *"How do you proceed with deciding which improvement opportunity to choose to address? What is important to see to decide on a specific*

improvement opportunity?") (analysis process, visualization). Thirdly, we focused on understanding whether the visualization helps them in communicating improvements: *"How do you proceed with communicating the chosen improvement opportunity to the business people/clients?"* (analysis process).

4.2.2 Data Analysis

To analyze the interviews, the transcriptions were processed in the same tool as previously, and edited manually afterwards. At this stage, we utilized the approach of thematic analysis again. Initial codes also focused on the areas of the research question described before. For example, there was a code named *"understanding the current problem"* which related to the participants explaining how they seek to understand what issue they are trying to solve in the process improvement project. Similarly, the code *"parameters for prioritization"* related to the area of analysis process and incorporated all interview parts where the participants spoke about the parts of the visualization that help them prioritize the identified improvements. There also were codes like *"interface"* for when the interviewees made general comments about the mockup. After composing the first themes and then iterating on them and the codes, as previously, some codes had to become themes and be broken down into smaller parts (the code *"interface"* became a theme combining other smaller codes related to different parts of the interface), or some themes were merged together. The themes and the interview findings are described in detail in subsection 7.1.

5 Requirements

This section covers the requirements for the visualization of the process improvement opportunities mockup. Particularly, subsection 5.1 elaborates on the interview findings, subsection 5.2 outlines the user stories developed based on the interview findings, and subsection 5.3 describes concrete requirements specifications.

5.1 Requirements Elicitation Interviews Findings

In this subsection, the themes from the thematic analysis of the interviews are introduced (subsection 5.1.1), and the findings from the interviews are given (subsection 5.1.2).

5.1.1 Themes

As described previously, we analyzed the interviews conducted with process analysts for requirements elicitation using thematic analysis (cf. subsection 4.1.2). As a result, eight themes were identified in the data (Figure 8). The full list of themes and codes can be viewed in the Google Sheets table[3].

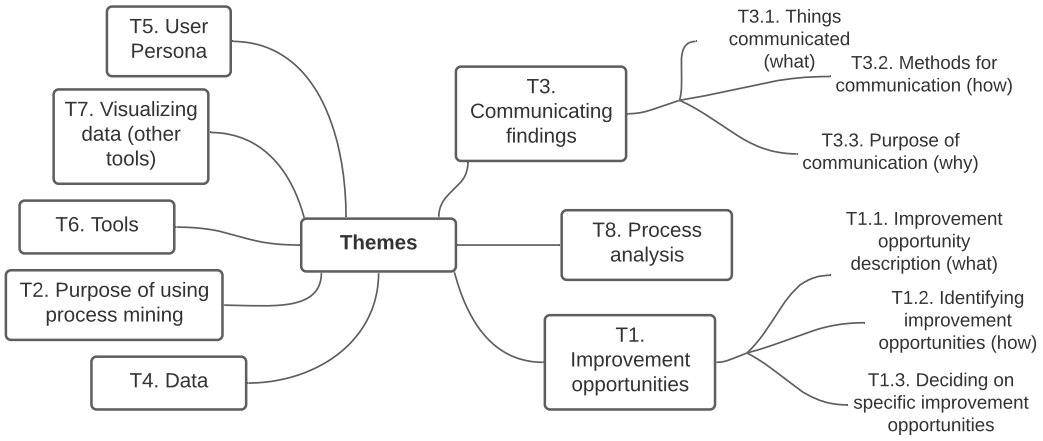

Figure 8. Themes (Requirements elicitation interviews)

Each theme captures a certain aspect of the data from the transcripts. As such, the themes are:

- **T1. Improvement opportunities.** This theme consists of three sub-themes that consolidate all data on the improvement opportunities. Thus, the sub-theme

[3]https://bit.ly/30N7Brz

T1.1. Improvement opportunity description (what) covers specific improvement opportunities that the interviewees identified in the process improvement projects they spoke about, e.g., reworks, bottlenecks around waiting times, etc. Next, the sub-theme *T1.2. Identifying improvement opportunities (how)* captures the data on how the specific improvement opportunities were identified, e.g., from the process mining software itself, by incorporating other methods. Finally, the sub-theme *T1.3. Deciding on specific improvement opportunity* describes what the process analysts need to know to prioritize improvement opportunities and decide which one to proceed with, e.g., financial criticality or dependability of the finding.

- **T2. Purpose of using process mining.** This theme incorporates the goals for which process analysts use process mining in their jobs, e.g., for audit, to design dashboards, etc.
- **T3. Communicating findings.** This theme consists of three sub-themes that consolidate the data on how and to whom process analysts communicate their findings from process mining. Thus, the sub-theme *T3.1. Things communicated (what)* covers the methods of representation of the process analysts' findings, e.g., technical reports, slides, etc. Next, the sub-theme *T3.2. Methods for communication (how)* captures the way in which the artifacts are communicated, e.g., interactively in the process mining software, and how the artifacts are adjusted, if required. Finally, the sub-theme *T3.3. Purpose of communication (why)* describes why the findings are communicated, e.g., for the end-users to see their performance, to confirm initial assumptions, etc.
- **T4. Data.** This theme captures the data-related aspects mentioned by the interviewees, e.g., issues with initial data.
- **T5. User persona.** This theme incorporates findings in the interviews that are specific to the user persona (c.f. Subsection 4.1.1.1), especially those concerning the responsibilities and characteristics of a certain user persona.
- **T6. Tools.** This theme captures what process mining and other analytical tools process analysts use in their process improvement projects and how they use them, e.g., process mining software, other analysis tools.
- **T7. Visualizing data (other tools).** This theme mainly contains the information from the additional materials provided by the interviewees (screenshots, videos, images shown directly during the interviews) on visualizations of the data in the tools that were mentioned in T6. As such, this theme contains descriptions of what is visualized (e.g., heat map with throughput times of different variants) with linked images of the visualization.
- **T8. Process analysis.** This theme contains additional data on process analysis that directly relates to the steps taken during the process improvement projects but not covered by the previous themes, e.g., approaches and analysis methods developed throughout the various projects.

5.1.2 Findings

This subsection gives an overview of the findings from several perspectives, such as how the process analysts identify improvement opportunities, how they assess them, how they prioritize them, and how they communicate them to their clients or business users. These perspectives derive from the research question areas described in subsection 4.1.1. Based on findings from each perspective, specific visualization needs are identified. For the most part, the findings are based on themes T1, T3, T5, and T7 since they directly relate to the topic of improvement opportunities and visualizations of those. Other themes, e.g., T2, T4, T8, provide an insight into the context around improvement opportunities projects the interviewees spoke about.

It is worth mentioning that while interviewee I-05 represents user persona Jane, they manage the team of consultants. Thus, while their input is mainly based on their role of a senior process analyst, there have also been multiple insights related to the tasks and responsibilities of user persona Lucy. Therefore, if the data from their interview was related to Lucy (i.e., describing how they approach the work with various clients), they were grouped together with other excerpts representing Lucy. Similarly, if the data from their interview was related to Jane (i.e., describing coordination of concurrent projects), they were grouped with other excerpts representing Jane.

Improvement Opportunity Identification

From the interviews, we found that the visualization of improvement opportunities should consider presenting the improvements from the perspective of relevant to the process metrics. Additionally, the findings suggest that the improvements should also be visualized on process maps. Lastly, it is evident that insights from both the process mining and the business side should be considered in the visualization.

Process analysts identify improvement opportunities based on the metrics relevant to their processes. For instance, cost savings are deemed as the most essential KPI when improving an order-to-cash process (I-01), while waiting times are important in a claim-to-resolution process (I-04), and processing and waiting times matter in an application-to-approval process (I-03, I-06). As stated by I-01 who works with an order-to-cash process, *"we of course want to focus on the ones that will bring us a lot of savings, for example."* (I-01). Similarly, I-06 who works on improving an application-to-approval process, states that *"the time between that you receive it and the first step is often long. It takes a long time before anything is getting done. And you have to remove this bottleneck because otherwise the cases might take very long time."* (I-06). Thus, it is evident that process analysts require being able to connect the possible improvement opportunities to the metrics that matter for their processes.

As to the usage of particular process mining visualizations to identify improvement opportunities, process analysts rely on process maps (I-02, I-03, I-04, I-06). For example,

one interviewee found their opportunities by comparing maps of the same process from different years: *"[I] did the analysis of the process and found out that the steps of the process in fact stayed the same. However, there was one difference. The duration between the steps changed."* (I-04). Other analysts give examples of comparing the process models for different attributes, such as *"different types of claims"* (I-03), and to *"compare the process across countries"* (I-01). However, process analysts mention having problems with understanding the data sources in situations when they do not have enough knowledge of the process (Lucy, I-02, I-03), or if there are problems with how data is recorded due to legacy systems (I-04, I-06). As a result, it can become challenging to understand the process map. As I-02 put it, *"[...] you always have challenges with the data sources. Not so much technical. About our understanding of the data, not messing up with too many assumptions. [...] And you have to decide how to interpret certain things."* (I-02). Based on that, cleaning the data and re-arranging it for it to be suitable for process mining is required (I-01, I-02, I-03). As highlighted by one interviewee, *"[...] the format they provided us, it was suitable for regulatory purposes, but wasn't ideal for process mining. So then we had to kind of mingle with the data, [...] and then we asked for additional historical information. And then we managed to reconstruct it to the point where it was useful."* (I-03). Therefore, process analysts utilize process maps in their analysis process which might indicate the need to see an improvement opportunity on the process map. However, when doing so, they report having issues with relying on data sources which might also imply that there can be lack of trustworthiness of the visualizations.

Process analysts also highlight the importance of approaching process analysis from both process mining and business side perspective to have a holistic overview of the findings (I-03, I-05). As one interviewee said: *"There are also some findings that may be very dominant and very interesting from process mining, but they have zero interest for business."* (I-05). Additionally, process analysts find it beneficial to understand the initial data (event log) to correctly understand visualizations of the process (I-02, I-03). As one process analyst motivates it, *"Oftentimes, if you don't understand how the data is generated and where the data comes from, you might misinterpret the visualization of the process mining software. And oftentimes, if you don't understand the business process, you might overreact to exceptions that are shown on the process map."* (I-03). Thus, it seems that the visualization of improvement opportunities should consider both aspects to the improvement, namely, representing possible process mining and business insights.

In conclusion, process analysts approach improvement opportunity identification from the perspective of what can potentially improve the type of the process they are working with, e.g., eliminating rework in the processes that are time-saving oriented. Thus, this indicates that the improvements in the process should be visualized from the perspective of metrics that are relevant to them. Additionally, process analysts utilize process maps in the analysis process to identify opportunities through comparing the

state of the process in different years or from different attributes. This finding shows that the identified improvement opportunities should be visualized on process maps. Moreover, when analyzing processes, process analysts consider both the process mining and the business perspectives of the process so that the opportunity is not perceived more critical than it actually is. This suggests that the visualization should consider insights from both perspectives.

Improvement Opportunity Assessment

Three aspects emerged from the findings on assessing improvement opportunities. Firstly, the findings indicate that it is important to include financial savings to visualize possible outcomes of addressing an improvement opportunity. Secondly, visualizing change in specific process KPIs is equally required to provide a holistic overview of the process. Thirdly, process analysts utilize several tools in addition to process mining software to create custom visualizations for their process. It thus might be considered to include customization and flexibility of provided graphs.

Process analysts consider what an improvement opportunity entails for the business in terms of the financial component. As put by Jane: *"[...] we associate the materiality and the importance of each one, so what does it mean if 1 million of your cases falls in this group? What does this have impact on? What does it mean for you as a business?"* (I-05). Additionally, process analysts focus on the change in process performance for specific KPIs. One process analyst gave an example of assessing *"[...] how the average throughput would change or how the resource utilization would change and that kind of things."* (I-03). As another interviewee summarized it, they need to understand *"how much the process would be improved from what performance measure; if the KPI is time, how much time would be saved"* (I-04). The findings indicate that to assess the identified improvement opportunity, it might be important for the process analysts to visually see the financial savings that it entails. Additionally, they seem to need to understand how certain KPIs can change if the opportunity is addressed.

Process analysts use additional tools together with the process mining software for a more detailed view of the data (I-03), or to combine process mining with business intelligence insights (I-06). As one process analyst put it, *"[...] that's why I like using BI tool. Because you need more than just process mining. You need selections and you need lists of cases, you need to go into details."* (I-06). Other motivations to use additional tools are constructing graphs and diagrams that are not supported by the process mining software (I-01), writing custom scripts (I-03, I-04, I-06), or creating more complex visualizations (I-03, I-04). As explained by I-03, *"But then for certain, more complex or not even not complex, but for certain analysis, I just wrote a Python script that used the tabular data."* (I-03). The recurring visualizations are heatmaps (I-01, I-04) and specific to the process histograms (I-01, I-03, I-04). These findings show that process analysts create other diagrams that are not standard to the primary process mining software they

use for a deeper insight into data. This might suggest that process analysts need to customize certain diagrams provided by the visualization of improvement opportunities.

User persona who have less domain knowledge require input from process experts or clients on whether certain findings in the process can be investigated further or implemented (I-01, I-02, I-03). For instance, Lucy commented on communicating with a client: *"[...] then we doublecheck this, and we send the preliminary findings to the [client] to confirm, because we are also learning how to use this and we are also trying to refine our own methodology."* (I-03). Another process analyst, Leo that works for multiple departments in the company, said: *"I'm not a process expert, right. [...] We have process experts in the company who design this process, [...], so we're then discussing it with market representatives, like, "Can we use it? How can they implement it?"* (I-01). Additionally, Lucy follows the same steps during the analysis of different clients' processes if the clients represent the same domain (I-03, I-05): *"Important information depends on the domain knowledge. So that could more or less stay the same across multiple clients."* (I-05). Thus, it seems that Lucy and Leo require intermediate confirmation of the findings from their clients or business users. Moreover, it seems to be important for Lucy to reproduce the same analysis steps for similar clients. However, while these findings are important for process analysis, they do not directly relate to visualization.

Therefore, to investigate further what an improvement opportunity entails, process analysts utilize various analysis possibilities of the process mining software in combination with other tools that provide more insight into the problem, e.g., BI tools. Process analysts use the tools primarily to create custom visualizations which might indicate the need to provide customization in the visualization. They also seem to take into account the financial component of the opportunity when assessing it. Its impact on the process performance is equally important. Thus, the visualization of improvement opportunities might have to include the financial savings and the process performance change associated with an opportunity. In certain situations process analysts also require input from the business users to have a holistic overview of the opportunity. This, however, does not directly relate to the visualization, despite being important for process analysis as such.

Improvement Opportunities Prioritization

When it comes to selecting an improvement opportunity to proceed with further in the process, process analysts take into account four conditions. Firstly, they want to understand what fraction of the process is affected by the improvement opportunity. Secondly, Leo and Lucy mention that it is important to adjust the process analysis according to the KPIs relevant to the process. Thirdly, they prioritize opportunities based on which one brings the most cost savings. Lastly, Lucy requires to understand whether the change can be implemented from within the system or it has links to entities outside of it.

Prioritizing the improvement opportunities from the perspective of their impact on the process, process analysts consider how many variants and cases are affected (I-01, I-03, I-05). This is motivated by the need to tackle business critical problems, i.e., those that involve a large percentage of the cases in the process. As explained by one of the consultants interviewed, *"So first of all, what is the importance of that finding, or so in terms of the impact on the overall process? So, for example, does it involve a large number of cases? Or does it involve a large population of the process?"* (I-05). Thus, it might be important to visualize the proportion of the process affected by an improvement opportunity.

Additionally, the opportunities are prioritized based on how much cost can be saved if they are addressed (I-01, I-05, I-06). In the previous subsection about assessing an improvement opportunity, it was mentioned that process analysts want to understand the financial component of the opportunity already while assessing it. It also becomes a prioritization criteria afterward. One interviewee explained it as following: *"We used to go with all of them [improvement opportunities], but now we would prioritize based on where we can gain the most, and where the biggest problems are."* (I-01). As such, the visualization might need to indicate which proposed improvement opportunity entails the most cost savings.

Another factor that Leo and Lucy consider is that of the ability of the improvement opportunity to change the KPIs that are relevant to the analyzed process (I-03, I-04, I-05). They motivate the need to understand that because *"it's important to know the concrete KPI from the beginning of the project but it happens that they need to be clarified with the managers and have to be readjusted, so what would change in the process using what specific KPI."* (I-04). As one interviewee explained, *"it depends a little bit on the opportunity as well; so under which angle you're looking at it? Are you looking at it from efficiency perspective? Are you looking at it from a risk perspective? Are you looking at it from a cash flow perspective?"* (I-05). Thus, when prioritizing the improvement opportunities, it is essential to keep in mind the reason for initiating the process improvement project. This might suggest that the visualization should specifically indicate which KPIs can be improved if the identified improvement opportunity is addressed.

There is also evidence that the user persona Lucy requires an understanding of whether an identified improvement opportunity can be addressed from within the organization or other entities have to be involved (I-03, I-05). It might be due to the fact that they have fewer links to the internal organization. As one interviewee explained, an improvement opportunity might be *"external to the process. So you can't do anything with that. So then you have to specify that the recommendation I'm making is intrinsic to the system or extrinsic to the system."* (I-03). Thus, in the case of having to contact additional entities, it may be not feasible or not possible from the perspective of the analyzed organization, and the improvement opportunity thus becomes obsolete. These findings suggest that the

visualization should specify whether the improvement opportunity has any connection to other parts of the organization.

In conclusion, process analysts prioritize their improvement opportunities based on several factors. They need to understand how big the opportunity is for their process, which might indicate the need to visualize the proportion of the process affected by the opportunity. Additionally, they prioritize the costs they can save from addressing an improvement opportunity. This suggests that the visualization should compare the opportunities from the cost aspect. Another criteria is ability of the opportunity to improve the KPIs relevant to the process that indicates that such KPIs might need to be explicitly specified in the visualization. Lastly, understanding whether an opportunity can be addressed from within the organization is deemed important and thus, the visualization should potentially indicate that.

Communicating Improvement Opportunities

The findings show that process analysts communicate their findings at several stages of the projects. When doing that, they use storytelling as a method of communication that helps create the flow and improve understandability. When communicating, process analysts adjust their means of communication to the client, and sometimes have to simplify their findings so that the end-users can comprehend them.

Communication of the findings from process improvement projects happens both as the last step of process analysis when the improvement opportunities are presented (I-02, I-03, I-04, I-05, I-06), and as an intermediate step when some initial findings have to be confirmed with the process experts and the business users (I-01, I-03, I-04). Process analysts adjust the means of communication based on the client or business users that they are showing their findings to, since *"some clients require more formal approaches"* (I-02), and some *"people in management who were more technical"* (I-03) require technical reports. These findings indicate that certain parts of the visualization might be utilized for communication purposes to users with different levels of insight into the problem and the technology.

Process analysts, for the most part (four out of six interviewees), simplify the way of presenting their findings to business users. For instance, using storytelling, first thinking about the story they want to tell with their findings and then selecting the images from the analysis tools to support that story (I-03, I-04). As described by one process analyst, *"And then you take that piece of information or data and try to put it in a simplified context that is, that works as a narrative and easy enough to understand in five minutes."* (I-03). Other ways include adding notes or comments if images from the analysis tools are used (I-01, I-03, I-04), and changing names and terms in a way that is more comprehensible for the business users of the process (I-04, I-05). As one interviewee put it, *"We didn't talk the jargon of the people who we were talking to, they had no clue what we are showing them. They would see a process, and a very complex process, but then the next question*

for them would be, "So what shall I do with this?" (I-05). Thus, the findings suggest that it might be important to provide an option to edit the visualizations to simplify them for understandability.

Table 3. Findings summary (Requirements elicitation interviews)

#	**Description**	**Based on (theme)**
F1	Process analysts improve processes based on the metrics that stem from the business objectives.	T1.1
F2	Process analysts rely on process maps to identify improvement opportunities.	T1.3
F3	Process analysts consider insights from both process mining and business sides of the analysis.	T8.2
F4	Process analysts manually create additional diagrams that are not standard to their process mining solutions but help them in the analysis.	T3.2, T7
F5	Process analysts require an understanding of the cost savings related to improving the process.	T1.3
F6	Process analysts consider process performance change when addressing improvement opportunities.	T1.3
F7	Process analysts sometimes require input from the business users when assessing an improvement opportunity.	T5.5
F8	Process analysts evaluate the proportion of the process involved in the improvement opportunity.	T1.3
F9	Process analysts prioritize more profitable improvements higher.	T1.1
F10	Process analysts need to understand which KPIs in the process an improvement opportunity affects.	T3.2, T1.3
F11	Process analysts consider whether an improvement opportunity depends on any entities external to the company.	T1.3
F12	Process analysts simplify their visualizations when presenting them to the clients or business users.	T3.2

In conclusion, process analysts adjust their means of communication mainly based on the clients' or business users' preferences. Nevertheless, it is a recurrent finding that process analysts use various simplification methods to what they communicate since only a limited circle of people has insight into process mining. Therefore, such people require a more standardized communication approach such as slides, simpler process maps, or additional notes and comments to the graphs used for analysis. It can be concluded then that providing a possibility to edit visualizations for simplification might be required.

Summary

All the findings described in detail in the previous sub-headings are summarized in Table 3. The table features a short description of the finding and specifies which theme the finding is based on.

5.2 User Stories

In order to develop requirements from the findings, first, a set of user stories was written. We chose this approach because user stories are defined as a method for describing functionality that will be valuable for a user of the system [12]. Thus, writing user stories allows for acknowledging that the visualization has to create value for the three developed user persona and therefore, prevents from adding irrelevant elements to it. User stories were composed for each of the three user personas. However, there are 15 user stories in sum for Leo, four for Lucy, and one for Jane. The reason for such difference is that all the interviewees, regardless of whether they are a consultant or a junior/senior internal process analyst, do process analysis in their daily job, as elaborated in subsection 4.1.2. Therefore, those stories written for user persona Leo capture the needs of all three user persona at the abstraction of process analysis. The stories developed for Lucy and Jane are unique to their role; however, they do not *contrast* the user stories written for Leo but *add* to them. For instance, user story US.16. *As a process analyst, I want to see the process maps of the old process and the improved process, so I can assess changes in the activities and the paths* is based on the finding F2 that relates to the interview excerpts of interviewees I-03, I-02 (both - Lucy) and I-04, I-06 (both - Leo). In contrast, user story US.17. *As a consultant, I want to understand whether the improvement is internal to the process, so I recommend changes that can be implemented from within* is based on the finding F11 that incorporates the data from I-02, I-03 (both - Lucy) which makes it unique to the role of Lucy and at the same time not relevant for Leo.

The developed user stories are grouped into five categories. The categories Improvement opportunity, Improvement prioritization, Communication, and Process analysis (other) derive from the question areas described in subsection 4.1.1.2 (improvement opportunity, analysis process, visualization). One other category, Business relevance, emerged as a separate category, while it initially was included in Process analysis (other). The latter category includes user stories that describe process analysis from various perspectives. However, the Business relevance category features multiple user stories that specifically highlight the importance of keeping the analysis in line with the business objectives of the process improvement project.

- **Improvement opportunity**. User stories related to the process of identifying and assessing process improvement opportunities. E.g.: US.5. *As a process analyst, I want to understand how many cases of the process are involved in the improvement, so I can assess its impact on the process.*

- **Improvement prioritization**. User stories that describe how the process analysts prioritize identified process improvement opportunities and decide which changes to implement in the process or recommend to the clients or the management. E.g.: US.1. *As a process analyst oriented at cost savings, I want to prioritize improvements, so I can choose the ones where the financial gains are the biggest.*
- **Business relevance**. Separate category of user stories that relate to assessing the identified process improvement opportunities against the business objectives while estimating and prioritizing them. E.g.: US.13. *As a process analyst reporting to management, I want to adjust KPIs when I clarify them with management, so I suggest relevant to the KPIs improvements.*
- **Communication**. User stories related to how process improvement opportunities are communicated. E.g.: US.11. *As a consultant, I want to assess the importance of the finding to the task, so I don't communicate irrelevant things to the clients.*
- **Process analysis (other)**. User stories covering other aspects of process analysis relevant to the process improvement projects but not covered by the previous categories. E.g.: US.7. *As a process analyst, I want to compare process performance statistics between different variants, so I can assess the uniqueness of the problem.*

The full list of user stories is available in Appendix II.

5.3 Requirements Specification

Concrete requirements specifications [34] are developed based on the aforementioned user stories (cf. subsection 5.2). As such, requirements are more detailed than the user stories and capture the "what" of the visualization rather than the "why" described extensively by the user stories. Three user stories are discarded in the requirements development process since they do not directly relate to the visualization of process improvement opportunities. While the omitted user stories assist process analysts in the analysis process, they are out of scope of the visualization of improvement opportunities. These user stories are: US.15. *As a process analyst working with many processes, I want to collaborate with the process experts, so that I better understand how the process is designed*, US.18. *As a consultant, I want to get input from the clients, so that I have a business-side understanding of the process*, and US.20. *As a lead process analyst, I want my team to assess the process from its domain perspective and the process mining perspective, so that both sides are included in the analysis.*

In total, 38 requirements were specified based on the remaining user stories. The requirements are grouped into categories that derive from those of user stories and are related to the question areas from subsection 4.1.1.2, too (improvement opportunity, analysis process, visualization). Some categories emerged as separate since they needed more detailing, e.g., the category Improvement assessment that is broken into smaller

categories of Implementability, Dependency, etc. As such, the categories represent various stages of process improvement projects and are as following:

- **R.II. Improvement identification**. Requirements related to displaying the general information about a single improvement opportunity. E.g.: R.II.1. *The system displays the improvements sorted into process performance dimensions*.
- **R.IA. Improvement assessment**. A set of requirements related to assessing an individual improvement opportunity from multiple perspectives, such as:
  - **R.U. Uniqueness**: assessing whether the identified improvement opportunity is prevalent in multiple variants and/or cases of the process. R.U.1.1.1. *The system displays differences between the process performance statistics of selected variants*.
  - **R.BR. Business relevance**: estimating the relevance to the objectives set by the business. E.g.: R.BR.2. *The system displays the improvements sorted into various categories of financial gains*.
  - **R.IP. Impact on the process**: assessing the differences between the original and the improved processes from the process map and process performance perspectives of view. E.g.: R.IP.1.2. *The system highlights the differences between the process maps of the original and the improved processes*.
  - **R.I. Implementability**: assessing whether the improvement can be implementable from the perspective of rules or laws. E.g.: R.I.1. *The analyst is able to specify that the change cannot be implemented*.
  - **R.D. Dependency**: displaying whether the implementation of the change to address the improvement opportunity depends on the other departments of the company or components of the system. E.g.: R.D.1. *The analyst is able to specify that the proposed improvement is extrinsic to the process*.
- **R.P. Prioritization**. Requirements related to prioritizing improvement opportunities to select the one(s) to proceed with in the process. E.g.: R.P.1. *The analyst is able to assign prioritization ranks to improvements*.
- **R.C. Communication**. Requirements related to communicating the improvement opportunities to the other users. E.g.: R.C.1. *The analyst is able to switch between an analysis and a reporting view*.
- **R.O. Other**. Other requirements not covered by the previous categories. E.g.: R.O.1. *The analyst is able to save prioritization options as a template*.

To decide which requirements to proceed with for the visualization, the developed requirements were prioritized. As defined by [7], requirements prioritization helps to define the most valuable requirements, i.e., distinguish the critical ones from the trivial ones. The requirements were prioritized using the MoSCoW method [27], according to

which requirements can be classified into four categories: must have (critical requirements that must be in the visualization), should have (high-priority requirements that are of high value but are not critical), could have (desirable but not necessary requirements), and won't have (requirements that will not be implemented in the visualization).

The requirements were classified into the four categories based on how frequently they were mentioned by the interviewees (must if >=3 interviewees), and by their relevance to the purpose of creating a visualization. As mentioned before, multiple user stories were disregarded before the requirements were written; however, a number of requirements still were found to relate to the functionality behind the visualization rather than the visualization itself, for instance, R.C.2. *The analyst is able to create a custom reporting view*, R.O.1. *The analyst is able to save prioritization options as a template*. Thus, such requirements (R.BR.2, R.C.2, R.C.2.1, R.O.1) were classified as "won't" during the prioritization.

In total, 17 requirements were classified as "must", where four of them are in the R.II. Improvement Identification category, seven belong to the R.IP. Impact on the Process category, and six are in the R.P. Prioritization category. Thus, identifying an improvement opportunity, assessing its impact on the process, and prioritizing the identified opportunities are perceived as the most important steps by the interviewed process analysts as these steps were mentioned by multiple interviewees, while others were specific to a separate user persona/interviewee. The visualization was created based on the "must" requirements as they are classified as critical and therefore, without them the visualization would not be valuable. "Should" and "could" requirements can thus be implemented in the next iterations of improving the visualization. The list of the "must" requirements is as follows:

- R.II.1. *The system displays the improvements sorted into process performance dimensions.*
- R.II.2. *The analyst is able to view the details of a single improvement opportunity.*
- R.II.3. *The system displays the process map of the process.*
- R.II.4. *The system displays the number of cases affected by the improvement opportunity.*
- R.IP.1. *The system displays the process map of the original process.*
- R.IP.1.1. *The system displays the process map of the improved process.*
- R.IP.1.2. *The system highlights the differences between the process maps of the original and the improved processes.*
- R.IP.2. *The analyst is able to select KPIs of interest.*
- R.IP.2.1. *The system displays the selected KPIs for the original process.*
- R.IP.2.2. *The system displays the selected KPIs for the improved process.*
- R.IP.2.3. *The system displays the differences between the KPIs of the original and the improved processes.*

- R.P.1. *The analyst is able to assign prioritization ranks to improvements.*
- R.P.2. *The system displays savings in the improvement for the selected KPI.*
- R.P.3. *The analyst is able to select the improvements for comparison.*
- R.P.3.1. *The system displays the process maps of the compared improvements.*
- R.P.3.1.1. *The system highlights the differences between the compared process maps.*
- R.P.3.2. *The system displays differences in the selected KPIs of the compared processes.*

The full list of requirements is available in Appendix III.

6 Development

This section describes the development of the artifact based on the requirements from section 5. In particular, subsection 6.1 outlines the visualization guidelines from the literature that are used in this thesis, and subsection 6.2 gives an overview of the mockup.

6.1 Visualization Principles

Based on the literature discussed in sections 2 and 3, certain visualization guidelines are considered when developing the artifact.

Firstly, Munzner's Rules of Thumb [36] are followed. "Function First, Form Next" suggests that the effectiveness of the visualization is required first, and the visual beauty can be adjusted after the effectiveness is ensured. "Responsiveness Is Required" prescribes that there should be interaction between the visualization and the users, for instance, that the user receives some visual feedback after completing an action. "Get It Right in Black and White" implies that the most important parts of the visualization have to be understandable even if they are presented in a black-and-white version. "Eyes Beat Memory" suggest to consider navigation in a single view to compare objects rather than use separate views for that since evaluating objects on a single view is perceived easier than to switch between them due to cognitive load on the user. One more rule is "Overview First, Zoom and Filter, Details on Demand" which is Shneiderman's visual mantra [43]. It provides guidance to "layering" the information: an overview should be given first, then the user should be able to zoom and filter to reduce complexity, and lastly, they can receive additional more detail information through a simple action.

Secondly, for the elements in the visualization that are represented as graphs, the guidelines from [58, 46] are followed to select the graph type that fits the task. According to the guidelines, to select the graph type, the message to convey and the kind of comparison should be considered. Therefore, as summarized by [46], the following graph representations suit such tasks: pie or area charts represent the data structure, clustered column or bar charts are used for value sets comparison , scatter and line charts help find correlations and see relationships, and line charts aid in analysis of time series and trends.

Thirdly, approaches to visual comparison [20] are considered for screens where comparison of elements is used. The approaches are juxtaposition, superposition and explicit encoding. Juxtaposition is used to represent the compared objects separately, superposition overlays the compared objects; explicit encoding calculates the differences between the compared objects and represents them using other visual elements. According to the authors, juxtaposition is best suited for comparisons where the user can match the objects themselves and thus notice various patterns and differences. On the other hand, if the data is compared on the same axis or if the visual comparison is created to detect

similar patterns between objects, superposition can be used. Lastly, explicit encoding suggests that the relationships between the objects are known beforehand and thus can be highlighted with certain visual elements. Combinations of the three approaches can be used if the data requires so. For instance, a combination of juxtaposition and explicit encoding represents the differences between the objects but shows them separately at the same time.

The outlined principles are summarized in Table 4.

Table 4. Visualization principles summary

#	Name	Description	Reference
VP1	"Function First, Form Next"	The effectiveness of the visualization is required first, and the visual beauty can be adjusted after the effectiveness is ensured.	[36]
VP2	"Responsiveness Is Required"	There should be interaction between the visualization and the users.	[36]
VP3	"Get It Right in Black and White"	The most important parts of the visualization have to be understandable even if they are presented in a black-and-white version.	[36]
VP4	"Eyes Beat Memory"	Consider navigation in a single view to compare objects rather than use separate view for that.	[36]
VP5	"Overview First, Zoom and Filter, Details on Demand"	Layer the information accordingly with overview first, then reducing complexity with zoom, filter, and more details.	[36, 43]
VP6	Graph type fits the task	Choose a graph for data representation according to the task that is supposed to be done.	[58, 46]
VP7	Comparison representation fits the objects compared	Choose a visual comparison type considering the types of objects that are being compared.	[20]

6.2 Mockup Description

Based on the requirements, static mockup screens were developed in the Balsamiq[4] tool. The static screens were then uploaded to the InVision[5] tool and connected with each other to provide its interactivity for future demonstrations. To do so, a hotspot was defined on the screen and linked with the other screen, e.g., clicking on the "Impact" button brings the user to the "Impact" tab. Thus, the static mockup was converted into a clickable, interactive mockup[6].

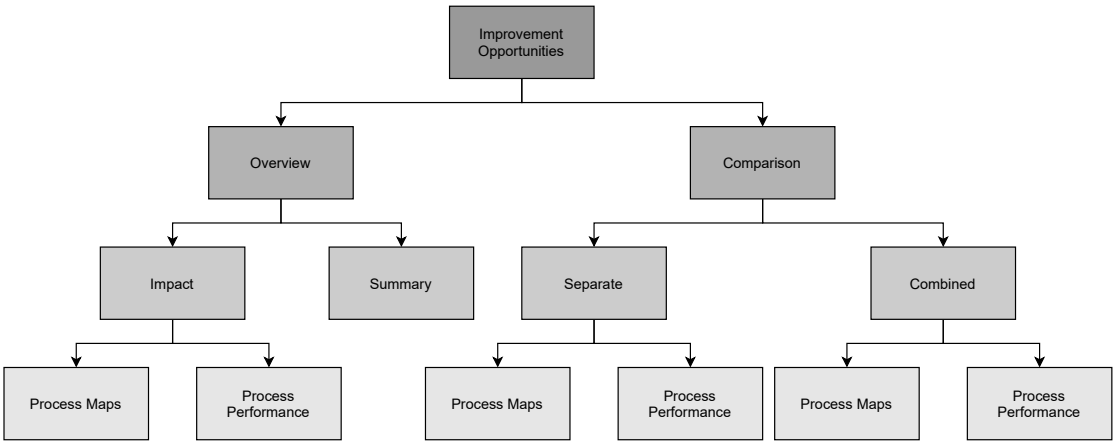

Figure 9. Hierarchy of the mockup screens

The visualization of improvement opportunities is presented as part of a web app of an imitation process mining tool to make it look more realistic. During the demonstration to the user study participants, it was also mentioned that we are using this hypothetical tool to identify the improvement opportunities (cf. subsection 4.2). Thus, every screen is visualized on a "browser window" element, and elements "user profile" and "help button" are also added to the header. All the mockup screens are developed using the principles described in subsection 6.1. For instance, the mockup does not use any colors except for white, gray, and black to ensure that the "Get It Right in Black and White" rule is followed (VP3). The only exception are the highlights on the process maps that are used to depict the improvement opportunities (described in more detail further). Additionally, the rule of "Function First, Form Next" (VP1) is considered too, focusing on implementing the requirements right before perfecting the visual appeal of the mockup.

The mockup is composed of 13 screens in total that together form a hierarchy (Figure 9). The highest element of the hierarchy, Improvement Opportunities, is a separate tab in the imitation process mining tool web app, where a user can explore

[4]https://balsamiq.com/
[5]https://www.invisionapp.com/
[6]https://invis.io/VR10FPZUY5FC

the improvement opportunities. All other screens are visualized as part of this tab. The Overview screen corresponds to the Improvement Identification and Improvement Assessment categories of the requirements. It is divided into two more screens, Summary and Impact, where the first one visualizes the Improvement Identification category, and the latter one visualizes the Impact on the Process subcategory. The Comparison screen corresponds to the Prioritization category of the requirements.

Overview - Summary

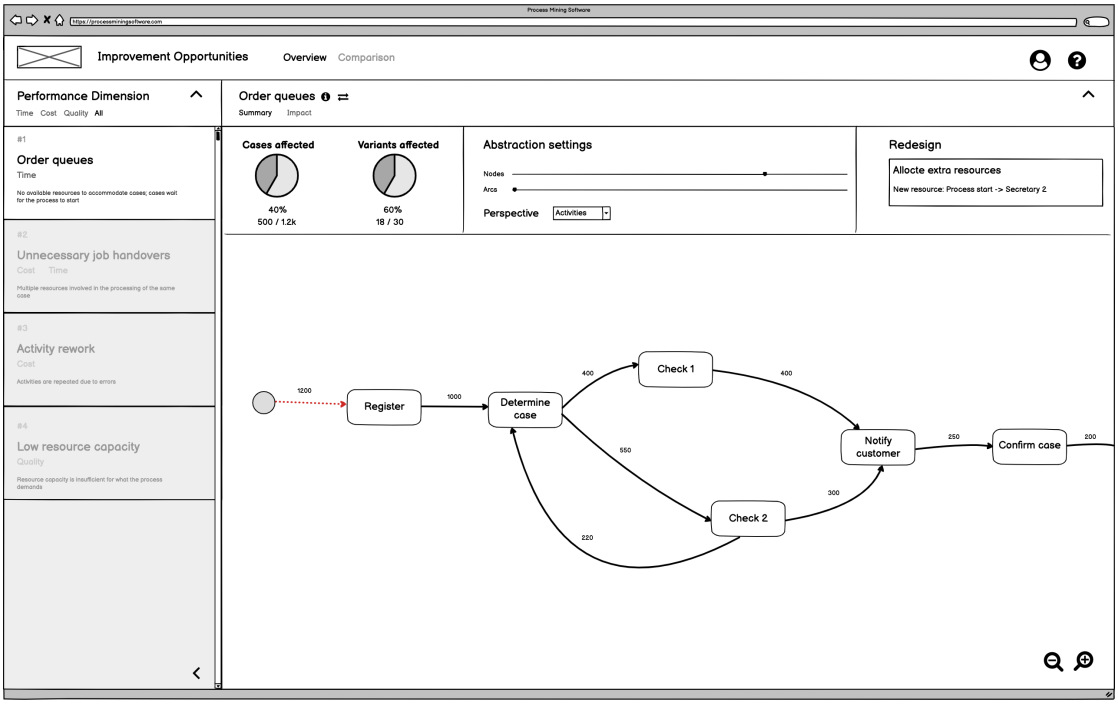

Figure 10. Screen 1: Overview - Summary

This screen covers the "must" requirements from the Improvement Identification category (Figure 10). Several principles discussed in subsection 6.1 are followed for this and the following screens. For instance, as prescribed by "Responsiveness Is Required" (VP2), certain elements in the mockup require an action from the user. This, however, is in line with another rule, the visual mantra, that suggests that details should be accessible on demand. Thus, it is ensured that the interactivity is used where it is actually required and not at the every step. As such, first, an overview of the improvement opportunity is presented to the user, and in the case of the process map on this screen, it is also highlighted with a different color on the screen; to get more details about it, the user has to hover on it. Similarly, to get more details about the impact of the improvement

opportunity on the process and its possible redesigns, the user has to select the respective options.

The screen consists of the following components:

- A list of identified improvement opportunities that are tagged with performance dimensions of time, cost, and quality [17], and are sorted according to the dimensions chosen (cf. R.II.1). Note: current selection in the given screen is "All", whereas all identified improvement opportunities are displayed together. The four improvement opportunities in this mockup, Order queues, Unnecessary job handovers, Activity rework, and Low resource capacity are selected from the Improvement Opportunity Framework [31].
- A separate screen with the details of the selected improvement opportunity (cf. R.II.2), where these details are outlined in the following requirements.
- A process map of the process for which the improvement opportunities are proposed (cf. R.II.3).
- A highlighted in red path that correlates with the improvement opportunity analyzed (cf. R.II.3.1). There are no activities correlated with this improvement opportunity in the given screen (cf. R.II.3.2) but it is highlighted in the same manner.
- Two pie charts that display the number of cases and variants affected by the improvement opportunity (cf. R.II.4). Pie chart was chosen for this kind of data as according to VP6, they are suitable for representing the size of each cluster as a part of the whole.
- Standard abstraction settings where the user can select how many arcs and nodes are displayed on the process map.
- Redesign option(s) available to resolve the improvement opportunity. The redesign option for the improvement opportunity of the given screen, Order queues, was also adopted from [31].
- Option to switch the perspective view from Activity to Resource. In the given figure, the Activity perspective is selected. The Resource perspective can be viewed in the interactive mockup[6].

Overview - Impact (Process Maps and Process Performance)

These two screens cover the "must" requirements from the Impact on the Process subcategory, where Figure 11 corresponds to the requirements R.IP.1-R.IP.1.2 and visualizes differences between the process maps, and Figure 12 corresponds to the requirements R.IP.2-R.IP.2.3 and visualizes differences between the KPIs of the process.

The first screen, Impact (Process Maps) (Figure 11), shares a number of same elements with that of Summary screen, such as a list of identified improvement opportunities, the pie charts that display the number of cases and variants affected, the abstraction

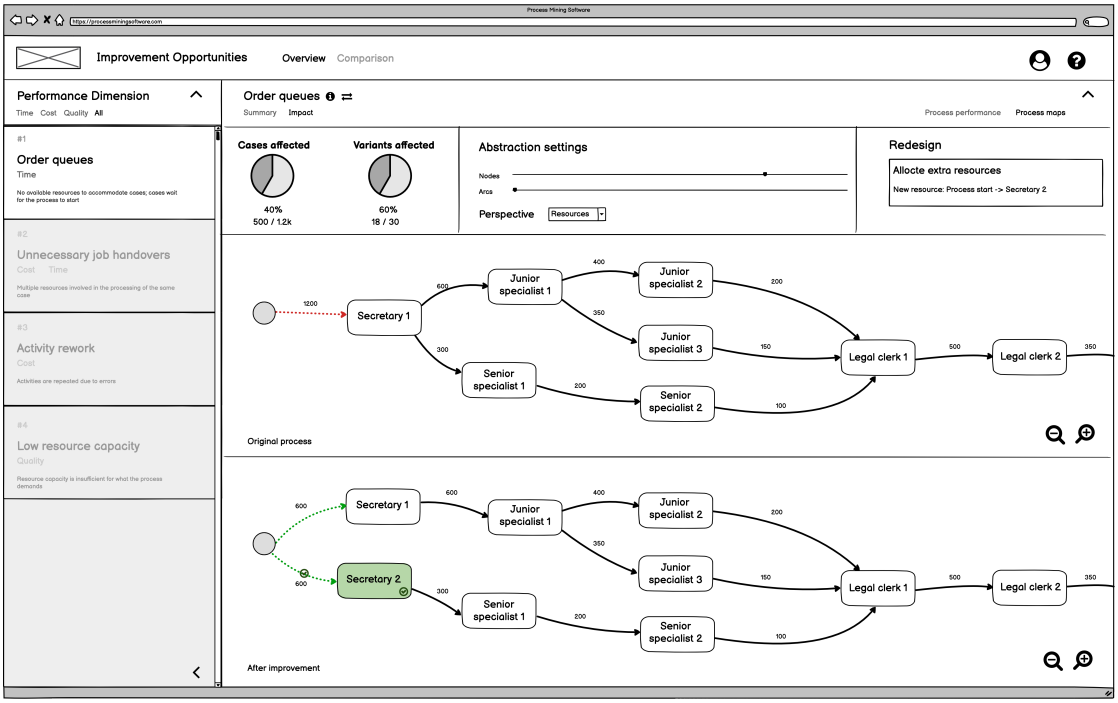

Figure 11. Screen 2: Overview - Impact. Process Maps

settings, the redesign option, and the Activity-Resource perspective switch. However, the purpose of the Impact screen is to visualize what changes in the process given that the proposed improvement opportunity is addressed. Thus, on the process maps, the user can see what activities, resources, and paths are added or removed from the process (cf. R.IP.1.2), and which activities, resources, and paths are updated compared to the original process (cf. R.IP.1, R.IP.1.1). The highlighted parts on the process maps are based on [19], where the authors use color coding and symbols to highlight the differences between the process models. As such, green color in combination with a checkmark are used to mark the addition of a new activity/resource/path, while red color in combination with a cross are used to mark the deletion of an activity/resource/path. A combination of juxtaposition with explicit encoding [20] (VP7) is used so that the user can explore both process maps but also explicitly see the differences between those. As the current figure displays the process maps from the Resource perspective, the Activity perspective can be viewed in the interactive mockup[6].

The second screen, Impact (Process Performance) (Figure 12), shares the same elements with the Summary screen as the Impact (Process Maps) does. However, instead of visualizing differences between the process maps, this screen visualizes the differences between the KPIs of the process. For the purpose of the mockup, four various KPIs are available in the drop-down list (cf. R.IP.2): average throughput time, average resource

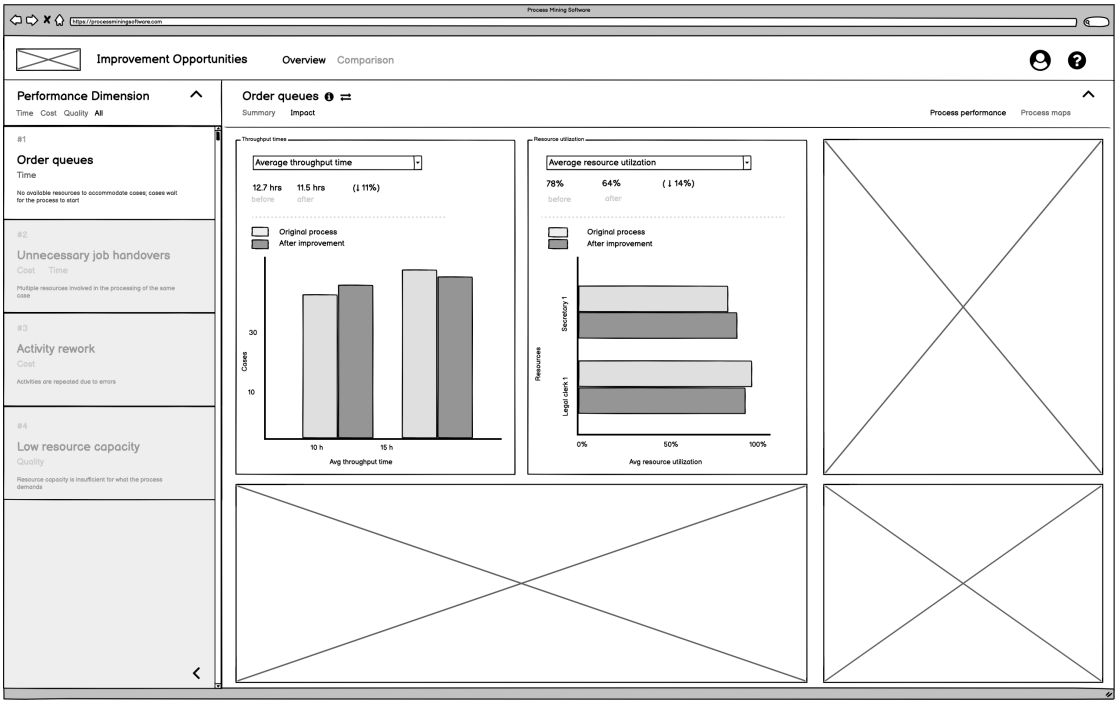

Figure 12. Screen 3: Overview - Impact. Process Performance

utilization, cost per execution, error rate (all adopted from [17]). Two of them, average throughput time and average resource utilization, are visualized. Such selection is justified by multiple interviewees mentioning throughput time (I-01, I-03, I-04, I-06) and resource utilization (I-01, I-03) as important KPIs for their processes. Both KPIs are visualized as a graph that depicts the differences between the original and the improved process (cf. R.IP.2, R.IP.2.1, R.IP.2.2), and additionally, consolidated difference statistics are given (cf. R.IP.2.3). Throughput time is visualized as a column chart, and resource utilization is visualized as a bar chart based on VP6. Moreover, juxtaposition is selected [20] (VP7) since by comparing the bars located next to each other, it is possible to notice the differences without additional visual elements. The throughput time chart thus displays how many cases in the process, for instance, have throughput time of 10 and 15 hours. The resource utilization chart displays resource utilization for Legal clerk 1 and Secretary 1. Consolidated difference statistics thus show average throughput time and average resource utilization, respectively.

Overview - Comparison (Separate and Combined)

These two screens cover the "must" requirements from the Prioritization subcategory (requirements R.P.1-R.P.3.2), where Figure 13 visualizes differences between process maps of the compared improvement opportunities, and Figure 14 visualizes differences

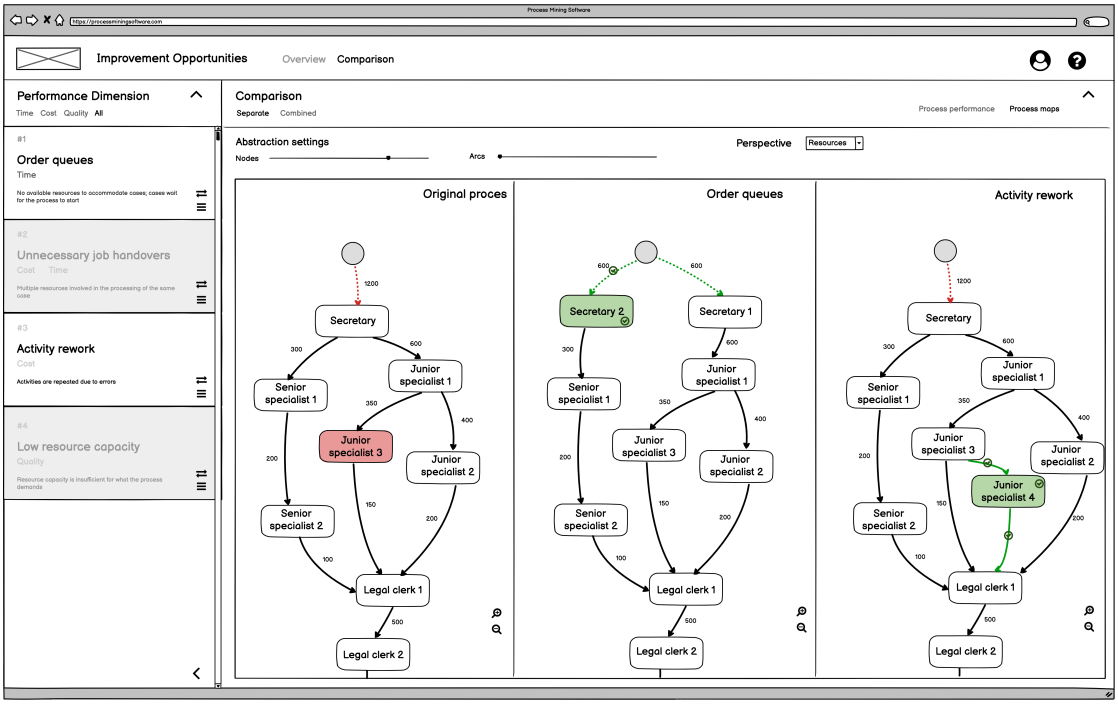

Figure 13. Screen 4: Comparison (Separate) - Process Maps (Resources)

between their selected KPIs. These two screens visualize the original process and the potentially improved process. The first potential improvement is that of #1 Order queues for which the other screens have been described previously. The second potential improvement is that of #3 Activity rework, which was also selected from [31].

The first screen that shows differences between the process maps of the process before improvement and process after improvement (with opportunity #1 or #3) (Figure 13) shares a number of same elements with that of Summary screen, such as a list of identified improvement opportunities, the pie charts that display the number of cases and variants affected, the abstraction settings, the redesign option, and the Activity-Resource perspective switch. The visualization of the three process maps is based on the same set of principles described for the Impact (Process Maps) screen (Figure 11) with color coding and symbols, and also a combination of juxtaposition and explicit encoding is used. The purpose of this screen is to visualize the differences in the process on its process maps (cf. R.P.3-R.P.3.2).

The second screen (Figure 14) thus shares the same elements and follows the same visualization principles with the Impact (Process Performance) screen (Figure 12). To visualize the differences between the process before improvement and the process after improvement with potential opportunity #1 or #3, additional bar/column with a different color are added.

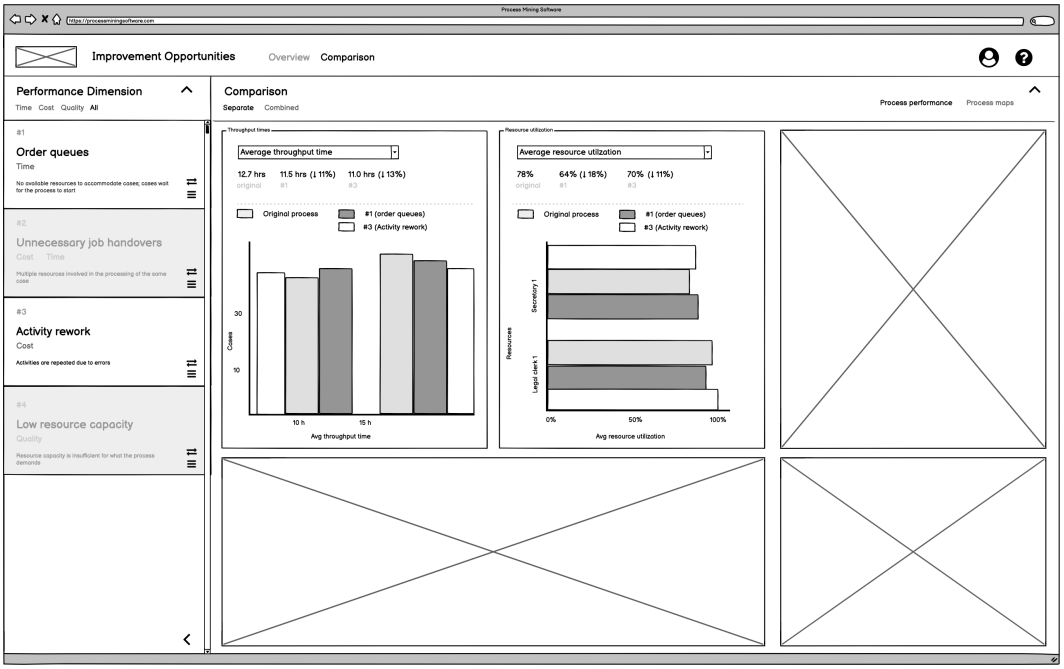

Figure 14. Screen 5: Comparison (Separate) - Process Performance

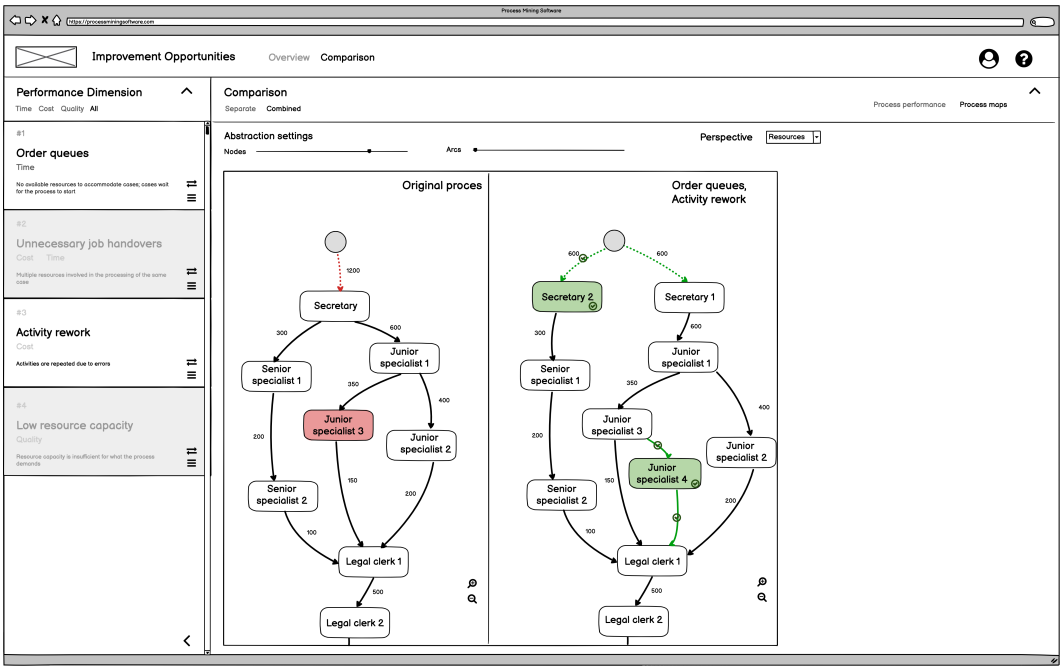

Figure 15. Screen 6: Comparison (Combined) - Process Maps (Resources)

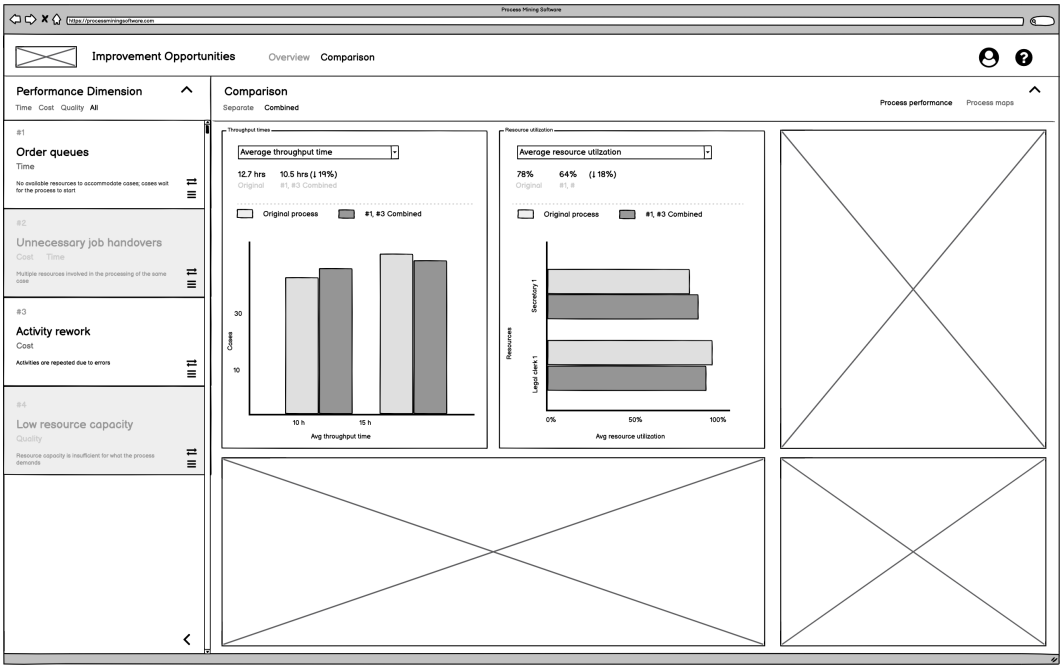

Figure 16. Screen 7: Comparison (Combined) - Process Performance

An addition to the visualization of a few improvement opportunities compared is that they can be viewed separately, i.e., to compare how the process changes if one of the proposed improvement opportunities is addressed (Figure 13, Figure 14), but also combined, i.e., if the process analyst decides to address together a few of the proposed improvements (Figure 15, Figure 16). As it can be seen from Figure 15, there are two process maps instead of three. The process map on the right combines the changes in the process if both improvement opportunity #1 and #3 are resolved. Thus, all the elements and symbols are repeated from the previous screens that visualize process maps, and the combined process maps contains new and updates resources and paths from respective improvement opportunities. The visualization of change in process performance for combined improvement opportunities thus contains the graphs where the original process performance is displayed, and the performance of the process where a few of the proposed improvement opportunities are addressed through a certain redesign suggestion. For instance, the average resource utilization graph contains bars of two colors, where one color represents the original process, and another color corresponds to the two potentially addressed improvement opportunities. The mockup visualization of such "combined" graphs (Figure 16) closely resembles the mockup visualization of an individual improvement opportunity (Figure 12). This is due to the limitation of the choice of graphs in the used mockup tool.

All described screens can be viewed in more detail in the interactive mockup[6].

7 Evaluation

This section outlines evaluation of the developed mockup with potential users. Section 7.1 elaborates on findings from the user study (user study design is given in section 4.2.1), section 7.2 contains additional requirements derived from the user study, and section 7.3 describes the implementation of the additional requirements in the artifact.

7.1 User Study Findings

In this section, the themes discovered in the process of data analysis are introduced in subsection 7.1.1. Next, the findings from the themes are outlined in subsection 7.1.2.

7.1.1 Themes

During the user study, we first introduced the mockup to the participants using a scenario. Then, they could interact with it while we were asking additional questions about it (cf. subsection 4.2.1). After analyzing the interviews (cf. subsection 4.2.2), we identified six themes in total (Figure 17). The full list of themes and codes can be viewed in the Google Sheets table[7].

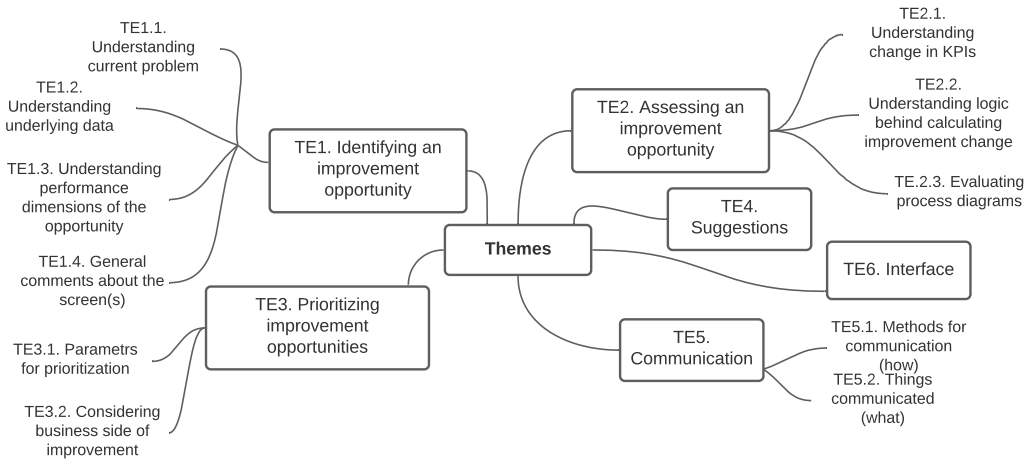

Figure 17. Themes (User evaluation study)

Each theme captures a certain aspect of the data from the transcripts. As such, the themes are:

[7]https://bit.ly/3mzBwxx

- **TE1. Identifying an improvement opportunity.** This theme consists of four sub-themes that consolidate all comments made about the screen corresponding to identifying an improvement opportunity. Thus, theme *TE1.1. Understanding current problem* covers comments from the participants regarding elements of the mockup that work well for them and the missing things for identifying an improvement opportunity. Sub-theme *TE1.2. Understanding underlying data* covers the comments specifically about the need to understand the background data when identifying an improvement opportunity. Next, sub-theme *TE1.3. Understanding performance dimensions of the opportunity*, similarly to the previous theme, captures the comments for specifically relating the opportunities to certain performance dimensions of the process. Finally, sub-theme *TE1.4. General comments about the screen(s)* captures overall feedback for the screen(s) of improvement opportunity identification.
- **TE2. Assessing an improvement opportunity.** This theme's three sub-themes describe the comments on the screens that aid in assessing an identified improvement opportunity. Thus, sub-theme *TE2.1. Understanding change in KPIs* contains the comments about the graphs in the mockup that visualize differences in KPIs between the processes before and after improvement, sub-theme *TE2.2. Understanding logic behind calculating improvement change* elaborates on the previous theme and captures the need to view the graphs' background data, and sub-theme *TE2.3. Evaluating process diagrams* captures the comments about visualizations of process maps of the processes before and after improvement and changes between them.
- **TE3. Prioritizing improvement opportunities.** Sub-theme *TE3.1. Parameters for prioritization* captures comments on the things that the participants deem important on the visualizations while prioritizing their improvement opportunities, and sub-theme *TE3.2. Considering business side of improvement* captures the comments of the importance of balancing different aspects of the problem when prioritizing improvements.
- **TE4. Suggestions.** This theme contains all direct suggestions from the participants on what to add, remove, or change in the mockup.
- **TE5. Communication.** This theme describes how the participants communicate the identified improvement opportunities with the help of the mockup, where sub-theme *TE5.1. Methods for communication (how)* describe how the participants utilize the mockup for communication, and *TE5.2. Things communicated (what)* contains the information on exact elements of the mockup the participants mention for communication.
- **TE6. Interface.** This theme consolidates all comments about the mockup interface as such (e.g., colors, naming, etc.)

7.1.2 Findings

In this subsection, user study findings are presented using the same sub-headings as in subsection 5.1. Namely, how process analysts evaluate the parts in the visualization related to improvement opportunity identification, improvement opportunity assessment, and improvement opportunities prioritization. Additionally, their comments on communicating improvement opportunities are captured. Such representation allows to see which needs were addressed and what aspects still have to be considered. General comments on the visualization are also presented separately at the end of this subsection.

Improvement Opportunity Identification

The findings demonstrated process analysts utilizing several elements of the visualization of improvement opportunities. Firstly, they found useful having an overview of all identified opportunities and seeing to which metric in the process they relate. They, however, found that the broader metrics that are displayed now do not provide the necessary level of insight and therefore should be more detailed. Secondly, our analysis suggests that seeing how much of the process is affected by an opportunity helped in the initial evaluation of it. Thirdly, the analysts demonstrated the need to see an overview of the current process performance in connection to the improvement opportunity, which is not presently available in the visualization. Lastly, they highlighted the need to be able to explore the data behind the identified opportunity which was also not presented in the mockup.

Before evaluating the impact of resolving improvement opportunities and prioritizing them, process analysts first require an understanding of the issue identified in the process (I-02, I-07, I-08, I-01, I-09, I-03). As such, before going into redesign options, process analysts wanted to understand the problem they are trying to solve: *"Before I start looking at solutions, I would want to know why. Why bother about it?"* (I-02). Therefore, process analysts commented that on the Overview Summary screen, it is *"[...] very good that I see how many cases, how many variants are affected. So basically I can see [...] quickly that if I resolve these improvements, then what amount of process will be affected."* (I-09). It is also demonstrated in theme TE3.1 where the process analysts name understanding how much of the process is affected by an improvement opportunity as one of the criteria for selecting which improvement opportunity to address. Additionally, they mention helpful seeing all the improvements together: *"I would say that it is very good that I can see all these improvements in one place with the description."* (I-09), *"I really like the the left sidebar, I can see time, cost, quality, all [...]"* (I-10). Process analysts also demonstrate usage of the process map. For instance, one interviewee said that *"I would check the context the neighboring nodes and connections in the process map."* (I-03). However, it is a recurring finding that process analysts also require an understanding of the process performance and how it is affected by the improvement

opportunity now before assessing redesign options (I-02, I-01, I-09, I-03). As one process analyst questioned expressed it, *"So, okay, it takes a long time. So what? Does anyone complain? Do we lose customers? Do we get our money later?"* (I-01). Another analyst gave a similar example: *"Does it lengthen the execution time of the overall process? And by, so what is that time?"* (I-02). Currently, the overview of the process performance is not presented on the Overview screen.

Another common question the process analysts were looking to understand was the underlying data of the improvement opportunity. Six out of eight participants commented on the importance of investigating the specific bottleneck data in more detail (I-02, I-07, I-08, I-03, I-10, I-11). This is motivated by two aspects. Firstly, there exist concerns that the data underlying the bottleneck may not be correct: *"[...] it's at least our experience, that some of the cases are pure garbage, it's not real data, that is connected with real cases that have been processed by our workers."* (I-07). Secondly, process analysts require more granularity of data, for example, to understand the work of a certain resource in more detail: *"We can maybe have the opportunity to put a click on the flow and see some some data, some details that make you, let you understand deeply the work of the secretary alone."* (I-10). Presently, the process map on the Overview screen visualizes only the improvement opportunity on it but does not provide any background data on it.

The participants also highlight improvement opportunities being sorted into process performance dimensions: *"So I like here that these are like tags, I could have two of them associated to any particular opportunity."* (I-02). However, process improvement opportunities are currently sorted into performance dimensions of time, cost, and quality. Process analysts point out that such sorting would have more value if the dimensions were more detailed because *"[...] it's about time. But more specifically, how's that? What kind of time, first of all, because time means a lot of things. [...] But only two of them matter because the others are not really aligned with you know, the, the strategic directives for this process."* (I-02). Another reason for that is that process analysts see the possibility to associate any of the performance dimensions with any improvement opportunity: *"I'm thinking that when I am eliminating rework, certainly it will decrease my time. So maybe time would be here as well."* (I-09), *"[...] I think that in the end, each of them can be linked to time, cost, and quality."* (I-01). Thus, this indicates that improvement opportunities should rather be sorted according to more detailed KPIs.

In conclusion, while the participants find the overview of improvement opportunities, how much of the process is affected by the improvement opportunity, and its process map useful, they also highlight the importance of understanding the process performance in terms of the current improvement opportunity. It is also evident that the interviewees prefer using more granular process performance dimensions than the general ones of time, cost, and quality. Additionally, the participants require an understanding of the data behind the bottleneck for a deeper insight into the problem.

Improvement Opportunity Assessment

While assessing an improvement opportunity, the participants demonstrated utilizing the visualizations of both the impact of the improvement opportunity on the KPIs of the process and on the process maps. As to the KPIs, some had issues with the diagram representations which is considered in the further improvement part. As to the process maps, we found that additional instructions or legend can be used to clarify them for the users. Similar to the previous section, process analysts also demonstrated the need to see the base data behind the diagrams and the process maps.

Firstly, in order to evaluate whether the improvement opportunity should be addressed in the process, the participants focus on understanding how the KPIs change (I-02, I-07, I-08, I-01, I-03, I-10). As they explained it: *"I would focus on changes in KPIs. So I would try to look what, what exactly changes? And if it's significant?"* (I-07), *"[...] obviously, you always need to have your KPIs and comparison between them. That's the only way it can make sense to do."* (I-01). Some participants had issues with understanding the current process performance graphs, especially that of average throughput time. One interviewee said: *"Okay, this is something that's a bit unclear, I would say. Because why there are four bars here. Not two?"* (I-07). Presently, the graph is located in a placeholder that is called "Average throughput time" where there are numeric values of the average throughput time itself and the graph that visualizes the throughput times of different cases in the process. However, the x-axis that represents the throughput times is also named as "Average throughput time" which might have created confusion. Additionally, as in the case with the underlying data of the bottleneck, process analysts also require understanding the basis on which the KPIs graphs were built. This is motivated by the fact that there can be mistakes in the data or the calculations, so the process analysts want to understand whether they can "trust" the graphs that they see *"because it's, it's very easy to get this wrong. So rather than just show me like we saw here, just a quick comparison - this is definitely nice - but the next question is, okay, why do you think it's going to go by 10%? Show me the underlying data."* (I-02). Currently, the graphs only visualize the KPI as it was in the original process, and how it is in the process after improvement.

Secondly, the participants also evaluated changes in process maps (I-08, I-09, I-03, I-11). As to the visualization, several participants (I-08, I-03) found them understandable: *"So I think that what is good that we just could see on the graph improvement, where will it be located, what will be impact of the specific improvements [...]"* (I-08). Several others (I-09, E-7) pointed out that they would require additional insight into what exactly changes on the process maps: *"Okay, I have original process, I have order queue, what is the difference? When I look at these process map, I can't see, I see just that, just this is path updated."* (I-09). This finding highlights the opportunity of adding instructions or legend to certain parts of the visualizations to ensure that all users understand them in the same manner. As previously, the question of the data behind the change in process maps

was raised by some participants (I-02, I-03, I-10). One particular area in the visualization the interviewees commented on was the process maps from the resource perspective on the impact screen. In particular, they needed to understand what the logic of distributing work between the resources is: *"But for this particular realization, I see a split and I would like to know on what basis."* (I-02), *"[...] what is the rule of distributing the work between the two Junior specialists here, because what would make me suspicious about this is while it looks good on paper, but there might be some complicated rules."* (I-03).

In conclusion, the participants evaluated both the change in KPIs and the process maps while assessing an improvement opportunity but highlighted the importance of understanding the data behind such changes in order to make an informed decision regarding a certain improvement opportunity.

Improvement Opportunities Prioritization

The findings indicate that when prioritizing proposed improvement opportunities, process analysts focused on the visualizations that helped them compare the improvement of KPIs that each of the opportunities entails. Additionally, they highlighted the need to assess cost and effort of addressing an improvement opportunity which currently is not a part of the visualization.

While prioritizing improvements, seven out of eight participants (I-02, I-07, I-08, I-01, I-09, I-03, I-11) mentioned seeing the change in KPIs on the graphs being a crucial decision point. As one interviewee said: *"When I am put like to decide okay, what is important for me, definitely, I will go directly to comparison, I will compare."* (I-11). Another one confirmed that *"for deciding on the specific improvement opportunity, the important thing for me to understand is how they change the process performance."* (I-01). However, the participants also mentioned that understanding the change in only KPIs is not enough to take decision because there can be complications in ease and/or cost of implementation of changes in the process. Thus, interviewees I-02, I-07, I-01, I-09 highlighted a lack of information about how easy it is to implement the change: *"That's the one thing the difference between the output or the result of improvements, but also I'm thinking of some effort that I have to put to resolve this issue."* (I-09). Interviewees I-08, I-11 also mentioned the importance of understanding the cost of implementation: *"The most important is to understand, okay, if I have these root causes, what will it cost, like, what will it cost me to implement things."* (I-11). Several participants (I-07, I-11) would assess both cost and effort in addition to change in KPIs. Another factor was the amount of cases affected by the opportunity (I-09, I-11) which was highlighted by the analysts as a criteria to initially assess the opportunity. It is evident that this criteria is also considered when prioritizing the opportunities: *"First, how many cases are affected? So when it's just 5%, maybe I will not do anything. So this is important. So when I go to overview, I will look at this."* (I-09). However, currently, this information is only accessible on the Overview screen of an individual improvement opportunity but not on

the Comparison. Additionally, user persona Jane mentioned that she would prioritize an improvement with addressing which she could also address another one: *"[...] if I can, by solving the problem two, probably I can already solve half of problem number one, yeah. So we always have to consider the solutions, if they could actually be solved, how to say, concurrently, in the same time."* (I-11). Such criteria was evident only for the user persona of Jane.

Several participants (I-02, I-07, I-03, I-10) raised concerns regarding including the business side of the problem during the analysis because *"[...] you might see, it may give you ideas, but you would not necessarily view them as a realistic improvement option."* (I-03), or *"it could be actually easy to do it from the human point of view, but the system is so inflexible that it ends up being high effort. So we have input from the business users and from technology."* (I-02). These findings were also identified in the first round of interviews. However, these concerns still relate to the process analysis as such and not to the visualization of improvement opportunities.

In conclusion, while prioritizing improvement opportunities, the participants utilized change in KPIs on available graphs but pointed out that such information is not enough to make the decision as the effort and/or cost of implementation also need to be considered.

Communicating Improvement Opportunities

The findings suggest that process analysts showcase the current state of the process, the KPIs improvement diagrams, and alternatives to the improvement opportunity when communicating their findings. These representations are currently present in the visualization.

Six participants (I-07, I-08, I-01, I-09, I-03, I-11) commented they would use the graphs of the change in KPIs when communicating their findings to the higher management or clients. As one interviewee commented, *"Oh, certainly use images from the application and therefore, it's to support my decision by data."* (I-09). The participants mentioned that *"what's important that you can just screenshot the graphs from here."* (I-01). As to the available graphs, the participants say that they should be simple to help convey the message and keep the focus on the important things. Therefore, the current representation suits that purpose: *"I think that KPI this could be just the screenshot because this is quite simple, and I think it's good."* (I-08). Apart from the graphs, the interviewees also highlight communicating the current process state and the alternatives to the improvement opportunity: *" First, you're telling them, here's your process as it is, here is your most, like, top root causes, what we could found, we could find. Here are opportunities, what we can do better."* (I-11). Another analyst described this process as such: *"[...] okay, there's improvement opportunities, and you have scenarios, if we improve this, then that's what we're going to achieve. If we improve that, that will be another result. And then you provide your recommendation."* (I-01). Both the current process state and the alternatives to the improvement opportunity are available in the

current visualization.

General Comments

Overall, during the evaluation, the participants had several points for minor improvements on different screens of the mockup but as they commented on the general impression, the feedback was positive, e.g., *"Even though it's early, it's very clear."* (I-07), *"No, not technically [no questions], because it seems to be quite straightforward."* (I-03), *"And with this visualization, I would say I'm actually fine. [...] you know, the easiest [...] the better, the less objects are better, because then you can stay focused. So for me, it's perfect."* (I-11). The interviewees highlighted the usefulness of such visualizations since they would make their daily jobs easier, e.g., *"Because such views, you really need [them], and it's sometimes, it takes too much time [...] to make such conclusions."* (I-11), *"because while you are thinking about how to redesign a process, you are very, very concentrated [...], you have a lot of [...] information in your brain. And to have everything in front of you is very, very useful."* (I-10).

One source of confusion was the naming of certain elements in the visualization. For instance, three interviewees (I-02, I-01, I-09) commented on usage of the word "impact" for visualizing both the change in process maps and process performance between the original and the improved processes, since for them, impact means how only the performance of the process changes, e.g., *"Why I asked you about impact, because [...] the word impact suggests that I will see [...] performance indicators. Because I would like to see what changes for me, what impact makes this change?"* (I-01).

Additionally, the interviewees proposed several direct suggestions for improvement of the mockup. Mostly, such suggestions were either individual to each participant or mentioned by at most two out of eight participants. For instance, I-07 and I-01 suggested presenting different KPIs by default on the Impact (Process Performance) screen and let only two or three others for the user to select themselves, *"And then we'll just show at least maybe three key KPIs, they are already visible here. And maybe you can leave room for two others, which the people can select."* (I-01). I-02 suggested having the improvement opportunity description always visible to ensure everyone working on it has the same understanding of it, *" I would want it to be always visible, and not have people interrupt and ask me or even worse, not ask at all and assume that they understood what order queues are but actually they didn't."* (I-02). I-08 suggested coloring different resources in different colors, *"maybe this is not quite important, but color for example, a senior specialist would be one color, and junior another, it would be easier to follow the graph."* (I-08).

In conclusion, general feedback on the mockup was positive, and participants did not demonstrate major struggles with understanding the visualizations. Minor confusion came from the naming of certain elements. The interviewees also proposed a number of concrete suggestions for improvement but they were either individual or mentioned by at

most two participants.

Summary

All the findings described in detail in the previous sub-headings are summarized in Table 5. The table features a short description of the finding and specifies which theme the finding is based on. The table includes only those findings that relate to the new visualization needs identified.

Table 5. Findings summary (Evaluation user study)

#	Description	Based on (finding)
F13	Process analysts need to understand the process performance in regards to the improvement opportunity before evaluating the redesign options.	TE1.1.1
F14	Process analysts require an understanding of the background data of the graphs included in the visualization of improvement opportunities.	TE2.2.1, TE2.2.3
F15	Process analysts need an overview of all KPIs that are changed if the opportunity is addressed.	TE1.3.1, TE1.3.4
F16	Process analysts want to understand the data behind the identified improvement opportunity.	TE1.2.1, TE1.2.2
F17	Process analysts consider cost and ease of implementation of change in the process together with the KPIs improvement it entails.	TE3.1.1, TE3.1.2
F18	Process analysts need to understand whether multiple improvement opportunities can be addressed at once.	TE3.1.7

7.2 Additional Requirements

Based on the new findings, we formulated a set of additional user stories (Table 6). The user stories were developed according to the same principles as in subsection 5.3. Only new findings were considered while writing the additional user stories, but other minor comments regarding interface were also addressed in the further development of the mockup (subsection 7.3). As previously, if a user story relates to multiple user personas,

it is written into the section of the table for Leo, and one specific to Jane user story is written in a separate category. In total, we developed eight new user stories.

Table 6. Additional user stories (after evaluation)

#	Group	Story	From (theme)
Leo			
US.22	Process analysis (other)	As a process analyst, I want to understand how the identified improvement opportunity influences my process performance.	F13
US.23	Process analysis (other)	As a process analyst, I want to understand the background data of the identified improvement opportunity.	F16
US.24	Changes in the process (improvements)	As a process analyst, I want to understand which KPIs in the process the identified improvement opportunity can change.	F15
US.25	Changes in the process (improvements)	As a process analyst, I want to understand the data behind a certain KPI improvement.	F14
US.26	Changes in the process (improvements)	As a process analyst, I want to understand the data behind a certain process map change.	F14
US.27	Prioritization (improvements)	As a process analyst, I want to understand the ease of implementation of the identified improvement opportunity.	F17
US.28	Prioritization (improvements)	As a process analyst, I want to understand the cost of implementation of the identified improvement opportunity.	F17
Jane			

Continuation of Table 6			
#	**Group**	**Story**	**From (theme)**
US.29	Prioritization (improvements)	As a lead process analyst, I want to understand whether I can solve several improvement opportunities with one redesign pattern.	F18

As the next step, we wrote specific requirements on the basis of the additional user stories (Table 7). The requirements also retained the same categories as those in the subsection 5.3. Three user stories were discarded while developing the requirements since they do not directly relate to the visualization. These user stories are: US.23. *As a process analyst, I want to understand the background data of the identified improvement opportunity.*, US.25. *As a process analyst, I want to understand the data behind a certain KPI improvement.*, and US.26. *As a process analyst, I want to understand the data behind a certain process map change.*

In total, we identified five new requirements in two categories, R.II. Improvement Identification and R.P. Prioritization. In the previous iteration, the requirements were prioritized based on how many participants related to them (must if >=3). Thus, for the additional requirements, those to which four or more participants relate are prioritized as "must" (exactly half of the participants, same principle as previously).

Table 7. Additional requirements (after evaluation)

#	Requirement	Based on	Prioritization
R.II. Improvement Identification			
R.II.6	The system displays current process process performance.	US.22	must
R.II.7	The system displays the KPIs which the improvement opportunity changes.	US.24	must
R.P. Prioritization			
R.P.5	The analyst is able to specify ease of addressing the improvement opportunity.	US.27	must

Continuation of Table 7			
#	**Requirement**	**Based on**	**Prioritization**
R.P.6	The system displays cost of addressing the improvement opportunity.	US.28	must
R.P.7	The system highlights improvement opportunities to which the same redesign pattern can be applied.	US.29	won't

7.3 Additional Development

To implement the new requirements, we edited an Individual Improvement Opportunity screen (R.II.6, R.II.7) and developed an additional screen for opportunities comparison (R.P.5, R.P.6). For the purposes of highlighting changes in screens, new additions are framed with green rectangles, and edited elements are framed with red rectangles.

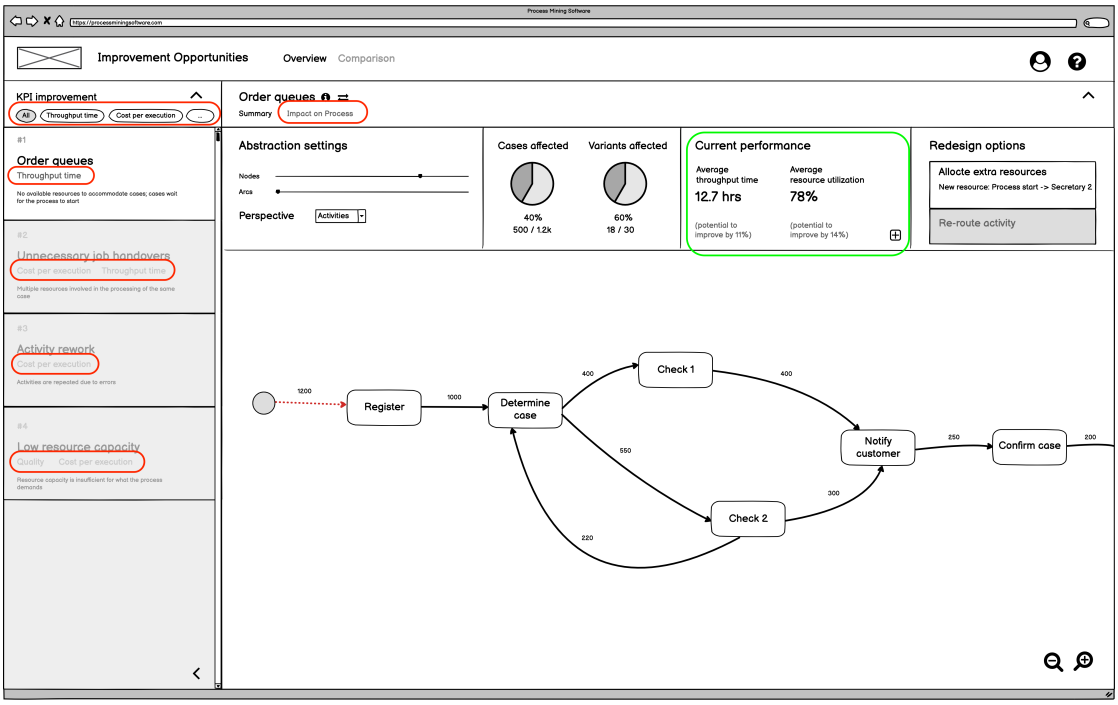

Figure 18. Screen 8: Overview - Summary (Additional Development)

Overview - Summary

Requirements R.II.6 and R.II.7 are from the *R.II. Improvement Identification* category and thus the respective screen (cf. Figure 10) was edited.

The screen with a summary of an individual improvement opportunity (Figure 18) thus has several of changes. Firstly, as to R.II.6, an overview of the current process performance was added. For the purposes of displaying the current performance, the same KPIs as in the other screens were used (average throughput time and average resource utilization). The elements were re-ordered so that the new element could fit. Secondly, performance dimensions of time, cost, and quality were transformed into concrete KPIs that the improvement opportunity has an influence on (R.II.7). Lastly, the "Impact" menu item was changed to "Impact on Process" to resolve the confusion that the interviewees had with it during the user study.

Comparison (Separate) - Prioritization

Figure 19. Screen 9: Comparison (Separate) - Prioritization (Additional Development)

To implement requirements R.P.5 and R.P.6, a new screen was developed in the Comparison view of the mockup (Figure 19). The purpose of such a screen is to allow the process analyst to prioritize identified improvement opportunities based on multiple factors, namely, not only change in KPIs but also cost and effort of implementation.

The link to this view (Prioritization) was added to two other existing links (Process performance, Process maps). The screen contains the following new components:

- KPI for comparison selection: the user selects a KPI to compare in more details. According to the selection, a detailed graph with the original process and the compared improvement opportunities is displayed (as in Figure 14).
- Other KPIs view: an overview of change in other KPIs that are not selected for a more detailed view.
- A graph with KPI change and cost of implementation: the graph visualizes potential improvement in KPIs (for instance, for the currently selected average resource utilization, the bars represent this KPI for the first and the third improvement opportunities). Two other bars represent respective costs of implementation. The graph keeps the use of juxtaposition as other graphs before, which is justified by the fact that for this graph's task it is possible to notice the differences (height of bars) without additional visual encoding.
- A sidebar on the right with summaries of each improvement opportunity, and elements to select the effort of implementing the change and its priority.

Performance dimensions transformed into KPI improvement (left sidebar) were also changed on the current screen. Additionally, elements on the left sidebar that previously were to aid in prioritization were moved as due to the addition of the right sidebar where an improvement can be assigned prioritization ranks, the previous elements became obsolete.

8 Discussion

In light of the RQ. *How can process improvement opportunities identified from the event log be visualized to support decision-making?*, the findings suggest that in order to evaluate and select improvement opportunities from visual representation, process analysts focus on understanding three main things: (1) the current problem and its influence on the process, (2) benefits of addressing an improvement opportunity, (3) alternatives for process improvement.

It is evident that process analysts require a holistic overview of the problem in the process that the improvement opportunity is highlighting. As expected, when exploring the problem, the analysts incorporate analyzing the process map and the current process performance, as previously reported by [35]. However, in addition to such information, process analysts also seek to see how much of the process is affected by the improvement opportunity as one of the first steps in their analysis. This is motivated by trying to understand whether the problem is worth being considered, where significance is determined by the number of cases and variants involved. Therefore, it might be suggested to visualize improvement opportunities not only on the expected elements as process maps and process performance dashboards, but also include other factors that aid in understanding the importance of the finding, such as cases/variants affection.

The findings also indicate that it is important for process analysts to understand the potential benefits for their process that come with addressing the improvement opportunity. As such, process analysts evaluate the improvement opportunity through its connection to specific KPIs, i.e., which KPIs of the process can be improved if the opportunity is addressed. That helps to evaluate whether the improvement opportunity is relevant to the business objectives of the company, as depending on the company and sometimes also on the specific process, KPIs to consider vary. Thus, when visualizing improvement in process KPIs, it might be suggested to focus not only on specific change in a KPI, but also to give an overview of others that are affected by the improvement and might be perceived important by the process analyst. This finding is in line with previous findings where researchers specify that the decision-makers need to understand the relationship between selecting the process change and its outcome [44] but it provides more insight as to what specific factors and representations to consider.

As to prioritizing alternative improvements in the process, the prioritization criteria are complex. Not only do the analysts compare KPI graphs of different improvement opportunities, but also seek to consider cost and effort of implementation of a certain change to address the improvement opportunity. This is motivated by the fact that while addressing an improvement opportunity can drastically change the process performance, it can also be unrealistic to address it due to the lack of financial, technological, or human resources. As such, while visualizing the change in process performance is essential, it might be advisable to also consider other factors like cost and effort of implementation.

Moreover, there is evidence that process analysts consider combinations of changes when improving the processes. This is motivated by the fact that addressing multiple improvements at once can bring more savings to the company. However, there are concerns that such changes may also be contradicting to each other or impossible to implement. Therefore, it might be suggested to consider visualizing changes in the process that may arise given that not one but several improvement opportunities are prioritized to be addressed.

Additionally, the findings provide indication that process analysts might have concerns about trustworthiness of the visualizations if they are not supported by respective background data. In other words, regardless of the visualization presented (for example, process map with the improvement opportunity highlighted, comparison graph of KPIs), it seems to be important for the process analysts to explore the data on which this visualization is built, e.g., an event log, calculation data. The motivation behind this need is two-fold. Firstly, process analysts tend to demonstrate concerns regarding the quality of the original data and thus the credibility of the identified improvement opportunity. Secondly, process analysts tend to double-check the findings of various software they use as they might have a certain insight into data or processes that software does not capture; hence, some automatic findings become obsolete. In conclusion, it might be suggested to include linkage to the original data behind each of the visualizations.

8.1 Limitations

In order to achieve the goal of this research that is to understand how to visualize process improvement opportunities, we followed a design science approach. There can be certain limitations to it [21]. To center our research on the users, we considered input from practitioners from industry. As such, six different process analysts participated in the requirements elicitation interviews. While theoretically founded case selections were made, different participants working in different domains on different projects might yield different results. Despite reaching the data saturation after the fifth interview, it is still a subjective perception. Thus, to obtain a broader outlook, more participants could be involved.

Moreover, there exists an internal threat as to misinterpreting qualitative data due to bias or subjectivity. To minimize this threat, the findings of the answers received in the first round of interviews and feedback and opinion collected during the evaluation were discussed between the involved researchers. We also used the additional materials that the interviewees presented (screenshots, images shown directly in the process mining and analysis software used, videos) which helped the interpretability. To mitigate the threat of misinterpretability, we also invited several interviewees from the first round to participate in the user evaluation as well. Thus, upon analysis of the data from the second stage, we did not find any obvious differences between the new participants and the repeated participants. Additionally, the findings from the repeated participation of

some interviewees confirmed that the visualization needs were elicited adequately in the first round. However, summaries of findings were made and therefore, limitations remain as to generalization of results.

9 Conclusion

This thesis aimed to research how improvement opportunities identified from an event log can be visualized. The visualization, a clickable mockup, was created considering input from process mining practitioners, several visualization principles and previous developments in the field.

To understand the user needs for the visualization, three user personas were developed that cover various backgrounds and levels of experience of the potential users. Based on these personas, six process mining practitioners from the industry were recruited for the first round of interviews, requirements elicitation. As the interviews were analyzed, 20 user stories were developed that captured various user needs identified during the interviews. As a result, 38 requirements were specified based on the user stories. The requirements were prioritized, and "must" requirements were implemented in the interactive mockup where screens are linked through clickable zones. At this stage, several visualization principles and previous developments were taken into account. The mockup was evaluated in the second round of interviews, where eight practitioners participated, three of them from the previous round. As the outcome of the evaluation, several points for improvement of the mockup were identified, which resulted in eight additional user stories and five additional requirements based on them. All the new "must" requirements were implemented in the new version of the mockup. As a result, it was found out that the elements that process analysts are looking for in a visual representation of improvement opportunities relate not only to process maps and process performance graphs but also to understanding the current problem and its influence on the process, benefits of addressing an improvement opportunity, and alternatives for process improvement.

There are several opportunities for future development of the visualization. Firstly, the visualization can be improved by further researching how to visualize certain parts of it in more detail, i.e., what are other ways of representing the change in process maps of the original and the improved process, the change in KPIs of the process. Secondly, it should be researched how to increase the trustworthiness of the visualizations through connecting them to the original data of the process, certain calculations, and recommendations for change.

Acknowledgements

We are thankful to all the interview participants for kindly finding time to provide valuable insights during the requirements elicitation and user evaluation stages of this research.

References

[1] AL-KASSAB, J., OUERTANI, Z. M., SCHIUMA, G., AND NEELY, A. Information visualization to support management decisions. *International Journal of Information Technology & Decision Making 13*, 02 (2014), 407–428.

[2] ANDERSEN, B. *Business process improvement toolbox*. Quality Press, 2007.

[3] AUGUSTO, A., CONFORTI, R., DUMAS, M., AND LA ROSA, M. Split miner: Discovering accurate and simple business process models from event logs. In *2017 IEEE International Conference on Data Mining (ICDM)* (2017), IEEE, pp. 1–10.

[4] AUGUSTO, A., CONFORTI, R., DUMAS, M., LA ROSA, M., AND BRUNO, G. Automated discovery of structured process models from event logs: The discover-and-structure approach. *Data & Knowledge Engineering 117* (2018), 373–392.

[5] BACHHOFNER, S., KIS, I., DI CICCIO, C., AND MENDLING, J. Towards a multi-parametric visualisation approach for business process analytics. In *International Conference on Advanced Information Systems Engineering* (2017), Springer, pp. 85–91.

[6] BASOLE, R. C., PARK, H., GUPTA, M., BRAUNSTEIN, M. L., CHAU, D. H., AND THOMPSON, M. A visual analytics approach to understanding care process variation and conformance. In *Proceedings of the 2015 Workshop on Visual Analytics in Healthcare* (2015), pp. 1–8.

[7] BERANDER, P., AND ANDREWS, A. Requirements prioritization. In *Engineering and managing software requirements*. Springer, 2005, pp. 69–94.

[8] BOLT, A., DE LEONI, M., AND VAN DER AALST, W. M. A visual approach to spot statistically-significant differences in event logs based on process metrics. In *International Conference on Advanced Information Systems Engineering* (2016), Springer, pp. 151–166.

[9] BOSE, R. J. C., AND VAN DER AALST, W. M. Discovering signature patterns from event logs. In *2013 IEEE Symposium on Computational Intelligence and Data Mining (CIDM)* (2013), IEEE, pp. 111–118.

[10] BRAUN, V., AND CLARKE, V. Using thematic analysis in psychology. *Qualitative research in psychology 3*, 2 (2006), 77–101.

[11] CHEN, C.-H., HÄRDLE, W. K., AND UNWIN, A. *Handbook of data visualization*. Springer Science & Business Media, 2007.

[12] COHN, M. *User stories applied: For agile software development*. Addison-Wesley Professional, 2004.

[13] CONFORTI, R., DE LEONI, M., LA ROSA, M., AND VAN DER AALST, W. M. Supporting risk-informed decisions during business process execution. In *International Conference on Advanced Information Systems Engineering* (2013), Springer, pp. 116–132.

[14] CORDES, C., VOGELGESANG, T., AND APPELRATH, H.-J. A generic approach for calculating and visualizing differences between process models in multidimensional process mining. In *International Conference on Business Process Management* (2014), Springer, pp. 383–394.

[15] CRAFT, B., AND CAIRNS, P. Beyond guidelines: what can we learn from the visual information seeking mantra? In *Ninth International Conference on Information Visualisation (IV'05)* (2005), IEEE, pp. 110–118.

[16] DI FRANCESCOMARINO, C., GHIDINI, C., MAGGI, F. M., AND MILANI, F. Predictive process monitoring methods: Which one suits me best? In *International Conference on Business Process Management* (2018), Springer, pp. 462–479.

[17] DUMAS, M., LA ROSA, M., MENDLING, J., AND A. REIJERS, H. *Fundamentals of Business Process Management*. Springer, Berlin, Heidelberg, 2018.

[18] ELKHAWAGA, G., ABUELKHEIR, M., BARAKAT, S. I., RIAD, A. M., AND REICHERT, M. Conda-pm—a systematic review and framework for concept drift analysis in process mining. *Algorithms 13*, 7 (2020), 161.

[19] GALL, M., WALLNER, G., KRIGLSTEIN, S., AND RINDERLE-MA, S. Differencegraph-a prom plugin for calculating and visualizing differences between processes.

[20] GLEICHER, M., ALBERS, D., WALKER, R., JUSUFI, I., HANSEN, C. D., AND ROBERTS, J. C. Visual comparison for information visualization. *Information Visualization 10*, 4 (2011), 289–309.

[21] GREGOR, S., AND HEVNER, A. R. Positioning and presenting design science research for maximum impact. *MIS quarterly* (2013), 337–355.

[22] GULDEN, J. Visually comparing process dynamics with rhythm-eye views. In *International Conference on Business Process Management* (2016), Springer, pp. 474–485.

[23] HAMMER, M., AND CHAMPY, J. *Reengineering the Corporation: Manifesto for Business Revolution, A*. Zondervan, 2009.

[24] HARRELL, M. C., AND BRADLEY, M. A. Data collection methods. semi-structured interviews and focus groups. Tech. rep., Rand National Defense Research Inst santa monica ca, 2009.

[25] HOMPES, B., BUIJS, J. C., VAN DER AALST, W. M., DIXIT, P. M., AND BUURMAN, J. Detecting changes in process behavior using comparative case clustering. In *International Symposium on Data-Driven Process Discovery and Analysis* (2015), Springer, pp. 54–75.

[26] HUANG, Z., AND KUMAR, A. A study of quality and accuracy trade-offs in process mining. *INFORMS Journal on Computing 24*, 2 (2012), 311–327.

[27] HUDAIB, A., MASADEH, R., QASEM, M. H., AND ALZAQEBAH, A. Requirements prioritization techniques comparison. *Modern Applied Science 12*, 2 (2018), 62.

[28] JORBINA, K., ROZUMNYI, A., VERENICH, I., DI FRANCESCOMARINO, C., DUMAS-MENIJVAR, M., GHIDINI, C., MAGGI, F., LA ROSA, M., AND RABOCZI, S. Nirdizati: A web-based tool for predictive process monitoring. In *Proceedings of the BPM Demo Track and BPM Dissertation Award (CEUR Workshop Proceedings, Volume 1920)* (2017), Sun SITE Central Europe, pp. 1–5.

[29] KOSARA, R., HEALEY, C. G., INTERRANTE, V., LAIDLAW, D. H., AND WARE, C. Thoughts on user studies: Why, how, and when. *IEEE Computer Graphics and Applications 23*, 4 (2003), 20–25.

[30] LA ROSA, M., WOHED, P., MENDLING, J., TER HOFSTEDE, A. H. M., REIJERS, H. A., AND VAN DER AALST, W. M. P. Managing Process Model Complexity Via Abstract Syntax Modifications. *IEEE Transactions on Industrial Informatics 7*, 4 (nov 2011), 614–629.

[31] LASHKEVICH, K. Business process improvement opportunities: a framework to support business process redesign. Master's thesis, University of Tartu, 2020.

[32] LOW, W. Z., VAN DER AALST, W. M., TER HOFSTEDE, A. H., WYNN, M. T., AND DE WEERDT, J. Change visualisation: Analysing the resource and timing differences between two event logs. *Information Systems 65* (2017), 106–123.

[33] MAGGI, F. M., DI FRANCESCOMARINO, C., DUMAS, M., AND GHIDINI, C. Predictive monitoring of business processes. In *International conference on advanced information systems engineering* (2014), Springer, pp. 457–472.

[34] MAGUIRE, M., AND BEVAN, N. User requirements analysis. In *IFIP World Computer Congress, TC 13* (2002), Springer, pp. 133–148.

[35] MILANI, F., AND MAGGI, F. M. A Comparative Evaluation of Log-Based Process Performance Analysis Techniques. 2018, pp. 371–383.

[36] MUNZNER, T. *Visualization analysis and design*. CRC press, 2014.

[37] PIKA, A., WYNN, M. T., FIDGE, C. J., TER HOFSTEDE, A. H., LEYER, M., AND VAN DER AALST, W. M. An extensible framework for analysing resource behaviour using event logs. In *International Conference on Advanced Information Systems Engineering* (2014), Springer, pp. 564–579.

[38] PINI, A., BROWN, R., AND WYNN, M. T. Process visualization techniques for multi-perspective process comparisons. In *Asia-Pacific Conference on Business Process Management* (2015), Springer, pp. 183–197.

[39] PLAIN, C. Build an affinity for kj method. *Quality Progress 40*, 3 (2007), 88.

[40] REDLICH, D., MOLKA, T., GILANI, W., BLAIR, G. S., AND RASHID, A. Scalable dynamic business process discovery with the constructs competition miner. In *SIMPDA* (2014), Citeseer, pp. 91–107.

[41] RUNESON, P., AND HÖST, M. Guidelines for conducting and reporting case study research in software engineering. *Empirical software engineering 14*, 2 (2009), 131–164.

[42] SAUNDERS, B., SIM, J., KINGSTONE, T., BAKER, S., WATERFIELD, J., BARTLAM, B., BURROUGHS, H., AND JINKS, C. Saturation in qualitative research: exploring its conceptualization and operationalization. *Quality & Quantity 52*, 4 (jul 2018), 1893–1907.

[43] SHNEIDERMAN, B. The eyes have it: A task by data type taxonomy for information visualizations. In *The craft of information visualization*. Elsevier, 2003, pp. 364–371.

[44] SIHA, S. M., AND SAAD, G. H. Business process improvement: empirical assessment and extensions. *Business Process Management Journal 14*, 6 (nov 2008), 778–802.

[45] SIRGMETS, M., MILANI, F., NOLTE, A., AND PUNGAS, T. Designing Process Diagrams – A Framework for Making Design Choices When Visualizing Process Mining Outputs. 2018, pp. 463–480.

[46] SITOVA, I., AND PECERSKA, J. Process data analysis using visual analytics and process mining techniques. In *2020 61st International Scientific Conference on Information Technology and Management Science of Riga Technical University (ITMS)* (2020), IEEE, pp. 1–6.

[47] STEIN DANI, V., DAL SASSO FREITAS, C. M., AND THOM, L. H. Ten years of visualization of business process models: A systematic literature review. *Computer Standards & Interfaces 66* (2019).

[48] TUFTE, E. R., GOELER, N. H., AND BENSON, R. *Envisioning information*, vol. 2. Graphics press Cheshire, CT, 1990.

[49] VAN DER AALST, W. Process mining. *Commun. ACM 55*, 8 (Aug. 2012), 76–83.

[50] VAN DER AALST, W., ADRIANSYAH, A., DE MEDEIROS, A. K. A., ARCIERI, F., BAIER, T., BLICKLE, T., BOSE, J. C., VAN DEN BRAND, P., BRANDTJEN, R., BUIJS, J., ET AL. Process mining manifesto. In *International Conference on Business Process Management* (2011), Springer, pp. 169–194.

[51] VAN DER AALST, W., ADRIANSYAH, A., AND VAN DONGEN, B. Causal nets: A modeling language tailored towards process discovery. In *International conference on concurrency theory* (2011), Springer, pp. 28–42.

[52] VAN DER AALST, W. M., LA ROSA, M., AND SANTORO, F. M. Don't forget to improve the process! *Business Process Management 58*, 1 (2016), 1–6.

[53] VAN DONGEN, B. F., CROOY, R. A., AND VAN DER AALST, W. M. Cycle time prediction: When will this case finally be finished? In *OTM Confederated International Conferences" On the Move to Meaningful Internet Systems"* (2008), Springer, pp. 319–336.

[54] VAN ZELST, S. J., VAN DONGEN, B. F., VAN DER AALST, W. M., AND VERBEEK, H. Discovering workflow nets using integer linear programming. *Computing 100*, 5 (2018), 529–556.

[55] WESKE, M. *Business Process Management Architectures*. Springer, Berlin, Heidelberg, 2012.

[56] WYNN, M. T., POPPE, E., XU, J., TER HOFSTEDE, A. H., BROWN, R., PINI, A., AND VAN DER AALST, W. M. Processprofiler3d: A visualisation framework for log-based process performance comparison. *Decision Support Systems 100* (2017), 93–108.

[57] YESHCHENKO, A., MENDLING, J., DI CICCIO, C., AND POLYVYANYY, A. Vdd: A visual drift detection system for process mining. CEUR Workshop Proceedings.

[58] ZELAZNY, G. *Say it with charts: the executive's guide to visual communication*. McGraw-Hill Education, 2001.

Appendix

I. Requirements Elicitation Interview Questions

1. First of all, please describe your position and responsibilities in this position in a few words.
2. We are interested in a specific process improvement project that you carried out. Please try to think about one specific project, for example, if you had to improve the performance of one concrete process in your company, or if there was something interesting that happened in the project.
3. Can you share some of the materials for the project with me? (e.g., a process model, a screenshot from the tool that you used, some dashboards, slides, etc.) Can you send them to me or share your screen and show them live?
4. What was the case for which the process improvement project was initiated?
5. Why was this case of interest?
6. Is it a typical case? If not, what is different between it and usual cases for such a project?
7. What data was used in the case? Where was the data extracted from?
8. What tool was used in this project? (i.e, specific process mining tool)
9. What was the specific improvement opportunity identified?
10. How was the improvement opportunity identified?
    (a) What were the criteria/measures to identify the improvement opportunities/bottlenecks?
    (b) Were any visualizations used to help identify the improvement opportunity? *(if yes, c)*
    (c) What was important to see visually?
    (d) What challenges did you have when you tried to analyze the process? (e.g., too much information to see, can't distinguish the types of bottlenecks, etc.)
11. Were there any alternatives to the selected improvement opportunity? *(if yes, a and b)*
    (a) How was it decided which one to select? Who made this decision?
    (b) Were any visualizations used to help decide on the improvement opportunity? What was important to see visually to compare them?
12. What happened during the course of this project that was unexpected?

13. What else do you wish you could know to make a decision regarding the improvement opportunity?
14. Generally, can you name one thing that worked well in this project and one thing that went wrong?
15. If you had to make one suggestion for your colleague who's working on a similar improvement project, what would you suggest?
16. *(Leo, Lucy, Jane)* Who were the results presented to?
17. *(Leo, Lucy, Jane)* How were the results presented (images, dashboards, etc.)?
18. *(Jane)* What data do you generally need to see to make a decision regarding a specific improvement opportunity?

II. User Stories

Table 8. User stories

#	Group	Story	From (finding)
Leo			
US.1	Improvement prioritization	As a process analyst oriented at cost savings, I want to prioritize improvements, so I can choose the ones where the financial gains are the biggest.	F9
US.2	Improvement prioritization	As a process analyst oriented at cost savings, I want to see different categories of possible cost savings in the process, so I can assess them.	F5
US.3	Improvement prioritization	As a process analyst oriented at time savings, I want to see waiting times between activities, so I can improve the long waiting times.	F1
US.4	Improvement prioritization	As a process analyst oriented at time savings, I want to see the processing times of activities, so I can investigate what slows processing times.	F1
US.5	Improvement opportunity	As a process analyst, I want to understand how many cases of the process are involved in the improvement, so I can assess its impact on the process.	F8
US.6	Improvement opportunity	As a process analyst, I want to see the statistics of the old process and the improved process, so I can assess how much a certain improvement would change the process.	F6
US.7	Process analysis (other)	As a process analyst, I want to compare process performance statistics between different variants, so I can assess the uniqueness of the problem.	F6
US.9	Communication	As a process analyst, I want to communicate the performance information in a self-explanatory way, so the users have an overview of how they improve.	F4

Continuation of Table 8

#	Group	Story	From (finding)
US.10	Communication	As a process analyst reporting to management, I want to include specific explanations of findings for management, so they don't make wrong conclusions from the charts and graphs.	F12
US.13	Business relevance	As a process analyst reporting to management, I want to adjust KPIs when I clarify them with management, so I suggest relevant to the KPIs improvements.	F10
US.13.1	Business relevance	As a process analyst reporting to management, I want to see the most valuable improvements for different KPIs, so I can present improvements for various KPIs.	F10
US.14	Process analysis (other)	As a process analyst, I want to compare process maps between different cases, so I can assess how different the cases are.	F2
US.15	Process analysis (other)	As a process analyst working with many processes, I want to collaborate with the process experts, so that I better understand how the process is designed.	F7
US.16	Improvement opportunity	As a process analyst, I want to see the process maps of the old process and the improved process, so I can assess changes in the activities and the paths.	F2
US.19	Business relevance	As a process analyst, I want to confirm findings with the process experts, so I can understand whether certain improvements can be addressed.	F7
Lucy			
US.11	Communication	As a consultant, I want to assess the importance of the finding to the task, so I don't communicate irrelevant things to the clients.	F1
US.12	Communication	As a process analyst, I want to understand what the recommendation for improvement is based on, so I can share the assumptions based on facts.	F7

Continuation of Table 8

#	Group	Story	From (finding)
US.17	Communication	As a consultant, I want to understand whether the improvement is internal to the process, so I recommend changes that can be implemented from within.	F11
US.18	Business relevance	As a consultant, I want to get input from the clients, so that I have a business-side understanding of the process.	F7
Jane			
US.20	Process analysis (other)	As a lead process analyst, I want my team to assess the process from its domain perspective and the process mining perspective, so that both sides are included in the analysis.	F3

III. Requirements

Table 9. Requirements

#	Requirement	Based on	Prioritization
R.II. Improvement Identification			
R.II.1	The system displays the improvements sorted into process performance dimensions.	US.1, US.3, US.4	must
R.II.2	The analyst is able to view the details of a single improvement opportunity.	US.5, US.3, US.4, US.6	must
R.II.3	The system displays the process map of the process.	US.16	must
R.II.3.1	The system highlights the activities correlated with the improvement opportunity.	US.12, US.16	should
R.II.3.2	The system highlights the paths correlated with the improvement opportunity.	US.12, US.16	should
R.II.4	The system displays the number of cases affected by the improvement opportunity.	US.5	must
R.II.5	The system displays the issue in the process for which the improvement is proposed.	US.12	should
R.IA. Improvement Assessment			
R.U. Uniqueness			
R.U.1	The analyst is able to select process variants of interest.	US.7	could
R.U.1.1	The analyst is able to select process performance statistics of interest.	US.7	could
R.U.1.1.1	The system displays differences between the process performance statistics of selected variants.	US.7	could
R.U.1.2	The system displays process maps of selected variants.	US.7, US.14	could

Continuation of Table 9

#	Requirement	Based on	Prioritization
R.U.1.2.1	The system highlights the differences between the process maps of selected variants.	US.7, US.14	could
R.BR. Business Relevance			
R.BR.1	The analyst is able to select a KPI of interest.	US.13	should
R.BR.1.1	The system displays improvements related to the selected KPI.	US.13	should
R.BR.2	The system displays the improvements sorted into various categories of financial gains.	US.2	won't
R.IP. Impact on the Process			
R.IP.1	The system displays the process map of the original process.	US.16	must
R.IP.1.1	The system displays the process map of the improved process.	US.16	must
R.IP.1.2	The system highlights the differences between the process maps of the original and the improved processes.	US.16	must
R.IP.2	The analyst is able to select KPIs of interest.	US.6	must
R.IP.3.1	The system displays the selected KPIs for the original process.	US.6	must
R.IP.3.2	The system displays the selected KPIs for the improved process.	US.6	must
R.IP.3.3	The system displays the differences between the KPIs of the original and the improved processes.	US.6	must
R.I. Implementability			
R.I.1	The analyst is able to specify that the change cannot be implemented.	US.19	could

Continuation of Table 9

#	Requirement	Based on	Prioritization
R.I.1.1	The analyst is able to select from the list of reasons why the change cannot be implemented.	US.19	could
R.I.1.2	The system excludes the improvement from the list of proposed improvements.	US.19	could
R.I.1.3	The system displays the list of improvements marked as unable to implement.	US.19	could
R.D. Dependency			
R.D.1	The analyst is able to specify that the proposed improvement is extrinsic to the process.	US.17	could
R.D.1.1	The system displays the list of improvements marked as extrinsic.	US.17	could
R.P. Prioritization			
R.P.1	The analyst is able to assign prioritization ranks to improvements.	US.1, US.3, US.19	must
R.P.2	The system displays savings in the improvement for the selected KPI.	US.1, US.3, US.4	must
R.P.3	The analyst is able to select the improvements for comparison.	US.1, US.3	must
R.P.3.1	The system displays the process maps of the compared improvements.	US.1, US.16	must
R.P.3.1.1	The system highlights the differences between the compared process maps.	US.16	must
R.P.3.2	The system displays differences in the selected KPIs of the compared processes.	US.6	must
R.P.4	The system highlights the improvement with maximum savings.	US.1	could
R.C. Communication			

Continuation of Table 9

#	Requirement	Based on	Prioritization
R.C.1	The analyst is able to switch between an analysis and a reporting view.	US.10	could
R.C.2	The analyst is able to create a custom reporting view.	US.9	won't
R.C.2.1	The analyst is able to select graphs to display in the reporting view.	US.9	won't
R.C.3	The analyst is able to add notes to the process map in the reporting view.	US.10	could
R.O. Other			
R.O.1	The analyst is able to save prioritization options as a template.	US.8	won't

IV. Licence

Non-exclusive licence to reproduce thesis and make thesis public

I, **Kateryna Kubrak**,

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to

   reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright,

   Visualizing Business Process Improvements,

   supervised by Fredrik Payman Milani and Alexander Udo Nolte.

2. I grant the University of Tartu a permit to make the work specified in p. 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 3.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.

3. I am aware of the fact that the author retains the rights specified in p. 1 and 2.

4. I certify that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Kateryna Kubrak
12/05/2021