GETTING YOUR RPA PRIORITIES STRAIGHT WITH PROCESS MINING: THE PLOST FRAMEWORK

Complete Research

Hilde E. Jongeling, Utrecht University, Utrecht, The Netherlands, mail@hildejongeling.com

Xixi Lu, Utrecht University, Utrecht, The Netherlands, x.lu@uu.nl

Iris Beerepoot, Utrecht University, Utrecht, The Netherlands, i.m.beerepoot@uu.nl

Inge van de Weerd, Utrecht University, Utrecht, The Netherlands, g.c.vandeweerd@uu.nl

Hajo A. Reijers, Utrecht University, Utrecht, The Netherlands, h.a.reijers@uu.nl

Abstract

Robotic Process Automation (RPA) has gained much attention both in industry and academia. One of the main challenges for a successful implementation of RPA is selecting which tasks should be automated. While different methods exist to identify RPA candidate tasks, they lack in providing objective evidence on why to automate that task. Such objective evidence can be gathered by applying process mining techniques to gain insights into the performance of a process and its tasks. Although this has multiple advantages, it can be time-consuming to analyze all potential processes. We conducted a literature review of existing methods to identify relevant criteria and method components, based on which we designed a framework for identifying and prioritizing suitable RPA candidate tasks: the Prioritized List of Suitable Tasks (PLOST) Framework. The framework includes both qualitative and quantitative assessment criteria and guides the analyst to focus on relevant processes before zooming in on the task level. It also takes into account a customized automation strategy. We conducted a case study to evaluate the applicability and effectiveness of the PLOST Framework and performed thinking-aloud sessions to evaluate its usability, practicality, and completeness. The results show that the framework is easy to apply and feasible.

Keywords: Robotic Process Automation, Process Mining, RPA Task Identification

1 Introduction

Robotic Process Automation (RPA) is an emerging form of automation for business processes and is seen as an advanced technology in the area of computer science and information technology (Madakam, Holmukhe, and Jaiswal, 2019). Its main goal is to replace human tasks with a virtual workforce or a digital worker performing the same work the human worker was doing (Choi, R'bigui, and Cho, 2020). For organizations, implementing RPA solutions can be costly and time-consuming (Lamberton, Brigo, and Hoy, 2017), especially if the selected process was not suitable or too complex. Therefore, it is important to identify suitable processes to automate with RPA (Syed et al., 2020). Once a process is selected, the identification phase is used to filter out highly complex tasks that could form a stumbling block for automation. Skipping this phase or not paying enough attention to it is one of the main reasons why RPA projects fail or lack behind expectations (Viehhauser and Doerr, 2021). Although the importance of the phase is highlighted, identifying where RPA is likely to provide significant value remains a challenge (Choi, R'bigui, and Cho, 2020).

Several methods have been proposed to select candidate tasks for RPA (Agaton and Swedberg, 2018; Berghuis, 2021; Burgess, 2017; Choi, R'bigui, and Cho, 2020; Leopold, Aa, and Reijers, 2018). However, these methods have several limitations. Firstly, they are time-consuming. Most methods rely only on interviews to understand a big set of processes, while this form of data collection and analysis is costly and time-consuming (Ryan, Coughlan, and Cronin, 2009). Secondly, the existing methods focus on either qualitative or quantitative criteria to select candidate tasks. Solely focusing on qualitative criteria results in a more subjective output, as the assessor has to judge whether a qualitative criterion is met. On the other hand, when focusing on quantitative criteria alone, one ignores qualitative criteria that are important to meet for a process to be suitable for RPA. Examples of such mandatory qualitative criteria are for a task to be mature, the existence of only a few exceptions and digital input (Agaton and Swedberg, 2018; Haan, 2021; Shafik Salah Elsayed and Kassem, 2022). A combined assessment of both qualitative and quantitative criteria could highly benefit the identification phase. Thirdly, every existing method focuses on one level of detail, namely high-level or low-level. The high-level is in this case the process level, whereas the low-level is the task level. When focusing on only one of these two, the methods lack in giving a full guide on how to select a task to automate from a certain process. Therefore, it should not be the question which process or which task to automate, but which task from which process.

Because of the limitations in existing identification methods, we conclude that there is a need for formal, systematic, and evidence-based techniques to determine the suitability of candidate tasks for RPA. In this study, we propose a framework to systematically identify tasks suitable for RPA. This framework is inspired by well-founded components of existing methods and takes into account both the characteristics of business processes as well as low-level tasks. The characteristics used in this framework, both qualitative and quantitative, are selected based on prior research. The framework proposes to start with a qualitative analysis to filter out the processes that are not aligned with the organization's automation strategies. Then, it uses process mining techniques to quantitatively evaluate the performance of the remaining processes. With these techniques, insights into the performance of processes can be extracted from collections of event logs (Van der Aalst, Weijters, and Maruster, 2004). With those insights, the performance of a process including its tasks can be evaluated objectively, rather than on the perception of process participants. The proposed framework is evaluated in a case study in an industrial setting.

The paper is organized as follows. Section 2 discusses the related work and a list of criteria collected from existing literature. Section 3 describes our research method. Section 4 presents the PLOST framework. Section 5 explains the evaluation, a case study and two thinking-aloud sessions, one with a domain expert and one with an RPA expert. Sections 6 discusses the results, while Section 7 concludes the paper.

2 Theoretical Framework

In this section, we introduce the background regarding the combination of process mining and RPA and discuss techniques to identify suitable tasks for RPA. This background is formed based on a literature review. This has been conducted following the, semi-structured snowballing approach (Wohlin, 2014). An initial set of literature was formed, after which more relevant literature was collected by both forward and backward snowballing.

2.1 Combining Process Mining and RPA

Identifying where RPA could provide the most value is challenging (Choi, R'bigui, and Cho, 2020), because it often relies on the study of process documentation (Jimenez-Ramirez et al., 2019). The lifecycle of RPA projects starts with the identification phase, in which the processes to be automated are analyzed and selected (Jimenez-Ramirez et al., 2019). Process mining can be used to identify promising candidate tasks (Van der Aalst, 2021) because the discovery of RPA possibilities is closely related to Automated Process Discovery, which is studied in the field of process mining (Augusto et al., 2018). To find out which tasks are suitable to be automated with RPA, a new class of techniques called Robotic Process Mining (RPM) has been proposed (Leno et al., 2020). RPM techniques focus on discovering automatable routines from logs of interactions between workers and digital applications. The RPM tools take as input logs of user interactions with the applications, which are called *user interaction (UI) logs*. These UI logs replace the traditional event logs typically used by process mining techniques. With such a UI log, an RPM tool aims to identify the tasks that can be automated and their boundaries. Additionally, variants of each task are collected, standardized, and streamlined. This helps to discover an executable specification that corresponds to the streamlined and standardized variant of the task. The identified tasks can be defined in a platform-independent language, which can be compiled into a script to be executed by an RPA tool.

Another application of process mining to the implementation of RPA is seen in the deployment phase, for example in the approach developed by Geyer-Klingeberg et al. (2018). In this approach, process mining is deployed to help find the most effective RPA implementation. First, they started with training robots with the existing workflow, while their activities are tracked by the underlying IT systems. After some executions, the generated process instances are evaluated by using process mining techniques. In this way, the performance of the different robot executions and the human-supported processes are compared to select the best-performing implementation.

After RPA has been implemented, process mining techniques can also contribute to the successful implementation. At the operation and maintenance stage, it can be used to monitor processes and systems, even if these use a combination of RPA bots, human employees, and traditional automation (Van der Aalst, 2021). This can be done by using real-time detection of process changes over time through process mining techniques (Geyer-Klingeberg et al., 2018). This ensures tracking the impact of the implementation and more importantly, the return on the investment. It can also help to detect when a process evolves in such a way that the robot needs to be adopted to an alternating business environment, reducing the chance of RPA failures.

2.2 Techniques to Identify Suitable Tasks for RPA

In earlier studies, four approaches have been proposed that identify RPA candidates. The first is the RPA Suitability Framework (Agaton and Swedberg, 2018), a qualitative method operating on the process level. The framework consists of five steps and uses the Business Process Model and Notation (BPMN) to model all candidate processes. It analyzes the risk level and business values using six mandatory criteria. Although the use of business process mapping is seen as an effective addition when identifying suitable processes for RPA (Seasongood, 2016), it is also time-consuming. The second framework developed by Haan (2021) combines both qualitative and quantitative analyses on the low-level of detail, i.e., the task level. Although the framework is said to give the advantage of providing a strong basis for a business case for RPA, no business metrics were taken into account. Without such metrics, the strategy of an organization cannot be aligned with the automation. Moreover, the authors argue that applying process mining to all processes is a time-consuming activity. A third method proposed by Choi, R'bigui, and Cho (2020) makes use of RPM and focuses on the task level. Because no qualitative analysis is conducted, the UI logs are based on recordings of all the actions a user performs. The result is that the method includes a log filtering step that is not only time-consuming, but it raises other challenges as well, such as calculating the duration of tasks in real-life situations. Last, the Framework for Process Suitability Assessment (FPSA) (Shafik Salah Elsayed and Kassem, 2022) applies quantitative analyses on the process level. The framework provides a customized approach by taking the organization's objectives into account, but the subjective approach goes against the idea of developing a data-driven objective framework. Moreover, it is only demonstrated with sample data.

In Table 1, we summarize the criteria that are used in the four methods. The meaning of each criterion is included in the corresponding thesis (Jongeling, 2022). Criteria that have a similar meaning but a different name are merged into one criterion. For example, the criterion *Structured Digital Data* is a combination of

Criterion	Method	Mandatory?	Scope		Analysis type	
			Process level	Task level	Qualitative	Quantitative
(1) Structured Digital Data	(Agaton and Swedberg, 2018; Haan, 2021; Shafik Salah Elsayed and Kassem, 2022)	✓	✓		✓	
(2) Easy data access	(Agaton and Swedberg, 2018)	✓	✓		✓	
(3) Standardized process	(Agaton and Swedberg, 2018)		✓		✓	
(4) Few exceptions / Low variations	(Agaton and Swedberg, 2018; Haan, 2021)	✓	✓		✓	
(5) Repetitive	(Agaton and Swedberg, 2018; Shafik Salah Elsayed and Kassem, 2022)	✓	✓		✓	
(6) Rules Based	(Agaton and Swedberg, 2018; Haan, 2021; Shafik Salah Elsayed and Kassem, 2022)	✓	✓		✓	
(7) Mature	(Agaton and Swedberg, 2018; Haan, 2021)	✓	✓		✓	
(8) Multiple systems	(Agaton and Swedberg, 2018)		✓		✓	
(9) Digital trigger	(Agaton and Swedberg, 2018)		✓		✓	
(10) Redeployable personnel	(Agaton and Swedberg, 2018)		✓		✓	
(11) High Standardization level	(Shafik Salah Elsayed and Kassem, 2022)		✓			✓
(12) Low exception handling	(Shafik Salah Elsayed and Kassem, 2022)		✓			✓
(13) Frequency	(Shafik Salah Elsayed and Kassem, 2022)		✓			✓
(14) Low Process complexity	(Shafik Salah Elsayed and Kassem, 2022)		✓			✓
(15) Low automation rate	(Shafik Salah Elsayed and Kassem, 2022)		✓			✓
(16) High number of FTE's	(Shafik Salah Elsayed and Kassem, 2022)		✓			✓
(17) High execution time	(Shafik Salah Elsayed and Kassem, 2022)		✓			✓
(18) Prone to human error	(Shafik Salah Elsayed and Kassem, 2022)		✓			✓
(19) Frequency	(Choi, R'bigui, and Cho, 2020; Haan, 2021)			✓		✓
(20) Duration	(Choi, R'bigui, and Cho, 2020)			✓		✓
(21) Time sensitive	(Haan, 2021)			✓		✓
(22) Human Productivity	(Haan, 2021)			✓		✓
(23) Human error prone	(Haan, 2021)			✓		✓
(24) Irregular labor	(Haan, 2021)			✓		✓
(25) Periodicity	(Haan, 2021)			✓		✓
(26) Cost reduction	(Haan, 2021)			✓		✓

Table 1. Overview of the 26 unique criteria from the existing four methods

the criterion *Digital and structured data* (Agaton and Swedberg, 2018), *Structured Readable Input* (Haan, 2021), and *Structured Digital Data* (Shafik Salah Elsayed and Kassem, 2022). Some other quantitative criteria were named with a subjective term, such as *high*, *low*, or *few*.

In this study, we build on the different criteria from the existing methods, and combine them into a framework that takes into account the different levels and analysis types. We describe the development of this framework in the following section.

3 Research Approach

To develop our framework, we adopt a design science research approach. Our goal is to develop an artifact that helps organizations to identify and prioritize suitable RPA candidate tasks, while explicitly taking into account the automation strategy of an organization. We follow the Design Science Research Methodology (Peffers et al., 2007) which starts with the *Problem identification and motivation*. In the context of our study, the problem is the difficulty of identifying RPA tasks in an objective and time-efficient manner. From the literature review outlined in the previous section, we define our *Objectives of a solution*. The proposed framework should take into account both the *process* and *task* level and combine the strengths of *qualitative* and *quantitative* approaches. For the *Design and development* of our artifact, we then transformed the list of criteria from Table 1 into different steps and synthesized our framework (see Section 4). More specifically, there are three steps in our final framework which use these criteria to conduct analysis: Step 3 focuses on *process*-level *qualitative* analysis, Step 6 focuses on *process*-level *quantitative* analysis, and Step 7 focuses on *task*-level *quantitative* analysis, as listed in Table 2. With these different focuses in mind, the list of criteria for each step is selected accordingly, see Table 3. Note that the names of the criteria in the final list are not exactly the same as the ones from Table 1, as subjectivity is removed from the names and the most logical word order is chosen.

	Criteria					
Steps	***Mandatory?***	***Process level***	***Task level***	***Qualitative***	***Quantitative***	***Using Process Mining?***
Step 3	✓	✓		✓		
Step 6		✓			✓	✓
Step 7			✓		✓	✓

Table 2. Selecting the criteria for the three steps in the PLOST framework.

Step 3	**Step 6**	**Step 7**
Digital and Structured Input	Cycle Time	Activity Frequency
Easy Data Access	Case Frequency	Case Frequency
Few Variations	Activity Frequency	Duration
Repetitive	Standardization	Automation Rate
Clear Rules	Length	Human Error Prone
Mature	Automation Rate	Irregular Labor
	Human Error Prone	

Table 3. The criteria used in Step 3, 6, and 7 of the PLOST Framework, adapted from existing criteria in Table 1

The steps form the basis of our proposed framework, which we consequently *Demonstrated* through a case study at ProRail[1], the company that manages and maintains the Dutch railway network. The objective of the case study was aimed at demonstrating the steps of the framework and improving it accordingly.

We started off by determining the automation strategy with six stakeholders. Then, interviews were conducted with different stakeholders to create the initial set of processes. The result was a collection of six processes. In the third step, these processes were analyzed regarding the mandatory process criteria. In the next step, data for the two processes was collected from the ITSM tool Marval[2] with help of the business intelligence software Xtraction[3]. Two event logs were made; one for each process. In the fifth step, dashboards were created with the process mining tool Celonis. For each process, a process and a task dashboard were created with all the available statistics that would be needed later in the framework.

After the demonstration of the framework, an *Evaluation* took place in the form of thinking-aloud sessions that were performed with an RPA expert and a domain expert. Thinking-aloud as a method is characterised by encouraging participants to express out loud what they are looking at, thinking, doing and feeling, while they perform certain tasks (Gupta, 2015), in this case the steps of the PLOST Framework. Both participants received a clear tutorial of the framework (see footnote 4), different templates where they could fill in the information of the different steps and two process mining dashboards. The focus of the evaluation was on the usability, practicality and completeness of the framework. During the thinking-aloud experiments, the participants walked through every step in chronological order with the goal to identify a prioritization of suitable RPA tasks. To get their opinions, questions were asked at the end of each step

[1] `https://www.prorail.nl`
[2] `https://www.marval.co.uk/`
[3] `https://www.ivanti.nl/products/xtraction`

and after they finished the session (see footnote 4). Because of the scope of the research, the experts did not execute step two, four and five of the framework themselves. In these steps respectively the processes are gathered, the data is collected and process mining is applied.

4 The PLOST Framework

In this section, we explain the PLOST Framework (Figure 1). The first six steps focus on the process level, after which there is a switch to the task level for the last two steps. The steps on the process level work as a funnel, which means that depending of the output of the step the decision can be made to skip some steps. This results in only applying quantitative analysis to the processes and tasks that are worth analyzing. In the following, we explain the eight steps of the PLOST framework in depth. More explanation of the steps can be found in the related thesis (Jongeling, 2022). To show some concrete examples, we sometimes refer forward to some figures in Section 5. A complete tutorial on the framework is made available via github [4].

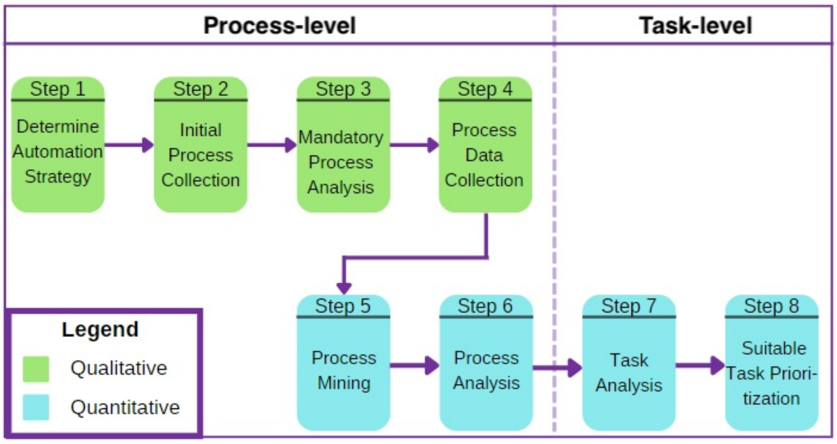

Figure 1. The PLOST Framework

Step 1 - Determine automation strategy. In the first step, the PLOST Framework starts with developing an organization's automation strategy. The strategy is based on two components that are also used in Agaton and Swedberg (2018): the prioritization of the business values (BV) and the determination of the risk level. These two parts are executed with the help of the stakeholders of the implementation. Because multiple opinions have to be taken into account when making a decision, the prioritization is done by using cumulative voting, also known as the 100-Point method (Dean and Don, 2003). Each stakeholder is given 100 points to divide over the three business values. In this way, the amount of points assigned to an issue represents the stakeholder's preference in relation to other issues. For each business value, the final score is the average of all the stakeholders. To create a steady automation strategy, it would be best to include at least three stakeholders. The prioritized business values are: *time saving*, *quality and accuracy improvement*, and *availability & flexibility increase*. An example of the business value assessment is shown in Fig. 2.

With the assessment of the risk level, the organization indicates how much risk they are willing to take with the RPA implementation. The three risk levels are inspired by the method outlined in Agaton and Swedberg (2018) and is based on two main factors: *Process Importance* and *Process Complexity*. The former determines whether or not essential processes are considered, and the latter determines which level of complexity the process can have. The framework includes three levels of risk, where the first one means the highest risk level and the last one the lowest. The highest risk level can be found in processes

[4] `https://github.com/HildeJongeling/PLOST-Framework/`

Question	Motivation
What is an example of a process that fits within the description?	This question provokes the interviewee to start telling about a new process.
How does this process start?	To understand the process, it is important to know if the process is manually started or triggered by another task.
What are the different steps of the process?	This questions helps to thoroughly understand the process.
Are these steps always the same?	This question is important because if the steps differ from time to time, the process is not a candidate for automation.
Which applications are involved?	With the answer of this question the consideration can be made if RPA is the right form of automation.
Which person is executing the process?	To understand the context of the process, it is good to know who is executing the process.
How often is this executed?	Only frequent processes are worth automating.
Is there an improvement going on with this process?	If someone within ProRail is already improving the process, then applying RPA is of no use now as it is not known how the future process will look like.
Is this process improved before?	Based on previous improvements and their results, a better estimate can be made.

Table 4. (Step 2) Interview questions to collect processes.

that score high on importance and complexity. At the medium risk level, processes score high on one of these factors. A low risk level is linked to processes that score low on both importance and complexity.

Step 2 - Initial process collection. In the second step, processes are identified at the organization with the help of semi-structured interviews. What the optimal amount of interviews is, is hard to say as quality says more than quantity. We advise to conduct at least two interviews with two different interviewees. The interviewees are domain experts within the organization. Such an expert can have any role or function that an employee within the scope can have, from manager to system administrator. This interview script consists of three parts: a general introduction, the questions about processes, and the closing questions. Table 4 show the questions asked about processes, the full set of questions is shown in the corresponding thesis (Jongeling, 2022). As the result of this step, the collected processes during these interviews form the initial selection of processes.

Step 3 - Mandatory process analysis. The third step analyzes the processes in the initial process selection made in step 2 based on six qualitative criteria. All these criteria are mandatory, meaning that a process should meet all of them to be worth analyzing further. If that is not the case, the process will be removed from the selection. This shows how this step works as a funnel. The six criteria based on which the mandatory process analysis is done are *Digital and Structured Input*, *Easy Data Access*, *Few Variation*s, *Repetitive*, *Clear Rules*, and *Mature*, listed in Table 3. The processes that meet all the mandatory criteria form the revised process selection in the framework.

Step 4 - Process Data Collection. In this step, process data is collected for the processes in the revised process selection, that was made in step 3. How detailed the data should be depends on the decision made in the automation strategy. More details about the requirements for the process data can be found in the related thesis (Jongeling, 2022). When the desired output is detailed, data with a high detail level should be gathered. With this process data, an event log is created for each process. This event log should be available in different formats such that the log can be used in any process mining tool.

Step 5 - Process Mining. From the event logs, visualizations of the processes are made using process mining tools. The process analyst decides which tool to use; examples of process mining tools are Celonis, Disco and ProM.

Step 6 - Process Analysis. In the sixth step, the output of step five is analyzed against seven quantitative criteria. This happens at the process level. With this analysis, a decision is made as to which process best matches the chosen risk level from the first step. With the help of the criteria overview in Table 1, the decision was made to use the following criteria: *Cycle Time*, *Case Frequency*, *Activity Frequency*, *Standardization*, *Length*, *Automation Rate*, and *Human Error Prone*. The output of this quantitative analysis can be aligned with the chosen risk level from the automation strategy. The higher the chosen risk level the higher the values of the criteria can be, while for a lower risk level it is the other way around. With this information, a decision can be made regarding the process, or processes, that remain in the set.

Step 7 - Task Analysis. The seventh step analyzes the different tasks in the remaining process, or processes, from step six with the help of six quantitative criteria. This analysis takes place on the task level. The criteria are *Activity Frequency*, *Case Frequency*, *Duration*, *Automation Rate*, *Human Error Prone*, and *Irregular Labor*. The values of the tasks are collected in a table.

Step 8 - Suitable Task Prioritization. In the last step of the PLOST Framework, all the components come together to result in the final output: a prioritized list of tasks that are suitable to be automated with RPA. First, all the tasks in the task analysis from step seven are ranked for each criterion where the highest value receives the score *N*. N is the number of tasks in the process. The second highest value gets the score *N-1* and so on. After that, the scores of the ranking are multiplied by the scores of step one. Each business value matches one or more criteria from the task analysis in step seven. If a criterion matches multiple business values, the highest score is used to calculate with. By adding up the scores for all the criteria for each task, the final scores are calculated. This leads to the final prioritization where the task with the highest score has the highest prioritization and the task with the lowest score has the lowest prioritization. This does not mean that the task with the lowest score is not worth automating; it rather means that other tasks generate more of the business values that the organization desires. It is also worth mentioning that the final score does not have a meaning, it is only used to make up the ranking.

5 Demonstration and Evaluation

In this section, we describe the steps that we took at the case study organization. Then, we reflect on the steps, describe the results from our evaluation with domain experts, and close with the lessons learned.

Step 1: Determine Automation Strategy. To determine the automation strategy, the business values need to be prioritized and the risk level has to be assessed. As for the first, different stakeholders at ProRail were asked for their opinion, which resulted in the prioritization in Figure 2. Six stakeholders gave their opinion about which value they thought deserved the highest prioritization.

After the business values prioritization, the risk level was assessed. Together with the stakeholders, it has been decided to choose a low risk level. ProRail does not have experience with RPA yet and wants its first RPA implementation to be an example for the rest of the organization to look into RPA. To increase the chances of a successful implementation, the low risk level is chosen. With the outcome of the business values prioritization and the decision on the risk level, the automation strategy is made and can be found in Figure 3. The average scores are calculated.

Step 2: Initial Process Collection. For this step, we conducted semi-structured interviews with four experts from ProRail, from departments that are related to the Central Service Desk (CSD). The identi-

Business value	S1	S2	S3	S4	S5	S6	Total
Time Savings	35	15	47	70	25	25	217
Quality & Accuracy Improvement	60	40	39	30	55	50	274
Availability & Flexibility Increase	5	45	14	0	20	25	109
Total	100	100	100	100	100	100	600

Figure 2. (Step 1) Prioritization of the business value as made by the stakeholders.

Business value	Score
Time Savings	36,17
Quality & Accuracy Improvement	45,66
Availability & Flexibility Increase	18,17
Risk level	Low

Figure 3. (Step 1) Automation strategy as made by the stakeholders.

Criteria	P1	P2	P3	P4	P5	P6
Digital and structured input	✓	X	✓	X	✓	✓
Easy data access	X	X	X	✓	✓	✓
Few variations	X	X	✓	X	✓	✓
Repetitive	✓	✓	✓	✓	✓	✓
Rules Based	✓	✓	✓	X	✓	✓
Mature	✓	✓	✓	✓	✓	✓
Filtered away	✓	✓	✓	✓	X	X

Figure 4. (Step 3) Mandatory process analysis of the ProRail case study.

fication and roles of the interviewees can be found in Table 5. The column *Outcome* shows how many processes were collected during each interview. Out of the sixteen collected processes, eight processes were existing processes that are executed. Five processes are descriptive without execution data. For the other three processes, an attempt is made to check whether they exist. For the resulting eleven processes that are being executed in the organization, we check if the event data are suited within the scope of this research. As a result, six processes are selected to keep in the initial process selection of the framework, listed in Table 6.

#	Role	Outcome
1	Process leader CSD	3
2	IT Advicer & Operations	6
3	Manager CSD	1
4	Process leader at department of IMA & Process Support	6

Table 5. Roles of the interview participants and the amount of processes that resulted from the interviews.

Process id	Process Description
1	Manually searching for the incident handling scenario most suited for an incident
2	Manually adding changes to the Marval ticket of an incident when a change is occurring or executed, while the change(s) and the incident are related.
3	Manually adding personal details for an access request for person related to a change when a change has been approved.
4	Sending e-mail to OS (Operations Support) when a change has not yet been executed, but the change is prepared and the end time has arrived.
5	Having a priority 1 incident, sending a SMS via a web form to related people.
6	Creating a Marval ticket and resolving the incident after receiving a NCSC notification by e-mail.

Table 6. (Step 2) The six processes in the initial process selection of the ProRail case study.

Step 3: Mandatory Process Analysis. Next, the processes are analyzed regarding the mandatory process criteria. Figure 4 shows the mandatory process analysis of the six processes. Because the first four processes fail some of the mandatory criteria, they are removed from the framework and only processes five and six are left. These two processes together form the revised process selection.

Step 4: Process Data Collection. The data of the two processes, that resulted from step three, is collected with the help of Xtraction [5]; business intelligence software developed by Ivanti[6]. ProRail uses Xtraction as the report tool for Marval, their ITSM software. Xtraction provides multiple rules, grids, and graphical representations of data. This can be summarized in reports and dashboards, which can be scheduled so that the real-time data is sent by email or placed on a server. All the fields from Marval are available in

[5] `https://www.ivanti.nl/products/xtraction`
[6] `https://www.ivanti.nl/`

Xtraction and new fields can be added as well. To collect the data for the two processes, different steps were taken. For every process, a work list was made in Marval with all tickets for the two processes occurring between 16 May 2021 and 16 May 2022. Other filters for this data were that the type of the ticket equals *Incident*, the current status equals *Closed* and they are not archived. After creating the work lists, they are exported as a CSV file. All the IDs for the requests are copied and used in Xtraction to filter on these IDs in a data source called *MSM12 X13 DM V1102 Request is/was Status Report*. All event data of these IDs is collected and again exported as a CSV file. The output is two CSV files for the two different processes.

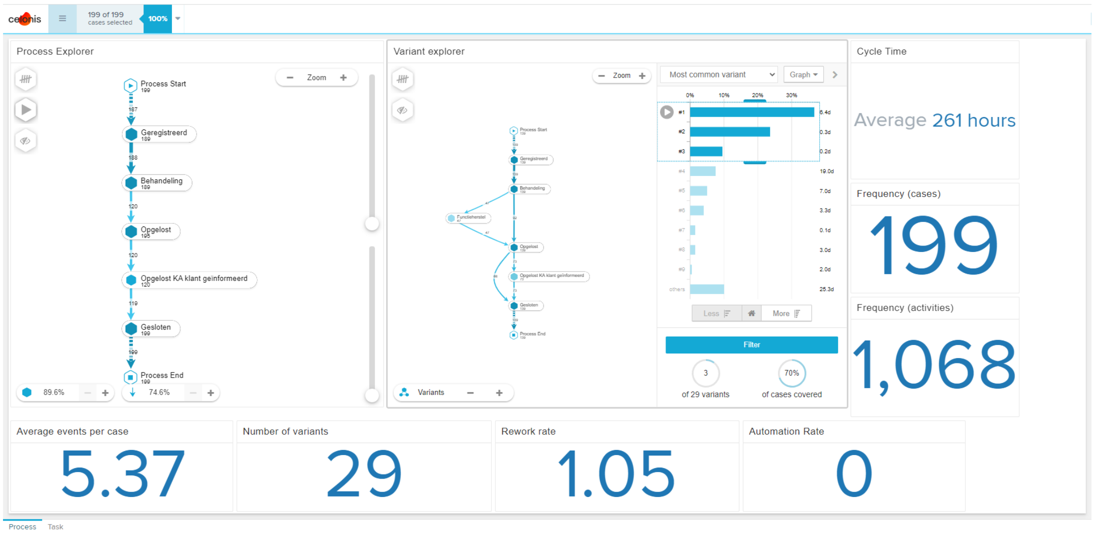

Figure 5. The process dashboard of the SMS Prio 1 process in the tool Celonis.

Step 5: Process Mining. The two data files from step four are used to apply process mining techniques. Because Celonis offers a free academic account with clear guidance and a user-friendly interface, this process mining tool is chosen for the case study. Two data pools were created through a manual file upload. With this data, two process data models are made. Figure 5 shows the process dashboard of the SMS Prio 1 process in Celonis. The two dashboards are also available online[7].

Step 6: Process Analysis. The partner organization has chosen a low risk level in step one, which means they aim to automate a non-essential process with low complexity. Based on the process dashboards made with Celonis, Table 7 was filled with the process analytics. For each criterion, the best value is marked in the table with a green color. Whether a higher or a lower value is better, depends on the chosen risk level. For the low risk level, a low value is marked as the best. Because process six has more green colored cells than process five, process six remained in the framework while process five was eliminated. Unfortunately, the performer of the tasks was not clearly described in the data. Therefore, the automation rate was zero for both processes. In this case study, it would not have made a difference in the outcome if one of the two processes had a higher or lower automation rate, but it is good to check already in the data collection step whether this data can be retrieved somewhere.

As can be seen in Table 7, the SMS Prio 1 process has two colored cells, while the NCSC process has six colored cells. This means the latter matches the chosen risk level the best and therefore is further analyzed in the next steps. This means the SMS Prio 1 process is eliminated.

Step 7: Task Analysis. In the seventh step, the task dashboard created in Celonis helped fill the task analytics of the sixth process, which resulted in the table shown in Figure 6. The colors represent the value of the cell regarding the other values in that row. This helps making the ranking in step eight.

[7] The SMS Prio 1 dashboard, see link, and the dashboard of process six is called the NCSC Process Dashboard, see link

Criteria	P5) SMS Prio 1 process	P6) NCSC process
Cycle Time	261 hours	273 hours
Case Frequency	199/year	100/year
Activity Frequency	1068	475
Standardization	29 variants	7 variants
Length	5.37	4.75
Automation Rate	0.00	0.00
Human Error Prone	1.05	1.01

Table 7. Quantitative process analysis for the two processes in the ProRail case study.

Criteria	T1	T2	T3	T4	T5	T6	T7	T8
Activity Frequency	100,00	105,00	4,00	47,00	100,00	12,00	103,00	3,00
Case Frequency	100,00	100,00	4,00	47,00	100,00	12,00	100,00	3,00
Duration	0,00	4,49	16,43	2,86	4,00	4,11	0,02	3,61
Automation Rate	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Human Error Prone	1,00	1,05	1,00	1,00	1,01	1,00	1,03	1,00
Irregular Labor	0,27	0,27	0,00	0,33	0,64	0,00	0,44	0,00

Figure 6. (Step 7) Quantitative task analysis for the tasks in the NCSC process.

Step 8: Suitable Task Prioritization. The final prioritization is made with help of Table 7 from step seven. The tasks in the table are ranked for each criterion from eight to one, as there are eight different tasks in the process. This ranking is shown in Figure 7. The scores of this table are multiplied by the output of step one. The result of this calculation is the final prioritization of the tasks in the NCSC process, which is shown in Figure 8. The first task has the highest business value to be automated and the eighth task the lowest value. The outcome does not necessarily mean that the eighth task is not worth automating. Especially when this task is needed to automate another task, it could be that it still has to be automated.

Criteria	T1	T2	T3	T4	T5	T6	T7	T8
Activity Frequency	6,00	8,00	3,00	5,00	6,00	4,00	7,00	2,00
Case Frequency	8,00	8,00	5,00	7,00	8,00	6,00	8,00	4,00
Duration	1,00	7,00	8,00	3,00	5,00	6,00	2,00	4,00
Automation Rate	8,00	8,00	8,00	8,00	8,00	8,00	8,00	8,00
Human Error Prone	5,00	8,00	5,00	5,00	6,00	5,00	7,00	5,00
Irregular Labor	5,00	5,00	4,00	6,00	8,00	4,00	7,00	4,00

Figure 7. (Step 8) Ranking of the tasks in the NCSC process.

Criteria	T1	T2	T3	T4	T5	T6	T7	T8
Activity Frequency	273,96	273,96	136,98	228,30	273,96	182,64	319,62	91,32
Case Frequency	365,28	365,28	228,30	319,62	365,28	273,96	365,28	182,64
Duration	36,17	253,19	289,36	108,51	180,85	217,02	72,34	144,68
Automation Rate	365,28	365,28	365,28	365,28	365,28	365,28	365,28	365,28
Human Error Prone	228,30	365,28	228,30	228,30	273,96	228,30	319,62	228,30
Irregular Labor	90,85	90,85	72,68	109,02	145,36	72,68	127,19	72,68
Total	1359,84	1713,84	1320,90	1359,03	1604,69	1339,88	1569,33	1084,90

Figure 8. (Step 8) The prioritization of the tasks in the NCSC process.

5.1 Reflection on the Steps

This case study showed that the PLOST Framework results in a ranking of RPA candidate tasks with which ProRail can decide where to start automating. One interesting lesson was learned during the case study. This lesson involves the level of detail of the event data. In the case study, the event log used for ranking and selecting the tasks contained only task-level information (e.g., the duration of a task, the resource who executed the task), which are less suited for automatically deriving RPA bots. The latter requires UI logs (Gao et al., 2019; Leno et al., 2020). Nevertheless, the tasks selected using the PLOST framework can be used as input for collecting specific UI logs, reducing the manual effort needed to collect the UI logs. By applying the framework to a real-life use case, the applicability of the framework is demonstrated.

5.2 Evaluation

The evaluation of the framework took place in the form of think-aloud sessions and an interview after the sessions. Overall, both experts mentioned that the framework can easily be applied in other organizations or use cases, which is perceived as a major advantage. Moreover, they appreciated that the framework first focuses on the business side, then carries out a qualitative check and finishes with a quantitative analysis, especially because such quantitative analysis is objective and helps convince top management of RPA. By observing the participants executing the framework, it was evident that they could finish all components, which validates the usability of the framework. When asking the participants how they rated the addition of process mining, they gave it an average of 7.5. This could be improved by changing the level of detail of the process mining data. This demonstrates the practicality of the framework.

Both experts thought the overall completeness of the framework was adequate, although they have two suggestions. The first one was to not remove processes from the framework if they do not yet meet the mandatory criteria, but rather find out if they can be redesigned to fit the criteria. However, this only applies when the criterion Repetitive is met. The second suggestion was to make the final ranking clearer in the sense of providing a ranking overview rather than a table with highlighted numbers. The experts mentioned as well that the last task of the prioritization can still be suitable for automation but has less priority. Apart from the two ideas, the experts advised to apply the framework with more detailed data.

5.3 Lessons Learned

In this section, we summarize the lessons learned. The first one is that the desired level of detail of the data has to be specified in the automation strategy of step one. This will then be used in step four as well to collect the desired data. In case this data is not available, UI logs can be created and gathered. For this creation, substep 4a is added to the framework. It is out of scope for this research to specify how to obtain UI logs, but this has been explained in the method by (Choi, R'bigui, and Cho, 2020). The second lesson is to add an additional data check in step four to make sure all the needed data for the metrics of the criteria in step six and seven is available. The third one resulted in substep 3a. If a process does not yet meet all the mandatory criteria, the process is saved and redesigned if possible. This only applies to processes that meet the criterion Repetitive. The preference is still given to processes that directly meet all the mandatory criteria, as redesigning will take extra time.

Threats to validity We discuss the threats to validity with respect to the PLOST Framework, the case study, and the thinking-aloud experiments. The PLOST Framework is only applied in one case study and should be tested more extensively in different organizations and case studies. The main limitation of the case study was that the outcome did not consist of concrete RPA solutions but rather a set of candidate tasks suitable to be automated. Because of this, the effectiveness of the framework could not be evaluated in terms of the effectiveness of the RPA implementations. That the output of the framework is not directly ready to be automated does not have to be a problem, if that is what the desired output is. This can be the case when there is limited time and the framework is just used to scan which tasks are the most suitable to investigate further. The thinking-aloud experiment was conducted with two experts. Although these gave interesting and helpful insights, the evidence of the usability, practicality, and completeness of the framework would have been stronger if the experiments had been conducted with more experts. Another limitation of the thinking-aloud experiments was that the two experts only executed specific steps of the framework. Due to practical and time constraints, the steps that included the process gathering, data collection, and process mining dashboards were already built for them.

6 Discussion and future work

This section addresses the main contributions of the PLOST Framework as well as its limitations, which will then lead to recommendations for future work. The main contribution of the PLOST Framework is offering a way to identify and prioritize where RPA could be used with the help of both qualitative and quantitative analysis steps. The addition of the automation strategy that determines the output is a second contribution made in this study. A third contribution is that the metrics of tasks from different processes can be compared with each other. For example, when conducting the last two steps of the framework with multiple processes, the framework can compare if task A from process one is more suitable to automate than task B from process two. This has not been discussed before in the studied literature.

Another evaluation method of the PLOST Framework is to implement the output of the PLOST Framework. This means automating the task that came out of the PLOST Framework as the best suitable task to automate with RPA. To do this, a RPA tool of own choice can be used. Due to data extraction constraints, this evaluation is not performed. More importantly, the output of the PLOST framework is a set of tasks selected for automation. However, when an organization desires to obtain the exact rules of a task, it is recommended to gather more detailed data. This step can be supported by creating UI Logs, as explained by Choi, R'bigui, and Cho (2021), or collecting low level event logs (also known as IO-log) to automatically derive automation rules (Gao et al., 2019). Due to time constraints, no UI logs were created in this research, which means it was not possible to implement an RPA solution. This means the effectiveness of the PLOST Framework could not be thoroughly tested. Future research could focus on obtaining UI logs to make it possible to test whether the output of the framework is automatable.

The insights into the limitations of this research offer an opportunity to propose possibilities for future work. First of all, future work could apply the revised version of the PLOST Framework and verify its effects. Applying the framework in different use cases and contexts than the one used in this research could lead to new insights in terms of the generalizability of the steps. This might include a reflection on the criteria used in the framework and the possible addition of relevant ones. Applications of the approach in new settings will most likely lead to new adjustments to the framework. Therefore, it is recommended to keep iterating back to the design phase so the framework improves with every application.

7 Conclusion

In this paper, we proposed the PLOST framework which systematically uses both qualitative and quantitative criteria to identify and prioritize candidate tasks to automate with RPA. The PLOST Framework consists of eight steps that start from establishing the automation strategy, to selecting processes using both qualitative methods as well as process mining techniques. The output of the framework is a list of prioritized suitable tasks for RPA. By applying the framework to an industrial case study, the applicability of the framework was demonstrated. The output of this case study was evaluated, and three lessons learned are discussed. During the case study, two thinking-aloud sessions were executed, one with an RPA expert and one with a domain expert. In these thinking-aloud sessions, the framework was evaluated on its usability, practicality, and completeness. The framework met the objectives; especially the addition of process mining to the identification of RPA candidate tasks was valuable. The combination of qualitative and quantitative aspects was mentioned as a benefit as well. Regarding the completeness of the framework, the experts had two suggestions, which were then incorporated into a revised version of the framework. In addition, an opportunity was identified to utilize RPM and UI log technologies.

References

Agaton, B. and G. Swedberg (2018). "Evaluating and Developing Methods to Assess Business Process Suitability for Robotic Process Automation-A Design Research Approach." MA thesis.

Augusto, A., R. Conforti, M. Dumas, M. La Rosa, F. M. Maggi, A. Marrella, M. Mecella, and A. Soo (2018). "Automated discovery of process models from event logs: Review and benchmark." *IEEE transactions on knowledge and data engineering* 31 (4), 686–705.

Berghuis, L. (2021). "Using the Wisdom of the Crowd to Digitalize: Designing a workshop-based process selection method for the identification of suitable RPA processes." MA thesis. University of Twente.

Burgess, A. (2017). *The Executive Guide to Artificial Intelligence: How to identify and implement applications for AI in your organization*. Springer.

Choi, D., H. R'bigui, and C. Cho (2020). "Robotic Process Automation Implementation Challenges." In: *International conference on smart computing and cyber security: strategic foresight, security challenges and innovation*. Springer, pp. 297–304.

— (2021). "Candidate Digital Tasks Selection Methodology for Automation with Robotic Process Automation." *Sustainability* 13 (16), 8980.

Dean, L. and W. Don (2003). *Managing software requirements: A use case approach*.

Gao, J., S. J. van Zelst, X. Lu, and W. M. P. van der Aalst (2019). "Automated Robotic Process Automation: A Self-Learning Approach." In: *OTM Conferences*. Vol. 11877. Lecture Notes in Computer Science. Springer, pp. 95–112.

Geyer-Klingeberg, J., J. Nakladal, F. Baldauf, and F. Veit (2018). "Process Mining and Robotic Process Automation: A Perfect Match." In: *BPM (Dissertation/Demos/Industry)*, pp. 124–131.

Gupta, S. (2015). "A comparative study of usability evaluation methods." *International Journal of Computer Trends and Technology* 22 (3), 103–106.

Haan, W. J. (2021). "How can process mining be used to identify Robotic Process Automation opportunities?" B.S. thesis. University of Twente.

Jimenez-Ramirez, A., H. Reijers, I. Barba, and C. D. Valle (2019). "A method to improve the early stages of the robotic process automation lifecycle." In: *International Conference on Advanced Information Systems Engineering*. Springer, pp. 446–461.

Jongeling, H. E. (2022). "Identifying And Prioritizing Suitable RPA Candidates in ITSM Using Process Mining Techniques: Developing the PLOST Framework." MA thesis. Utrecht University.

Lamberton, C., D. Brigo, and D. Hoy (2017). "Impact of Robotics, RPA and AI on the insurance industry: challenges and opportunities." *Journal of Financial Perspectives* 4 (1).

Leno, V., S. Deviatykh, A. Polyvyanyy, M. La Rosa, M. Dumas, and F. M. Maggi (2020). "Robidium: automated synthesis of robotic process automation scripts from UI logs." In: CEUR Workshop Proceedings.

Leopold, H., H. van der Aa, and H. Reijers (2018). "Identifying candidate tasks for robotic process automation in textual process descriptions." In: *Enterprise, business-process and information systems modeling*. Springer, pp. 67–81.

Madakam, S., R. Holmukhe, and D. K. Jaiswal (2019). "The future digital work force: robotic process automation (RPA)." *JISTEM-Journal of Information Systems and Technology Management* 16.

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

Ryan, F., M. Coughlan, and P. Cronin (2009). "Interviewing in qualitative research: The one-to-one interview." *International Journal of Therapy and Rehabilitation* 16 (6), 309–314.

Seasongood, S. (2016). "A case for robotics in accounting and finance." *Financial Executive*.

Shafik Salah Elsayed, N. and G. Kassem (2022). "Assessing Process Suitability for Robotic Process Automation: A Process Mining Approach."

Syed, R., S. Suriadi, M. Adams, W. Bandara, S. Leemans, C. Ouyang, A. ter Hofstede, I. van de Weerd, M. T. Wynn, and H. Reijers (2020). "Robotic process automation: contemporary themes and challenges." *Computers in Industry* 115, 103162.

Van der Aalst, W. (2021). "Process Mining and RPA: How To Pick Your Automation Battles?" *Robotic Process Automation: Management, Technology, Applications. De Gruyter*, 223–239.

Van der Aalst, W., T. Weijters, and L. Maruster (2004). "Workflow mining: Discovering process models from event logs." *IEEE transactions on knowledge and data engineering* 16 (9), 1128–1142.

Viehhauser, J. and M. Doerr (2021). "Digging for Gold in RPA Projects–A Quantifiable Method to Identify and Prioritize Suitable RPA Process Candidates." In: *International Conference on Advanced Information Systems Engineering*. Springer, pp. 313–327.

Wohlin, C. (2014). "Guidelines for snowballing in systematic literature studies and a replication in software engineering." In: *Proceedings of the 18th international conference on evaluation and assessment in software engineering*, pp. 1–10.