Universiteit Leiden

ICT in Business and the Public Sector

Incorporation of automated process discovery in the RPA lifecycle

Name: Aditi Vaishampayan
Student-no: S2379724

Date: 11/08/2020

1st supervisor: Dr. Joost Visser
2nd supervisor: Dr Werner Heijstek
Company supervisor: Jean-Louis Colen

MASTER'S THESIS

Leiden Institute of Advanced Computer Science (LIACS)
Leiden University
Niels Bohrweg 1
2333 CA Leiden
The Netherlands

Incorporation of automated process discovery in the RPA lifecycle

Aditi Vaishampayan

Leiden Institute of Advanced Computer Science (LIACS)
P.O. Box 9512, 2300 RA Leiden, The Netherlands

August 14, 2020

Abstract

Background. Organizations in many industries, in a constant pursuit to gain operational efficiencies, introduce automation using Robotic Process Automation (RPA). Their experiences show that manual process assessment for automation has a negative impact on the success of automation initiatives.
Aim. We study how automated process discovery using process mining enhances the RPA implementation process. Further this study aims to identify possible impediments that organizations can face while implementing such mining techniques. In addition, the study aims to provide a contrast in process mining and task mining approach to automated process discovery.
Method To this end, an integrated research methodology of case study review and exploratory research was adopted. Using qualitative research approach, 4 empirical case studies using process mining and RPA were reviewed. Additionally, we conducted stakeholder interviews at a large energy company to understand various phases in the RPA lifecycle. The data collected from multiple sources was analyzed, interpreted further.
Results. In all the case studies, automated process discovery emerged as a common method adopted to provide automation guidelines. All the case studies derived operational benefits by automating using the automated process discovery recommendation. Based on this knowledge, an updated RPA lifecycle roadmap is built upon the existing RPA lifecycle by Guðrún et al.
Conclusions. Incorporating automated process discovery in the process selection of RPA implementation, improves the success of overall RPA lifecycle.

"The first rule in any technology used in a business is that automation applied to an efficient operation will magnify the efficiency. The second is that automation applied to an inefficient operation will magnify the inefficiency" - Bill Gates

Contents

List of Figures

List of Tables

List of Abbreviations

ABPD Automated Business Process Discovery

AI Artificial Intelligence

APD Automated Process Discovery

BLM Bot Lifecycle Management

BPM Business Process Management

BPR Business Process Re-Design

RPA Robotic Process Automation

CoE Centre of Excellence

FTE Full Time Equivalent

KPI Key Performance Indicator

PM Process Mining

POC Proof Of Concept

RPM Robotic Process Mining

SDLC Software Development Lifecycle

Chapter 1

Introduction

Organization's approach to explore new digital technologies in order to maximize value creation for all stakeholders, is one way to implement digital transformation strategies [21]. Within digital transformation, novel and evolving automation capabilities has now become common place to improve efficiency, reduce cost and innovate. Robotic Process Automation (RPA) amongst them is becoming increasingly a strategic transformation lever to improve business performance [52].

Today's leadership understands how automation such as RPA derives economic gains with an increase in cash flow that benefits the entire organization [30]. According to Gartner, a leading research and advisory company, RPA demand has increased rapidly across the world with a software revenue of $846 million in 2018 to almost $1.3 billion in 2019, making it the fastest-growing segment of the global enterprise software market[36]. However, there are organisation who find it difficult to derive benefits from automation beyond the initial start-up phase. Until recently, 30-50% of the first RPA projects within an organization failed and were unable to scale-up [53].

The goal of this study is to investigate how automated process discovery techniques using process mining improves overall RPA implementation. The study on RPA lifecycle is conducted in a multinational energy utility company. However, the research is much broader than that and therefore empirical case studies are also looked into in order to understand different approaches organizations take to achieve process efficiency. Linking the case study analysis and the RPA lifecycle within the energy company, this study will draw conclusion about incorporating automated process discovery in the overall RPA lifecycle. Additionally, the study also aims to identify the challenges an organization comes across while implementing mining techniques.

1.1 Problem Statement

Organisations seeking operational efficiency have taken keen interest in RPA which

emulates human interactions with the information systems as part of particular business process [4]. This consequently, reduces the time taken to perform the business process thereby giving financial benefits.

For an energy company, who started implementing RPA as early as 2013 to keep business operating costs low and to improve customer service, RPA projects on an average, yielded 200 percent return on investment reduction compared to a manual process [31]. Despite the fact that these figures sound unreachable, the utility company had to go a long way to build the maturity level to do it the right way. It was through trial and error, that the organization learned to identify processes or sub-processes to automate[31]. Similar to the utility company, many organization undergo the same challenge of process selection in order to benefit from automation. Failure to do so leads to insufficient or misguided automation that doesn't lead to any service improvements, let alone cost savings.

Similarly, the challenge to assess the business process suitability for RPA that clearly generates a business case and utilizes the value of RPA is not self-evident [2]. Many researchers until now have ascertained RPA-suitable characteristics which help to evaluate the task for automation [44]. The tasks for automation should be Transactional, Rule-based, High volume, Standardised, Repetitive and with low-exception handling; to name a few suitable characteristics. Organisation typically, take into account the aforementioned characteristics to identify the possible processes or sub-processes for automation. Others build a customised questionnaire which the process expert or business users answer based on their knowledge and understanding of the processes to arrive at a choice of selection of processes [4]. Even then, lack of investigation and appropriate process assessment before initiating RPA, can lead to failure of such digital initiatives [2].

However, scientific methods like Process mining offers a data-driven approach to unearth and visualize the actual business workflow on the basis of transaction logs from the IT systems of the organization [49]. Industry implementation of process mining techniques, can be used in RPA landscape to automatically discover and identify processes with automation potential [17]. Despite the fact that, utilizing process mining in RPA context is a known feature, it has not yet been widely implemented by the organizations.

The void of scientific literature about automated process discovery aspect and its place in the RPA lifecycle demands more research. This study will draw a theoretical lens over automated process discovery that can be carried out for RPA and how it will further value-add to the overall automation lifecycle which is currently to a large extent limited to using traditional process assessment methods. By addressing this research gap, this study will delve into different approaches of automated process discovery the organization can engage into, providing practical recommendations for successful digital transformation using RPA.

1.2 Research Question

Concluding form the above, specific research questions are formulated as follows:
How can automated process discovery enhance the overall RPA lifecycle?

RQ1 - ***What are the challenges organisations face to adopt process mining for Robotic Process Automation?***

RQ2 - ***what are the trade-offs between Process Mining and Task Mining, the organization can consider while adopting automated process discovery?***

1.2.1 Objective of the Research

RPA is a rapidly emerging technology used in business process automation, although very little scientific research is available presenting overall RPA lifecycle within an organization. The lifecycle presented by the RPA vendors themselves and some demonstrated by the big consulting firms such as KPMG and Delliotte does not provide a complete picture [6, 50]. Therefore it is seen that the RPA roadmaps currently available are from different sources, but they disagree on certain aspects, lack details or no specific information about the evaluation of their models. Similarly, though many companies have adopted process mining, it is typically used to produce a "As-is" model of a company's business process and identify bottlenecks in the process. Significantly fewer companies have realised the benefit of using complementary technologies such as Process mining and RPA as part of the automation lifecycle.

Therefore, the objective of this study is -

1. To build upon the understanding of existing RPA roadmap presented by [43] and if necessary, modify it to incorporate automated process discovery as part of the RPA lifecycle.

2. The second objective is to present the challenges encountered by the organization in adopting mining techniques.

3. The last objective of this study is presenting the mining approaches an organization can consider with respect to RPA.

We believe that such scientific explanation will provide practical guidance to the organization to make a conscious decision of selecting the right mining approach for the organization.

1.3 Research Scope and Limitations

The focus of this research is on the overall RPA lifecycle within an organisation and

specifically the process assessment methods companies employ to plan automation. The scope under consideration is any organisation irrespective of the industry they operate in. Therefore, an organization from any industrial vertical like finance, energy, utility or telecommunication is suitable for this research. Given that this study investigates the phases within the RPA lifecycle, the research will consider organisations who have already began their RPA implementation.

The study focuses on process selection and assessment methods engaged by the organization. The researcher believes that the key to RPA success lies in the process selected for automation. This paper focuses on Process mining as a part of process assessment for RPA initiatives. However, other solutions of process selection found in scientific literature will be discussed in the later chapter. Correspondingly, the study explores the influence of having the mining approaches along with RPA. The researcher considers that determining the right mining approach for RPA, can be a decisive factor too in the overall RPA lifecycle. Due to the rising interest of organisations in RPA as well as PM, it has drawn lot of interest within the scientific circles. Hence during the study cycle, the possibility of further progress in this field cannot be neglected. To this end, it is possible that the observations and conclusions of this study to a certain extent outdated.

Additionally, there are certain limitations in terms of how the research was conducted. Firstly, this thesis takes the business perspective of RPA and PM as a precedent and therefore doesn't delve into the technological aspect such as coding, environment setup or testing for that matter. Likewise, from a business outlook process mining will be viewed as a tool or business application as a whole, details about specific process mining algorithms or their performances will not be discussed as it falls beyond the realms of this research. Secondly, due to the time and resources constraints, there are certain limitations to the number of interviewees. However, the researcher was able to interview experts and people working within the RPA project from different Business Units to gain better understanding of the situation from different perspectives. Out of the 9 interviewees conducted during the study, a fair amount of information and significant data points were gathered to arrive at a conclusion.

1.4 Structure of the Thesis

The remainder of this research undertaken will be structured as follows:

- Chapter 2, Conceptual foundation will set the knowledge base for this study. The first section presents how literature review was performed for this study. Further sections introduce terminology, identify important topics surrounding RPA which will illuminate over previous research in this field and contextualise RPA lifecycle.

- Chapter 3, Research Methodology will address the method adopted in this research.

- Chapter 4, RPA Stakeholder Interviews will present the exploratory research con-

ducted by the researcher. Based on the interviews with the stakeholders in a large energy company, we will study the current RPA lifecycle followed by the organization.

- Chapter 5, Review of past Case Studies will help us to analyze how organizations use automated process discovery in their process lifecycle.
- Chapter 6, Automated Process Discovery approach and challenges will concentrate on answering the sub RQ1 and RQ2. Based on the interviews with the mining expert and the RPA experts, this study will identify the barriers faced by the organization in adopting mining. Moreover, it will also provide the contrast in process mining and task mining approach of automated process discovery.
- Chapter 7, Enhanced RPA implementation lifecycle answers the main research question. Based on the case studies analysis and stakeholder interviews, we draw conclusions and propose an updated RPA implementation lifecycle. The section outlines the design method chosen for building the lifecycle as well as evaluation performed over the proposed lifecycle.
- Chapter 8 Conclusions and Recommendations will summarise the research thesis. This will be done by summarising the main findings, highlighting the scientific and practitioner's contributions and also suggest opportunities for further research.

Chapter 2

Conceptual Foundation

In order to build the conceptual foundation of our study, we explain the core concepts that play a role in our thesis. This first section explains how literature review was carried out in this study. The following sections introduce terminology, identify topics related to RPA and their inter-relation in this research.

2.1 Literature Review

The research objective entails the study and understanding of literature about RPA and overall lifecycle of RPA implementation. Further topics researched include current methods followed to identify suitable processes to automate as well as different mining techniques. To this end, Leiden University Database and Google scholar repositories were used primarily to find out various scientific literature available currently. The main search criteria and keywords used to access these repositories were RPA, Robotic Process Automation, BPM, RPA lifecycle, Automated Process Discovery, Process Mining and Task mining.

The number of scientific papers associated with RPA is relatively small. However, with the rising demand for automation in general, there has been an increasing trend since 2018. In order to broaden the understanding and the practical knowledge about this topic, we also accessed various blogs, articles, whitepapers of big consulting firms. In-contrast, there are a significant number of research paper about process mining in general although, most of the studies discuss different algorithms used for process mining rather than their business applications and benefits. Furthermore, to tie the whole study together, we studied a subset of relevant BPM literature.

The general process adopted during the study was reading specific text like the abstract, introduction and conclusion of the applicable literature and if necessary delve deeper. Backward reference searching was invaluable as it opened doors to further research and references. Relevant material was analyzed further as part of literature review while oth-

ers were used for establishing conceptual understanding for this paper.

2.2 Robotic Process Automation (RPA)

RPA is currently creating wide-spread interest in businesses and becoming a common phenomenon across various industries. To understand the origins of RPA as technology, one has to look back to three of its key predecessor technologies that include screen scraping software, workflow automation and management tools and Artificial Intelligence (AI) which is currently gaining a lot of attention in the research world [38]. RPA technology is hugely reliant on screen scrapping as well as workflow automation and certain RPA vendors also utilise optical character recognition (OCR) technology to adapt to changing websites without requiring intervention from a human employee. Moreover, AI is a technology that makes decisions about data, RPA vendors are increasingly using AI to recognize images on a screen more reliably [38].

Numerous definitions have eventually evolved since Robotic Process Automation term was first coined by Patric Geary in 2012, the Marketing Director in Blue Prism [23]. According to the IEEE Standards Association, RPA is "A pre-configured software instance that uses business rules and predefined activity choreography to complete the autonomous execution of a combination of processes, activities, transactions, and tasks in one or more unrelated software systems to deliver a result or service with human exception management" [1].

Many overlapping RPA definitions can be found within the realms of scientific literature as well as RPA software vendors; each one defining the term depending on the suitability of their needs. However, extensive and detailed analysis of RPA definitions has been collectively presented by Syed et al in their paper 'Robotic Process Automation: Contemporary themes and challenges' which contributes towards a solid understanding on RPA. Hence, the paper by Syed et al is the primary source of RPA literature used in this study. Based on the various characteristics of RPA, Syed et al interpret RPA as, "This term amalgamates robotics, referring to software agents acting as human beings in system interactions, and process automation, i.e. workflow management systems or, more generally, systems that are process-aware" [44].

In practice, the software agents or programmes that perform these system interactions are also termed as 'robots' or 'bots' for short. Lastly, big consulting firms have also coined their own terms for RPA such as 'Digital Labor' by KPMG or 'Digital Workforce' by Deliotte [19, 50]. Drawing conclusion from the comprehensive analysis of RPA characteristics using the paper by Syed et al [44], the table 2.1 is accomplished.

Finally, taking into account all the various RPA definitions, RPA can be portrayed as ***"A software-based solution that emulates human interactions spanning multiple applications to automate workflow management systems involving rule-based routine tasks with standardised data."***

Highly Rule-based	RPA requires a prescribed rule for every eventuality, which are unambiguous.
High volume	Sufficient routine transaction volumes that help to maximise benefits from RPA implementation
Mature	Mature tasks are those that have been in place for a while and are stable
Easy to achieve and show impact	Tasks performed within processes with the best return(clear understanding of manual costs) and simplest delivery
Has digitised structured data input	All input data must be digital and in a structured format
Highly manual	"Swivel chair"-like processes/tasks, which do not require much human intervention
Transactional	RPA suited for transactional work, as it increases accuracy and can perform many transactional activities at once
Standardised	Processes that execute consistently and follow a predefined path, especially in the initial RPA implementation phases
Low-levels of exception handling	Processes targeted for RPA should not have to deal with a lot of exceptional behaviours
Highly repetitive	Automating tasks that are 'repeatable enough' will help to yield a better return on investment
Less complex processes	Processes should be simple enough so that bots can be implemented quickly
Well-documented	Process descriptions that accurately detail processes are essential for bots to be taught behaviours at the keystroke level.
Interacts with many systems	Good candidates for RPA are processes that need access to multiple systems for high performance

Table 2.1: *Characteristics of RPA as elaborated in RPA : Contemporary Themes and Challenges [44]*

Moreover, to answer additional questions about RPA for instance 'Why', 'When', 'How' and 'What' is RPA, information from various scientific literature is consolidated in the following figure 2.1. The reason to implement RPA is the capability to virtually eliminate processing errors providing accuracy and removal of manual gaps between otherwise digitized systems, as mentioned by the Institute for Robotic Process Automation [24]. Yet another reason for companies to implement RPA was the increased compliance because software robots were configured to follow regulations precisely [30].

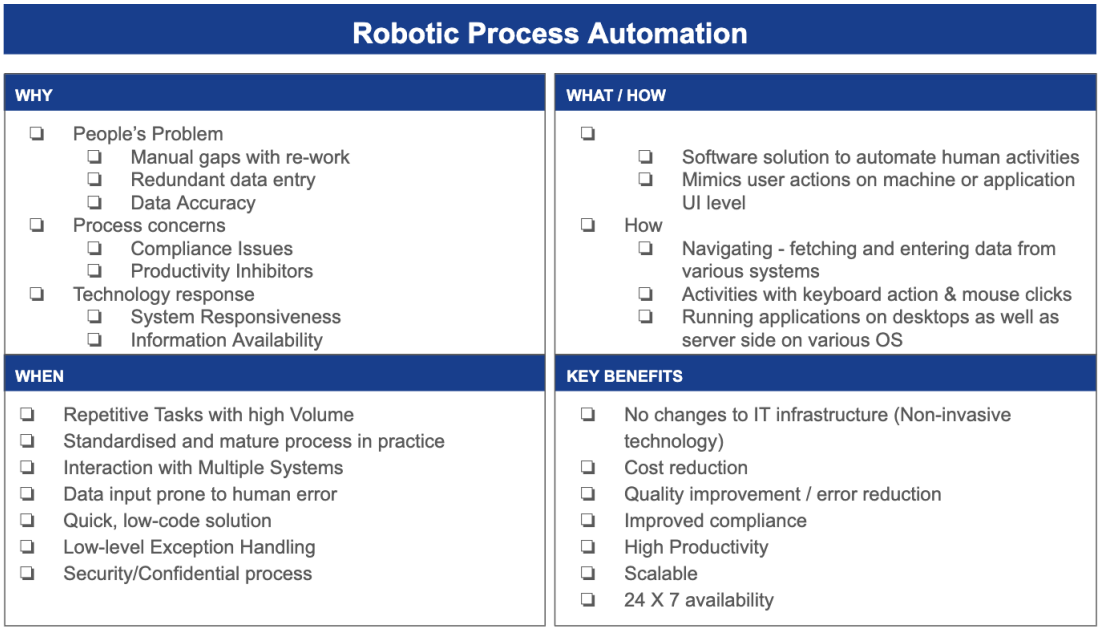

Figure 2.1: *RPA Overview compiled from RPA literature review [14, 30, 44, 52]*

It is also seen that RPA implementations are carried out when there are repetitive and high volume transactions in order to save the human resources for more creative tasks thereby increasing productivity and employee morale [44]. RPA involves a robot developed using screen scrapping concept, which records the user actions and mouse clicks spanning multiple systems, logging in wherever necessary [43]. Based on the task requiring manual decision making, 2 types of robots can be deployed namely; attended and unattended. Further, these robots are then scheduled and managed from a central point called orchestrator and can be assigned to various individual machines performing the tasks otherwise assigned to human resources.

RPA Benefits

Apart from quality improvement and accuracy, there are other benefits of RPA. As emphasized in the table above, the noticeable benefit of adopting RPA is cost-reduction due to the FTE savings accrued by the companies. In a particular energy company where RPA was implemented to find any anomaly within the digitized meter reading entered and the normal consumption range, the savings were considerable from 30 FTEs to 12 FTEs. The 12 FTEs were necessary to handle truly unusual cases were human intervention was required [30]. Not only were they able to reduce the cost in terms of money, the energy company was able to improve quality, consistency, and speed of problem resolutions with RPA initiative [30]. Moreover the most remarkable aspect of such automation is that RPA can be implemented without changing the existing IT landscape of the organization. Inadvertently, rapid implementation of RPA automation can be achieved as the RPA tools are

easy to use and designed for subject matter experts to configure the automation [52]. However, even so in practice RPA developers are involved in the development of these robots as other aspects such as security compliance and multiple system connectivity needs to be attributed in the automation. Ultimately, organization benefit from RPA as robots can work 24/7 non-stop, is a huge contributing factor to not only improving productivity but also reliability and continuation of service to the customer [44].

RPA Shortcomings

Even though RPA, as described above shows a sizeable number of benefits, many experts are still skeptical about what RPA actually delivers against its claims. One argument is that RPA is only automating the process and not truly transformational as it does not change existing processes. Often somewhat ironically experts remark about RPA that, "if used appropriately, RPA can be a very useful tool in a strategic transformation initiative. But if approached incorrectly, they can perpetuate legacy system problems" [46]. To name a few of prominent shortcomings are as follows:

- Change management can become a big problem if IT within the organization is not involved early on [30]. RPA though a non-invasive technology, needs to undergo change management process with changes to the respective target systems. Any change that might change the screen or steps performed on a business application implies a direct impact on the robots that actual execute those activities.

- One more shortcoming of RPA is that, software robots can't deliver worthwhile transformation on their own i.e they can only automate tasks and not re-design processes [46]. That is why, if organizations do not adopt a holistic view of processes, they might end up refining only at a very basic level without actually streamlining the overall operations.

- Lastly, interface integration with RPA is still inferior at the back-end machine-to-machine integration and incidentally, RPA then becomes a temporary arrangement to circumvent the manual process until the newly redesigned processes is up and running [4].

In addition, there are other RPA shortcomings that may lead to a negative outcome for a RPA adopter. These mostly originate within the organizations such as having unrealized expectations about RPA, automation islands within the business unit and therefore low maturity of RPA implementation, lack of sourcing skills and so on.

RPA and Business Process Management

According to RPA providers themselves, RPA technology is a tool in the whole game of business process automation. To understand process automation, it is imperative to first get familiar with the terms such as Business Process and Business Process Management.

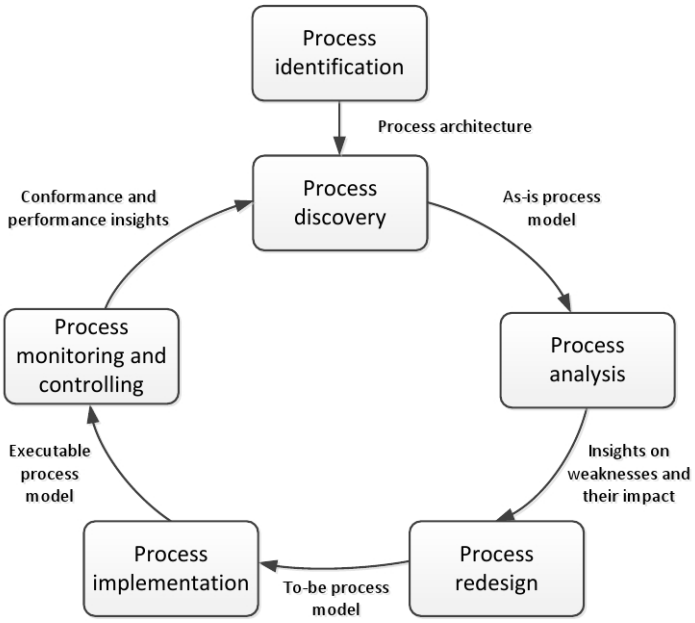

Figure 2.2: *BPM Lifecycle of Dumas et al [13]*

Business Process can be defined as activities encompassing a number of atomic events or activities performed by various actors collectively producing a specific outcome in an organization providing value to the end customer [13]. Business Process Management (BPM) as explained by Dumas et al is "The art and science of overseeing how work is performed in an organization to ensure consistent outcomes and to take advantage of improvement opportunities." [13]. Yet another definition of BPM mentions it as a "management discipline that integrates the strategy and goals of an organization with the expectations and needs of customers by focusing on end-to-end processes" [8]. To summarize, BPM can be regarded as a continuous process improvement cycle, consisting of phases - Process identification, Process analysis, Process redesign, Process implementation and process monitoring and controlling as highlighted by Dumas et al in their BPM lifecycle [13]

Using the BPM lifecycle shown in the above figure 2.2, one can come to a conclusion that BPM engages people as well IT systems to analyze, design, implement, control, and continuously improve end-to-end processes in order to ascertain process governance. This end-to-end process improvement may in turn lead to automation in order to obtain cost and time benefits due to workload reduction. However, it needs to be noted that automation is a subsequent phase of process improvement which will accentuate efficacy and not the other way around. As rightly said by Bill Gates, "The first rule in any technology used in a business is that automation applied to an efficient operation will magnify the efficiency. The second is that automation applied to an inefficient operation will magnify the inefficiency" [10].

Based on the conceptual understanding of RPA and BPM, one can conclude that though

both work towards the goal of business process improvement, they differ in their approach. BPM focuses more on building software to improve processes and establish process excellence by cost and time reduction thereby increasing productivity. BPM also aims to re-engineer processes by introducing new components or sub-processes to facilitate functioning of the business. On the other hand, RPA focuses more on using a robotic software to automate already existing processes. In addition to that, BPM takes into account the whole Business process to improve and be efficient whereas RPA considers automating single or couple of activities. Yet another differentiating factor in terms of their approach is that BPM is a typical "inside-out" approach entailing possible changes to the information systems whereas RPA is an "outside-in" substituting the manual labour by task automation [48].

2.3 RPA Process Lifecycle

RPA, though emerging as the common choice to automate processes, has no specific international standard of implementation. A typical RPA lifecycle begins with defining the process to automate, designing and developing the robot and finally testing and executing it in production. It is reasonable to consider RPA implementations as any other IT projects as most of the steps are similar [28].

Most of big RPA vendors like UiPath, Automation Anywhere, Blue Prism, NICE and Pegasystem, have their own set of RPA implementation methodologies. In addition, these RPA vendors present their own distinctive guidelines and development strategies to the organization.

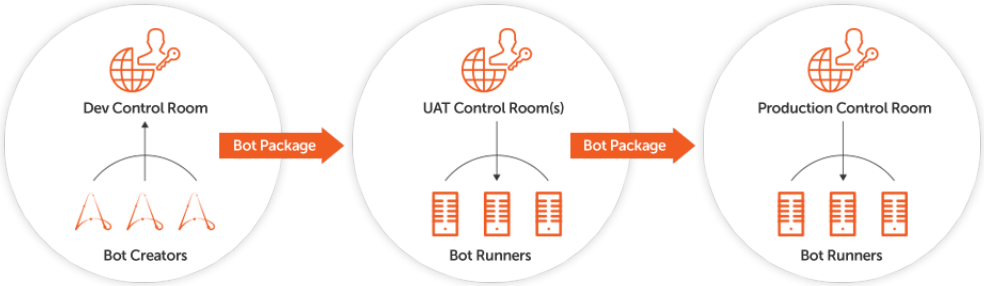

Figure 2.3: *"Bot Lifecycle Management" - Automation-Anywhere, [6]*

For example, Automation Anywhere has its own Bot Lifecycle Management (BLM) guideline to manage RPA as illustrated in the figure [2.3]. BLM is based on DevOps best practice providing a framework for continuous testing and deployment of bots in separate Software Development Lifecycle (SDLC) environments and lets bots transition between lifecycle stages smoothly as per the organization's policy even before they are released into production [6].

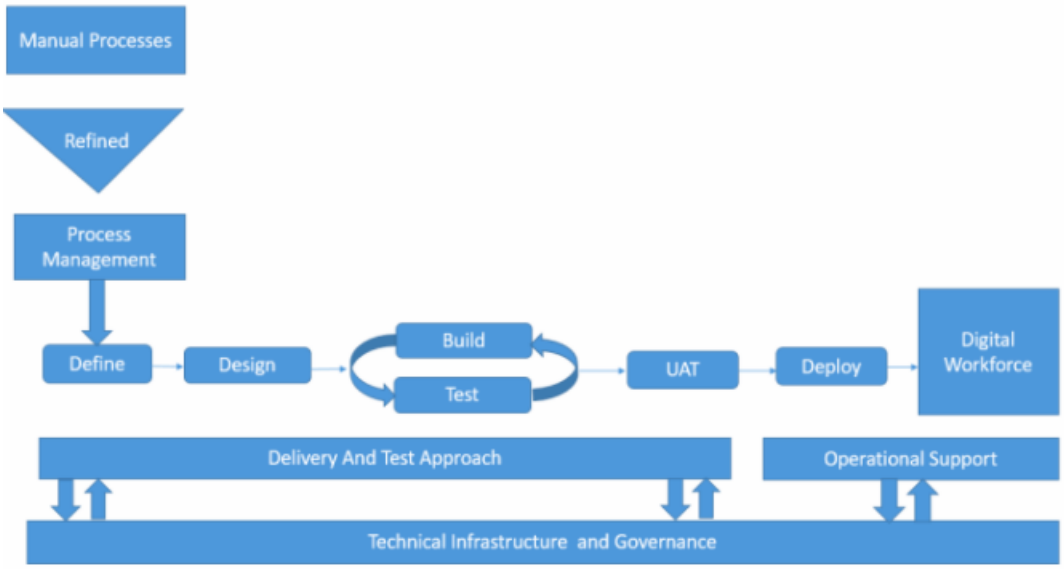

Figure 2.4: *"Delivery Methodology" - Blue Prism, [41]*

Yet another RPA vendor, Blue Prism offers its "Delivery Methodology" which presents the lifecycle from the starting phase of defining and refining manual processes followed by finalising the task to be automated and then completing it by usual SDLC phases. Furthermore, there are RPA implementation guidelines contributed by internationally operating consulting firms such as KPMG, Deliotte among others. KPMG has come up with its own RPA governance framework, which explains how organizations should adopt a long-term vision and strategy for RPA adoption across the organization, once the Proof Of Concept (POC) phase is completed successfully [25]. However, the overall RPA lifecycle involves many more phases and steps than just a project lifecycle like any other IT project.

To begin with, many companies would prefer to test the technical feasibility and financial viability of RPA software within the organization before introducing any new technology [29]. The next phase is to select a process for automation which is an important phase in the overall RPA lifecycle. Process selection is an activity in itself which will be discussed in detail in the following section. However, observations point that organizational failure in selecting the suitable processes for automation creates additional learning curve within the RPA lifecycle [31]. None of the above methodologies or framework provide end-to-end RPA lifecycle phases starting with POC, process selection up until the adoption of RPA as a regular software automation cycle in the business processes. The models presented by the RPA vendors begin with Robot designing for the selected manual process whereas the governance framework talks more about organization's RPA strategy and the various levers that help in executing that strategy.

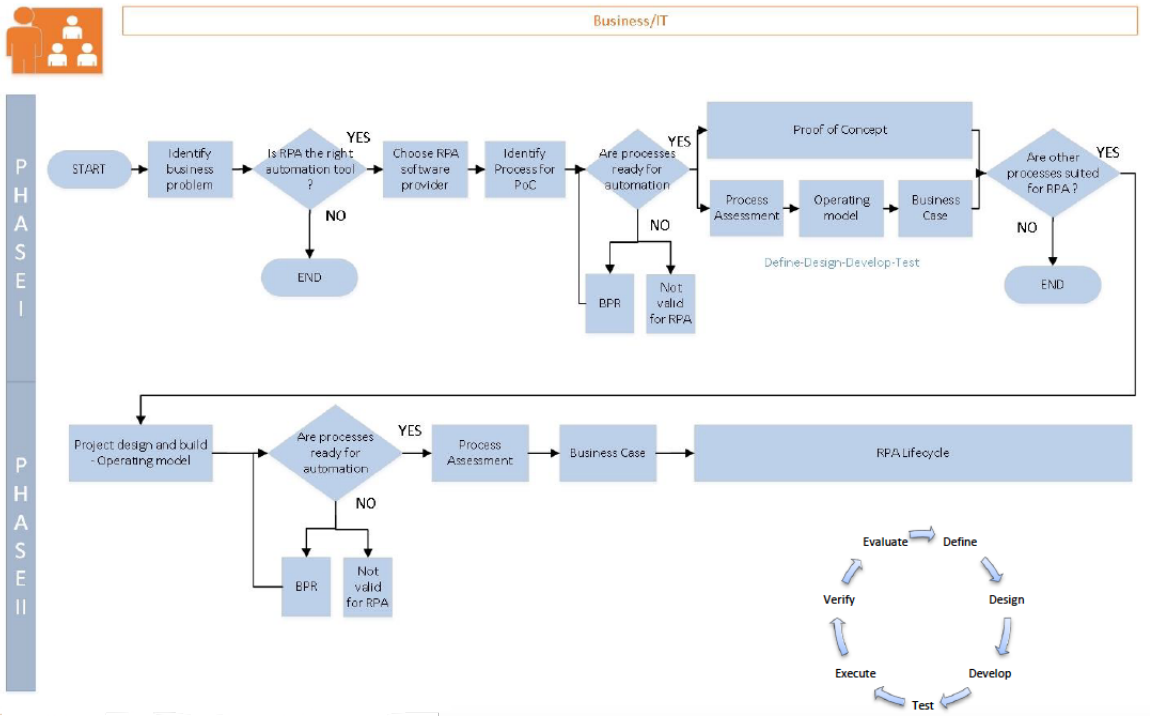

Figure 2.5: *RPA : Dynamic Roadmap for successful implementation, [43]*

To this end, Guðrún et al proposes a dynamic roadmap for successful RPA implementation which addresses the issues mentioned above. The dynamic roadmap is established on the basis of various case studies in diverse industries and the kind of implementation processes followed by those companies [43]. Therefore, the roadmap doesn't confine itself to a single RPA automation cycle instead provides step-by-step actions covering different phases in RPA, the organization undergoes in order to adopt RPA. It presents business problem identification, process assessment alongwith standardization and business process re-design (if necessary), executing a POC with a business case in the first phase [2.5]. The first phase of RPA helps the organizations to understand their technical compatibility of RPA software in the existing IT landscape as well as choose the right RPA vendor.

According to the roadmap, second phase begins with creating an operating model within the organization which will help in deployment and roll-out of process automation. Similar steps of process assessment, standardization and building a business case for RPA automation are performed for every automation project which runs in parallel to the existing development of the IT systems. To summarise, the dynamic RPA roadmap provides detailed RPA stages initiating right from identifying a business problem up until having an operating model to implement and manage software bots within the organization. In this thesis, the RPA Lifecycle of Guðrún et all will be taken as a point of departure for further improvement.

2.4 RPA Process Identification

According to the BPM pundits, process discovery can be done in 3 typical ways namely; Evidence-based, interview-based or workshop-based [13]. Dumas et al further argue over the strengths and limitations of each method of process discovery on the basis of objectivity, richness, time consumed and immediacy of feedback. These qualitative comparisons in turn imply that though evidence-based methods have upper-hand by highlighting the problems, they might not give solutions or direct feedback [13]. Having said that, process analyst and RPA enthusiast find it hard to choose the potential candidate for automation. Identifying a right process to automate leads to a good business case even then, whether a process or sub-process can or should be automated is not always apparent, causing in practice a considerable time investment [54]. Undoubtedly, process selection therefore plays a crucial role in successfully implementing RPA within organizations [43].

In the case of a utility company when RPA was in a nascent stage, RPA team began finding processes to automate, learned from their failures as initial choice of processes were based on certain unrealistic expectations and subsequently picked processes that didn't fall under automation criteria. The company learned through trial and error to better identify processes or sub-processes for automation [31]. Similarly past experiences of RPA implementation suggest that it is of paramount importance that business operations rather than IT lead the automation initiative as they are in a better position to select suitable automation processes as well as prioritize projects that will provide value to the customers [30].

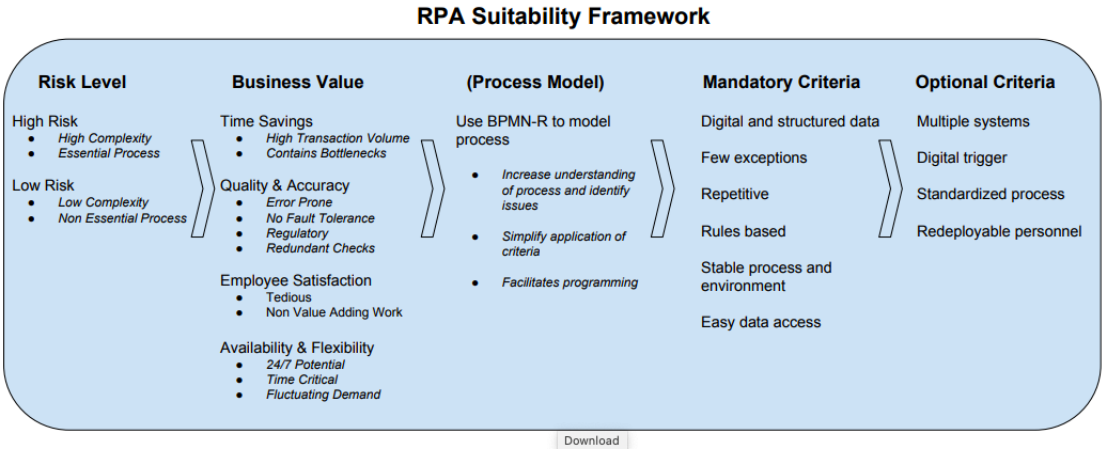

Figure 2.6: *Framework assessing Business Process suitability for RPA [2]*

But then again, process identification can be at various different levels within a process landscape. Topmost level being the end-to-end process such as Purchase-to-order scenario

or it can be a specific (pivot of the end-to-end process) or support processes and the lowest level can be a sub-process level or even further deep down containing individual tasks within the process [13]. Regardless of the circumstances, the organization should be able to identify the right process to automate be it at sub-process level or at a task level. Apart from a few exceptions, there are not enough scientific papers available outlining the selection for RPA processes in the current scientific literature scenario. The recent study by Agaton et al demonstrates a framework assessing business process suitability for RPA, in the figure 2.6 illustrated above [2].

This RPA suitability framework provides a step-by-step guide to assess the business process based on associated risk level and business value along with RPA based suitability criteria [2]. Though this paper accounts for various factors such as risk level, business value, process model as well as other mandatory and optional criterias in process selection, it still remains a manual assessment. A lot depends on how the analyst using this framework categorises processes and hence a different person analyzing the same process can come up with different responses. Recently, an extension to the above study was published presenting evidence based process discovery for the purpose of finding RPA suitable processes. This particular research proposes the use of Process Mining techniques to discover processes and further improve the reliability of the RPA suitability framework [18]. The research further goes on to suggests that by using evidence based insights to identify the potential process to automate can increases the chance of having a good business case and a successful RPA implementation within an organization.

Yet another equally recent scientific study, motivates and claims a full-scale use of process mining technique to assist businesses select suitable RPA candidates. This paper in particular, focuses on how process mining supports organizations throughout the lifecycle of RPA initiatives by providing process visualization, fact-based automation prioritization possibilities accompanied by automation monitoring [17]. The paper further ascertains that, use of event logs (also termed as digital traces) from business application is supported by data science rationale to discover processes to identify bottlenecks and recommend improvements if any. Finally, some automation experts argue that RPA does not necessarily operate on a process wide level, and it is important to discover a process on a task level [48]. In practice, discovering processes at task level is called as 'Task Mining' which will be elaborated further in section 2.6, while some present it as 'Desktop Mining' or 'Robotic Process Mining' [34, 35]. Therefore it is important to understand what type of process identification method an organization must adopt in order to utilise RPA. Following section elaborates further on automated process discovery using process mining and task mining technique.

2.5 Process Mining

The genesis of Process mining is a Dutch computer scientist, Wil Van der Aalst. According to Van der Aalst, Process Mining is defined as "A technique that uses event data to discover processes, check compliance, analyze bottlenecks, compare process variants, and suggest improvements." [49]. The scientists further adds that the process perspective of process mining can be a means to bridge the gap between data science and process science. Data science is an interdisciplinary field aiming to turn data into real value whereas the term process science combines understanding of information technology and management sciences to improve and run operational processes [49]. Process mining helps to uncover the discrepancies between the observed behaviour in reality i.e. the event data and the processes planned and modelled by the process architect. Apart from the replay or play-out as it is termed, of observed actions, process mining with the help of event data is replayed on the existing process model to analyze compliance and performance checking to monitor business processes [49]. In order to produce the real process model, process mining uses sequential event records of activities performed or the digital traces from different systems such as CRM, ERP or a billing system and produces a complete workflow from start to finish based on the input events. Likewise, there are 3 pillars of process mining or 3 different types aimed to gain some knowledge from the process logs being analyzed as highlighted by Van der Aalst [49]. Firstly, process discovery; A method that takes an digital event log and produces a model without using any a-priori information. Secondly, conformance where the present process model is compared with an event log of the same process and lastly, enhancements to optimise an existing process model using data collected from the actual process.

The need to provide such realistic process analysis instead of assuming that the underlying process works fine, instituted the development of various process mining tools. In order to bring this process mining theory into practice , there are many such process mining tools, that offer mining algorithms for businesses. These business oriented process mining tools present a visual reconstruction from a single activity to an aggregated global view of a business process based on the transactional logs provided as inputs [17]. The PM tools display a complete picture of processing time as well as chronology of the event as they were executed, allowing to trace process flows, get to know any setbacks as well as determine complexities. Process Mining is an innovative digital tools that provides holistic insights into actual processes complexities, thereby allowing to spot inefficiencies [40]. The benefits derived from process mining include process efficiencies, optimization of inventories, check compliance, Key Performance Indicator (KPI) monitoring as well as waste reduction inevitably transitioning the businesses towards operational efficiency.

2.5.1 Automated Process Discovery using Process mining

In business practice, process discovery is also known as Automated Process Discovery (APD) or Automated Business Process Discovery (ABPD) as it uses digital traces from various underlying information systems. Although in practice, the term Process Discovery is used interchangeably along with Process mining; in theory, Process discovery is the very basic and demanding task within process mining as it constructs a process behaviour model based on an event log within enterprise systems [49]. Automated process discovery presenting process analysis using the event logs has been widely studied in the process mining field and a significant subset of these algorithms focus on discovering process models from the control-flow perspective [5]. Therefore as mentioned earlier in the section, process discovery with its transparent view into process complexities is one way of getting to know the business processes within an organization.

Automated Process Discovery

Process discovery with its data visualization ability based on the event logs, presents real business process along with possible deviations. Deviations from a standard path cause delay to the overall throughput time, and can represent certain activities being carried out manually or re-work which essentially can have improvement potential [17]. Moreover, process mining with its integral capability to learn from input data, can facilitate in detecting delays or anomalies with respect to key success factors like efficiency, speed, agility and compliance within the processes [17]. The process discovery visualization and analysis can be used to complement the time-consuming interview process with the process experts to spot these deviations and analyze further to scout potential candidates for RPA [18]. Additionally, feasible business case for RPA can be built based on the supporting process discovery evidence. Using the mining analysis, RPA can therefore be used to automate processes and overcome delays or to adhere to process compliance, thereby optimizing the overall process.

2.6 Task Mining

It is common knowledge that in recent years, business processes management relies heavily on the information systems. Yet in reality, in order to complete the whole business process, there are activities which are not performed within the information systems. These activities can be time-consuming and user performance dependant leading to delay in completion of the overall process. Similarly, some automation experts argue that not all automation can be accomplished on a process wide level, therefore it is important to discover a process on a task level [48]. To address this gap and capture such activities, mining can be used to produce a process map with the help of user-interface recording

such as click-stream or keystroke captures [33]. The main goal of such process discovery is to capture user interactions that are not logged by information systems and obtain all possible process variations to derive a process model with a maximum level of detail [35].

Many commercial or open source mining tool developers also refer it as 'Task mining' and describe it as gathering useful information from low-level event data through the chronological actions performed using the workstation [7]. Unlike process mining, it is important to note that the final aim of Task mining is to obtain a list of activities executed that can be subject to automation, rather than coming up with complete end-to-end business process understanding [7]. To this end, in one of the recent papers, mining User Interface logs is termed as Robotic Process Mining (RPM) due to its association with RPA [34]. According to Leno et al, RPM is "A class of techniques and tools to analyze data collected during the execution of user-driven tasks in order to support the identification and assessment of candidate routines for automation and the discovery of routine specifications that can be executed by RPA bots." [34]. Some mining vendors like Kryon [27] address User-Interface based mining as automated Process discovery while others address it as Task mining as it mines individual tasks performed by the user. For the purpose of simplicity, this thesis will use 'Process mining' term for general mining of processes whereas 'Task mining' term will be used to address user-level mining activities.

2.6.1 Automated Process Discovery using Task mining

The primary goal of task mining as cited in [35] is to:

- Capture user actions that are not logged by information systems, that include mouse clicks, a document upload or editing, navigate a website or any other active desktop application. The main aspect of such recording is to capture the execution sequence of activities performed along with the timestamp and is referred to as a task trace. Furthermore, screenshots using the OCR technology are also captured to realize the functionality of the sequence of activities performed.

- The captured sequence data is consolidated and pre-processed further to obtain a complete set of process variations. The captured sequence is further transformed from different process instances by assigning transactional ids representing actions performed by different users. Similarly, the captured traces could have actions which have no impact in the final process therefore making it redundant. One example can be, pressing CTRL-C twice has the same impact even if performed once or performing actions not related to the expected task. To this end, consolidation and pre-processing encompasses noise reduction, removal of unnecessary data and create a logical structure by assigning transactional id.

- Finally using the process mining technique, process map is derived with maximum level of detail depending on the captured data. The user interactions recorded, represent different level of data from the normal event logs in the enterprise systems.

Therefore, the discovered process model also represents different detail levels within the process. The generated process map helps in understanding the deviations within the tasks performed.

Similar to process mining, the deviations found by task mining can be considered as possible automation candidates or an opportunity to standardise the activity. Standardisation of these activities can be achieved by reducing the variance across the users executing the same tasks [47]. Certain commercial task mining software provides process map based on captured data which can be further analysed to identify possible activity to automate whereas in other cases the mining tool goes a step further by creating an executable routine of the recorded activities for automation [34]. This routine is compiled and can then be included in the RPA script using specific command in the language of the target RPA tool [34].

2.6.2 Differences between Process mining and Task mining

	Desktop Activity Mining	Process Mining
Data Source	The user-application interaction during a user's process execution path. Additionally, the user as source of knowledge. Furthermore, information systems which log transactional data.	Information systems which log transactional data.
Data Structure	Uniform over one DAM implementation.	Potentially heterogeneous data structures through heterogeneous information systems.
Scope of Data	Complete list of user-application interactions and user defined steps that are part of the process.	Transactional data which is recorded by Information Systems
Goal	Automation, analysis, optimization, documentation	Analysis, optimization, conformance checking
Requirements	DAM Software. Optionally information systems which provide transaction logs.	Information systems which create and store fairly high quality event logs.
Detail Levels	Two dimensional data structure consisting of user defined process steps and user-application interaction. Additionally a third, optional layer, the transaction log layer.	One-dimensional level of activities.

Figure 2.7: *Comparison between Task mining and Process mining from [35]*

Having described both process as well as task mining concepts and their execution, we now move onto highlighting the differences between the two. Figure 2.7, highlights further differences between process mining and Desktop Activity mining, however it is commonly know as task mining in practice.

The biggest difference between the process and task mining is their scope. Process mining scope is much broader, typically involving the whole business process than task mining which focuses more on lowest level activities performed within the business process. Process mining captures and analyzes event logs from the enterprise systems, providing higher level of abstraction. On the other hand, task mining enables to capture the data at the user level which are not captured by the transactional systems and more likely to involve manual activities [34]. Difference in the data structure i.e. event logs lack data uniformity or incompleteness of data as the system was designed to have data in a specific manner, may have issues in further process analysis. However, the user interaction data is recorded specifically with automation as the purpose and hence all the relevant actions are captured subject to the recording capability of the tool [35]. Lastly, Process mining provides a top-down approach to understand the business processes whereas task mining provides a bottom-up approach to realize the activities executed within the system.

Based on the aforementioned capabilities and features of individual mining techniques, many tools are introduced in the market offering stand-alone mining as well as in certain cases, both task and process mining. Task mining with automation as its primary objective, in most cases is incorporated as part of automation tool by the RPA vendors. Product such as Kyron and edgeverve have their own automated process discovery mainly task mining integrated alongwith the RPA software [15, 27].

2.7 Literature Review - Conclusion

The general understanding of RPA provides insights into this technology; its benefits as well as shortcomings. Given the RPA advantages, it can be considered as a value-driven automation within business process management. However, implementation of RPA in the organization involves different phases starting with RPA tool selection, process assessment, performing the POC and finally adopting RPA for business processes automation. Of all the phases involved in RPA implementation, we draw attention to the fact that process selection phase is important and can be an integral part of overall RPA implementation. Firstly, the literature review points out that organizations mainly follow manual (conducting meeting with process experts) or bespoke process selection using RPA specific selection criterias. Since process selection in RPA context can be a deciding factor or value generation to the organization, it becomes apparent that a scientific approach of process selection is essential to improve RPA implementation.

Secondly, RPA is described as an automation tool that yields quick wins but for it be effective, transparency regarding overall processes and understanding the interactions between the processes is vital. Therefore, it becomes significantly important to have a complete understanding of the process which forms the basis of selecting the right process to automate. Implementing RPA on a wrong process would not provide value to the organization.

To this end, few methods exists in the scientific literature over the automation process selection and they have been argued in the aforementioned section. However, two out of the three papers studied, point towards using the mining technique to identify processes. Both the papers in favour of using mining technique, present mining as an automated way to discover processes to automate. But one of them points out using mining as an additional reliability check for the chosen process along with the normal interview-based process selection [18].

Drawing conclusion from the above arguments, guiding in process selection becomes one of the advantages of using process mining out of the other broad benefits in terms of process optimization, inventory management and waste reduction to gain productivity. Going a step further, the literature also describes two approaches of automated process discovery using process mining and task mining. Despite the fact that task mining is considered as an extension or complementary process mining [49], there are considerable differences in terms of scope and data collected for process analysis. Though both are mining techniques, the main principle of PM is 'As-is' process analysis, optimization and conformance check with the expected model, while task mining goal is to achieve workforce optimization and bench-marking of workflow as well as standardisation by automation. However, using mining technique to discover processes may prove to be a pivotal stage in the overall RPA implementation. Having motivated the use of automated process discovery, this research will continue further based on the above conclusion.

Chapter 3

Research Methodology

This chapter describes and provides rationale for the methods, procedures and approaches employed in this study. The section begins by understanding what is meant by 'Qualitative Research', and why we adopt such methodology. Building upon the selection of research methodology, we focus on integrated approach of the research methodology. Further down the section, it argues about the validity of the methods used, together with the measures that were undertaken to strengthen validity.

This study makes use of qualitative research. It is common knowledge that qualitative research findings are not based on statistical procedures or other quantitative methods. The distinctive features of any qualitative studies is to provide a contextualized perspective and understanding of human experiences contributing new insights or formulating new concepts by intensive study of multiple sources [56]. In addition, even though qualitative research is a subjective assessment of attitudes, opinions and behaviour Kothari points out how interpretative nature in qualitative research leads to a frequent desire to develop hypotheses rather than to generalise to larger populations [26]. The objective of this research is to investigate and understand whether automated process discovery improves automation implementation within an organization as well as the challenges encountered by the organizations to implement such process. Firstly, opinions and the background as to how and why the organization followed certain steps in RPA implementation will help to understand the context of the events unfolded within the organization. Secondly, to get a broader perspective of PM and RPA inter-relation and to establish clear description of impediments faced by the firms, it is necessary to interact with the experts working within the firm and gain knowledge about their experiences, attitudes and beliefs. All the above reasons and the uniqueness of captured data for a particular organization which cannot be easily gathered through surveys and closed questions led to choosing qualitative research for this study.

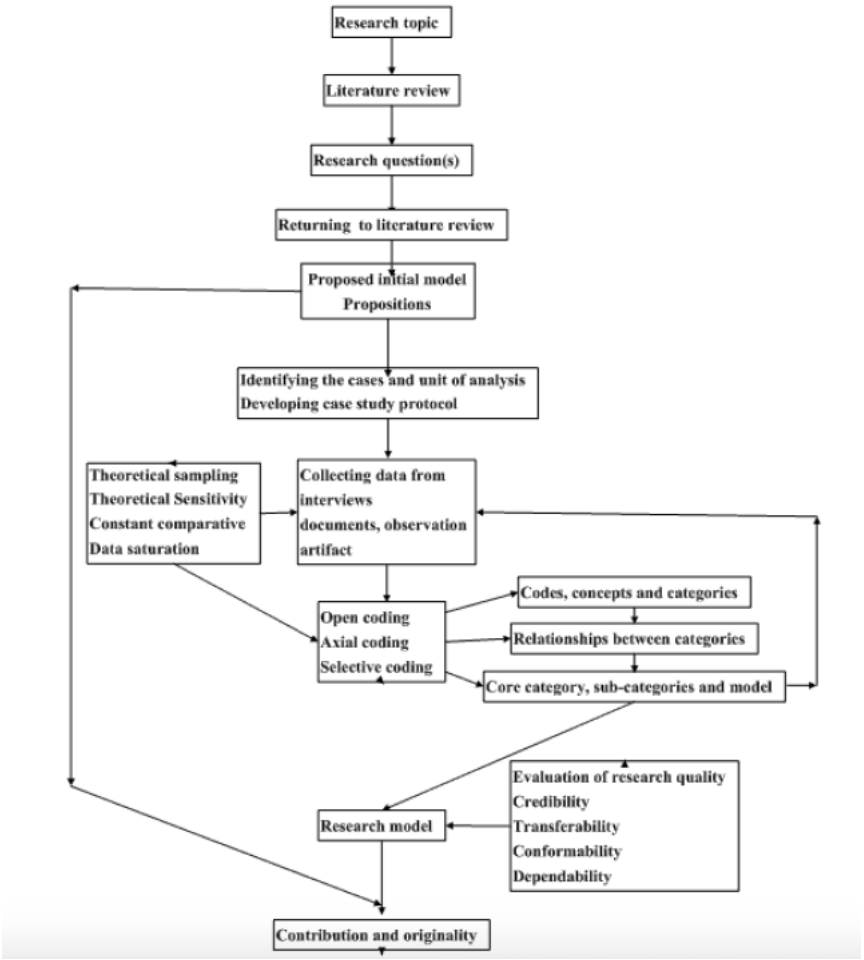

Figure 3.1: *Case Study - Grounded Theory Methodology [20]*

To understand and find solution to these research questions, we adopt an integrated research design methodology using case study review and grounded theory. As mentioned by Andrade, case study research and grounded theory complement each other and jointly they serve the purpose of theory building by interpretive researchers [3]. Other similar finding suggest that grounded theory (as a method) can be combined with case study method to build upon a compatible or an integrated research methodology within Information System research [20]. As illustrated in the above figure [3.1], Halaweh et al outlines the steps starting with case study selection and data analysis followed by data collection through interviews. Next step is to perform coding and categorizing from the theoretical sampling done on the data collected from case study as well as interviews, to arrive at a concept or theory building.

To this end, this study uses empirical case studies in the last couple of years to assess the impact of PM on automation initiatives. Inductive reasoning will be applied to the case

studies considered as part of the data analysis to arrive at a generalized understanding. Exploratory study will be performed by conducting interviews within an organization, to understand various stages of RPA implementation and the steps followed to select suitable RPA initiative. Correspondingly, to respond to the other research questions, we will engage in conducting expert interviews of process mining and RPA software vendors as part of this study. The main objective to do so is to look for broad patterns and interpret the interview data to arrive at theories and answers to the remaining research question. Grounded theory will be applied to these interviews as well as the data collected from various artefacts. In this regard, empirical case study approach is used to give validity and strengthen the objective of incorporating automated process discovery in the overall RPA lifecycle stages. While on the other hand, the exploratory research based on the interviews aims primarily to understand the overall RPA lifecycle and process assessment method followed in the organization. Secondly, they were used to interpret the underlying process mining challenges faced by the organization. All reasons mentioned above, justify the need for both case study and grounded theory research methodology integrated under the umbrella of qualitative research [20].

3.1 Exploratory Approach

In general under qualitative research, and especially involving case-study, use of multiple sources of data is stressed significantly [56]. Similarly, exploratory approach depends mainly on literature search, expert interviews with open-ended questions and records from fields notes [11]. The objective of conducting exploratory research for this study is multilateral; firstly to gain information on various stages of RPA lifecycle and the steps followed to select suitable RPA initiative. Secondly, to understand the barriers and bottlenecks for implementing mining within the organizations. Lastly, to comprehend mining approaches organizations can adopt based on their situation. Interviews can be one such source of information upon which rigorous process of ”explanation building”; a salient feature of Grounded theory can be applied [55]. To this end, the study will conduct a series of interviews to gather qualitative data and further analyze to find broad patterns in their responses. The interpretations based on these interviews can then be generalised to a larger population as challenges faced by the case-company would not be largely different.

3.2 Case Study Approach

Building theory or a concept using a case study, is a unique inductive process starting by specifying a research question, performing within-case analysis and replication logic to arrive at a generalized theory [16]. Paré et al, build further upon Eisenhardt’s case study approach presents a positivist view of research and asserts the use of empirical data and cases study to build theories incrementally, especially in the research study of IT imple-

mentation [39]. A substantial body of empirical case studies with a base theme of process mining implementations followed by process improvements using RPA, are available within the literature. These case studies will be analyzed objectively to understand to what extent has the use of process mining truly support process selection and accentuates RPA implementation.

Using integrated research methodology, we can consider the possibility of applying the existing case study findings to resolve the concern of identifying suitable processes for automation and improve the RPA lifecycle. As a result, this study will combine the empirical case study analysis and exploratory analysis gathered via interviews to provide answers to the research questions.

Chapter 4

RPA Stakeholder Interviews

This chapter elaborates over the data collected from the stakeholders involved in the RPA implementation at a large energy company. As mentioned in the earlier chapters, the purpose of conducting exploratory research is to understand stages of RPA lifecycle and the steps followed to select suitable RPA initiative as well as get to know the barriers and bottle-necks for implementing mining within the organizations. The analysis and discussion of the semi-structured interviews help us to answer the main research question and RQ1. Similarly, semi-structured interviews with the experts from Process mining and RPA software vendors allows us to find answers to RQ2.

4.1 Energy company selection

In order to achieve the purpose mentioned in the introduction of this chapter, an energy utility company is chosen to understand the stages of RPA implementation it follows. This organization is one of Europe's largest producers and retailers of electricity and heat operating in 5 main countries with approximate employee strength of 20,000. The organization's core business is power production and energy trading, but customer offerings include heat and electricity. The company has adopted RPA since 2018 and slowly acquired certain maturity level in RPA implementations. Over a period, the company has setup an RPA Centre of Excellence (CoE) which is a common RPA platform in the organization. The RPA CoE provides a common RPA infrastructure in terms of IT resources including server space, RPA developers, analysts and other best practices to execute RPA implementation. Within the organization, there are 6 Business Units based on their primary business function and they operate independently. Each unit functions autonomously and hence are independent to implement RPA within their unit and join the RPA CoE platform anythime. 3 units have already embarked onto their journey of digital transformation using RPA, but are at different stages of RPA lifecycle.

Given that the energy company is involved in RPA since 2018 with currently 13 robots

in production within 3 business units and more being developed as well as the RPA CoE platform formed in the organization providing a governance model, has led to achieving certain RPA maturity level. The reason to select this organization multilateral; firstly, the organization has adopted RPA over two years, progressing well into the RPA lifecycle and therefore is suitable for the scope of this research. Secondly, the organization has its own RPA CoE team which will help us to collect relevant data about the RPA stages followed within the company and the challenges faced. Lastly, with a respectable RPA maturity level of this organization, there is an opportunity to interview RPA experts as well as people in the RPA management position to get insights into their working model of RPA. All the above mentioned explanations, justify the choice of this organization for the interview.

4.2 Interview

Qualitative research in general, involves studying multiple data sources where qualitative interview is one of the most common and important data gathering tools [56]. Expert interviews are conducted from two different sources. First set of interviews is conducted within an organization who has adopted RPA to automate business process and will be the focus of this chapter. Second set of interviews is with process mining and RPA software vendor and will be discussed in the automated process discovery chapter 6.

The intrinsic goal of doing these 2 separate types of interviews was to get answers and explanations to the research questions. The first set of interviews represent the experts associated with actual RPA implementation in different roles within a single organization. The latter interviews are of professionals from the RPA and PM software vendor and were in a customer-facing role for the software implementations. Selecting at least 1 stand-a-alone PM software company and other with an integrated (i.e mining as well as RPA software provider) software solution was the rationale behind the 2 alternatives. The criteria to choose a specific software companies such as UiPath was based due to the fact that UiPath was a chosen RPA software in the energy utility organization and therefore the researcher could reach out to those experts swiftly. Celonis being a market leader in process mining technique, was an ideal choice to make for this study.

4.2.1 Interview Selection

One has to rightly determine a strategy of interview selection in a qualitative research. A purposeful sampling in interview selection process will intentionally sample a group of people that can provide information in depth about the research problem under investigation [12]. On the other hand, sampling strategy for a phenomenological study i.e. finding people who have experienced the phenomenon leads to rewarding data collection [12]. However the downside of such strategy is, researcher may interview only people of high status thereby disregard broader understanding of the situation. This is termed as 'Elite bias' by Myers et al [37]. Other pitfalls highlighted by the author are lack of trust or lack

Business Unit	Code Name	Function & knowledge	Date/Time
BA Heat	Person A	RPA Programme Manager (NL- Management)	10 March 2020, 60 mins
BA Heat	Person B	RPA Programme Manager (Germany - Management)	3 April 2020, 60 mins
BA Heat	Person C	Process Owner (Process Expert)	28 March 2020, 45 mins
BA Heat	Person D	Process Owner (Process Expert)	30 March 2020, 45 mins
Customer & Solutions (B2B)	Person E	RPA Lead (Process & RPA Technical Knowledge)	5 April 2020, 60 mins
Customer & Solutions (B2C)	Person F	RPA Lead (Process & RPA Technical Knowledge)	5 April 2020, 60 mins
IT	Person G	IT Project Manager (RPA Technical Knowledge)	5 April 2020, 60 mins
UiPath	Person H	RPA- Task Mining (NL - Customer Success Manager)	6 April 2020, 90 mins
Celonis	Person I	Process Mining (Germany - Customer Solution Engineer)	13 April 2020, 60 mins

Table 4.1: *Overview of participants*

of constructing knowledge. This study specifically follows a purposeful sampling strategy by selecting interviewees who are actually part of RPA implementation within the organization and would be able to provide most insightful information. Moreover, we adopt maximum variation in interview selection process as the goal is to understand their experiences and expectation of the organization regarding RPA implementations. To this end, by interviewing a diverse group of people who have performed or are associated with the RPA initiatives, increases perspectives and gives a better picture. Having said that, the selection process is confined in terms of availability of people for interview, time and author's restricted knowledge about the organization.

Above table provides the details of all the interviews conducted in this study 4.1. We

conducted 7 interviews with experts from different business units within the energy organization as well as in different roles related to process management and RPA. Additionally, the last 2 interviews fall in the second category and those involve the professionals from RPA and PM software suppliers. All the interviews except 1 were conducted using online meeting tools such as Skype. A single interview was conducted face-to-face as it happened before the COVID-19 pandemic situation. All the interviewees were based in the Netherlands except 2, where the experts were from Germany. To maintain the privacy of the interviewee, their personal information has been kept anonymous by assigning code names.

4.2.2 Interview Protocol

As mentioned in the introduction above, the interviews were conducted in a semi-structured manner, with pre-determined group of questions related to a the theme. Inspite of having a structure of questions prepared, semi-structured interviews may have additional inquiries to explore the research question and objectives pertaining to a particular situation within the organization [45]. In the same vein, adopting a semi-structured interview technique, the researcher has more freedom to perform the role of an explorer rather than an expert [11].

Hence, the questions were grouped in a particular sequence in order to probe the interviewees to open up the discussion and give detailed information about the implementation and management of software robots. In contrast, Creswell et al points out that interview protocol should have questions narrowing to the central question and the sub-questions under examination [11]. During the actual interview, the real order or the flow of the interview was changed for different interviewees depending on their responses. However, the primary objective of the interviews to discuss RPA implementation and identifying suitable process for automation with the experts in their field was safeguarded. Lastly, semi-structured interview can be exploratory, explanatory as well as an evaluative, as per the study by Thornhill et al [45]. The interviews conducted in this study were *Explanatory* in the sense, interviewees were asked to explain the RPA implementation stages followed within their organization. *Exploratory* in terms of investigative questions about the challenges faced to adopt Process mining. Finally, *Evaluative* in nature as we attempt to evaluate the problem of process assessment for RPA faced by the organization.

Considerable factors need to be accounted for before conducting interviews such as scheduling interviews, giving some outline about possible questions, by allowing participants to reflect on the theme thereby provide a fuller account and increased information accuracy [45]. Following the procedure, we were introduced officially to the interviewees in person within the organization and in some cases via email using the corporate email address briefing them about the research conducted.

After the preliminary introduction, we reached out to participants via email to schedule an interview along with an agreement to record the interview. Consequently, an email in-

cluding general theme and direction of the interview question, along with an attachment of RPA successful implementation road map [2.5] was included for their referral. The foundation of every interview was set with initial introduction of either parties followed by explaining the purpose of the research. Introduction was preceded with different types of questions categorized as RPA phases, RPA implementation and post-challenges, overall RPA lifecycle and understanding of process mining and its challenges. Lastly, All interviews conducted were recorded followed by transcription promptly after the conclusion of the meeting. Any observations penned during the interview were also reviewed and aggregated along the transcription.

4.3 Data Analysis and Presentation

In the field of qualitative research, data is collected through multiple sources followed by its analysis and therefore data analysis becomes continuous as well as an iterative process happening simultaneously. Unquestionably, the researchers in general observe and analyze patterns through the entire data collection phase [11]. This data analysis can range from factual to conceptual or for that matter can also span from a straightforward descriptive account to using grounded theory with an outcome of insightful answers or building a theory [42]. To perform data analysis for this research, an inductive approach i.e. infer conclusions from the collected data - "let the data speak for itself" is followed. Grounded theory approach is implemented by interpreting and understanding the broad patterns visible in the data collected through interviews.

Halaweh et al further demonstrates that in order to implement grounded theory approach to theory building, coding is key activity and involves 3 steps namely; *'Open Coding'*, *'Axial Coding'* and *'Selective Coding'* [20]. The resulting process after coding will look as mentioned in the figure below:

Codes $\Rightarrow$ Concept(s) $\Rightarrow$ Categories $\Rightarrow$ Model(Theories)

In accordance with the above method by Halaweh et al, similar process of coding was followed in this study.

1. ***Open Coding*** was performed by assigning codes with focus on main ideas and expressions within the sentences of the interview transcription. Codes with similar properties were singled out to be grouped together under the heading of concepts to give it a meaningful form. Finally, the concepts were collected together and classified under different categories which usually represent logical groups within the data. The resulting categories resembled the main topics of research aiming to capture the information and insights provided in the interviews.

2. ***Axial Coding*** was accomplished with the intention of systematically comparing various categories and concepts to find relationships or contrasts in the collected data. As a

result, according to the 'paradigm model', varied casual relation, action-interaction, consequences or contextual relationship were found among the categories. These correlations and interdependence between the categories, assisted in arriving at findings from these interviews namely; "Use of manual Process Assessment", "Initiate POC for RPA" or "Create an Operating model for RPA implementation"

3. ***Selective Coding*** carried out the process of integrating the relationships found in the second step and refining them further. Finally, the "whole picture" was arrived by using 'paradigm model' which involves the concepts appearing repeatedly in the collected data form various interviews. For example, the findings mentioned above shed light on the central theme of RPA lifecycle stages adopted by the company. Use of manual process assessment leading to lack of prioritization and increased assessment time seems to appear number of times.

In order to explain the above mentioned process applied in this study, we present a working example of how a particular category or in our case finding was arrived at from the raw data of an interview transcript.

Raw data: *"As a process expert, we would present the process and explain the as-is process and the activities that we do manually. As far as I know, the decision was purely made from a questionnaire from the RPA manager. He had a few questions about the process as well as finance. What are the savings in terms of time ? How many people are involved in these activities etc? Then, it is RPA specialists' decision, whether we can automate these activities or not."*

Code derived: *Questionnaire-based assessment (Manual selection)*

Code derived: *Annual man hours of saving (RPA selection criteria)*

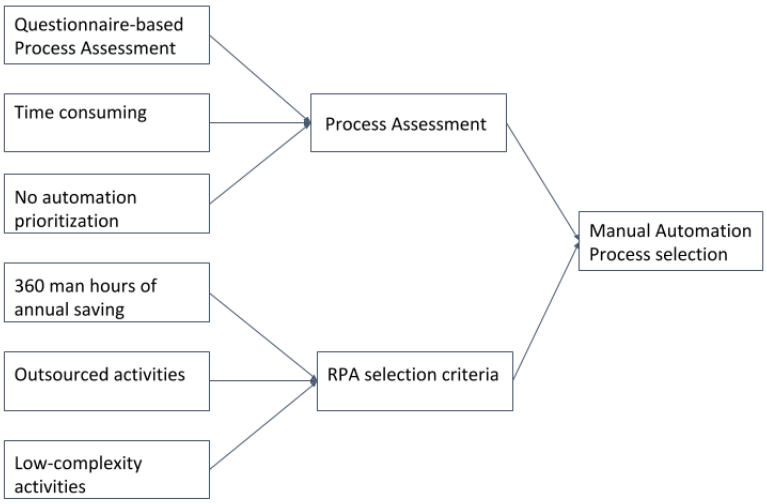

Figure 4.1: *Category arrived from codes and concepts*

In the example above, 'Questionnaire-based assessment' is the code derived from the interview transcript which highlights the process assessment method performed in the

organization. On the other hand 'Annual man hours of saving' code is derived from the mention of savings in terms of time and people. Similar coding is performed on all the interviewees and compared with the other codes within the group. Following open coding method, all the codes with similar properties are then grouped to produce higher level of abstraction called *concepts*. The codes used in the example above are grouped under the concept of *'Process Assessment'* and *'RPA selection criteria'* respectively. The figure 4.1 illustrates how codes like *'low-complexity'* and *'outsourced activities'* from other interviewees are also grouped under the same concept of *'RPA selection criteria'*.

Categories	Frequency of mentions
RPA initiative introduced as part of Digitization Strategy	7/7
Implemented RPA POC to test the concept within the organization	7/7
Built a Centre of Excellence platform to establish a common operating model	5/7
Process standardization is a prerequisite prior to an automation initiative	5/7
Manual Process Assessment	7/7
Automation prioritization	5/7
Time Consumed	3/7
Unable to progress further	3/7
Unclear Purpose	4/7
Data completeness	4/7
Data preparation	4/7
Data privacy	3/7
Lack of funding	5/7
Overwhelmed by the transparency	3/7
Resourcing and knowledge management	3/7

Figure 4.2: *Categories built on the basis of coding of interviews and frequency of mentions*

Finally, comparison of the concepts found across the interviews are then logically classified further to produce a third level of abstraction called *categories*. *'Manual automation process selection'* category signifies the method and criteria used for RPA process selection in the organization. All the codes that point us in that direction are therefore grouped under this category. Similarly, other categories are identified by repeating the process explained above and they are presented as findings from the RPA stakeholder interviews. *'Axial coding'* was applied on the derived categories and build correlation and interde-

pendence between them. *'RPA as digitization strategy'*, *'Implemented RPA POC'* or *'Built an operating model'* all these categories and their inter-relation supported in building the RPA lifecycle followed in the energy company.

Subsequently, not all the codes identified in the first round are selected or classified further. Even though these codes highlight an important aspect, they might not contribute towards the concepts created or was pointed out by a singular person. For example, the identified code *'low RPA familiarity'* was discarded as it was mentioned only by one of the process expert and therefore of less significance in actual stages of the overall RPA lifecycle.

Lastly, following the method of selective coding mentioned by Halaweh et al, we could arrive at the RPA lifecycle followed in the energy company using the distinct categories established on how often these appeared in the interview transcript. In addition with careful interpretation and coding the answers provided by the stakeholders also assisted us in re-creating the problem of manual and bespoke process assessment performed in the organization. The table 4.2 displays the main categories identified and its relative importance.

There are indeed many ways to present the data analysis and thereby the findings as pointed out by Saldana, "selecting the most appropriate representational and presentational modes that will best describe and persuade for your readership" [42]. The strategy selected for this thesis is to put forward the interviewees' views and opinions in interspersing quoted passages and then further describing and expanding the contexts in which the opinions were expressed. Furthermore, any kind of disagreements or conflicting views displayed are highlighted as well.

4.4 Findings and Discussion

This section presents the outcome from the data collected and analyzed through the interviews and contextualized further to arrive at findings. The individual findings depict the final categorization and the inter-relations, arrived after performing selective coding. The initial findings depict a welcoming picture for RPA technological scenario but it is hard to generalize the outcome throughout the organization. In line with the decentralised tradition of the organization, the individual units have independence over implementing new technologies and inevitably the maturity level of RPA adoption differs drastically. One unit is way beyond a successful RPA implementation whereas the other unit has just begun its POC stage while there are units where RPA has not yet begun.

The main finding is, that the organization has achieved basic maturity level by building a common RPA platform and adopting initial POC phase followed by further automation implementation projects. The other findings arrived from the coding performed on interviews are as follows:

1. RPA initiative introduced as part of Digitization Strategy

RPA was introduced in this organization as part of strategic initiative for automation

Version of August 14, 2020– Created August 14, 2020 - 15:32

and was acknowledged by all the participants in the interview. The principal driver for initiating emerging technologies like RPA was to fulfill the digitalization agenda. Similarly, considerable attention in the media as well as the global strategy consultancies spreading the importance and prospects of these new technologies further made them more keen to adopt these technologies in their work. Consequently, awareness and curiosity about automation and that these new technologies like RPA could be of value to the company led to begin the initiative.

"All this [RPA initiative] started because there was a specific request as part of the digitalization strategy from the top management to come forward with some activities within our business process for an automation Project. So adopted RPA as part of Digital strategy and transformation for efficient processes." - Person D

2. Implemented RPA POC to test the concept within the organization

All the participants unanimously said that the organization began the journey by automating small and less important process as part of the trial. Like any other new technology, the organization was hesitant to implement it straight away. With RPA's increasing acceptance as a technology on one hand and no existing best practices on the other, the organization initiated the POC for RPA, to understand the technical viability within the IT landscape. Robotizing a small process also helped them to realise its potential within their business process.

"Proof-of-Concept was a learning curve where we selected small and easy process with just three and half hours of savings every month to start with. The robot had to simply read this [meter reading] excel data from a specific location, extract the data and store the data into our database" - Person C

3. Built a Centre of Excellence platform to establish a common operating model

After the success of initial process automation, the organization set up a RPA CoE; a common platform across the organization whereby any business unit can join the platform with an upfront initial charges. Five participants out of seven acknowledged the formation of RPA CoE with two participants actually involved in the initiation of such platform. The immediate advantage of such CoE is to have shared RPA licenses and skilled resources as well as a common infrastructure to maintain the costs. Even then, the corresponding business units had full autonomy and control over their own decision making about the prioritization and maintenance of the robot to the development and delivery of robot in line with their decentralized organizational structure.

"Our organization has a RPA competence center, with development outsourced. This platform unifies technical aspects, code quality, security standards for RPA." - Person A

"We are currently in the moving out of POC stage. We now have the budget to develop certain RPA projects, but to join the RPA platform, you need more budget. Currently, there is a bigger process which will go live and then we can really say that we are out of the POC phase." - Person F

Taking into account that the organization is a large international energy utility company, it involves a great deal of governance and organizational complexities to manage such projects in silos or individual business unit. A common IT function catering to the needs of all decentralized business units makes it more difficult to manage the automation projects. Therefore, the organization benefits in having a Centre of Excellence platform where all the automation projects converge and have shared best practices.

4. Process standardization is a pre-requisite prior to an automation initiative

During the interviews, process standardization came out strongly as 5 out of 7 interviewees from the organization voiced standardization to be of utmost priority. Standardization of the activities within the processes produces reliability and accuracy. In a particular case in business unit Heat, they wanted to implement RPA to improve data accuracy entered into the SAP system. But the process expert realised that each site engineer was following his/her own way of entering data creating discrepancy in the system. However for the process to be automated, it expected standardized data and therefore the process expert suggested to first standardise the data and then go ahead with the automation initiative. To do so, the engineers had to all agree on the process of data entered into the system.

"We had to make small changes in how the 30 engineers collect their data. Previously, the Engineers used to collect data in a certain format of their own but for the RPA implementation, the engineers share the format which was used to collect the data and now they have one standardized way of collecting. This has made our process more accurate." - Person B

One of the interviewees came forward that as a process manager, he doesn't consider automation as the very first option to make the process efficient. Prior to automation, one should consider standardising the process or make required changes in the target system to make the process efficient. Outsourcing the activity or automating it further can be the next best option to gain operational benefits.

"Standardise the process, Make changes in the target system to make the process right, or outsource the activity or robotise the activity. With these 4 options, I see an opportunity to make the process more efficient. " - Person E

5. Process Assessment & RPA selection Criteria is manual

The very first stage of the RPA initiative begins with *Process Assessment* phase to identify if RPA is the right solution for the process inefficiency. To perform process assessment, the process owner fills an intake form with 15 questions in order to identify the process fitness vis-a-vis the expected RPA effort required to achieve this automation. Process fitness typically signifies RPA criteria's such as structured input data, transactional process, digital data, accuracy level and last but not the least any process change expected in the immediate future. On the other hand, expected RPA effort indicates the number of systems the process spans, is the process standardized, level of exception rules expected and lastly the lead time and time consumed to perform the action manually. The answers provided to the intake questionnaire provides RPA fitness and effort complexity level which is further considered for process standardisation and it's financially viability. Taking cognizance of the manual process assessment in the organization under study, following observations can be formed:

- ***Manual Process Assessment*** to identify the potential candidate for automation is manual merely backed by the person filling the form. All the participants mentioned that questionnaire-based process assessment was the norm to identify eligible RPA candidates. The person filling the intake form focuses specifically on that particular manual activity rather than the process as a whole which is then verified if financially feasible. Evidently, the process assessment depends on the robustness of the questions formulated in the intake form. Secondly, form filling becomes person dependent and not supported by the process data.

  "Process owners with the help of intake questionnaire decides potential RPA candidates and then the business owner decides if there is an attractive business case. Business is driver for RPA." - Person A

- ***Lack of automation prioritization*** due to automation initiatives not supported by systematic process assessment. The choice of the process to assess is done by process owner based on his/her personal experience. Moreover it is seen that in certain situations in order to create the automation initiatives, process owners are asked to highlight any manual or time consuming activity. These processes are then considered for RPA by filling out the intake form leading to no prioritization of automation. Consequently, these processes are not supported by any evidence as to why a particular process is prioritized over the other for automation.

  "We[Process owner] explained to them [RPA lead] which activities are being done manually as well as which are time consuming. Then, they decide if it is good for

automation." - Person D

- ***Time Consumed*** in some cases the whole process assessment activity can be completed in a short time but many times it becomes a laborious task. Additionally, it is seen frequently that low complexity processes that are outsourced become immediate candidates for RPA, in order to reduce the outsourcing cost. In such situations, the process assessment activity otherwise executed is not considered again leading to not taking into account the whole process workflow.

  "Sometimes the whole process assessment absorbs a month. Also, a lot of low-complexity processes have been outsourced to our partner company, which provided us with a solid preselection of candidates for RPA." - Person G

- ***Unable to progress further*** with the automation. Three out of seven interviewees quoted about having any structured way to progress in identifying the processes for automation. One of the business unit faced a problem of people grabbing opportunity to automate any 'low-complexity' or 'low-hanging fruit' activity without giving much consideration to the process as a whole. Similarly, another interviewee mentioned using their 'gut-feeling' to select processes while, another process owner in contrast faced an issue where they are unable to find any activity to automate further. It can be seen that the process owners face such situation, due to the fact that they are unable to substantiate the potential RPA initiative with process data.

  "After the Proof-of-concept RPA implementation, we were not able to identify suitable activities for RPA further." - Person C

  "With the automation initiative in my business unit, people are jumping to the opportunities of automating any low-complexity activity." - Person E

The above interview analysis and findings illustrate an encouraging picture of RPA implementation in the energy company, given that they introduced RPA just over two and half years ago. First and foremost finding is that, the organization introduced RPA as part of it's digitization strategy and within a short span of time established RPA CoE platform. Secondly, it can be seen that they have largely followed RPA lifecycle, starting with a POC phase and progressing into further RPA implementation, which is commonly followed by most organization. Moreover, process harmonization and standardization prior to automation are considered as prerequisites for automation and has emerged frequently in the

interviews. Lastly, use of manual process assessment to find suitable automation activities in the organization, leads to an unstructured way of implementing process improvements.

4.5 Energy company's RPA implementation lifecycle

Energy company's adoption of RPA has been mostly bottom-up, where the business comes up with areas and ideas for automation with the help of RPA process leads to guide them through the implementation. Typical stages of RPA project implementation within the organization are as follows and are similar to those presented in the research paper on RPA - Dynamic roadmap of successful implementation [43].

Process Assessment ⇒ Business Case ⇒ Proof of Concept ⇒ Project Design ⇒ Build - RPA Lifecycle

However, the most important factor to highlight is that, the company follows manual process assessment to find a suitable process for RPA. As mentioned in the findings above, manual process identification, time taken for performing the process assessment and finally unable to move beyond the initial quick wins or savings, all justify the need to have a scientific approach to identify an eligible process for automation. In addition, with multiple processes to automate, they are unable to prioritize the automation due to lack of supporting data. All this leads to an inconsistent way of process selection and thereby the organization fails to increase the RPA maturity level.

Chapter 5

Review of past Case Studies

This chapter focuses on the existing case studies that we reviewed to provide answers to the research question accompanied by analysis and discussion. This involves review of empirical case studies from different organizations who have implemented process mining as well as RPA in order to gain process efficiencies. This entails an extensive search and in-depth understanding of the context of case studies, before including them in the research.

5.1 Case Selection Criteria

The primary focus of this study is to understand the value of using automated process discovery, which in turn can be done by employing process mining technique, to find RPA eligible processes. Accordingly, diverse multiple case studies will be studied to get in-depth perspective and answers to the ”why” and ”how” questions with regards to PM and RPA. Yin presents both sides of using multiple case studies; it can allow the study design to replicate the phenomenon across the cases or can be used to showcase the contrasts among the outcome of the cases [55]. Number of case studies to be selected for a research is at researcher's discretion and a judicial choice on one hand but on the other hand they provide a stronger substantiation to the proposed theory [16].

We are of the opinion, that use of multiple case studies will enable us to demonstrate broader conformity and consistency in using Process mining as a rationale for RPA initiatives. Additionally, information-oriented selection is used as a case selection strategy because the choice of case studies are based on the implementation of PM and RPA within the organization as part of digital transformation. Diverse selection and appropriate number of case studies controls incongruous variation and help produce results as well as helps to sharpen external validity [16]. Therefore, we focus on four empirical research case studies in diverse markets like telecommunications, manufacturing and service industry. The diverse selection increases external validity and at the same time taking into account the

key themes of PM and RPA implementation within these organizations, gives a holistic approach to the case selection.

5.2 Case Description and Analysis

Multiple case studies are selected from various companies where PM and RPA were used to gain process optimization. To this end, main source of such case studies is a book namely ; "Process mining in action : Principles, Use Cases and Outlook" [40]. The table below provides an overview of the selected case studies followed by their description.

Case	Company	Process Mining Implemented	Reference
Case 1	Uber	Identify opportunities to improve the E2E process of customer Support contact handling with the goal to create efficient and harmonised processes with less errors and re-work loops	By Martin Rowlson, [40]
Case 2	Siemens OM4T	The global programme of Order Management for Tomorrow (OM4T) was to have a standardised ERP system across the countries all over in Siemens. Further goal was to automate sales back office processes in order to reduce manual re-work in turn to lower cost of regional order management of Siemens digital Industries	By Gia-Thi Nguyen, [40]
Case 3	Siemens P2P	Introduced Process mining to boost operational efficiency and optimize working capital by reducing the cost towards external suppliers	By Khaled El-Wafi, [40]
Case 4	Vodafone P2P	Introduced Process mining to boost operational efficiency and optimize working capital by reducing the operational cost	[32]

Table 5.1: *Case Studies Overview*

The book presents 12 use cases of process mining in diverse industries. The main objective of the book is to showcase business benefits of implementing process mining which is a broader concept than RPA. This study chooses only 3 use cases in particular that take into account process mining and RPA implementation. These case studies reflect the scenario wherein process mining guides the automation of processes and hence represents the selection criteria. Other cases were not considered as it involved only process mining context. On example of such case study is of BMW, where they introduced PM to identify

the friction and its root cause in the innovative paint shop production. In BMW case, data was collected from the machine sensors and analyzed to find bottlenecks to react quickly and minimize production problems [40]. This case can be a good selection for process mining, but doesn't involve any RPA and hence rejected. The Vodafone case study shown in the above figure 5.1 is also incorporated in the list as it increases the diversity of case selection as well as it is an independent scientific paper explaining the importance of mining technique to select suitable processes for RPA [17].

- ***Case 1 - Uber*** Uber's fast rise in the service industry, resulted in high process variation leading to a void of uniformity in handling customer support. This inconsistent response in turn increased not only process waste as well as cost in agent handling time but also reflected badly in terms of customer experiences due to variation in service levels. Uber was in a dire need to be able to understand how the processes actually performed even before rectifying the problem of unnecessary process variations.

  The solution to this problem was found by deploying process mining to identify opportunities based on the data provided to improve the end-to-end process of customer support contact handling with the goal of creating efficient and harmonized processes leading to increase service level responses. With the implementation of PM tool, Uber gained $ 20 million in operational efficiencies through handling time improvements. With the evidence-based process visualization empowered Uber to uncover the process inefficiencies and rework loops thereby helping them to realize process improvement opportunities and introduce RPA automation. According to Uber, to automate the right processes one must have a good understanding of processes on which the Mining techniques will transparently illustrate how these process perform.

- ***Case 2 - Siemens*** Earlier in Siemens, regional offices had their own process landscapes referring to the processes in the regional companies from order entry, order processing, follow up, billing, and cash collection in the customer facing realm. Each region strived for process improvement, but couldn't come up with any breakthrough improvements due to the lack of speed. In order to bring in process efficiencies, in 2016 Siemens initiated a centralized consolidation of ERP systems across the organization worldwide and was termed as Order Management for Tomorrow (OM4T). The aim was to collect all the data from each ERP system and transfer into a global data model across all regions of all organizational units of Siemens. With the use of process mining technology implemented on this data, Siemens could visualize how the process flow was transparently displayed as well as ability to view the manual activities performed within.

  Based on the event logs, Process mining also provides the automation rate which is arrived by the automation of activities against the rework rate around the manual

steps around items. Therefore, Siemens came up with a KPI to have a Digital Fit Rate, which is calculated by number of manual activities divided by the number of sales order item. With the use of this information, Siemens introduced RPA automation to reduce the digital fit rate by initiating "low-hanging fruit" automation in the first year and gradually automating further decreasing the rework. Continuous automation insights, the Digital Fit Rate has now become an integral part of process excellence in every region giving improvements at a global level. The Digital Fit Rate continues to decrease from 2.1 to 1.0 and in some countries even down to 0.3, which means that the amount of human interventions in the process is decreasing.

- ***Case 3 - Siemens*** Siemens in order to remain competitive in the market, it was of utmost importance to focus on purchase to payment (P2P) process to have efficient supply chain. To enable effectiveness of the overall process, the supplier costs needs to be kept in check as well monitor timely delivery and service level quality. Process mining was introduced in this area to boost operational efficiency and optimize working capital by reducing the cost towards external suppliers.

  With the help of process visualization and the rework rates, Siemens could identify overall high approval activities for validating the information within the process, increased re-work. The constant validation checks were in fact due to the high level of manual data entry which was obviously error prone. Therefore, implementing RPA automation to validate transactional data against master, avoided any error or miss out on preferred suppliers, pricing, and other terms. Overall the automated Purchase Order led to faster processing times, cost savings thereby optimized working capital. According to Siemens, process mining gave them data transparency enabling them to identify process weaknesses whereas RPA generated operational efficiencies due to reduced errors and cost and delivery time savings.

- ***Case 4 - Vodafone*** As of 2018, Vodafone has more than 300 million customers and mobile operations in 25 countries with broadband and fixed phone in 18 countries, it is one of the world's largest telecommunications companies generating voluminous data on a daily basis. It is operating in the market for about 30 years in different regions initiated a common shared SAP system but also able to handle the local requirements. The key challenges at the time of rolling out process mining were inconsistent processes and lack of adherence to global policies of processes. Without real clarity regarding process operations became difficult to adapt to as their businesses changed. The Vodafone process excellence team struggled to identify and prioritise process improvement as many purchase orders had excessive throughput times before they are sent out to the supplier.

  After introducing Process mining, Vodafone could visualise the variations within the same process due to high number of deviations from the standard process. This triggered the fact-based discussion of solutions and taking immediate action of imple-

menting RPA on the repetitive activities in the process. The primary win of process mining and RPA together achieved a key milestone of a perfect order rate of 92% and and at the same time reduce operational costs due to faster throughput time.

5.3 Case Study Discussion

This section will examine the case studies mentioned in the previous section and present process mining capabilities that can be beneficial for process automation. Although the selection criteria for choosing the case study was both Process mining and RPA implementation, the main underlying fact is that all the organization's primary aim was process efficiency and to that end all of them adopted process mining first. RPA automation has been a by-product to achieve process excellence and thereby operational cost benefits. Therefore, it can be convincingly said that RPA is not a goal in itself, but a tool on the path of achieving process excellence. At the same time, all the case studies highlight the fact that process mining assisted them to automate suitable processes, be it 'low-hanging fruit' or re-work tasks. It is important to note that though process mining has much more capabilities and features that lead to better processes, this study will highlight the capabilities that guide towards automation.

- ***Process Transparency -*** All the case studies adopted process mining for the main reason that it produced a process map of how activities flow in the business process. The main defining feature of the process mining technique is to display "As-is" process workflow visualisation that enables the business user to view all the activities performed in reality as opposed to modelled process. As the mining algorithm uses application event logs, it produces various possible paths the system has apart from the standard or expected path.

- ***Faster process Assessment -*** The claim that process mining performs fast process assessment is true to a certain extent. Using the process visualization along with variants in processes with throughput time in each deviation, business users can spot the inefficiencies within the process workflow and then further investigate with the concerned employee to clarify its details. In other situations, the user can drill down the visualization to zoom into a particular process variant to understand the delay or bottleneck for that particular activity. Process Mining facilitates spotting these irregularities faster rather than a process expert holding workshops or interviews to understand and identify process redundancies.

  Uber in case study 1, utilized this capability to identify possible process improvement opportunities to remove redundancies. However, the case study doesn't quantify the time saved due to faster process assessment. Given that visualization is based on the event logs, in most cases it can be said that assessment can be faster as opposed to

Process Mining Claim	**Evidence found in Case Study**	**Statistical Evidence**
Fast process Assessment	Case 1 - Analyze process data with speed that allows us to accelerate the realization of process improvement opportunities to remove rework	
"As-Is" Process Transparency along with process variants	Case 1,2,3 & 4 - Process Mining showed us a new level of Transparency. It was Transparency beyond imagination.	Case 2 - Overall Automation rate increased by 24% and manual rework reduced by 11%
Process Optimization	Case 1,2,3 & 4 - Process mining showed us unknown wastes and process inefficiencies.	Case 1 - Identified around $20 million efficiency savings.
Evidence based Automation Intelligence	Case 1,2,3,&4 - Process Mining based on event-log evidence highlights automation rate for individual activities. Automation rate is based on automation of activities against the manual or rework activities	Case 2 - Digital Fit Rate reduced to 1.0 from 2.1. In some region was reduced to 0.3
Prioritize Automation (Good Process understanding)	Case 2 & 4 - "We discovered that traditional analytics tools could not uncover hidden inefficiencies or provide full transparency into our data, and we needed unbiased analysis to solve problems we did not even know we had, and to identify the most impactful opportunities for improvement."	Case 4 - Process mining helped us to identify the correct process to automate. 92% perfect PO (from 73%), 85% touchless invoices, 20% faster process
Operational value	Case 1,3 & 4 - "Increased process automation" is a lever for efficiency, improvement and savings and therefore increased efficiency along the E2E process.	Case 2 - Mining enabled automation led to $ 10M+ savings Case 4 - An correct purchase order rate of 92% and reduced operational costs
Monitor Automation	Case 2,3 & 4 - Automation can be monitored via process mining	

Table 5.2: *Case Study Analysis*

arriving at an assessment based on interviews with the process experts. In certain non-standard cases, human intervention and discussion can be attributed to further investigate the problem.

- ***Process Optimization*** **-** Similar to previous capability, process variations showcased by mining tool helps process experts to perform root cause analysis and standardize the processes. All the cases selected categorically mentioned about how mining helped them to optimize their operations. Particularly citing the Uber case study, mining technique helped them to optimize the customer experience by removing inefficiencies in the customer support time taken as well lack as of consistency in the process across the locations. Discovering these inconsistencies and automating those tasks provided roughly $ 20 million in savings to Uber.

- ***Evidence based Automation Intelligence*** **-** Based on the digital traces, mining algorithm produces visualization that can be drilled down to a single variant or display multiple paths of the same process. At the same time the visualization also highlights the time taken by each activity thereby calculating median time taken to complete the process by different paths. These process variances can be further filtered based on pre-defined attributes like delay, re-work or effort driver which form the basis for selecting these activities for automation. Moreover, process mining algorithm also presents Automation rate for the activities within the visualized process. The automation rate is calculated by number of automated activities against the manual activities, also guides the user to choose a possible processes for automation. Lastly, the tool offers an option to define overal KPIs to set a performance goal for the processes. A separate KPI dashboard displays target KPIs vis-a-vis goals achieved for a certain process. Using the combination of aforementioned parameters, the mining algorithm presents its analysis about possible process improvements to reach the target.

  In case of Siemens, the process excellence team set a KPI of Digital Fit Rate which is calculated as the number of manual activities divided by the total number of line items (Sales order items for the case). Keeping in mind the Digital Fit Rate KPI, Siemens used this automation rate for their activities to trigger RPA initiatives. By diligently following this strategy, they could reduce the Digital Fit Rate from 2.1 to 1.0

- ***Automation Prioritization*** **-** Process mining can give guidance to RPA initiatives and direct towards possible processes for automation. Automation rate for the activities within a process are displayed alongwith process visualization. By further analysing the activities, business user can make a rational decision to automate 1 activity before the other.

  In Vodafone's experience, they wanted to implement RPA automation with processes displaying low automation rates. However, they were unable to resource all the

potential automation candidates at given time, and therefore implemented it in a staggered way. They could see the value in no time by automating the "low-hanging fruits" to generate quick wins. In fact, Vodafone achieved a 97% perfect Purchase order from initial 73% in a span of very little time.

- ***Monitor Automation*** **-** An RPA initiative results in quick benefits, it requires continuous checks on its utility value in the future. Process mining can provide an added advantage in monitoring these automations implemented using KPIs like throughput times as well as dedicated measurement criterias for process improvements can be defined and its impact can be continuously measured. This also means that the organization has to constantly benchmark projects with successful prior initiatives to gain maximum value from automation [17]. Most of the case studies mentioned about the advantage of monitoring the KPIs and RPA automation using process mining and how it helped them to achieve the organizational level goals.

Lastly, the outcome of exercising process mining capabilities, generates ***Operational Value***, an indirect benefit to the organization. Optimized working capital derived from cost reduction and time savings with automation initiatives using the mining algorithm, all lead to increasing the operational value. Siemens in particular, adopted process mining with a purpose of increasing operational value and presented $ 10 million in cost savings due to automation. Unquestionably, all the case studies seemed to enjoy the benefit of increased operational value as they could gain savings in cost as well as overall improved end-to-end processes.

Based on case study analysis and discussion of 4 empirical case studies, it can be concluded that mining technique has certainly proved beneficial to the organization financially as well as gaining process efficiencies by increasing their automation. On the other hand, it is clear that process mining was implemented first to achieve process optimization and RPA was adopted later in order to achieve this goal. Therefore, the case study analysis point towards making use of process discovery, the integral part of mining technique to perform process assessment. In all the case studies automated process discovery emerges as a single way of discovering processes to automate. Fast process analysis due to process map visualization, guidance for automation process selection and prioritization as well as taking into account the whole process, all suggest a rewarding outcome of a process discovery phase. By drawing on the successful approaches adopted by the four case firms, automated process discovery can be incorporated in the RPA lifecycle of the energy company analyzed in the chapter 4. In conclusion, incorporating automated process discovery in the process assessment phase can accentuate the automation lifecycle for the energy company.

Version of August 14, 2020– Created August 14, 2020 - 15:32

Chapter 6

Automated Process discovery Approach and Challenges

This chapter presents the challenges in implementing process discovery using mining technique within the organization. Further in the chapter, we present the contrasts between process mining and task mining approach of Automated Process discovery.

The concept of automated process discovery and two different techniques used in practice namely process mining and task mining are elaborated in conceptual foundation chapter 2 in section 2.5 and section 2.6. Similarly, as mentioned in the earlier chapter 5 of past case studies, it is now established that automated process discovery can assist in determining the right process to automate. Therefore, incorporating it as part of process assessment can improve the rate of successful automation. The following section will describe various challenges encountered by the organization in implementing the mining technique.

6.1 Challenges to adopt process discovery

RQ1 - What challenges can organisations face to adopt automated process discovery for Robotic Process Automation?

Organization aim to use an innovative technology like process mining to encourage continuous improvement with better process visibility. Unsurprisingly, investments of this kind can lead to an array of challenges to account for. This section addresses these barriers and pitfalls that many organization encounter while adopting mining technology. In order to find answers to RQ1, exploratory study was performed by conducting interviews with RPA stakeholders within the organization. The semi-structured interviews with the stakeholders mentioned in table 4.1 included questions over challenges faced in adopting process discovery. Similarly, analysis of empirical case studies helped us to un-

derstand the pitfalls and failures experienced by large organizations while implementing process discovery. Grounded theory approach was applied, by identifying patterns in the interview responses as well as the empirical case studies examined. Furthermore, coding process similar to that mentioned in section 4.3 was performed and the result is presented by interspersing quoted passages accompanied with context explanation.

Although many interviewees agreed to the fact that process discovery does provide benefit in a distinct manner, it is not easily assimilated into the IT landscape of the organization. The most candid response over process mining software given by one of the interviewees highlight the possible issues the organizations come across.

"Tools are not always plug and play as they say. Need to take a lot of things into consideration." - Person G

By and large the challenges identified are applicable to both process and task mining and therefore are presented collectively. Based on the interviews conducted in the energy company as well as with the experts from mining vendors and the literature analysis of empirical case studies, following impediments were mentioned prominently:

1. Unclear Purpose

According to Reinkemeyer et al, not having a clear idea or purpose of using Process mining and how it can contribute towards improving business processes can lead to failure [40]. Process mining capabilities can be put to use in various different areas and in different industries but it needs to be adopted properly to serve its purpose. Some examples can be - in a manufacturing unit, process mining purpose is used to identify optimized inventories, reducing working capital or bottleneck in delivery process or in a Supply-chain unit, process mining purpose is to identify hurdles in the logistic delivery process.

The interviews with professionals from the mining vendor company also addressed the concern about having a defined business goal to introduce mining techniques in a organization.

"The end result is achieving business goals. Organization must start with a business goal and in order to achieve your aim, begin with Process mining and include RPA if necessary. Lots of organisations struggle to start, lots of departments are not sure as to how they should start with mining." - Person H

Similarly, two participants in the stakeholder interviews acknowledged their familiarity of process mining tool based on their experience of implementing a POC. However, to a certain extent they were unaware about having a specific goal to utilise the value of process mining beyond visualizing the 'as-is' processes.

Therefore it is imperative that the organization should be able to define the main goal

to implement mining. Whether the organization wants to achieve, operational efficiencies by having faster response time in a support service or the purpose is to merely substitute the mundane and manual activities performed outside of the transactional systems.

2. Lack of data quality

Unreliable data can be a big hurdle in adopting process mining by the organizations. Process mining presents process map based on the event logs from the informational systems and therefore if the collected data is unreliable or incomplete, the proportionate failure rate is high too. All the interviewees with process mining knowledge, without any exception, agreed to the challenge created by quality of data extracted from the enterprise systems. The energy company used for this study, undertook process mining POC couple of years back and faced data mismatch issues. According to the organization, it was a cumbersome task to match the data collection and cleaning according to the process mining specification. In certain situation, customization within the target systems caused mismatch with the mining data specifications.

Following data-related issues created barriers in process discovery implementation :

- **Data completeness** as well as its quality can prove an enormous task in itself before even implementing process mining. A simple mismatch between the formats of the target system and the PM tool to unavailability of the data due to lack of permission to use the data by the workers council can impede the progress of process mining implementation.

  "Data quality is a challenge. PM expects certain quality of data and that is a mismatch to what we have in our systems. A simple date format created a big problem" - Person G

  "We have a lot of customization in our target systems such as SAP. So Data is not as per the requirements of the PM software. PM tool requires high quality data with timestamps and the registered users for it to be successful." - Person F

- **Data preparation** or cleaning the data to provide input to process mining software can be a difficult task. More than 50% of the interviewees acknowledged the fact of time involved in data preparation before actually using mining analysis. Despite the fact that process mining can connect to multiple systems with its APIs, to bring it into reality taking into account various interfaces is an impediment. Security compliance for various systems within the organization too can be a barrier to cross even before the organization starts using process mining.

  "Data preparation is yet another challenge. two-third of total implementation time

is spent on preparation and collection." - Person F

- **Data privacy** can be yet another problem for implementing mining capabilities for the required purpose. The event logs generated from the system have timestamp as well as registered users who performed the activity and requires an extra step to anonymize the data before loading data onto the process mining servers. This is especially a bigger issue in task mining where screen capture records activities executed by the user. Anonymizing the data is a way to protect the privacy of the customer, even so organizations have to be prudent in finding the right process/task mining tool that supports confidentiality preserving approaches. In countries with strict data privacy rules, even after anonymizing the records, it cannot be used without the consent of workers council which in itself can be a long drawn process.

3. Lack of funding

Yet another paramount challenge is procuring budget for implementation of any new technology within the organization. All the stakeholder interviewees from the organization raised funding as an important issue as one of the impediments to introduce mining technique. Firstly, organizations have to present a business case even to carry out a POC. Secondly, many times such technology is part of a central initiative within the organization and hence getting consensus of all the units to agree on a common tool and share the costs proves to be a difficult task. Likewise, budgetary constraints sometimes make it difficult for the companies to maintain the technological support for process mining's data requirement.

"Funding main challenge even for POC. We do not get any budget unless we build a business case"- Person F

4. People attitude not to change

Even when the organization is ready to implement such technology, possible opposition or reluctance from the business users can further cripple the goal to achieve data-driven process efficiency.

Many factors contribute to this attitude displayed by the users:

- **Resistance to change** and learn new technology or fear of losing the job they have been performing for number of years results in such attitude within the organization.

  "People are resistant to change the current as-is situation. Mindset to accept change and willingness to make changes within process to generate business value." - Person H

Version of August 14, 2020– Created August 14, 2020 - 15:32

- **Overwhelmed by the transparency** displayed by process mining tool. The process mining tool provides detailed multi-level process map displaying all the possible paths and deviations an event undergoes within the system. It lays bare, the activities executed within the process thus highlighting any discrepancies within it. This can be challenging to accept for many employees and at the same time unwilling to have their performance compared with others.

  "lot of confronting or overwhelming to people to see what's going on within the process minute to minute as everything becomes transparent. Also people are resistant to change." - Person I

- **Resourcing and knowledge management** of mining activity is yet another impediment with regards to people issues. Either train the existing employees within the organization to step into the new role or recruit new resources with the required skill set. The situation becomes unusual, where company is not able to see value of mining without having right people know-how to use process mining. On the other hand hesitant to recruit new resources without seeing any benefit. This has to do with knowledge building about process mining among the people using or supposed to be using the tool to make these decisions. For example, Business analyst has to analyze processes visually using mining tool and based on the various automation initiatives displayed on the dashboard, analysts needs to comprehend those initiatives before choosing those improvements.

  "Then come the analysts who will come up with improvement initiatives. Lot of initiatives will pop-up, but then decisions need to be taken about what initiatives to be considered. You need to have the ability to compare various possibilities and knowledge involved in understanding how it works in process mining" - Person H

6.2 Choosing process or task mining

Even though task mining is complementary to the process mining technique, the input data to both techniques are different and so is the automation guidance given by them. Process mining produces visualization based on event log data whereas task mining captures user interaction data recorded by the screen capture tool. Therefore, this section will focus on answering RQ2, from the organizational perspective which process discovery approach can be useful to adopt.

RQ2 - What are the trade-offs between Process Mining and Task Mining the organization can consider while adopting automated process discovery?

Based on the key differences between process discovery using process mining and task mining, organization can decide on the process discovery approach they want to consider. However, aspects such as ease and cost of implementation or the time taken to implement these tools can cause one approach favourable over the other.

Following section highlights these differences between PM and TM:

Areas of application

Firstly, the fundamental difference between process mining and task mining is that they differ in their areas of application. PM presents an operational view of the business process in the information system whereas task mining presents individual activities performed in a process, both outside and inside of the informational systems. An operational view of business processes is a workflow visualization through the information systems with comprehensive depiction of all possible paths. On the other hand, task mining produces a narrow assessment of the specific processes or tasks.

If the goal or the use case of the organizations is to optimize the overall business process, process mining can help the organization to discover the bottlenecks in the process. Based on the process analysis guidance, they can further automate certain activities and eliminate or reduce bottlenecks. In the case of Siemens, in order to decrease the number of human interventions or human touches in the overall order management process, process mining was implemented to identify manual processes and automate them as much as possible [40]. Similarly, for Uber the goal was to have faster response time by improving the end-to-end customer support process [40].

On the other hand by capturing user level interaction, task mining introduces a new level of detail like the user actions not logged by information systems which can be suitable for automation [35]. If the goal of the organization is to automate user activities or standardise workforce behaviour performed within and outside of information systems, then process discovery using task mining can give better results. Typical use case of task mining for an energy company can be copying the meter reading collected by the field engineers from the excel spreadsheet on to the information system. However, it is important to remark that such automation at the bottom most level, can collectively benefit the overall business process in terms of decrease in throughput time.

Tool Implementation

Task mining and process mining differ in terms of the implementation process followed. Firstly, task mining runs on the user desktop or a workstation, with a common server to analyze the mined processes whereas process mining needs to connect with respective information systems like ERP, CRM to collect the transactional data. The data

further needs to be uploaded onto the mining server whereupon, the business users can perform their analysis. Based on the capability of process mining tools, the data can be collected with real-time API connection or organizations need to perform data extraction from the informational systems and loading onto the mining server.

Secondly, process mining implementation is time-consuming as it involves various steps to get initial data uploaded and ready, real-time or otherwise. Process mining has certain specific data criterias with respect to event logs, even before user performs process analysis. For example, every event log should have timestamp and a unique case-id throughout the business process in order to be able to model the whole process based on the event flowing through the system. In order to have correct data, the organizations therefore either need to change the target system or implement an intermediate process to set the event data right for process mining. Some of the responses from the interviews with the process expert in the energy company also point to the same direction.

"We have a lot of customization in target systems such as SAP. So Data is not as per the requirements of the PM software. PM tool requires high quality data with timestamps and the the users registered to be successful." - Person F

However, Task mining implementation is simpler and therefore less time consuming than process mining. But as mentioned in the challenges above, implementing task mining can have data privacy issue. Task mining goal is to achieve workforce optimization and that entails bench-marking of activities performed by various employees. Hence, can be opposed by the workers union in certain countries having strict privacy laws.

6.2.1 Trade-offs between PM and TM

Taking into account various factors involved in using process discovery to implement RPA initiatives, it can be concluded that the right approach should be primarily based on the organizational goal with respect to automation. If the business goal is to achieve end-to-end process optimization then automated process discovery using process mining technique can be adopted. On the other hand, if the goal is to achieve workforce optimization and workflow bench-marking, then process discovery using task mining is suitable.

Having said that, the organization can possibly consider certain trade-offs which can have an impact on the choice of process discovery approach. Ease of implementation or time taken to implement can be few trade-offs the organization can consider to take up task mining approach instead. Automating manual activities and standardising them at the task level to begin with, can collectively reduce the cost and processing time contributing towards overall process efficiencies.

Recently, certain commercial vendors offer Process discovery tools with both functionalities namely process mining as well as task mining features incorporated into one software [34]. Celonis software offers such capability of integrating the user interaction data

with the business data(event logs) to visually identify tasks that have relatively simple internal process and can, therefore be potentially selected for automation. According to Celonis, the connection between user interaction data and the transactional data enables to build the correlation between different process steps and the related business operation data [9]. However, future case studies can further illuminate and clarify the actual implementation and use of process mining and task mining in a single integrated software.

To summarise, unclear purpose, data compatibility and people attitude towards the technology all contribute to the barriers encountered by the organization to implement mining. However, certain barriers in adopting mining technique can be resolved and will be clarified further in conclusion and recommendations chapter 8. In addition, the chapter has broadly explained the contrasts between PM or TM for automated process discovery and what are the trade-offs between the two approaches. Additionally, the recent capability of having both PM and TM within a single software can contribute further in integrating the task level data with enterprise transactional data and provide further opportunities of optimizing business processes. To this end, we have provided conclusive answers to RQ1 and RQ2.

Version of August 14, 2020– Created August 14, 2020 - 15:32

Chapter 7

Enhanced RPA lifecycle using Automated Process Discovery

This chapter leads to the design of an enhanced RPA implementation roadmap using automated process discovery to find suitable RPA initiatives. It has been established in chapter 5, incorporating automated process discovery in the process assessment phase can accentuate the automation lifecycle. With this new outcome of including process discovery in the process assessment phase, contribution is made by modifying the earlier RPA lifecycle. To this end, an enhanced RPA implementation roadmap using automated process discovery is presented grounding the RPA lifecycle more solidly in business practice.

7.1 Design science research

According to Henver et al, design science, a research paradigm in Information System domain, constitutes three cycle view namely; relevance cycle, design cycle and the rigor cycle [22]. The relevance cycle sets the contextual environment by addressing the requirements and/or opportunities of the research project whereas rigor cycle draws conclusion from the evidence found in scientific literature and experience and experts. Finally, the core design cycle iterates between constructing and validating the design artefacts and processes of the research [22]. However Henver et al, gives importance to contributing to the body of scientific research by creating new artefacts or design. On the other hand, Wieringa et al proposes to design or re-design an artefact to improve a problem context to answer knowledge questions about the artefact in context [51].

Wieringa's design cycle treatment mentions building or improving an existing design artefact in order to answer the knowledge question. The main goal of this study is to improve upon the existing problem context of process assessment within the RPA lifecycle. By addressing the knowledge question whether process discovery can be used for RPA process assessment, the study aims to improve the RPA lifecycle. Therefore, this study

will make use of this design cycle proposed by Wieringer et al that leads to an enhanced dynamic roadmap for RPA implementation.

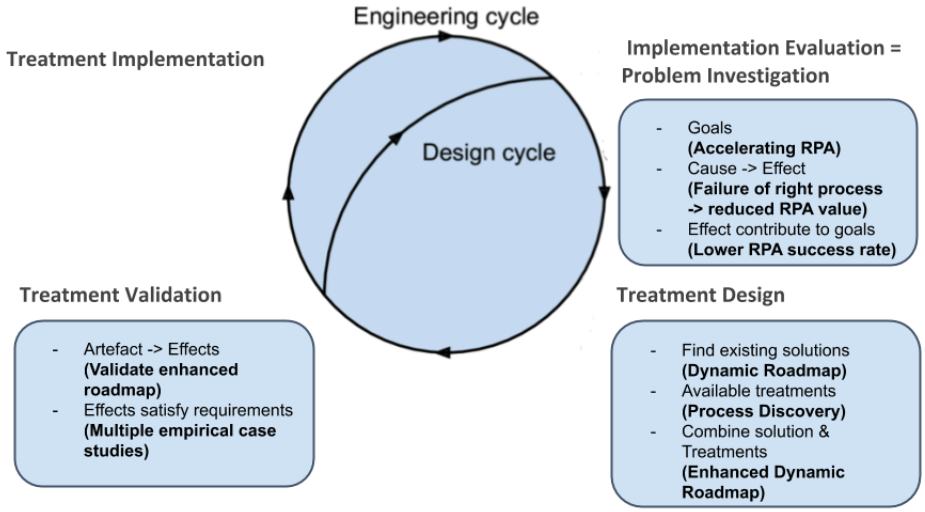

Figure 7.1: *Design Cycle implemented in this research based on Wieringa Design Cycle model [51]*

Figure 7.1 illustrates the main phases of this design cycle and its implementation for this study. The main phases followed in this engineering cycle are namely; Implementation evaluation, Treatment design, Treatment validation and final phase of Treatment implementation. The different phases within this design cycle show an iterative improvement of roadmap developed in the information system research. Having said that, even though the design cycle depicts an engineering cycle as part of the design treatment, this research does not implement actual software robots and therefore doesn't fall under the scope of this study.

The following section explains the individual phases implemented as part of this design treatment cycle:

1. Implementation Evaluation

The main objective of this study is to evaluate how organizations can improve their RPA lifecycle using a scientific approach to identify processes to automate. As highlighted in chapter 1 and section 1.2, the organizations in order to make their processes efficient have adopted robotic process automation to reduce execution time and free up resources for other activities. However, they constantly face challenge of selecting the right process to automate. The failure to choose the right process can have a negative impact on the success of automation initiatives. In order to understand different phases in the RPA lifecycle, we conducted interviews of RPA stakeholders in a large energy company. These

interviews 4, enabled us to understand the overall RPA lifecycle in a organization as well as contextualise the problem of process selection. Unable to identify any process to automate beyond POC or automate any manual process without taking into account the whole process were the main points that featured in the interviews. Moreover, it was found that manual process selection based on customised RPA criterias was time-consuming and person dependant. The effect of manual process selection to identify potential RPA candidates transpired into reduced rate of automation initiatives and low maturity level.

2. Treatment Design

With the objective of understanding overall RPA implementations within organizations, we studied various existing RPA lifecycle approaches to automate process and are explained in the chapter 2 and section 2.3. The lifecycle processes described in the section are of interest, but they are generally context specific restricted to actual RPA implementation only. However, there are lot more phases and steps involved prior to executing RPA initiatives. Therefore, from the operational viewpoint on RPA lifecycle, the Dynamic roadmap for RPA implementation presents comprehensive phases starting with identifying a business problem, performing a POC until robot deployment [43]. Even though this roadmap presents individual steps within RPA implementation, the process selection or process assessment phase is very generic. Considering the fact that process selection is pivotal to the success of RPA implementation, it is important to have more guidance in that respect.

Correspondingly, in order to identify the potential RPA candidates based on the evidence, various scientific literature was studied and discussed in the chapter conceptual foundation 2 and section 2.5. Automated process discovery using the mining technique provided evidence-based scientific approach towards RPA process selection and therefore provides a solution to identify right process for automation initiatives. Goris et al, proposes automated process discovery using mining technique as a reliability check along with the usual interview based process selection [18]. However, Geyer-Klingeberg et al mentions process mining as a guide that organizations can use throughout the lifecycle of RPA initiatives from identifying automation, prioritizing until monitoring the deployed robots[17]. Moreover, the case studies discussion in chapter 5, further help in conforming the use of automated process discovery that led to RPA initiatives in practice.

In the subsequent chapter 6, we further discuss two different approaches organizations can adopt in order to implement automated process discovery to select RPA potentials. Primarily based on the business goals of an organization, process mining or task mining approach can be used to determine the mining choice. Drawing conclusion from the existing solutions (to use automated process discovery) and finding the treatment to address the problem of manual process assessment, the dynamic roadmap of RPA is improved by incorporating automated process discovery and an enhanced roadmap of RPA lifecycle is presented in the figure [7.2]. Based on the different ways of performing automated pro-

cess discovery, the process assessment phase is expanded further to include task mining and/or process mining.

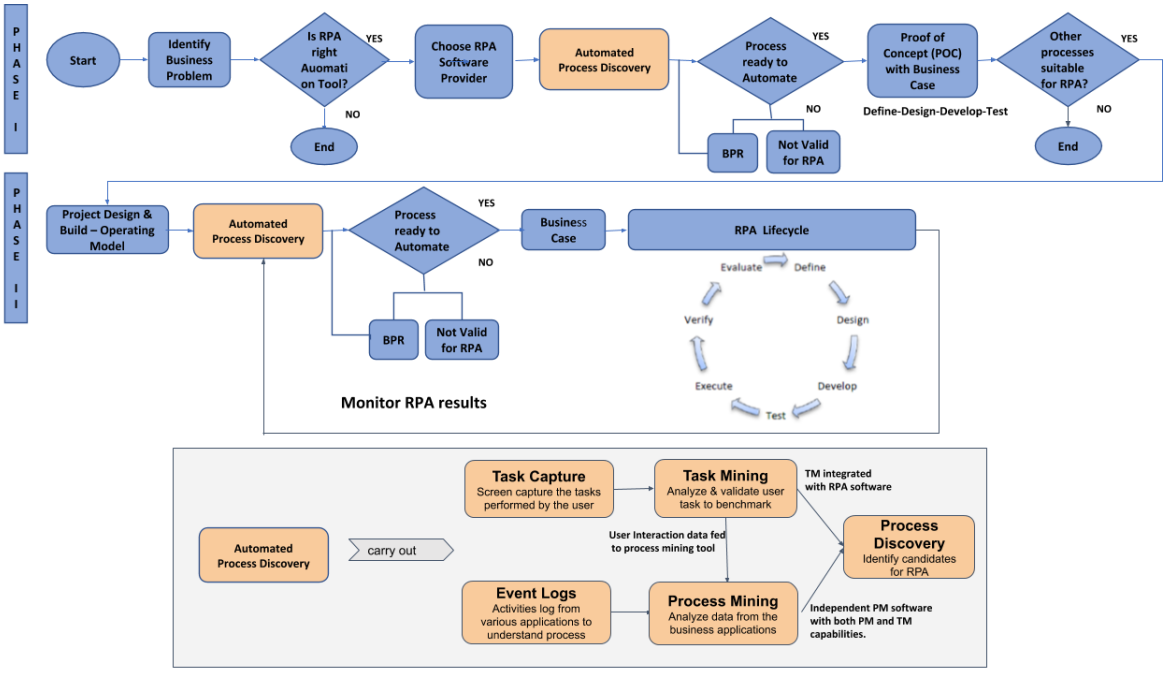

Figure 7.2: *Enhanced Dynamic roadmap of RPA Lifecycle using Process Discovery*

The main elements in the enhanced roadmap are :

- Phase 1 : Identification of business problem, choosing a RPA vendor, choosing a process for POC, process discovery using integrated or standalone PM software to create a business case, re-design or standardization if necessary and then execute the POC. The first phase ends with the completion of POC and based on the result of POC, the organization makes the decision of continuing RPA initiatives or not. The modified lifecycle includes the use of automated process discovery to identify the process to automate.

- Phase 2 : Based on the outcome of POC, the organizations decides to go ahead with further automations. Forming an operating model for automation projects, further executing process discovery, process re-design or standardization followed by actual RPA lifecycle execution. The performance of the robots can be monitored by mining technique as new process data is fed back into the mining server to analyze. This is shown by RPA lifecycle connecting back to automated process discovery in the diagram.

- Lastly, the proposed roadmap also illustrates the automated process discovery approach the organization can adopt. The diagram clearly presents two independent

mining approach namely, task mining for a bottom-up approach and process mining for a top-down approach to gain knowledge about the existing business processes. The significance of task and process mining is discussed in detail under chapter 6 and section 6.1. Moreover, the enhanced roadmap, also highlights the possibility of having the benefit of task mining as well as process mining combined to analyze and provide automation potentials.

Complete enhanced roadmap of RPA lifecycle using process discovery is presented in Appendix B.1 along with description of individual activities carried out throughout the RPA lifecycle. This modified RPA lifecycle can be used as a model template by the organization who want to adopt RPA as well as those who have begun their RPA initiative. The model can be used as a reference point of how to implement RPA in a structured manner using automated process discovery.

3. Treatment Validation

Treatment validation is to check the validity of the built artefact and to check its effect in a practical context. In order to determine the correctness of the model and to make any possible changes further in the roadmap, testing the roadmap in practical setting is the best possible way. However, the time involved in implementing the complete cycle of RPA as well as other constraints within the organization to initiate a complete RPA lifecycle in its entirety was not achievable. Nonetheless, we adopted model evaluation approach to check its validity. Model evaluation is described in detail in the following section and the final artefact is described in detail in appendix with adjustments made from evaluation interviews feedback.

7.2 Evaluation of proposed roadmap

Evaluation is needed to check the efficacy and reliability of the enhanced roadmap for RPA implementation. Nonetheless, to achieve the evaluation of the enhanced roadmap, evaluation interviews were conducted to gain views and any additional recommendations to the roadmap.

To facilitate the evaluation of this roadmap in an open manner, semi-structured interviews were brought about allowing both the interviewer and the interviewee to provide suggestions or share new ideas regarding the RPA roadmap. In order to proceed with the evaluation interviews, similar approach as earlier was employed to get to the right people. However, earlier it was possible to interview 7 people within the organization, but due to the time limitations 3 interviews were performed to receive feedback over the evaluation of the new roadmap. Table 7.1 provides the details of all the evaluation interviews conducted in this study. The interviewees were selected on the basis of senior position within the company and who have on overall view of the automation scenario within the organi-

Business Unit	Code Name	Function & knowledge	Date/Time
BA Heat	Person 1	RPA Programme Manager (NL- Management)	3 June 2020, 45 mins
Customer & Solutions (B2B)	Person 2	RPA Lead (Process & RPA Technical Knowledge)	4 June 2020, via email
BA Heat	Person 3	RPA Programme Manager (Germany - Management)	29 May 2020, 45 mins

Table 7.1: *Overview of participants in roadmap evaluation*

zation. Moreover, the selection of interviewees was also based on different location of the interviewee as it provides a diverse perspectives to same artefact. As done in the previous interviews to maintain the privacy of the interviewee, their personal information has been kept anonymous by assigning code names.

Evaluation outcome and final roadmap

The enhanced roadmap as illustrated above in figure 7.2 was presented to the concerned experts in the organization. The overall response and understanding of the model was consistent to the real-life RPA lifecycle in general with minor changes in terms of naming conventions used in the roadmap. The evaluation response and opinions received through the interviews or email responses can be summarised as follows:

- According to Person 1, the phases should be clearly labelled in order to understand the lifecycle. Consequently, the phases were renamed as ”POC PHASE” and ”RPA - APPLY PHASE” in order to satisfy with the overall understanding of the roadmap. Moreover, addtional minor changes asked like expand the abbreviation such as ’BPR’ were not taken into consideration given the limited presentation area for the whole roadmap. Additionally, the Business Process Re-Design (BPR) convention is universally accepted to the audience of process expert, process automation, business user or even an IT personnel in general.

- After completing RPA proof of concept in phase 1, the next step in the enhanced roadmap was to go for further suitable processes for RPA. Though this was not changed from the original dynamic roadmap of RPA lifecycle, the logical path for any organization would be to conclude if the POC was beneficial for them or not. Based on that reasoning, the organization can arrive at a decision over GO/NO-GO of RPA automation itself. Considering the fact that, the use of this roadmap is primarily to assist the potential businesses as well as the existing businesses to traverse

through the RPA journey smoothly, usage of business jargon seemed appropriate.

- The most invaluable feedback given by Person 1 interviewee was about the termination of robots. Even though the robots are used for the repetitive manual and/or high volume tasks, there can be a situation where due to the process change, the activities performed are no longer valid. In that case, the robots are terminated by removing them from the orchestrator schedule. This is an important activity and should be part of the overall RPA lifecycle in general. Accordingly, the activity was incorporated in the updated roadmap.
- The process discovery activity, expanded further primarily presents 2 different approaches that an organization can identify eligible processes for automation using mining namely, task mining and process mining. Additionally, a possible third approach to decide on potential RPA candidates can be to combine task and process mining approach. According to the Person 2 interviewee, possible third approach is not self-explanatory and the flow depicted on the roadmap doesn't signify any such thing. Therefore, to incorporate the feedback, approaches were numbered and a small legend explaining the 3 approaches was inserted in the final roadmap.

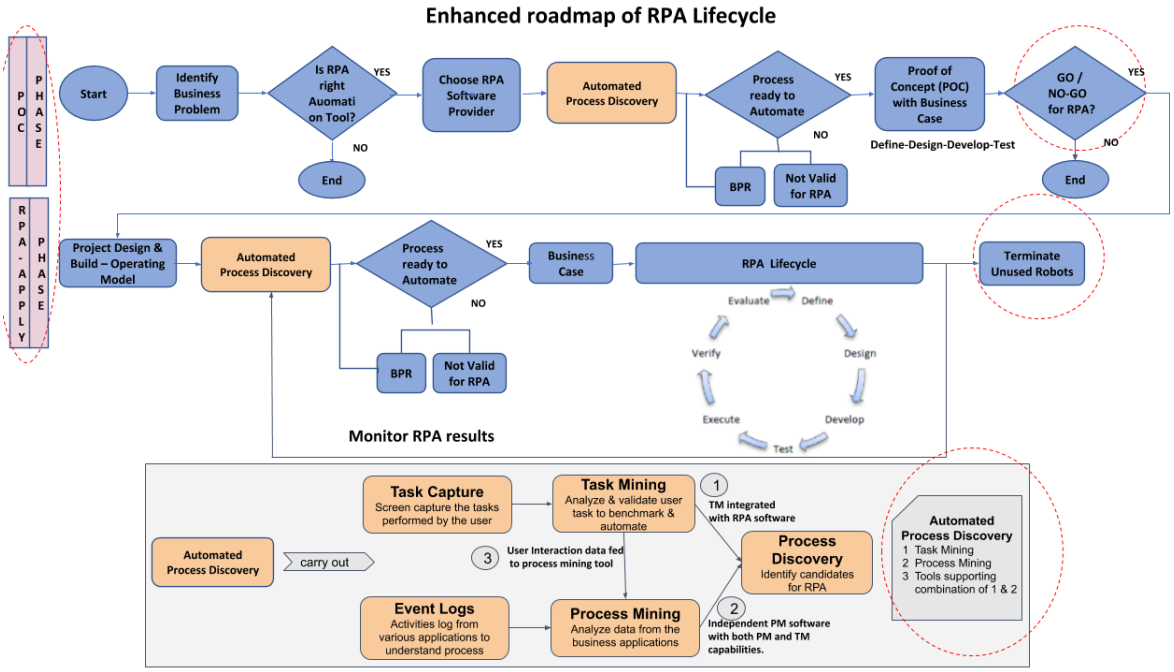

Figure 7.3: *Final Enhanced Dynamic roadmap of RPA Lifecycle using Process Discovery*

Concluding the general response over the usefulness of the model within the organization, person 1 interviewee opined that the model will be beneficial overall, however comparative analysis of mining approach can assist them further to make right choice. The

enhanced roadmap was modified according to the feedback given and following final roadmap is illustrated in figure[7.3] and changes highlighted in the final version. The complete enhanced roadmap of RPA lifecycle using process discovery is presented in Appendix B.1 along with description of individual activities carried out throughout the RPA lifecycle.

7.3 Research Validity

Reliability and validity are the terms used primarily for quantitative paradigm while they are conceptualized as credibility and trustworthiness in the case of qualitative research when collecting, analyzing the data and presenting the findings [42]. Furthermore, even triangulation of information or data where researchers converge their understandings from varied and multiple source of information in order to validate the concept is also a standard norm [42]. This study has adopted a common strategy of triangulation and credibility in order to build this research. Firstly, the research involved study of credible scientific literature from well-known sources in process mining. Secondly, going for multiple case-studies enabled us to validate the findings in diverse industries namely; manufacturing, telecommunication and service industry. Thirdly, interviewing various process experts and RPA managers assisted in re-creating the problem faced by companies to identify eligible processes to automate. Additionally, interviewing people from business process area, RPA management as well as IT increased credibility and provided more data points. Lastly, we also conducted interviews with the actual process and task mining software vendors, increasing the knowledge beyond mere reading of the manuals and watching demo tutorials. Therefore, multiple case-studies, interviews within the organization along with the credible scientific literature available, the researcher could corroborate necessary evidence for the study.

Though the thesis was conducted with utmost integrity and rigor, there are certain instances, wherein we faced limitations and constraints to take into account all the possible factors. Firstly, not all the key stakeholders involved in RPA projects were covered in this thesis. There was no representative interviewed from RPA developer community who could have given a different point of view on RPA. Secondly, the evaluation interviews were conducted with a small subset of interviews to check for its understanding and validity based on their role and experience within the organization. Even though the enhanced roadmap did undergo design iterations based on the feedback given by the interviewees, it has not been implemented or exhibited actual working corroboration. As a result, future research can substantiate this enhanced RPA roadmap by actual implementation in practical settings.

Chapter 8

Conclusion and Recommendations

This section concludes the study by providing answers to all the research questions followed by presenting scientific and practitioner's contributions made by this study. We further give recommendations for further studies and finally reflect over the masters course and the personal knowledge gained during the time of this study.

8.1 Conclusions

This section presents answers to the research question. The main research question is as follows:

How can automated process discovery enhance the overall RPA lifecycle?

Automated process discovery helps to identify suitable processes to automate, prioritize them and monitor their performance thereby improving the overall RPA lifecycle. Building upon the qualitative research, this study presented an updated roadmap of RPA lifecycle in chapter 7. As part of the qualitative research, stakeholders from a large multinational energy company were interviewed to understand the RPA lifecycle in the organization. Majority of the stakeholders expressed their views about use of manual process assessment followed for RPA candidate process selection. Correspondingly, 4 case studies from different industrial backgrounds with process mining and RPA context were analyzed. The analysis revealed the use of automated process discovery guiding automation initiative and emerged as a common approach followed by all the case study companies investigated. Further evidence from the case studies also led to the conclusion that process discovery not only helped in selecting the processes, but also to prioritize them based on the visualised information. The case studies also highlighted the fact that using process mining recommendations, provided added benefit in terms of increase in operational efficiency.

Making the connection between the empirical case studies and the RPA stakeholder interviews, it can be concluded that automated process discovery enhances the RPA lifecycle and increases operational efficiency. Therefore, using the design cycle methodology, the existing dynamic roadmap for RPA was modified to incorporate automated process discovery. The new roadmap was evaluated to ascertain its reliability and efficacy by conducting evaluation interviews. Following the iteration process by incorporating the relevant suggestions and feedback gathered from the evaluation interviews by RPA experts, final roadmap was constructed. A complete enhanced roadmap of RPA lifecycle using process discovery is presented in Appendix B.1 along with description of individual activities carried out throughout the RPA lifecycle. This RPA roadmap can serve as a reference model to company's invovled in process automation. The model also offers explanation about mining approach organizations can consider based on their business objective. In addition the model provides a brief description of individual steps to be performed based on task mining or process mining approach gives a broad understanding of how to incorporate automated process discovery in the overall RPA lifecycle.

The successive research question is as follows -

RQ1 - *What challenges can organisations face to adopt process mining for Robotic Process Automation?*

Lack of definite purpose to implement, data issues and employee attitude are the leading barriers in adopting process mining. Answer to this research question is discussed in chapter 6, and is based on the interviews conducted and the case study analysis performed.

Lack of understanding the purpose of implementing mining technique in the organization came out to be the biggest impediment from the findings. Despite this, the organization can mitigate this barrier to a large extent by clearly identifying the business problem. For example; if a company wants to optimise overall end-to-end business process in order to gain operational benefits, then process mining approach proves to be prudent option. The enhanced RPA lifecycle B.1 presents these approaches and can assist the organization to make a justifiable selection.

Followed by quality of data required for process mining becomes yet another bottleneck as process mining expects data from the system in a particular manner. Therefore data preparation and cleaning of data due to lack of completeness can lead to a long drawn process of actually executing process mining. Another important issue brought forward was the problem of data privacy and the level of confidentiality the tools offer. However, the empirical case studies implemented process mining a couple of years back, the process and task mining technologies have developed further to handle these situations in a much better way. The organization, therefore must be aware of these data and privacy issues and make judicious choice of confidentiality preserving mining tool.

Funding is yet another problem encountered by the organization in order to introduce supporting tools. Nevertheless, organization can follow the systematic method of

enhanced RPA lifecycle rather than performing automation initiatives in an adhoc manner. In doing so, the organization can reap operational benefits which will automatically lead to sustained funding.

Finally, the research also discovered that employee attitude towards introducing process or task mining into their daily routine can be a barrier. Resistance to change, overwhelmed by process transparency or even resourcing issues can add to the impediments in implementing process mining. Training employees in analyzing mining data as well as involving them in selecting the automation activities can build confidence in the employee leading to their attitude change. To summarise, the information provided by the enhanced roadmap of RPA lifecycle can help in mitigating certain barriers whereas in other cases the organization can be cognizant of these issues.

The last research question is as follows -

RQ2 - ***What are the trade-offs between Process Mining and Task Mining the organization can consider while adopting automated process discovery?***

Task mining and process mining provide 2 different ways to discover processes. Organizations based on their primary objective towards business process improvements, can adopt task mining or process mining approach. However, certain factors such as ease of implementation, considerably less costs involved as well as having task mining capability incorporated long with the RPA tool gives benefit to task mining. Similarly, if the business goal of the organization is to reduce operational cost in the business process, then process mining approach can be considered supported by automation initiatives. Process mining on one hand provides an end-to-end perspective of business processes and therefore serves broader benefits than just automation initiatives. However on the other hand, time-consuming and costly implementation process are its drawbacks. Therefore, the choice between TM or PM is a trade-off between the purpose (business goals) it serves and time or cost of implementation it demands.

8.2 Contribution

Collectively, the thesis has made following contribution to the scientific literature and provided practitioner's implications.

8.2.1 Scientific Contribution

This study produced an enhanced roadmap of RPA lifecyle with automated process discovery addressing the issue of manual process selection to decide automation candidates. By doing so, we have contributed to the existing research over RPA lifecycle, grounding the roadmap more solidly in business practice. Though use of automated process discovery in the identifying eligible processes for RPA has been suggested in the recent studies,

further binding this technique into the overall RPA lifecycle to provide data-driven process assessment adds to the scientific contribution of RPA knowledge.

Further, this study lists the contrast between process mining and task mining process discovery approach. To this end, we produce a brief summary of trade-offs between process and task mining implementation. The outcome is the choice between Task mining or Process Mining based on the purpose(business goals) it serves against the time or cost of implementation it demands.

8.2.2 Practitioner's Implications

From a managerial point of view, this study provides a model template an organization can refer to when introducing RPA within the organization. The enhanced roadmap of RPA lifecycle presents all the phases an organization can follow in their automation lifecycle leading to digital transformation. Similarly organizations already executing RPA initiatives can also gain from this model as it provides further explanation about different automated process discovery approaches organizations can employ. Overall, following a structured RPA lifecycle based on evidence-based process assessment, provides more sustained operational value to the organization.

This study also produces a list of challenges in introducing process discovery in the organization. By doing so, the organization can be aware of these barriers and can have a better approach to mitigate the challenges in introducing process discovery.

8.3 Recommendations for future works

The limitations of this study bring up prospects for future studies. The enhanced roadmap of RPA lifecycle using automated process discovery has been evaluated on the basis of expert interviews. Hence, further research can be done by actually testing this roadmap in real-life scenario for its effectiveness and usability.

As discussed in the case studies chapter 5, automated process discovery provides faster processes assessment supported by event data. However, it is seen in the case study analysis that, organization believe that automated process discovery provides faster process assessment. But how large are the actual benefits in terms of the time gained by automated process discovery, needs to be quantified further. Therefore, further research can be done to quantify the benefits provided by automated process assessment over the traditional process assessment like workshop or interviews.

Finally, this study talked about task and process mining approach of automated process discovery to support automation initiatives in the organization. Recently, certain commercial vendors offer an integrated approach using both task and process mining capabilities in a single tool. The combination of capturing both user interactions alongwith the business data increases the transparency of activities performed inside the organization. Further connecting these processes to user interactions by matching the event logs to transac-

tions recorded by the user activities enables organizations to optimize business processes, rather than just productivity. In future, research over case study in practice can further illuminate and clarify the use of tool with integrated mining capabilities.

8.4 Reflections

This specific master thesis is part of Master of Science for ICT in Business offered by Leiden Institute of Advanced Computer Science of Leiden University. The programme is designed specifically to help participants adapt evolving technologies as well as learning how to effectively and responsibly apply these in an organisational context. The thesis in particular studies RPA, an emerging technology whereas process mining, is a data-driven science in the realm of business process management primarily aimed to facilitate process excellence. To that end, the inter-disciplinary nature of the course reflects in every aspect of the research as expected from the master of ICT in business to demonstrate. The course such as managing digital business, process modelling on one hand and applied research methodology at the other end provide a perfect blend of handling evolving technologies like RPA and managing organizational perspectives to adopt them.

Secondly, to reflect upon my individual experience on the thesis, I can say that it was a complete learning curve for me. To start with, RPA as a technology was new to me besides general knowledge on the topic and therefore followed and completed RPA development course which taught me to build and deploy a software robot. By gaining hands-on knowledge, I realised the potentials of RPA as well as I could visualise the challenges accompanied by benefits the organization can possibly gain. Process mining knowledge was gained during the process modelling course and further increased my understanding by following various online tutorials created a strong base. Furthermore, courses like applied research methodology prepared me for qualitative research as well as coding analysis. A mock interview during the coursework assisted me not to fall prey of certain pitfalls while conducting interviews. However, I can now reflect upon the fact that the semi-structured interview could have been designed better to channel the answers towards specific areas of concern.

On the whole, I believe that the research over RPA roadmap does contribute scientifically in the RPA area and at the same time presents practical recommendation to the organization adopting automated process discovery in their process automation lifecycle.

Bibliography

[1] IEEE Guide for Terms and Concepts in Intelligent Process Automation. *IEEE Std 2755-2017*, pages 1–16, Sep 2017. doi:10.1109/ieeestd.2017.8070671.

[2] B. Agaton and G. Swedberg. Evaluating and developing methods to assess business process suitability for robotic process automation-a design research approach. Master's thesis, 2018.

[3] A. D. Andrade. Interpretive research aiming at theory building: Adopting and adapting the case study design. *The qualitative report*, 14(1):42, 2009.

[4] A. Asatiani and E. Penttinen. Turning robotic process automation into commercial success–Case OpusCapita. *Journal of Information Technology Teaching Cases*, 6(2):67–74, 2016.

[5] A. Augusto, R. Conforti, M. Dumas, M. La Rosa, F. M. Maggi, A. Marrella, M. Mecella, and A. Soo. Automated discovery of process models from event logs: Review and benchmark. *IEEE Transactions on Knowledge and Data Engineering*, 31(4):686–705, 2018.

[6] Automation Anywhere. Bot lifecycle management. URL https://www.automationanywhere.com/solutions/enterprise-bot-lifecycle-management.

[7] A. Berti. Task mining, Sep 2019. URL http://pm4py.pads.rwth-aachen.de/task-mining/.

[8] B. CBOK. Guide to the business process management common body of knowledge. *Versão 2.0. 2009. Disponível em: www. abpmp. org. Acesso em: 25 nov 2012*, page 21, 2009.

[9] Celonis. Process discovery: Celonis intelligent business cloud: Process mining solutions. URL https://www.celonis.com/intelligent-business-cloud/process-discovery/.

[10] K. Chatterjee, K. Sicinska, and H. J. Ehrlich. Tempted to rewrite bill gates' rules on automation?, Dec 2017. URL https://www.capgemini.com/2015/01/tempted-to-rewrite-bill-gates-rules-on-automation/.

[11] J. W. Creswell and J. D. Creswell. *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications, 2017.

[12] J. W. Creswell and C. N. Poth. *Qualitative inquiry and research design: Choosing among five approaches*. Sage publications, 2016.

[13] M. Dumas, M. La Rosa, J. Mendling, H. A. Reijers, et al. *Fundamentals of business process management*, volume 1. Springer, 2013.

[14] D. Dutta, A. Gillard, and G. Kaczmarskyj. Get ready for robots - ey, 2016. URL https://www.ey.com/Publication/vwLUAssets/Get_ready_for_robots/%24FILE/ey-get-ready-for-robots.pdf.

[15] Edgeverve. Process discovery: Assistedge discover. URL https://www.edgeverve.com/assistedge/process-discovery/.

[16] K. M. Eisenhardt. Building theories from case study research. *Academy of management review*, 14(4):532–550, 1989.

[17] J. Geyer-Klingeberg, J. Nakladal, F. Baldauf, and F. Veit. Process mining and robotic process automation: A perfect match. In *BPM (Dissertation/Demos/Industry)*, pages 124–131, 2018.

[18] V. Goris. Robotic process automation: An assessment of process discovery techniques with the purpose of finding rpa eligible processes. Master's thesis, 2019.

[19] M. Gough and D. Kirk. Capitalizing on robotics driving savings with digital labor, 2017. URL https://assets.kpmg/content/dam/kpmg/my/pdf/capitalizing-robotics-digital-labor-savings.pdf.

[20] M. Halaweh, C. Fidler, and S. McRobb. Integrating the grounded theory method and case study research methodology within is research: A possible'road map'. *ICIS 2008 proceedings*, page 165, 2008.

[21] T. Hess, C. Matt, A. Benlian, and F. Wiesböck. Options for formulating a digital transformation strategy. *MIS Q. Executive*, 15, 2016.

[22] A. R. Hevner. A three cycle view of design science research. *Scandinavian journal of information systems*, 19(2):4, 2007.

[23] J. Hindle, M. Lacity, L. Willcocks, and S. Khan. Robotic process automation: Benchmarking the client experience. *Knowledge Capital Partners*, 2018.

[24] IRPA. Introduction to robotic process automation a primer, 2015. URL https://irpaai.com/wp-content/uploads/2015/05/Robotic-Process-Automation-June2015.pdf.

[25] J. Juttman and M. V. Doesburg. Robotic process automation: how to move on from the proof of concept phase?, 2018. URL https://www.compact.nl/articles/robotic-process-automation-how-to-move-on-from-the-proof-of-concept-phase/.

[26] C. R. Kothari. *Research methodology: Methods and techniques*. New Age International, 2004.

[27] Kryon. Introduces Process Discovery - The Start of an Enterprise Revolution - removing barriers to organization-wide Automation across any vertical, May 2020. URL https://www.kryonsystems.com/kryon-introduces-process-discovery/.

[28] T. Kyheröinen et al. Implementation of robotic process automation to a target process–a case study. 2018.

[29] M. Lacity, L. Willcocks, and A. Craig. Robotizing global financial shared services at royal dsm. *The Outsourcing Unit Working Research Paper Series*, 2016.

[30] M. C. Lacity and L. P. Willcocks. A new approach to automating services. *MIT Sloan Management Review*, 2017.

[31] P. M. Lacity, P. L. Willcocks, and A. Craig. Robotic process automation: Mature capabilities in the energy sector. Oct 2015.

[32] S. Lazarus. Achieving a successful robotic process automation implementation: A case study of vodafone and celonis, 2018. URL https://spendmatters.com/2018/06/07/achieving-a-successful-robotic-process-automation-implementation-a-case-study-of-vodafone-and-celonis/.

[33] V. Leno, M. Dumas, F. Maggi, and M. La Rosa. Multi-perspective process model discovery for robotic process automation. In *CEUR Workshop Proceedings*, volume 2114, pages 37–45, 2018.

[34] V. Leno, A. Polyvyanyy, M. Dumas, M. La Rosa, and F. M. Maggi. Robotic process mining: Vision and challenges. *Business & Information Systems Engineering*, pages 1–14, 2020.

[35] C. Linn, P. Zimmermann, and D. Werth. Desktop activity mining-a new level of detail in mining business processes. In *Workshops der INFORMATIK 2018-Architekturen, Prozesse, Sicherheit und Nachhaltigkeit*. Köllen Druck+ Verlag GmbH, 2018.

[36] S. Moore. Gartner says worldwide robotic process automation software market grew 63% in 2018, Jun 2019. URL https://www.gartner.com/en/newsroom/press-releases/2019-06-24-gartner-says-worldwide-robotic-process-automation-sof.

[37] M. D. Myers and M. Newman. The qualitative interview in is research: Examining the craft. *Information and organization*, 17(1):2–26, 2007.

[38] N. Ostdick. The evolution of robotic process automation (rpa): Past, present, and future, Oct 2019. URL https://www.uipath.com/blog/the-evolution-of-rpa-past-present-and-future.

[39] G. Paré and J. J. Elam. Using case study research to build theories of it implementation. In *Information systems and qualitative research*, pages 542–568. Springer, 1997.

[40] L. Reinkemeyer. *Process Mining in Action: Principles, Use Cases and Outlook*. Springer Nature, 2020.

[41] P. b. rpanerds. Delivery road-map or life cycle of blue prism process, Mar 2018. URL https://rpanerds.wordpress.com/2018/03/01/delivery-roadmap-or-life-cycle-of-blue-prism-process/.

[42] J. Saldana. Fundamentals of qualitative research: Oxford university press. 2011.

[43] G. L. Sigurðardóttir. Robotic process automation: Dynamic roadmap for successful implementation. 2018.

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

[45] A. Thornhill, M. Saunders, and P. Lewis. *Research methods for business students*. Prentice Hall: London, 2009.

[46] A. Trefler. The big rpa bubble, Dec 2018. URL https://www.forbes.com/sites/cognitiveworld/2018/12/02/the-big-rpa-bubble/.

[47] UiPath. Introduction to task mining. URL https://docs.uipath.com/task-mining/docs/introduction.

[48] W. M. van der Aalst, M. Bichler, and A. Heinzl. Robotic process automation, 2018.

[49] W. M. van der Aalst, Wil M. P. *Process Mining Data Science in Action*. Springer Berlin, 2nd edition, 2018.

[50] J. Watson and D. Wright. www2.deloitte.com, 2017. URL https://www2.deloitte.com/content/dam/Deloitte/tr/Documents/technology/deloitte-robots-are-ready.pdf.

[51] R. J. Wieringa. *Design science methodology for information systems and software engineering*. Springer, 2014.

[52] L. Willcocks, M. Lacity, and A. Craig. Robotic process automation: strategic transformation lever for global business services? *Journal of Information Technology Teaching Cases*, 7(1):17–28, 2017.

[53] L. Willcocks, J. Hindle, and M. Lacity. Keys to rpa success. Technical report, Executive Research Report. Knowledge Capital Partners, 2018.

[54] L. P. Willcocks and M. Lacity. *Service automation robots and the future of work*. SB Publishing, 2016.

[55] R. K. Yin. Case study research: Design and methods. In *Applied Social Research Methods Series*, volume 5. Sage Publications, 1989.

[56] R. K. Yin. *Qualitative research from start to finish*. Guilford publications, 2015.

Acknowledgements

I would like to dedicate this work to my parents and family, especially my mother who has encouraged me to never stop studying. To my husband and my daughters who have always been there and spared no effort to support me and continuously encouraged me throughout my master studies.

My special appreciation goes to my supervisor, Dr. Joost Visser for his wonderful mentorship, insightful guidance, thorough review and patience from the very beginning until the end. My thank you to the second supervisor Dr Werner Heijstek who provided guidance on qualitative aspects of the thesis. I would also like to thank Esme Caube for helping me navigate through the administrative process in my masters programme.

I also extend my gratitude to Mr Paul van Leeuwen for a helping hand to get an internship in Vattenfall. My sincerest gratitude and a special mention of my supervisor within Vattenfall, Mr Jean Louis Colen for his honest and practical feedback that was valuable to this research. His guidance and extended support throughout the research was commendable; ensuring access to all the necessary information even during the strange times.

Finally, my gratitude towards my friends who believed in my ability to accomplish this goal and all those within Vattenfall and outside who helped me in data collection and interviews during my research.

Appendix A

Interview Protocol

Interviews were conducted twice, once in the beginning to understand the RPA journey as well as mining tool interviews to gain knowledge about the various tools available in the market with reference to the process automation selection.

A.1 Interviews in the Organizaton

Following was the overall setup of the semi-structured interviews conducted with the process expert, RPA programme manager and ICT manager.

Introduction

1. What is your role and responsibility within the organization?

RPA Lifecycle

The following questions were asked to understand RPA journey in the organization. Out of which some of the question were one-liners providing statistical answers and others needed elaboration about the process followed.

1. As part of business process management(BPM), was there any Process automation done within the business process before RPA?
2. Were you familiar with RPA, before it was considered as part of Digitalization Strategy?
3. What motivated you to adopt RPA in your process?
4. Was there any POC implemented for RPA within the organization?
5. What was the selection criteria for potential automation using RPA?
6. How was the process assessment carried out, to identify processes for automation? Any consideration given to business process standardization before considering activity/activities for RPA?

7. Amount of time spent in process assessment exercise to identify processes for automation?
8. Did the robots fulfill the expectations initially thought they would, explain? (in terms of Process Benefit, Financial, RPA complexity, any other)
9. What challenges do the Robots pose now, post implementation?
10. Does your organization follow any specific Operating model to implement and manage RPA, Please explain?

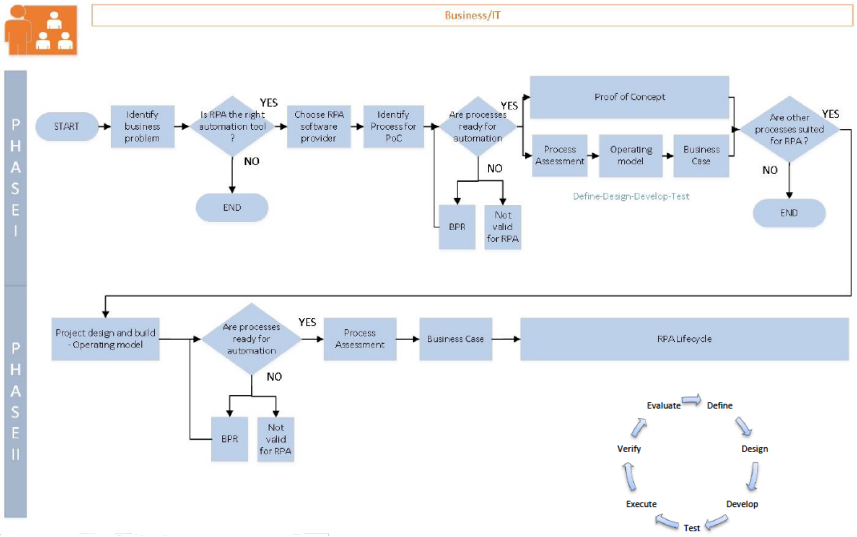

Figure A.1: *RPA : Dynamic Roadmap for successful implementation, [43]*

11. Above figure is a roadmap of RPA implementation, a reference from scientific literature . To what extent does this roadmap conform to your organization's RPA processes?

Process Mining

Questions pertaining to process mining were asked to find out existing knowledge over PM within the organization. Also, some questions were specific to a process expert and ICT manager who were involved in mining POC performed independently couple of years back.

1. Currently, is there any kind of tool used to analyze business processes? If yes, What kind of information or output does the tool provide you?
2. Do you see any benefits of using Process mining tool as a part of BPM activities?

Version of August 14, 2020– Created August 14, 2020 - 15:32

3. Would you consider a process mining tool to analyze your Business processes and possible automation suggestions?
4. How ready do you think are you to implement process mining software in terms of Input data(event logs), Knowledgeable Resources?
5. What challenges/ impediments do you think in implementing process mining as a part of process Lifecycle?

A.2 Interview with mining software vendor

Semi-structured interviews were conducted with 2 Mining tool vendors namely; Celonis and UiPath. One of them offered an independent mining tool whereas the other offered task mining along with RPA tool. Following questions were discussed as part of the after the initial introductory questions.

1. How does Process Mining help in identifying automation processes and reduce time in process assessment activity?
2. What challenges/ impediments do you think your clients/organization face implementing process mining as a part of process lifecycle?
3. Does independent process mining software have the capability to integrate/connect with any RPA software to execute the bots directly?
4. What is the advantage to have a standalone PM software like Celonis over task mining software integrated with RPA tool? or vice-a-versa?
5. What is the license model for the mining tool, process based , user based plus maintenance?

A.3 Roadmap Evaluation Interviews

Semi-structured interviews were conducted for roadmap evaluation by presenting both original as well as the enhanced RPA roadmap to RPA programme Manager and Process Expert. The feedback given on the enhanced roadmap was incorporated to arrive at the final version.

1. Is the presented roadmap consistent to the RPA lifecycle in general?
2. Can you understand and follow the roadmap just by going through the individual entities?

Version of August 14, 2020– Created August 14, 2020 - 15:32

3. Do you think that certain elements are irrelevant or overlapping or even the order of the process?
4. Will this model be useful in Vattenfall? Why?
5. Do you see any benefits of the model?
6. What according to you are the limitations and weaknesses in the model?
7. Do you have any other suggestions?

Appendix B

Enhanced roadmap of RPA lifecycle

The description of the model; enhanced roadmap of RPA lifecycle will be characterized in this section :

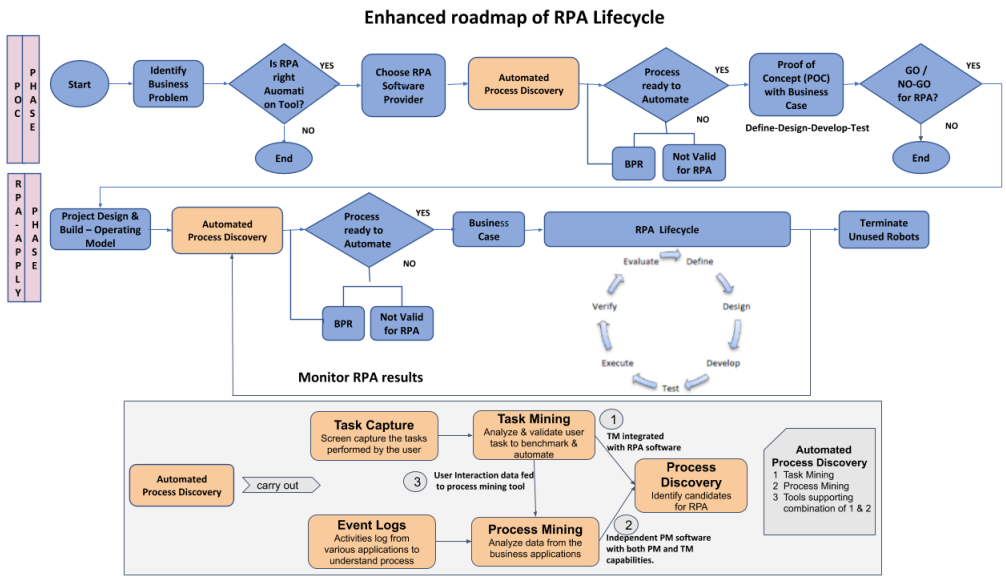

Figure B.1: *Enhanced roadmap of RPA lifecycle*

POC Phase

This is the initial phase where organizations want to test RPA feasibility and viability within the business. Following steps are performed in this phase :

- **Identify business problem :** The first step involves identifying the business problem, challenges, obstacles or tasks that make the business process inefficient. One way of identifying business problem is having a periodic review of respective business units.

- **Is RPA right solution :** Once organization identify the problem, they can begin looking into how to solve the problem. Based on the characteristics of the activities performed, the organization can consider automation, making RPA a potential solution.

- **Choose RPA software vendor :** The next step is to select a RPA software vendor from the array of RPA software available in the market. Commercial RPA tools are available with to build and execute software robots and automate the processes. In certain cases, additional capabilities like automated process discovery and building a execution routine based on the captured user actions add value to the software.

- **Automated process discovery :** This particular step involves performing process discovery in order to identify the process for automation. However, the activities to be performed in the automated process discovery phase are explained in detail in the section below.

- **Process ready for automation :** Based on the guideline given by the automated process discovery, the process expert makes the decision whether the process should be automated or can be improved by business process redesign. Redesigning processes can be done by eliminating steps that are not needed or performed in completely different manner. Therefore, the process may no longer require further automation.

- **POC with business case :** At this stage, the process expert presents a business case for the automation initiative based on the recommendation given by automated process discovery. Business case can be based on the potential FTE saved and cost reduction that can be achieved by automating the task or activities. On approval, the actual proof-of-concept is executed following the define, design, develop and testing RPA initiative.

- **Go/No-Go for RPA :** Based on the success of executing automation POC, organization can decide to continue with further automation implementations or not. This completes one cycle or the first phase of RPA implementation.

RPA Apply Phase

This is the second phase of RPA where PoC meets the required business value and the organization see an opportunity for other processes to be automated with a robotic software. Following steps are performed in this phase :

- **Project Design & Build operating model -** This step is the beginning of phase 2 in the overall RPA lifecycle. Further automation initiatives creates the need to have a suitable operating model in order to govern the its activities. It can be accomplished by forming RPA centre of excellence, a common platform to carry out automation initiatives or execute the automation independently on project basis.

- **Automated process discovery :** This step is similar to the automated process discovery activity in phase 1. It becomes a routine activity to perform process discovery in order to find the eligible process to automate.
- **Business Case :** Any potential automation initiative demands a business case presenting the need for automation. Business case also presents the cost-benefit analysis that can be achieved by automating the task or activities.
- **RPA lifecycle :** Once the business case is approved, the automation project can follow similar activities like any other software development lifecycle of define, design, develop, test and execute. Moreover, this stage also includes monitoring the performance of the software robots deployed into production to evaluate automation benefits.
- **Terminate unused robots :** In certain situations, there is no further utility value of the software robot due to process change or the activities not performed at all. In such cases, the unused robots needs to be terminated.

The second phase in the model is in a continuous loop where the organizations identify automation activities using process discovery. Additionally, the software robots in operation are monitored to evaluate their performance and necessity.

Below the RPA phases, the model also presents automated process discovery approaches, the organization can adopt. Primarily based on the business goal of the organization, the company can decide on using process or task mining approach for discovering processes in the RPA lifecycle. Expansion of the activities carried out within automated process discovery are briefly explained below :

1. Automated Process discovery using Task mining

- **Task capture :** In order to capture the tasks performed by individual users, Task mining client is installed on individual user desktops or workstations. The task capture program records all the actions performed by the user as per the applications listed in the recorder. The task capture tool is linked to the Admin Console where the data from all the users can be compiled together.
- **Task mining :** Process expert can converge the user recordings captured for a particular process and visualize the process map to identify deviations from the standard process or any inefficiencies in the activities performed.
- **Process Discovery :** The Task mining dashboard presents detailed information about the captured tasks in terms of minimum and maximum time taken, number of users, tasks performed. It also displays automation suitability for a particular process, considering the complexity of the automatization and the value it might bring.

2. Automated Process discovery using Process mining

- **Event Logs :** Transactional event logs is collected real-time API or uploaded through ETL software various enterprise systems within a organization. The event logs collected from various business applications are stored onto a server and further curated as per the process mining requirement.
- **Process mining :** Based on their user access, process analyst can access and view the process visualization on the process mining dashboard. The process flow displays all the possible deviations for a particular process. Various other details such as re-work activities, process statistics, automation rate and other aspects related to the process are displayed on the dashboard.
- **Process discovery :** Based on the event log, the mining analysis displays automation suggestions and it's benefit it terms of time saved or deviations mitigated. These inputs can be used by the process expert to automate processes and overcome delays or to adhere to process compliance, thereby optimizing the overall process. Moreover, based on the automation rate for the activities within the visualized process, process expert can prioritize the activities to be automated.

3. Automated Process discovery using Process and Task mining

Certain commercial tools offer an integrated approach using both task and process mining capabilities in a single tool. The combination of capturing both user interactions along with the business data increases the transparency of activities performed inside the organization. Connecting processes to user interactions by matching case ids in event logs to transaction ids in recorded user activities enables organizations to optimize business outcomes, rather than just productivity.

Version of August 14, 2020– Created August 14, 2020 - 15:32