BELINDA MYTEBERI

GRADUATION PROJECT

2020

HOW PROCESS MINING AND RPA HELP IN OPTIMIZING BUSINESS PROCESSES

GRADUATION PROJECT

BELINDA MYTEBERI

EPOKA UNIVERSITY

FACULTY OF ECONOMICS AND ADMINISTRATIVE SCIENCES

DEPARTMENT OF BUSINESS ADMINISTRATION/BUSINESS INFORMATICS PROGRAM

TIRANA, ALBANIA JULY 2020

HOW PROCESS MINING AND RPA HELP IN OPTIMIZING BUSINESS PROCESSES

BELINDA MYTEBERI

Thesis submitted in Fulfillment of Requirement for the Degree of Bachelor's in Business Informatics

EPOKA UNIVERSITY

2020

APPROVAL PAGE

Student Name & Surname: Belinda Myteberi

Faculty: Faculty of Economics and Administrative Sciences

Department: Business Administration (Business Informatics Program)

Thesis Title: HOW PROCESS MINING AND RPA HELP IN OPTIMIZING BUSINESS PROCESSES

Date of Defense: 21.07.2020

I certify that this final work satisfies all the requirements as a Graduation Project Thesis for the degree of Bachelor in Business Informatics.

……

Head of Department

This is to certify that I have read this final work and that in my opinion it is fully adequate, in scope and quality, requirements as a Graduation Project Thesis for the degree of Bachelor in Business Informatics.

…….

Supervisor

Examination Committee Members

Title / Name & Surname	Affiliation Signature
1-	
2-	
3-	

DEDICATION

To my family and friends who have always been there for me!

HOW PROCESS MINING AND RPA HELP IN OPTIMIZING BUSINESS PROCESSES

ABSTRACT

Business Process Optimization has always been the main goal of every organization worldwide no matter its nature. With the innovation in technologies, the amount of data has increased exponentially and the old data-driven approaches for business optimization analyses have no capability to deal with it. Reality is different from expectations. We can plan how a specific process must happen, when, who is supposed to do what and under which condition, but we can not plan other factors that affect the process such as human capability, changes of different factors from the environment and so on, making it impossible to have a clear view of the organization processes.

Process Mining is the solution. It analyses event data as a whole, taking into account every detail and creating a realistic map of process flow. For further digitalization of internal processes, Robotic Process Automation is being used, creating a virtual workforce to automatize manual, repetitive and error-prone tasks.

This paper provides insights on what is Process Mining, how it is being used by companies worldwide to uncover hidden problems such as inefficiencies, unusual transactions, deviations, bottlenecks, risk and weakness by presenting objective insights and how it can be used to help in every stage of RPA implementation from identification of what task is needed to be automated, to training of the robots and continuous monitoring of their execution.

Keywords: Process Mining, Process Optimization, Robotic Process Automation, Process Flow

SI NDIHMOJNË MINIMI I PROÇESVE DHE AUTOMATIZIMI ROBOTIK I PROÇESEVE NË OPTIMIZIMIN E PROÇESEVE TË BIZNESIT

ABSTRAKT

Optimizimi i Proçeseve të Biznesit ka qenë dhe do mbetet gjithmonë qëllimi kryesor i çdo organizate pavarësisht natyrës së saj. Me inovacionet në teknologji, sasia e të dhënave është rritur eksponencialisht dhe metodat e vjetra të bazuara në të dhëna për optimizimin e proçeseve nuk mund ta arrijnë këtë gjë.
Realiteti është i ndryshëm nga pritshmëritë. Ne mund të planifikojmë sesi një proçes specifik mund te ndodh, kur, kush supozohet të bëj çfarë dhe në çfarë rrethanash, por ne nuk mund të parashikojmë faktorët e tjerë që ndikojnë në proçese si aftësitë njerëzore, ndryshimet në mjedisin e jashtëm etj, duke e bërë të pamundur krijimin e një pamje të qartë të proçeseve.

Minimi i Proçeseve është zgjidhja. Me të mund të analizohen të dhënat e eventeve duke marrë parasysh çdo aspect dhe mund të krijohet një hartë realiste e rrjedhës së proçeseve. Për dixhitalizim të mëtejshëm të proçesve të brendshëm përdoret Automatizimi Robotik i Proçeseve, i cili krijon një krahë pune virtual për të automatizuar punë manuale dhe të përsëritshme.

Kjo temë ofron një prezantim të konceptit të Minimit të Proçeseve, si po përdoret nga kompanitë kudo nëpër botë për të zbuluar problemet e fshehura si devijimet, risqet dhe pikat e dobëta dhe sim und të përdoret gjatë çdo faze të implementimit të RPA nga identifikimi i çdo pune që duhet automatizuar, tek trajnimi i robotëve dhe monitorimi i vazhdueshëm gjatë ekzekutimit të tyre.

Fjalë kyçe: Minimi i Proçeseve, Optimizimi i Proçeseve, Automatizimi Robotik i Proçeseve, Rrjedha e Proçeseve

ACKNOWLEDGEMENTS

For the realization and the making of this thesis, I would sincerely like to thank my advisor of the thesis, professor Igli Hakrama for introducing me to the field of process mining and helping me throughout all the way from the beginning to the very end of it with his valuable directions and feedback. His effort and continuous encouragement pushed me to work hard in order to achieve the results that I did.

The last words go to my family and best friends for supporting and helping me successfully finish this thesis.

DECLARATION

I hereby declare that this graduation project, titled "How Process Mining and RPA help in optimizing Business Processes", is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that this thesis has not been previously or concurrently submitted for any other degree, at Epoka University or any other university or institution.

Belinda Myteberi

July 2020

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

PM	Process Mining
RPA	Robotic Process Automation
AI	Artificial Intelligence
DigitalOps	Digital Operations

CHAPTER 1

INTRODUCTION

We are living in the era of digitalization which means we are generating data all the time. In the last 10 minutes, we have generated more data than from prehistoric times until 2003. (van der Aalst, n.d.) The expenditure of this digital universe is well-aligned with processes in every organization, making it possible to record and analyze events. Events may vary from a scanning of barcodes in a supermarket, a student applying for a scholarship, booking of a plane ticket, answering a phone call, and so on.

The world runs into processes and all customers' experiences and business outcomes are products of invisible data flow and interactions of IT systems, people, and assets. Data is the new oil and the main challenge of organizations nowadays is to be able to extract value and information from data that are being stored in their information systems. The exponential growth of data is not the only reason why information systems are a must in organizations, but also the role of them in business processes as the physical universe is becoming more aligned with the digital one.

Businesses have opened their eyes and are coming to the conclusion that their most important competitive advantage they can have is creating a clear view of their operational details. They should be able to use event data in the most meaningful way in order to create an overall overview of their organization, to be able to identify bottlenecks, to see if everything is going according to their plans or if there are deviations happening.

Frederick Taylor Winslow was the first person to research and find ways to optimize workplace productivity. In 1911 he proposed that the business's core

operations should be analyzed, standardized, and improved on, so process improvement is not a new idea but the complexity of the modern business process environment has accelerated beyond the capabilities of traditional tools. Many people use spreadsheets and macros to analyze data but they forget that spreadsheets are not suitable to analyze event data. Process mapping software, interviews or post-it notes cannot deliver continuous process excellence and are lengthy, costly, offer one-time understanding and are subjective.

Process Mining is the solution. It is purpose-built to handle the inherent complexity and dynamism of the modern process environment. It delivers deep visibility and control into the minutiae of individual processes, the relationships between them, and the outcomes they deliver. That's the true value of Process Mining, not just in understanding how to make discrete parts of the business more efficient, but how to calibrate the individual components to optimize the whole operational engine for specific outcomes and also create propositions about improvements such as which processes need to be automated introducing us with RPA possibilities.

According to Gartner predictions the RPA market will top $1 billion by 2020 with a 63% growth rate from 2018 to 2019 and 40% of large enterprises will have integrated robots into their business processes making RPA the fastest-growing segment of the global enterprise software market. (Kerremans, 2019) However an Ernst & Young study estimated that 30% to 50% of all first time RPA initiatives fail and that is why I have decided to make a detailed report on how companies can use PM in order to avoid RPA failures. (Young, 2016)

My main research questions are:

What is process mining and how companies can use process mining to achieve frictionless processes?

and

How Process Mining helps to have a successful RPA implementation?

This paper presents an analysis of process mining techniques developed by academia and commercial vendors and it also represents an analysis of real use cases of PM and RPA in the real world. The research was based on several studies about Process Mining, RPA, surveys, different webinars where companies shared their experience in PM and RPA implementation.

1.Process Mining

According to W. van der Aalst, the godfather of PM, Process Mining is a young research discipline that brings together traditional model-based process analysis like simulation and data-centric analysis techniques like data mining and machine learning as shown in figure 1.1. Statistics, data mining, machine learning alone do not consider end-to-end process models meanwhile process science approaches are process-centric, but they tend to focus on modeling not on real event data.

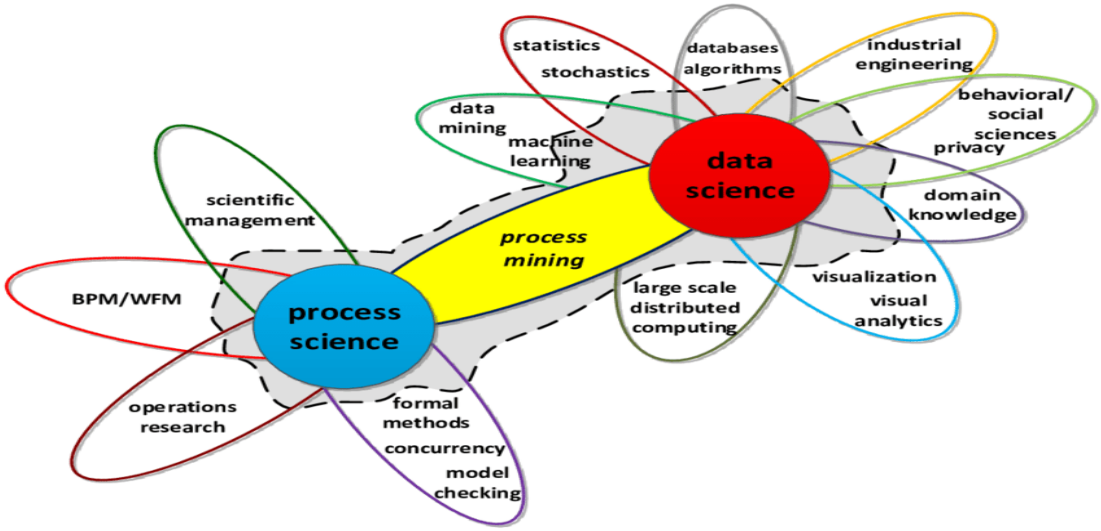


Figure 1.1. Process Mining, the bridge between data science and process science

A process is a series of actions or steps taken in order to achieve a particular end, a repeatable pattern which we try to execute over and over again, always trying to optimize to achieve the best outcome but in real life, assumed processes are very different from the AS-IS processes. The idea behind process mining is discovering,

monitoring and improving the AS-IS processes by analyzing event logs data available in the company information systems.

Event logs used during this thesis are from a loan application process of a Dutch financial institution. (van Dongen, B.F. (Boudewijn) (2017) BPI Challenge 2017 - Offer log. Eindhoven University of Technology. Dataset .https://doi.org/10.4121/uuid:7e326e7e-8b93-4701-8860-71213edf0fbe) More details about the process would be explained in the next chapter. In the figure 1.2 the model process of offering a loan is represented. As we can see there are no deviations but in real life (Figure 1.3) the process is being executed in a different way.

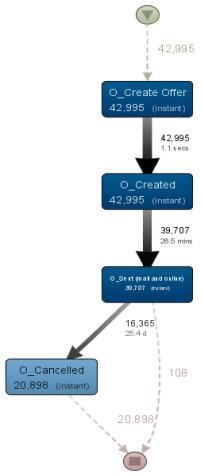


Figure 1.2 Modeled Process of Loan Offer Process

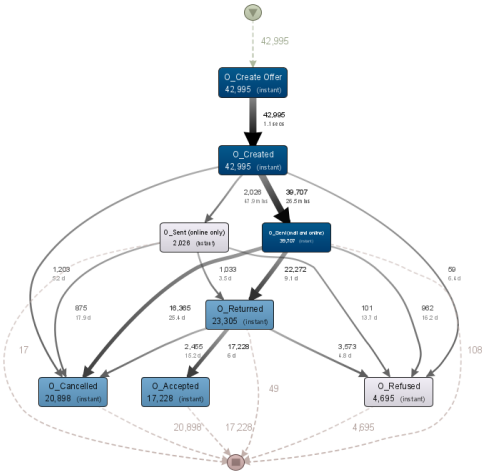


Figure 1.3 The AS-IS Process using event logs

With Process Mining questions like the ones below and many more can find the answer.

- What are the paths of my processes, how do they deviate from the reference process, what are the frequencies of these deviations?
- What have in common processes that last longer than 2 months?
- Where are the bottlenecks behind the delays?
- Where and why do people/machines deviate from expectations?

According to MarketWatch, “the global process analytics market size is expected to grow from $185.3 million in 2018 to $1.42 billion by 2023 and process mining is set to skyrocket.” (MarketWatch)

There are different tools in the market providing Process Mining such as Celonis, Fluxicon, Prom, Minit, Uipath, Perceptive, Rialto, QPR, SNP, Worksoft and so on.
Process Mining early adopters are retailers, telecommunication companies, finance companies (Rosik) and they all have admitted that they saved millions thanks to learnings mined from process data as will be explained in the chapter 4.

CHAPTER 2

STAGES OF PROCESS MINING

2.1 Event Logs

Starting point for process mining is an event log. As W. van der Aalst states, we should treat event logs as first citizens because they are the most valuable asset we have in our hands. Every interaction inside the transactional systems that an organization runs on (i.e SAP) leaves digital footprints, which are raw data that can be turned into a record of actions known as event logs. (van der Aalst, Process Mining Data Science in Action)

Each event should have a case id to identify the uniqueness of the case, activity name or description, timestamps and other valuable extra information such as name of an employee or other information. One example of an event log is shown in the figure below (figure 2.1)

```
Case ID,  Activity,  Resource,  Complete Timestamp,  Variant,  Variant index,  (case) Accepted,  (case) ApplicationID,  (case) CreditScore,(case) FirstWithdrawalAmount,(case) MonthlyCost,(case) NumberOf
Offer_247135719,O_Create Offer,User_17,2016/01/02 10:17:05.720,Variant 8,8,true,Application_196483749,0,10000.0,201.76,57,10000.0,false,Created,Offer_247135719,Offer,,complete
Offer_247135719,O_Created,User_17,2016/01/02 10:17:08.762,Variant 8,8,true,Application_196483749,0,10000.0,201.76,57,10000.0,false,statechange,OfferState_124849367,Offer,Offer_247135719,complete
Offer_247135719,O_Sent (online only),User_17,2016/01/02 10:19:21.330,Variant 8,8,true,Application_196483749,0,10000.0,201.76,57,10000.0,false,statechange,OfferState_440662877,Offer,Offer_247135719,complete
Offer_247135719,O_Cancelled,User_17,2016/01/02 10:21:26.034,Variant 8,8,true,Application_196483749,0,10000.0,201.76,57,10000.0,false,statechange,OfferState_591416028,Offer,Offer_247135719,complete
Offer_941964966,O_Create Offer,User_17,2016/01/02 10:21:42.022,Variant 1,1,true,Application_196483749,0,4100.0,201.76,57,10000.0,false,Created,Offer_941964966,Offer,,complete
Offer_941964966,O_Created,User_17,2016/01/02 10:21:43.573,Variant 1,1,true,Application_196483749,0,4100.0,201.76,57,10000.0,false,statechange,OfferState_275382868,Offer,Offer_941964966,complete
Offer_941964966,O_Sent (mail and online),User_17,2016/01/02 10:22:09.421,Variant 1,1,true,Application_196483749,0,4100.0,201.76,57,10000.0,false,statechange,OfferState_563158836,Offer,Offer_941964966,complete
Offer_941964966,O_Cancelled,User_1,2016/02/29 08:01:05.256,Variant 1,1,true,Application_196483749,0,4100.0,201.76,57,10000.0,false,statechange,OfferState_293765867,Offer,Offer_941964966,complete
Offer_1148420274,O_Create Offer,User_17,2016/01/02 10:26:43.598,Variant 1,1,true,Application_196483749,0,0.0,120.0,58,6000.0,false,Created,Offer_1148420274,Offer,,complete
Offer_1148420274,O_Created,User_17,2016/01/02 10:26:44.925,Variant 1,1,true,Application_196483749,0,0.0,120.0,58,6000.0,false,statechange,OfferState_1323744692,Offer,Offer_1148420274,complete
Offer_1148420274,O_Sent (mail and online),User_17,2016/01/02 10:26:57.389,Variant 1,1,true,Application_196483749,0,0.0,120.0,58,6000.0,false,statechange,OfferState_1151067844,Offer,Offer_1148420274,complete
Offer_1148420274,O_Cancelled,User_1,2016/02/29 08:01:05.248,Variant 1,1,true,Application_196483749,0,0.0,120.0,58,6000.0,false,statechange,OfferState_584641125,Offer,Offer_1148420274,complete
Offer_1365106765,O_Create Offer,User_17,2016/01/02 10:55:46.369,Variant 2,2,true,Application_1120819670,1021,6900.0,150.73,53,6900.0,true,Created,Offer_1365106765,Offer,,complete
Offer_1365106765,O_Created,User_17,2016/01/02 10:55:47.693,Variant 2,2,true,Application_1120819670,1021,6900.0,150.73,53,6900.0,true,statechange,OfferState_1700055433,Offer,Offer_1365106765,complete
Offer_1365106765,O_Sent (mail and online),User_17,2016/01/02 10:59:50.268,Variant 2,2,true,Application_1120819670,1021,6900.0,150.73,53,6900.0,true,statechange,OfferState_16123997,Offer,Offer_1365106765,complet
Offer_1365106765,O_Returned,User_116,2016/01/08 11:00:24.474,Variant 2,2,true,Application_1120819670,1021,6900.0,150.73,53,6900.0,true,statechange,OfferState_608790984,Offer,Offer_1365106765,complete
Offer_1365106765,O_Accepted,User_95,2016/01/11 10:42:07.001,Variant 2,2,true,Application_1120819670,1021,6900.0,150.73,53,6900.0,true,statechange,OfferState_1067718689,Offer,Offer_1365106765,complete
```

2.1 Event Log example

Using Prom Tool to analyze event logs of the dataset we have chosen to use for demonstration we can understand every step taken for every 42995 offers individually and also percentage for each start event and end event. For example, an offer with number 2471357189 contains 4 events such as O_Create_Offer, O_Created, O_Sent (online only) and O_Cancelled. We can also see the first withdrawal amount which is 10000 with monthly cost of 201.76, credit scores and number of terms and so on. (Figure 2.2)

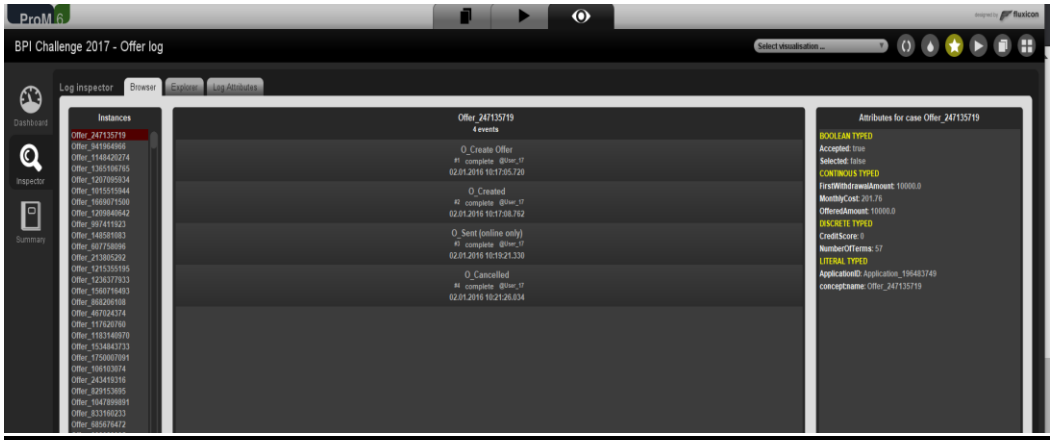


Figure 2.2 Details of a specific log from ProM Tool

2.2 Process Discovery (Process visibility, Play in)

Process Discovery means learning de facto process models from observed behavior by taking event logs and producing a model without using any prior information. It analyzes event logs of the entire organization in every detail, every step of every process, every time it has been executed in order to come into conclusion about points of friction and their root causes, optimal pathways and their deviations, business outcomes for initial actions or for corrective action taken.

Every process has a most efficient and effective route from initial input to final output and during process discovery we are able to understand and analyze all the deviations and find a way to correct them in the next process mining technique, process enhancement. In figure 2.2.2, we have created a process model based on event logs using Disco Tool which represents all 42935 cases with all appropriate frequencies and also shows even the least frequent case for example from O_Sent _online only to the finish with a frequency of only 17 cases. In figure 2.2.3 are all the deviated paths recorded by events data of another dataset with a lot more deviations. That kind of representation is known as spaghetti representation.

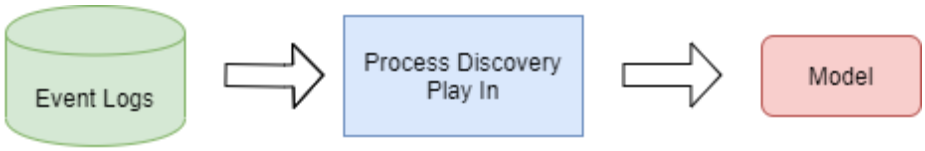


Figure 2.2.1 Process Discovery

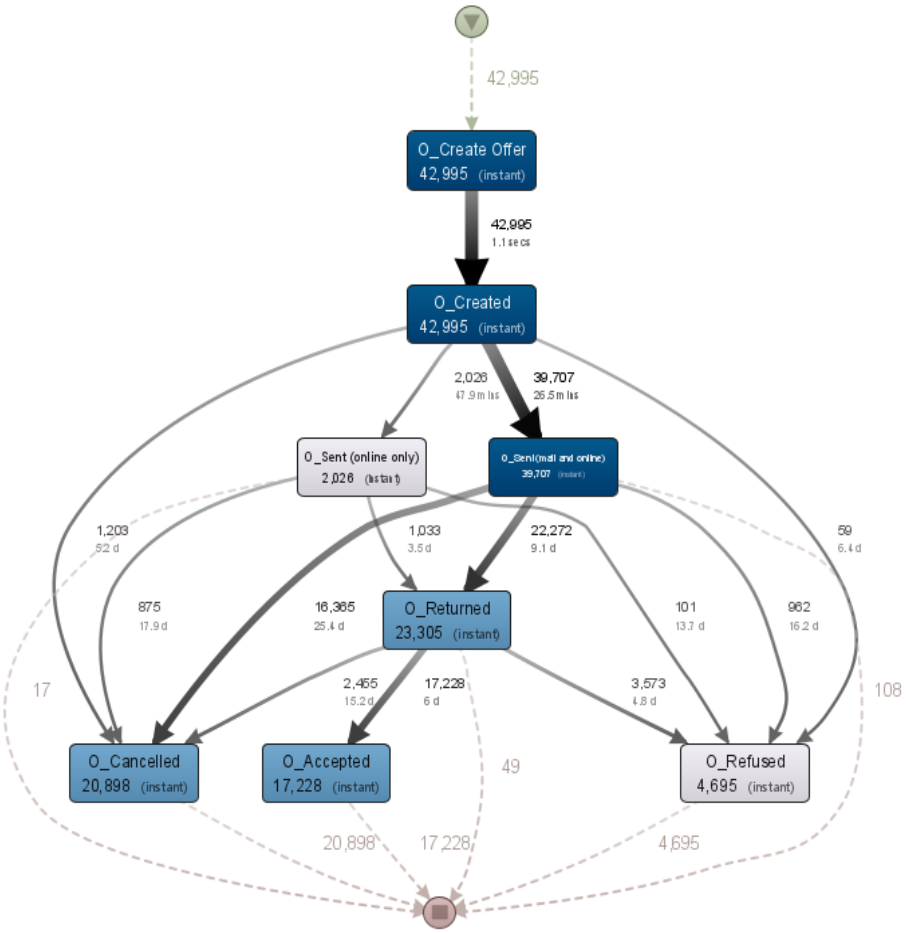


Figure 2.2.2 Process Model using Event Logs

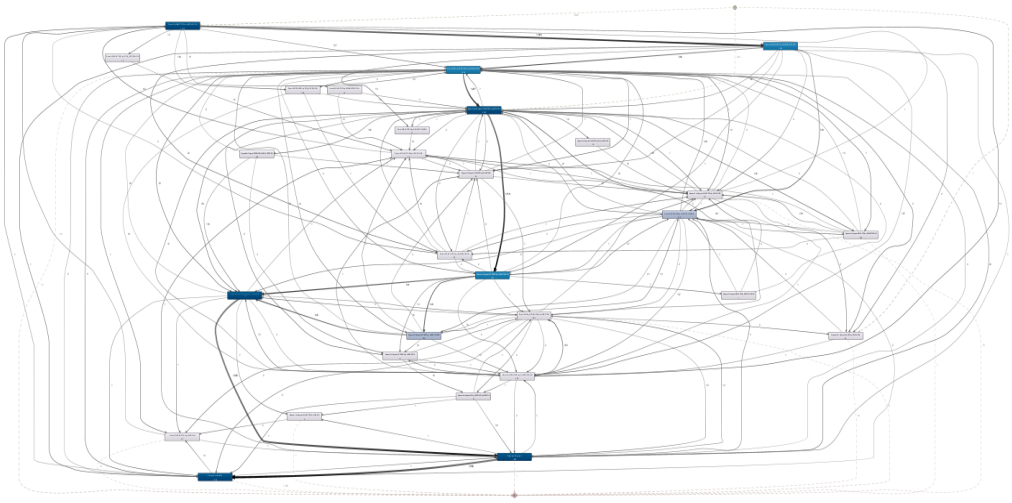


Figure 2.2.3 Spaghetti Process Model

A short example of a bottleneck analysis using Disco tool

Using the Disco tool we can understand from our 42995 event cases for one process we have 16 different variants and the mean of case duration is 19.1 days meanwhile the median is only 15.9 days from the outliers of our event logs. (Figure 2.2.4)

Figure 2.2.4 Overview of Cases and their duration

After performing a performance filter for cases which lasts minimum 33 days and 20 hours to 168 days and 19 hours which means 7% of cases, we can analyze the process map and try to see the activities problem causing these delays. (Figure 2.2.5)

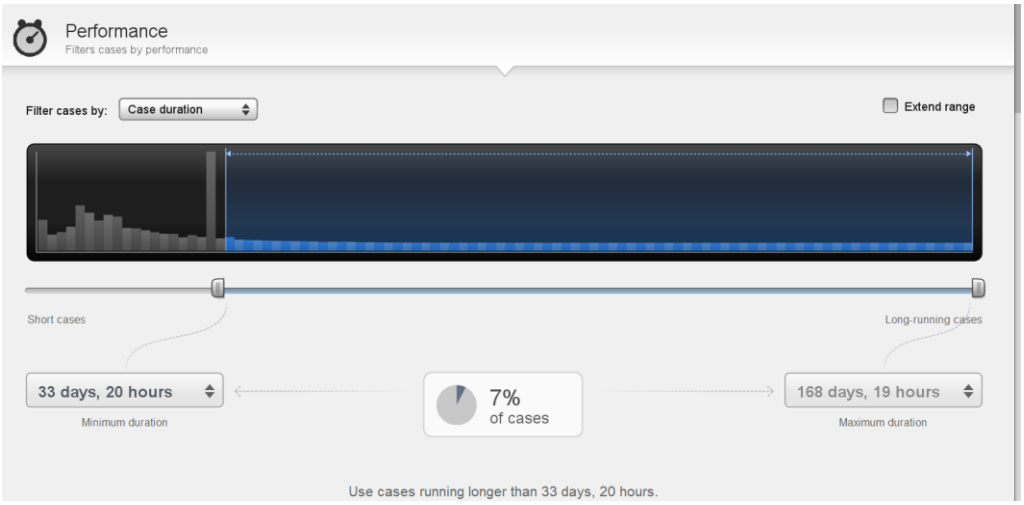


Figure 2.2.5 Performance Filtering according to case duration

Using the filter event logs and making a performance check with mean duration we can understand the processes which have a high mean duration such as O_Sent to O_Refussed with 57.6 days or O_Created to O_Cancelled with 45.9 days but we should now forget about the frequence of each case so we should take into account also the big picture and after conducting a bottleneck analysis for all the process, we came into conclusion that the bottleneck process was O_Sent from mail or online to O_Refused and it should be discussed with the employees of the bank to improve these steps in the process.

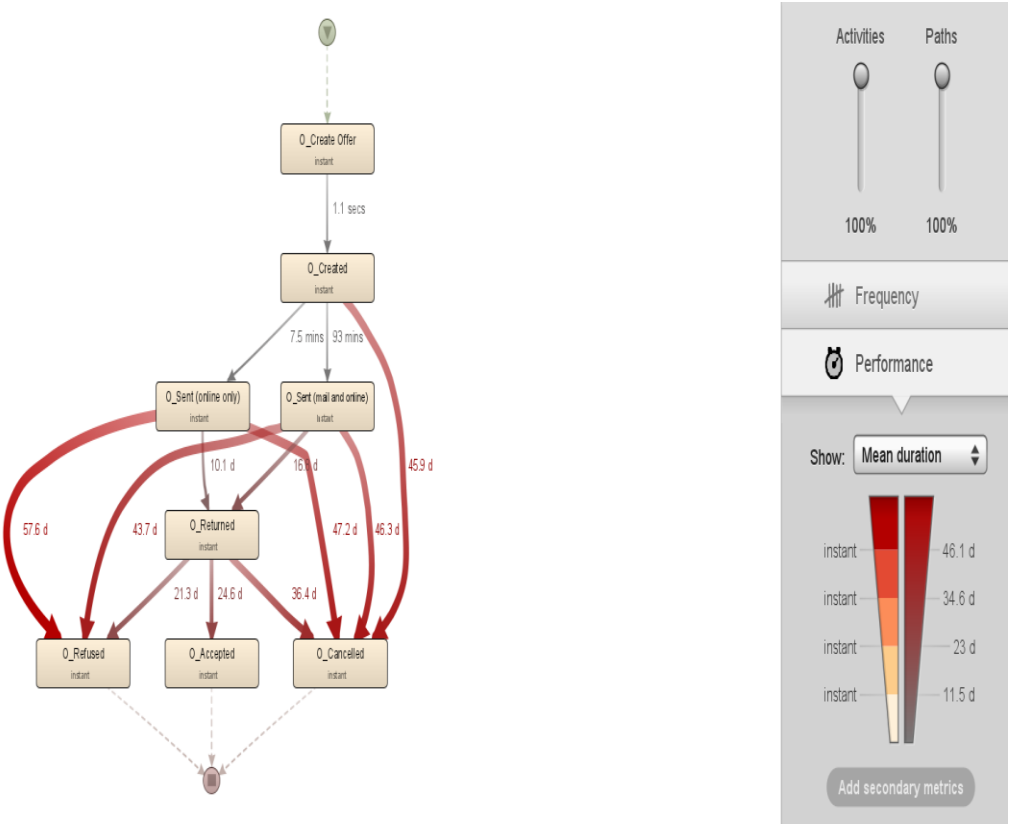


Figure 2.2.6 Bottleneck analyses using 7% of cases with mean duration metrics of Performance Checking

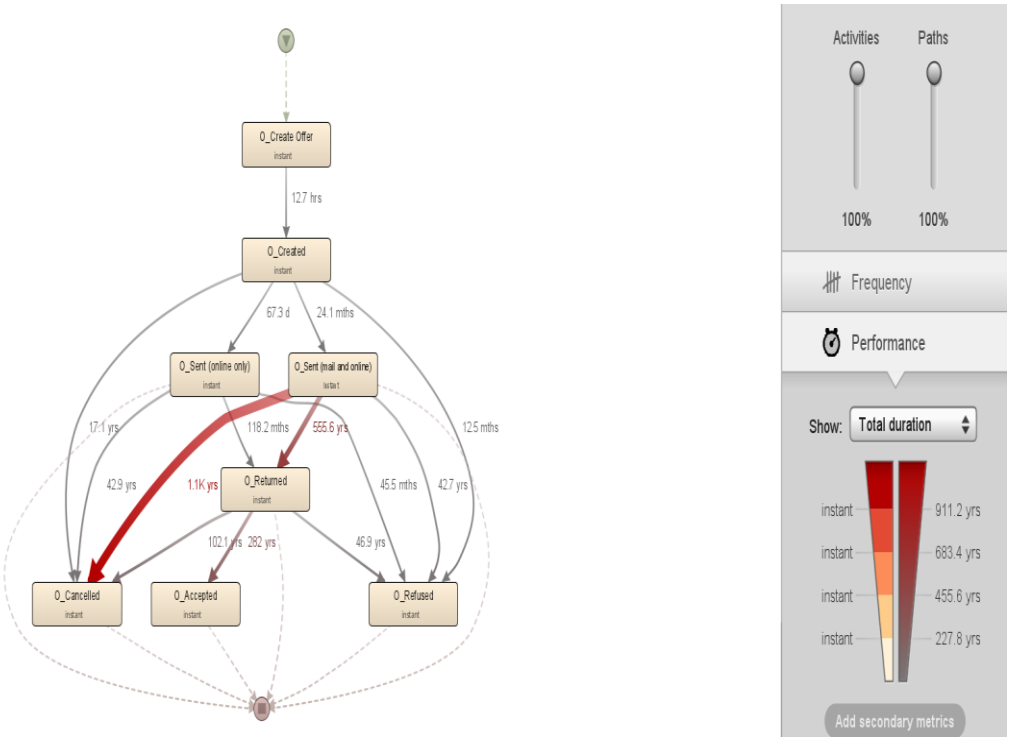


Figure 2.2.7 Bottleneck analyses using 100% of cases with total duration metrics of Performance Checking

2.3 Conformance Checking (Process monitoring, replay)

Conformance Checking compares an existing process model with the event log of the same process confronting the modeled process with reality. It helps in identifying and correcting new points of friction as they occur and ensure that everything is going in the correct way. For example, in figure 2.3.2 we can see that the processes which are taking a lot of time and creating delays are shown with red color and the density of it changes with the increase it takes for each process phase to be finished. In our case the process stage which creates bottlenecks is O_Sent (mail and online) and O_Cancell so the organization should analyze what are the causes of these delays in the organization and how can it be improved.

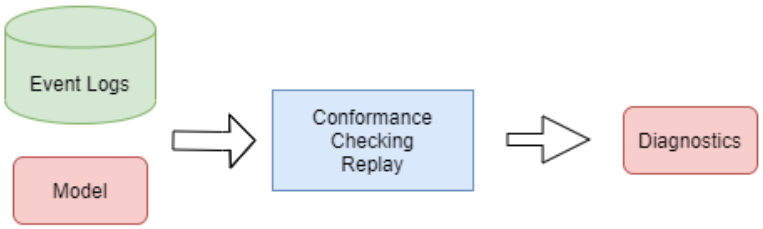
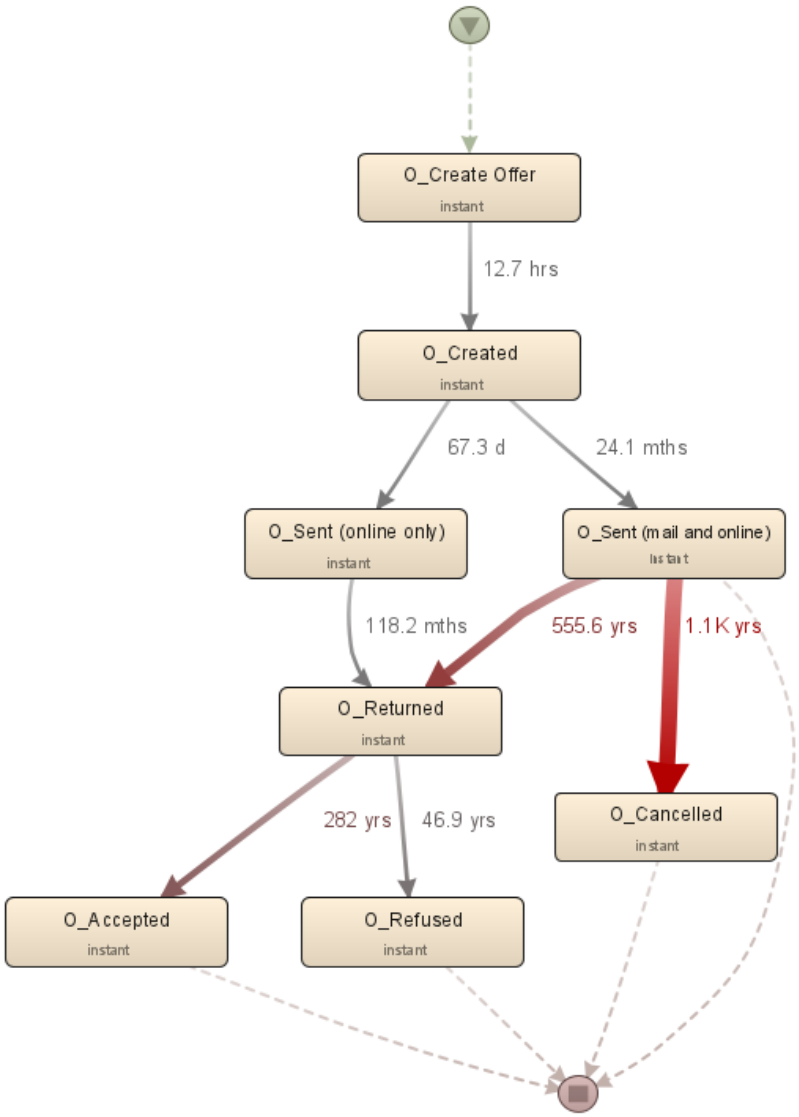


Figure 2.3.1 Conformance Checking



2.3.2 Process Model According to Performance

2.4 Enhancement (Process improvement)

Enhancement improves an existing process model using information about the actual process recorded in some event logs. The variations and root causes identified during the Conforming checking are then improved by practical actions that remove friction and automate flow for human and digital workforces.

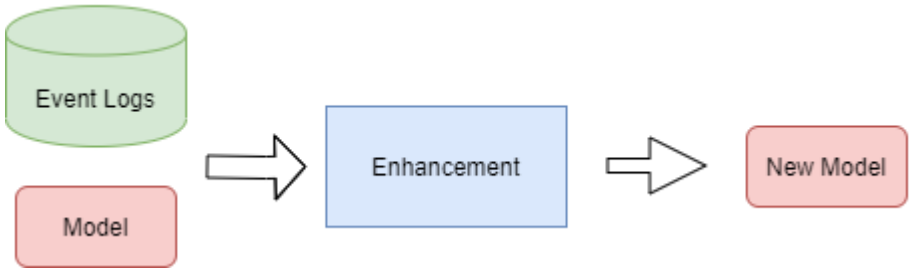


Figure 2.4 Enhancement Phase

CHAPTER 3

ALGORITHMS OF PROCESS MINING

There are various process mining algorithms divided into local and global approaches.

Local approaches such as Heuristic Miner studies local relations of activities in event logs. On the other hand, Genetic Mining or Fuzzy Miner are global approaches which are used to build a model based on the whole event logs.

3.1. Heuristic Miner

Heuristic Miner studies local relations of activities in event logs. It uses alpha algorithm without neglecting the frequency of event logs, enabling the mining of control flow of the process model.

Steps of Heuristic Miner are :

1. Event log reading
2. Analyze the ordering relations based on the frequency
3. Build the net based on these results
4. Output the net

Heuristic Miner is used to express the main behavior but not with all details and exceptions that are registered in an event log. It is recommended for real-life data that does not have too many different events.

3.2 Genetic Mining

Genetic Miner uses a genetic algorithm in order to mine a petri net representation of the process model from event logs. It is an approach that reassembles natural evolution. It begins with an initial population of individuals which in our case are process models and then the population evolves by selecting fittest individuals and generating new process models by using genetic operators such as crossover and mutation. Duplicate activities, hidden paths, noise can be detected using Genetic Miner.

Steps of Genetic Miner Algorithm are:

1. Event log reading
2. Creation of initial population
3. Fitness calculating for the individuals of the population
4. Stops and return the fittest individuals
5. Create next population (next process model)

It is recommended for mining logs that contain noise, handling invisible tasks and duplicates.

3.3 Fuzzy Algorithm

Real life processes are less structured than what people expect them to be and traditional process mining approaches are not welcoming. The discovered models that show all details without distinguishing the most important events or without any filter are “spaghetti-like”. In order to simplify the process model presentation we use a fuzzy algorithm which is created from the concept of a roadmap. (van der Aalst, Process Mining: Data science in Action, n.d.)

It removes unimportant edges, clusters highly correlated nodes into an individual node, and discards isolated node clusters. (Khalaf, 2013)

Steps of fuzzy miner algorithm:

1. Event logs reading
2. Fuse similar behavior attributes
3. Generate Meta rules
4. Generate frequent fuzzy item sets
5. Make fuzzy association rules

Fuzzy miner calculates a significance weight for each log element and an additional correlation weight of every edge. The element with the highest weight is the element chosen by the algorithm. It also uses thresholds and only the elements which meet the thresholds are picked. Fuzzy miner is recommended for mining when you want to depict desired traits, eliminating irrelevant details, complexity reduction and comprehensibility improvement. Figure 3.1 is the representation of the Fuzzy Model of our event logs using the Fuzzy Model Plug In of ProM tool. Disco, the process mining tool we are using for this thesis uses fuzzy algorithms to create process models.

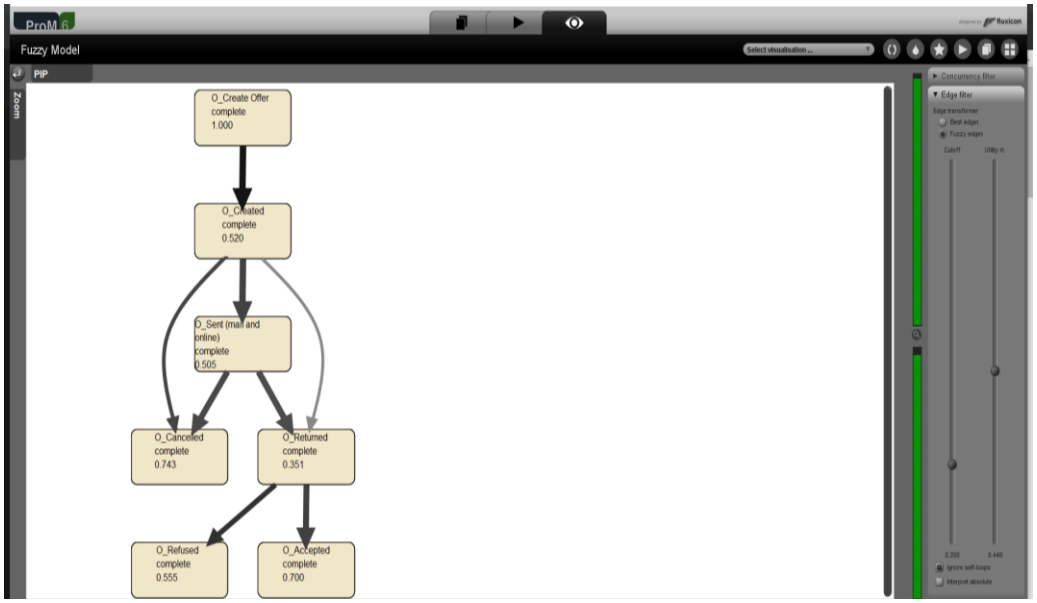


Figure 3.1 Fuzzy Algorithm Prom Plugin

CHAPTER 4
PROCESS MINING IN REAL WORLD

4.1 Main Drivers for businesses to adopt Process Mining

According to Gartner Market Guide for Process Mining, there are 4 main drivers for companies to utilize Process Mining in their day-to-day operations such as:

1. Digital Transformation

stimulates attention toward the advantage of analyzing and having a clear view of their processes, operations to know how and where to improve in order to create business value.

2. Development of Algorithms and AI

new important insights can be derived from event logs and with analyzing those results competitive differentiation can arise.

3. Hype of RPA

have raised the importance of identifying tasks that need to be automated and more other benefits which will be explained in chapter 5 in more detail.

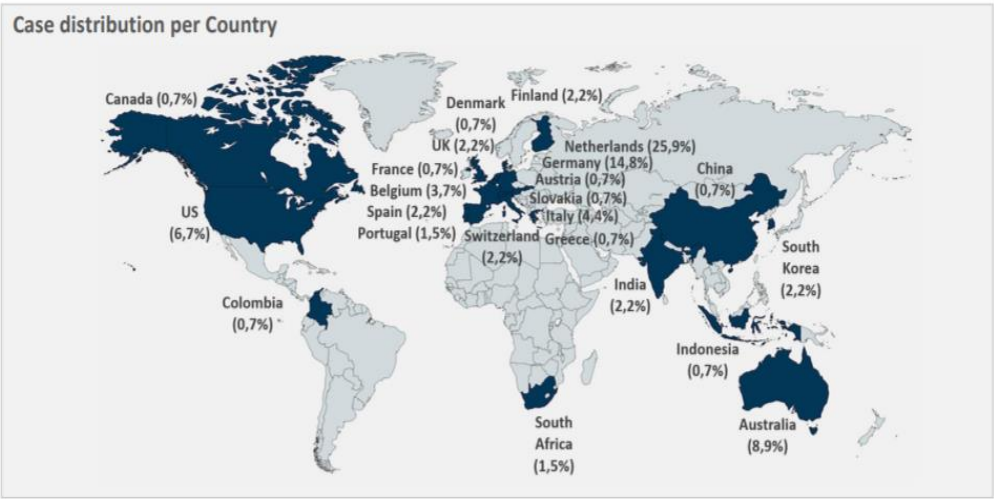
4. Emerging of Digital Operations known as DigitalOps

PM helps in process modeling and continuous performance monitoring of Digital Business Platforms.

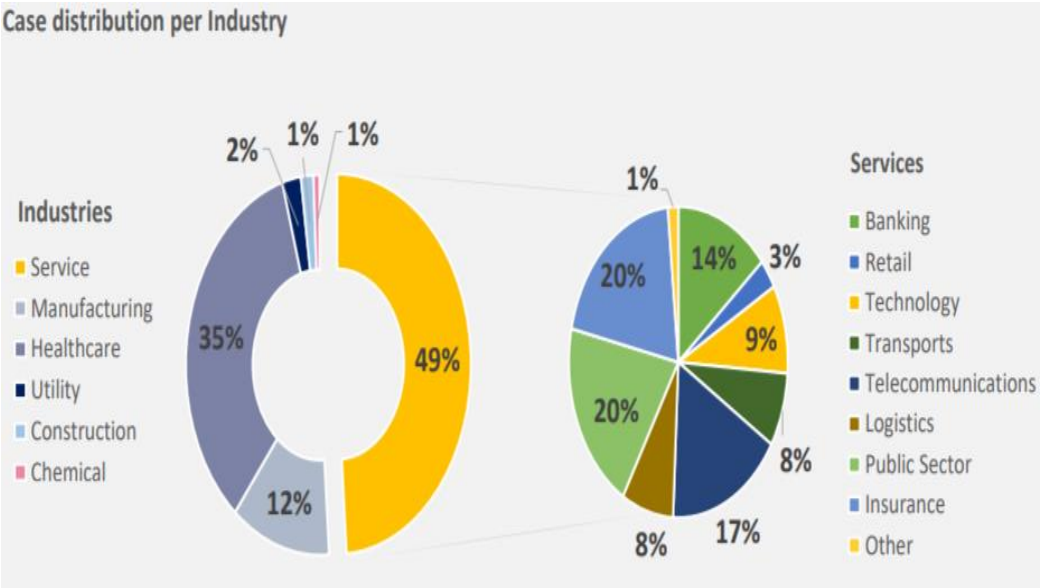
4.2 Use Cases of PM in real world and their results

Whether you are a hidden champion or a multinational industry giant, process mining transforms the way companies operate by taking friction out of the process in order to achieve the best results. According to a research conducted by HSPI in 2017, main users of process mining are located in Europe and in Australia. More than 40% are located in Germany and in Netherlands and 8.9% in Australia due to promotion of research about process mining in Universities and Excellence Innovation Centers. (figure 4.1.1) Process Mining main adapters were organizations operating in services (49%), in healthcare (35%) and manufacturing (12%). (HSPI, 2017) For more detailed information you can observe the figure 4.1.2.

In 2019 retailers, telcos and finance companies are the ones who have gained most benefits from PM. (Rosik) I have attended a three day webinar of Process Mining from Celonis and during that webinar I have gathered information about how real companies use process mining and what is their feedback. (Celonis, 2020)



4.1.1 Case Distribution per Industry of PM, HSPI, 2017 (Philippe Gonella)



4.1.2 Case Distribution per Industry, (Philippe Gonella)

4.2.1.Retailers

Retailers main goals are effective inventory management, an optimized purchasing strategy and swift delivery of product innovation. All these goals are process-oriented. Retailers mostly use PM to further optimize their areas of expertise such as “Fast Order to Cash” or “Seamless Customer Experience” , to strengthen areas of weakness with “On-time-delivery”. According to the research that Celonis conducted, the average on-time shipment rate is 42.8% which has huge consequences in customer satisfaction, in supply chain and so on. (Klenk, 2020)

4.2.2.Telecommunication

Telecommunication industry requires rapid innovation and fast adaptation towards changing customer demands and technology innovations. They use PM for service procurement, client acquisition, sales and billing management and many more.

Average rate of touchless invoices is 26.6% which is very low considering the technologies available nowadays. Vodafone touchless invoices went from 13% to 83% after using PM and RPA. Vodafone relies on process mining to optimize processes such as source-to-pay and procurement activities management. They manage more than 800000 purchase orders, 5 million invoices and 40 million assets producing more than 10 terabytes of data in various storage systems and applications. Process mining found out what processes had low automation rate that could be improved by RPA and also helped in eliminating errors in Procurement activities resulting in maximizing catalogue buying and increasing the speed of ordering to their suppliers. With the 85% increase of correctness in purchase orders, time to market improved by 20%.

Deutsche Telekom wanted to achieve being a data-driven process organization and to stop manually calculating KPIs due to increased complexity of internal processes. Process Mining ensured them significant savings in operational costs, more than $40 Million by realizing in time cash discounts, while reducing points of friction throughout their Procure-to-Pay process. (Tasev, 2020)

4.2.3 Finance

Being highly regulated businesses in need of strong compliance management, PM helps a lot by uncovering risks, optimizing fraud management, identifying upsell and cross-sell opportunities, consulting and so on. Based on Celonis ABB use case, an example of how process mining can be used in finance shown in the table below. (Hartmann, 2020)

ABB	PM Benefits					
PM Phases	Optimize Working Capital		Increase Productivity		Manage risk and compliance	
Discovery	Early payments triggered by premature baseline dates	Late payments caused by late entries into ERP, missing information, customer delays	Rework rates because there is excessive payment blocks and failed 3-way matching between PO, invoice, goods receipts	Automation rates should change due to failure of OCR technologies and non-standard format across different ventures leading to high rates of manual effort	Duplicate payments due to character mismatches on forms or by multiple countries vendor invoices	Segregation of duties by accidental or conscious compliance variations from lack or ERP enforcement or resource constraints
Enhancement	Postpone baseline dates of early invoices	Prioritize discrepancies between invoice data and on-time date	Postpone baseline dates for early invoices	Switch vendors to Electronic Data Interchange	Highlight and block duplicate payments	Highlight and diagnose violations
Conformance Checking	Check DPO-days payable outstanding	Check operational Cost per Purchase Order	Touchless Purchase Orders	Touchless Purchase Orders	Falling duplicate payments	Check violation rates

4.2.4 Non-Traditional Industry Usage and Heterogenous Organizations

The minimum requirement for applying PM is having an information system from which event logs can be produced. So PM can be applied in healthcare, manufacturing, logistics and so on.

PM can be used to decrease the food loss and waste crisis. According to the Celonis Research over 1.6 Billion tons of food wasted every year which is causing the economy $1.2 Trillion and 8% of greenhouse gas emissions. 80% of this food waste is created before it reaches the consumers. It is created in the supply chain by overproduction, late delivery, miscalculations of the stock and PM can help to eliminate this problem. (Klenk, 2020)

Martin Rowlson, Global Head of Process Excellence at Uber, declared that PM made it possible to diminish friction in order to increase customer satisfaction in every country they operate. Massive scale, growth and autonomy for meeting demand in a particular area created for Uber complex processes and gave a massive data pool. In order to get reliable, accurate and frequent data, their only solution was process mining which saved them around $20 million. Before if they wanted to improve a process, they observed, analyzed and thought for a solution but it is impossible to continuously observe million of processes and the number of observed processes were insignificant to give shareholders to convince them to approve the change or to identify unknown behaviors of process in contrast of PM which takes all event data and create an exact mirror of all processes happening. (Rowlson, 2020)

Bosch has more than 400000 employees worldwide and is present in almost every country in the world. Process mining becomes important in Bosch because it has a heterogeneous organizational structure and the business processes have increasing complexity and massive increase of data volume by

IoT activities which is impossible to be analyzed by traditional workflow analysis. (Buhrmann & Gottschall, 2020) Process mining also helps them to check implementation success and identify manual process steps for further process automation such as RPA.
PM project in Borsch started during 2017 with initial proof of concepts and was implemented in 2018 with a team set-up by inhouse consulting, central IT and business divisions. They used PM in Finance, IT Services, Logistics, Production, Purchasing, Sales. Borsch used PM in Global Business Service which deals with HR-Services, Logistics, Finance and Purchasing because of the high volume of indirect processes and to achieve their main goal which was the transformation from a functional to process oriented organization. They use PM to create a generic overview about end-to-end processes including development of process KPIs or to have an immediate reaction in case of process issues.

Lufthansa CityLine has set-up the PROMOTE (PROcess Mining for OperaTional Excellence) in order to increase punctuality for customers and enhance internal efficiency. Philipp Grindemann, Head of Business Development and Project Management at Lufthansa CityLine, declared that their main focus in 2019 was to increase customer satisfaction and process mining helped them. They identified a bottleneck in their process which was Ground Operation Process so decided to implement a new process in Munich Airport. They replaced transport of the passengers with walk boarding which made it possible to increase flight punctuality by 17% and increase available boarding time by leaving the gates open longer and also reduced fuel consumption. (Grindemann, 2020)

Chart Industries Inc after implementing PM come into conclusion that they have 120 unique sets of payment terms across 3 major regions such as North America, EMEA, APAC and only in APAC has 63 unique payment terms despite low invoice volume so they decided to take action by borrowing best

practices from top performing business units, standardizing processes, reducing complexity and speeding up operations and accelerating migration to a global shared service model. (Spillers & Puetz, 2020) PM made them to reduce manual touches, refocusing staff on value-add activities. PM also discovered that the A/P automation tool that they were using was not correct in the calculation of the payments so they decided to automate that process ensuring that the mistakes would not repeat again. Automation of reporting concluded in a saving of 25% of Category Manager's monthly time and $10.1 Million in negotiated savings.

This table is an abbreviation of the information written above.

Organization	Benefits	Goal
Uber	Saved around $20 million	increase customer satisfaction, get a full view of all the organization, a unified tool for all, communication tool for stakeholders, monitor processes in real time
Bosch	Convert into a process-driven organization	check implementation success and identify manual process steps for further process automation such as RPA
Lufthansa City Line	increase flight punctuality by 17% and increase available boarding time, reduced fuel consumption	increase punctuality for customers, enhance internal efficiency, to increase customer satisfaction, identified a bottleneck in Ground Operation Process replaced with walk boarding,
Chart Industries Inc	Automation of reporting concluded in a saving of 25% of Category Manager's monthly time and $10.1 Million in negotiated savings.	have 120 unique sets of payment terms across 3 major regions, PM made them to reduce manual touches, refocusing staff on value-add activities. Discovered the A/P automation tool that they were using was making errors and replaced with RPA bots
Vodafone	touchless invoices went from 13% to 83% after using PM and RPA, With the 85% increase of correctness in purchase orders, time to market improved by 20%.	Optimize processes such as source-to-pay and procurement activities management. PM found out what processes had low automation rate that could be improved by RPA, helped in eliminating errors in Procurement activities resulting in maximizing catalogue buying and increasing the speed of ordering to their suppliers.
Deutsche Telekom	saved more than $40 Million by realizing in time cash discounts	being a process-driven organization and to stop manually calculating KPIs due to increased complexity of internal processes

CHAPTER 5

ROBOTIC PROCESS AUTOMATION

5.1 Introduction to Robotic Process Automation

Robotic Process Automation (RPA) is a fast emerging process automation technology which allows organizations to automate repetitive, time consuming, manual and error-prone tasks previously performed by humans. According to PWS, RPA is not only making a difference in many industries but it is transforming the way we work. RPA tools are an important boost used by businesses to improve efficiency and effectiveness of their operations faster and at a lower cost than other automation approaches. Traditionally business processes were automated using programming languages, application programming interfaces (APIs), and system integrating techniques but these techniques involve application integration at a database or infrastructure level resulting in an implementation which takes months to complete concluding in irrelevant data, time delay and high expenses. RPA is the quick-fix solution by achieving frontend automation through robots known as bots which mimic human interactions with IT systems. (figure 5.1)

Discovering which process is the best suitable one to automate is the main challenge in an RPA implementation and process mining is the solution for it. (Van der Aalst et al.2018)

Figure 5.1 RPA compared to traditional process transformation approaches Deloitte Survey "Next generation automation Transform your business processes with robotic and intelligent automation" (Deloitte P. S., 2016)

5.2 Process Suitable for RPA

Gartner defines RPA as follows:

"RPA tools perform [if, then, else] statements on structured data, typically using a combination of user interface interactions, or by connecting to APIs to drive client servers, mainframes or HTML code. A RPA tool operates by mapping a process in the RPA tool language for the software robot to follow, with runtime allocated to execute the script by a control dashboard." (Tornbohm, 2017) RPA tools aim to reduce the burden of repetitive, simple tasks on employees (Aguirre and Rodriguez 2017).

In figure 5.2.1 is illustrated where RPA can be implemented. If most of the cases have the same business process structure then traditional process automation is attainable. RPA is feasible for business processes which are repetitive and have a low frequency of repetition are not considered due to the high cost of automation.

However, if the business process has multiple complex expectations and deviations and the case frequency is low, manual work of humans is needed such as entering information frequently and decision making.

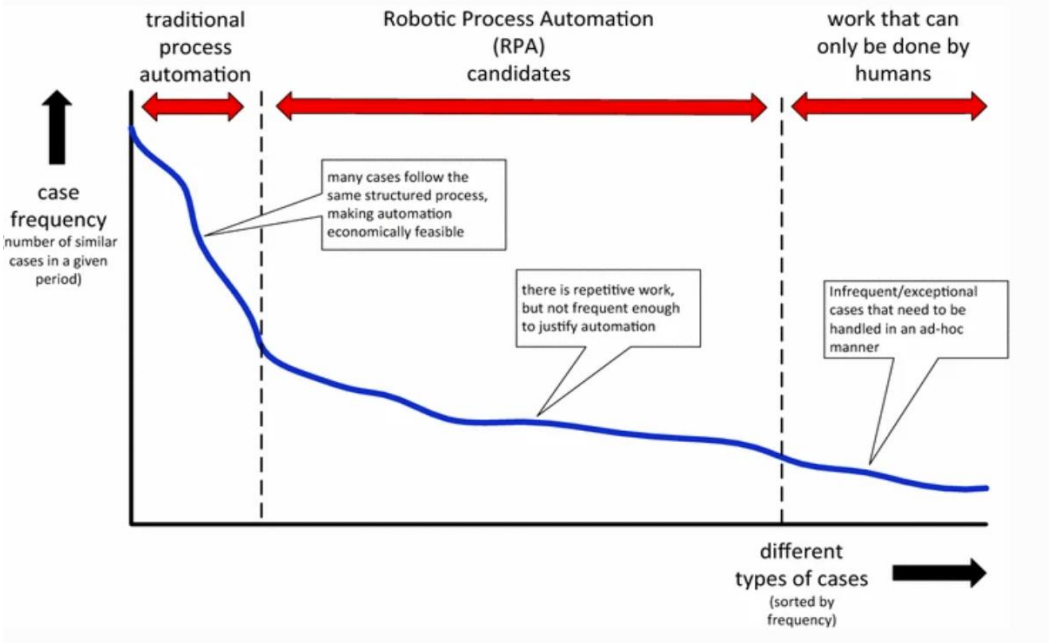


Figure 5.2.1 RPA feasibility to automate business process cases (Wil M. P. van der Aalst, 2018)

Willcocks et al propose that for a business process to be a candidate for RPA it should share these characteristics : (Willcocks, L., Lacity, M., & Craig, A. (2017))

1. Standardized
2. Rule based
3. Stable
4. High frequency of usage
5. Clear business value
6. Multiple systems
7. Clear costs

Accenture, a RPA vendor has developed a checklist with 5 criteria as shown below: (Accenture, n.d.)

1. Independent from Human judgment and rule based process
2. Originated by digital trigger and relying on digital data
3. Functioning and stable process
4. High frequency of execution
5. Key systems of the company should be leveraged by the process

Deloitte's criteria for RPA are : (Deloitte)

1. Processes prone to error such as data entry and data migration
2. Digital data involvement
3. Rule based processes
4. Repetitive processes

5.3 RPA Capabilities

RPA market in 2016 was worth $250 million and would grow to $2.9 billion in 2021. (Craig Le Clair, n.d.) RPA is a software-based technology that interacts with the presentation layer of applications and emulates human execution of a business process. (Hawkins I., 2018 PEXO) It does not require adaptation with the information system making it easier to be implemented and requiring less time. (Jerome Geyer-Klingeberg, 2018) In workplace bots are used to automate labor intensive day to day processes with high repetition, freeing up high value employees time for activities such as problem solving, exception handling or trouble shooting. (Wil M. P. van der Aalst, 2018)

RPA bots as shown in figure below (5.3.1) are able to execute a lot of activities but the list of their abilities increases everyday due to integration of RPA with artificial intelligence offered by some vendors.

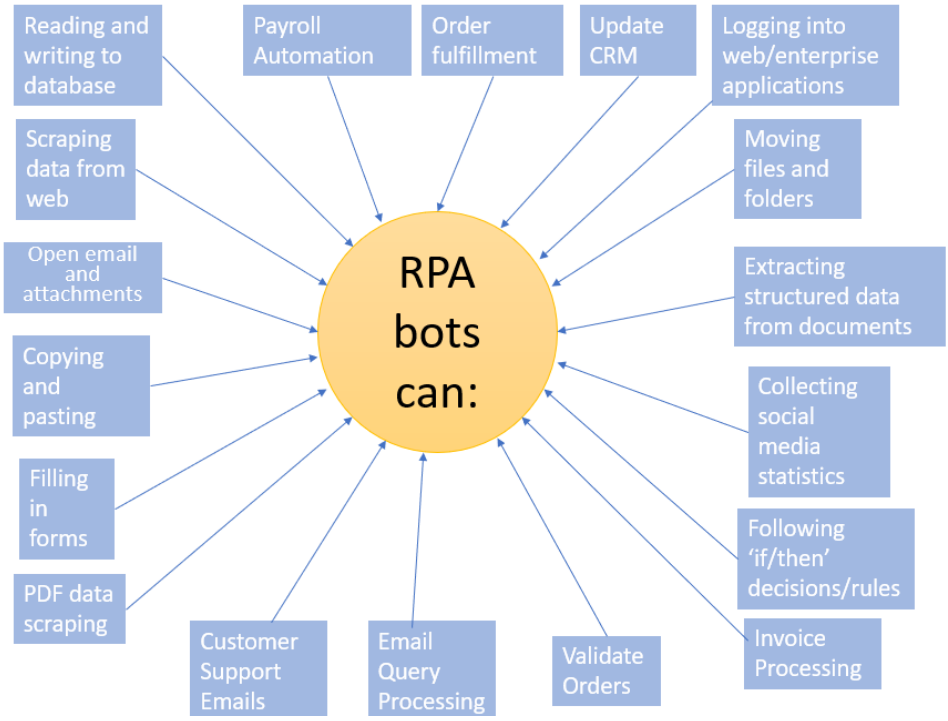


Figure 5.3.1 RPA capabilities

According to Accenture, financial institutions operate in a highly regulated industry which require data security and quality, auditability, operational resilience. (Accenture, n.d.) RPA helps them to achieve their goals by reducing costs by 80%, increasing quality by decreasing human errors, reducing time to perform tasks by 80-90% and easy to implement.

The main vendors in the market for RPA are UiPath, Automation Anywhere (AA), Blue Prism, Kryon RPA, and Microsoft Power Automate. (ITCentralStation) According to Everest Group Research for evaluating RPA vendors for market impact, overarching vision and product capability and support in 2018 UiPath and NICE were the star performance followed by Blue Prism, Automation Anywhere and Thoughtonomy.

5.4 RPA Phases and Process Mining

RPA is data-driven and data-dependent. An unplanned RPA puts an organization at a great risk. (Guide PM,2020) Van der Aalst considers the selection of the most suitable processes for automation a huge challenge. An Ernst & Young study came into conclusion that 30-50% of all first time RPA initiatives do not succeed. (Young, 2016)

PWC declares that:
"Automated systems are game changers but without effective controls, they can cause trouble in a hurry. If you do not spend time in risk and control assessments before implementing RPA, you are risking to make a big problem happen much faster." (PWC, 2017) With the increase of complexity of organizations operations, with processes that have different characteristics such as their frequency, length, involvement of various departments, stakeholders, variety of inputs and outputs making RPA not a one-size-fits-all solution but requires a careful analyses and informed decision making (Jerome Geyer-Klingeberg, 2018)

According to SSON Analytics, 29% of the respondents admitted that their process was not mature enough or was not fit for the solution, 23% due to insufficient change management, 15% solution provider ended up not being a fit for business needs. (SSON, n.d.) (figure 5.4) Automation of not optimized processes creates a lot of inefficiencies and errors.

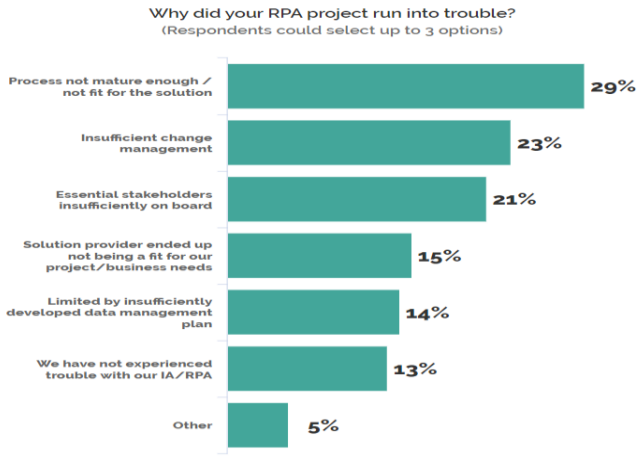


Figure 5.4 SSON Analytics (SSON, n.d.)

5.4.1.Pre-deployment of RPA

As explained in Chapter 1 and Chapter 2 business processes usually deviate from the modeled process and, for RPA a clear representation of business processes is required in order to understand the automation potential of it. Usually after choosing candidates, a long process of several interviews with the employees are conducted in order to understand the steps of the process and its decision points. This phase takes a long time because we have to consider the part that humans are subjective and each of them have a different way of performing a process.
We should take into account the processes between different departments, where different individuals may not have knowledge about the steps taken by the other departments employees. After many interviews a process map is created, a list with clear and defined activities is generated and the sequence of activities is translated into a logical basis for the bots. (Bisceglie, Taghiabadi, & Aklecha)

To be more efficient and objective in finding out which process should be automated, PM should be used to map and understand the AS-IS processes. PM would take event logs from the information system and generate the correct process map in the process discovery stage.

PM helps in understanding of which process should be prioritized due to significant improvement that RPA would yield. PM identifies which processes have inconsistencies, bottlenecks or redundant steps during process discovery and presents the process which is required to be improved by reengineering the process. Due to analyzing all event logs, hours of long interviews and different points of views from different departments are avoided. As mentioned before attempting to use RPA on unsuitable activities is a huge risk. The configuration of RPA bots should start only after finding the optimized workflow of the process and then these sequences of activities are turned into logical basis for the bots. (Lipovsky, 2019)

5.4.2. Mid-deployment RPA

During this stage of the RPA life cycle the process logic which we identified in the Pre-deployment phase should be turned into a script for the configuration of the bots. The bots should be tested and then the process owner and the team of RPA decide if the bots are working properly or not before their implementation.

No matter how professional their opinion is, it still remains subjective but if they use PM, they can simulate an exact environment of the organization and implement bots giving them a replica of the exact working environment and then analyze if the behavior of the bots were accurate or not. By using PM during this phase, we can also gather data to compare the bots with non-RPA supported executions of the process by creating a clear representation of process changes and case coverage. (Bisceglie, Taghiabadi, & Aklecha)

Various unexpected undesirable alterations caused by automation can be identified and then be improved before the implementation of the bots.

5.4.3. Post-deployment of RPA

After finishing the testing part, the bots are deployed and they start to work on their activities like a digital worker but it does not end here. It needs to be monitored even after testing because various unpredictable errors such as software updates can happen. Continuous monitoring is crucial in order to ensure correctness and efficiency. Conformance checking can be used to ensure that no deviations are happening after the bot is deployed and also PM can provide different components such as throughput times to verify the benefits of RPA. With conformance checking an existing process model is compared with the event logs produced by the same process in the past being able to compute the return of investment of the RPA.

CHAPTER 6

CONCLUSIONS

Business Optimization has been the main focus of every organization and there are different techniques available to model and to analyze business processes.

There are Process Science approaches such as Scientific Management, BPM, Operations Research and on the other hand there are Data Science approaches such as Data Mining, Statistics, Stochastics and so on but none of these approaches are capable of studying the end-to-end processes. The bridge between them is Process Mining.

PM aims to visualize the real business process and compare them with plans and theory. Reality often differs from the ideal world and this is also applicable in the case of business processes. In practice business processes are more complex and less structured than the documented and expected process flow. The way people think the processes are conducted and how they really are differs quite often. Employees devote a lot of their time to repetitive work, unnecessary activities and other types of rework but this extra time is not taken into account when the ideal process was designed. Likewise business processes are described as they should be conducted according to the ideal model but exceptions happen deviations happen every time.

With the help of innovation in technologies, organizations have the ability to collect data from events happening within their organization by creating event logs. Using these event logs in the process discovery phase, organizations create

a process model which is not based on subjective opinions but on real data. Process Mining is not only about process discovery but it is also being used for conformance checking to compare event logs derived processes against the ideal processes and mismatches between them helps to identify bottlenecks, deviations. During process enhancement phase of PM, the process map from our event logs is analyzed to improve a process by using different performance parameters.

According to the research, we came into conclusion that there are 4 main drivers for businesses to adopt process mining such as digital transformation, development of Artificial Intelligence, emerging of Digital Operations and also Hype of RPA. After analyzing many research papers, case studies and attending different conferences about Process Mining and RPA, we came into conclusion that Process Mining can be used in every type of organization which thrive for optimization from retailers to telecommunication, to finance, to hospitals and so on.

As we mentioned before RPA is a fast emerging process automation technology which allows organizations to automate repetitive, time consuming, manual and error-prone tasks previously performed by humans. RPA tools are an important boost used by businesses to improve efficiency and effectiveness of their operations faster and at a lower cost than other automation approaches. But another important aspect that was discovered during my research was that RPA can cause a lot of problems rather than fix problems if it is not implemented in the correct way and this paper shows how PM helps in a successful RPA implementation from the first stage of RPA in deciding which tasks to automate, in creating a AS-IS process map to use for the script of the bots, in testing of the bots and the last but not the least in measuring the results of RPA.

References

Aalst, v. d. (2004). Process mining: a research agenda", Department of Technology Management. Eindhoven University of Technology.

Accenture. (n.d.). *RPA - The future of technology in financial services*. Retrieved from Accenture: https://www.accenture.com/no-en/insight-financial-services-robotic-process-automation

Bisceglie, C. D., Taghiabadi, E. R., & Aklecha, H. (n.d.). DATA-DRIVEN INSIGHTS TO ROBOTIC PROCESS AUTOMATION WITH PROCESS MINING.

Buhrmann, C., & Gottschall, J. (2020). Retrieved from https://www.celonis.com/celosphere-live/bosch-a-successful-enterprise-rollout-and-eminent-importance-in-the-area-of-global-business-services/

Celonis. (2020). *Celosphere*. Retrieved from https://www.celonis.com/celosphere-live/recordings/

Craig Le Clair, A. C. (n.d.). *Forrester*. Retrieved from https://www.forrester.com/report/The+RPA+Market+Will+Reach+29+Billion+By+2021/-/E-RES137229#

Deloitte. (n.d.). Deloitte Global RPA Survey . Retrieved from https://www2.deloitte.com/bg/en/pages/technology/articles/deloitte-global-rpa-survey-2018.html

Deloitte, P. S. (2016). *Next generation automation Transform your business processes with robotic and intelligence automation.* Zurich: y The Creative Studio at Deloitte. Retrieved from https://www2.deloitte.com/content/dam/Deloitte/ch/Documents/technology/ch-en-technology-next-generation-automation.pdf

Grindemann, P. (2020). *Celonis*. Retrieved from https://www.celonis.com/celosphere-live/lufthansa-cityline-soaring-customer-satisfaction-with-on-time-departures/

Hartmann, U. (2020). *Celonis*. Retrieved from https://www.celonis.com/celosphere-live/360-process-view-in-a-complex-company/

ITCentralStation. (n.d.). *Robotic Process Automation (RPA)*. Retrieved from IT Central Station: https://www.itcentralstation.com/categories/robotic-process-automation-rpa#:~:text=The%20best%20Robotic%20Process%20Automation,Central%20Station%20reviews%20and%20rankings.

J. Wang, R. K. (2013, October-December). Efficient Selection of Process Mining Algorithms.

Jerome Geyer-Klingeberg, J. N. (2018). Process Mining and Robotic Process Automation : A Perfect Match. Retrieved from https://www.researchgate.net/profile/Jerome_Geyer-Klingeberg/publication/326466901_Process_Mining_and_Robotic_Process_Automatio

n_A_Perfect_Match/links/5b4f787ea6fdcc8dae2b378c/Process-Mining-and-Robotic-Process-Automation-A-Perfect-Match.pdf

Kerremans, M. (2019). *Market Guide for Process Mining.* Gartner.

Khalaf, G. T. (2013). *"Leveraging Process mining techniques".* IBM T.J. Watson Research Center.

Klenk, M. (2020). *Celonis*. Retrieved from https://www.celonis.com/celosphere-live/keynote-celonis-technology-strategy/

Lipovsky, R. (2019). *Why Process Mining is Necessary at all Stages of RPA*. Retrieved from Minit: https://www.minit.io/blog/why-process-mining-is-necessary-at-all-stages-of-rpa

MarketWatch. (n.d.). Global Process Mining Software Market Size 2020 by Manufacturers, Regions, Type, Application, Forecast to 2026.

Philippe Gonella, M. C. (n.d.). Process Mining: A Database of Applications 2017, HSPI.

PWC. (2017). *Who minds the bots? Why organisations need to consider risks related to RPA?* Retrieved from https://www.pwc.com.au/publications/assets/rpa-risk-controls.pdf

Rosik, M. (n.d.). Process minig trends in 2020. Retrieved from https://www.minit.io/process-mining-trends-2020-guide

Rowlson, M. (2020). *Celonis*. Retrieved from https://www.celonis.com/celosphere-live/how-uber-improves-global-customer-satisfaction-with-insights-from-process-mining/

Sarah Burnett, A. M. (n.d.). *Robotic Process Automation (RPA) – Technology Vendor Landscape with Products PEAK Matrix™ Assessment 2018*. Retrieved from Everest Group: https://www2.everestgrp.com/reportaction/EGR-2018-38-R-2595/Marketing

Sarah Burnett, A. M. (n.d.). *Robotic Process Automation (RPA) – Technology Vendor Landscape with Products PEAK Matrix™ Assessment 2019*. Retrieved from Everest Group: https://www2.everestgrp.com/reportaction/EGR-2019-38-R-3217/Marketing?SearchTerms=rpa

Spillers, N., & Puetz, B. (2020). *Celonis*. Retrieved from https://www.celonis.com/celosphere-live/chart-industries-how-process-mining-helps-optimize-their-business-services-backbone/

SSON. (n.d.). *Process Mining & Mapping as a Basis for Automating Performance*. Retrieved from SSON Analytics: https://workbook.sson-analytics.com/ProcessMining2019KOQOXFDM#/barriers

Tasev, P. (2020). Retrieved from https://www.celonis.com/celosphere-live/deutsche-telekom-how-they-gained-significant-value-by-becoming-a-data-driven-company-using-process-mining/

Tornbohm. (2017). *Gartner market guide for robotic process automation software.* Gartner.

Van der Aalst, M. B. (2018). *Robotic Process Automation.* Retrieved from https://link.springer.com/article/10.1007/s12599-018-0542-4

van der Aalst, W. (n.d.).

Van der Aalst, W. (n.d.). *Process Mining Data Science in Action - Wil van der Aalst*. Retrieved from https://www.youtube.com/watch?v=kIeLaNzw9hI

van der Aalst, W. (n.d.). *Process Mining: Data science in Action*. Retrieved from Coursera: https://www.coursera.org/learn/process-mining

Young, E. &. (2016). *Get ready for robots. Why planning makes the difference between success and dissapointment.* UK: E YGM Limited. Retrieved from https://www.ey.com/Publication/vwLUAssets/Get_ready_for_robots/$FILE/ey-get-ready-for-robots.pdf

Willcocks, L., Lacity, M., & Craig, A. (2017). Robotic Process Automation: Strategic Transformation Lever for Global Business Services? Journal of Information Technology Teaching Cases, 7(1), 17–28. https://doi.org/10.1057/s41266-016-0016-9