Tampereen yliopisto

Pavel Kaarnijoki

INTELLIGENT AUTOMATION

Assessing artificial intelligence capabilities potential to complement robotic process automation

Faculty of Engineering and Natural Sciences
Master of Science Thesis
January 2019

ABSTRACT

PAVEL KAARNIJOKI: INTELLIGENT AUTOMATION – ASSESSING ARTIFICIAL INTELLIGENCE CAPABILITIES POTENTIAL TO COMPLEMENT ROBOTIC PROCESS AUTOMATION IN BUSINESS PROCESSES
Tampere University of Technology
Master of Science Thesis, 71 pages, 2 Appendix pages
January 2019
Master's Degree Programme in Information and Knowledge Management
Major: Knowledge and Competence Management
Examiner: Professor Hannu Kärkkäinen, Professor Henri Pirkkalainen

Keywords: robotic process automation, intelligent automation, artificial intelligence, robotics, intelligent, automation

The aim of this study was to find out how artificial intelligence capabilities could complement robotic process automation when automating business processes. The purpose of this thesis was to find out what robotic process automation can do, what limitations does it have, what is artificial intelligence and what kind of different capabilities it has as well as how these capabilities could improve the automation rate of business processes when combined with robotic process automation. In addition, also challenges related to using artificial intelligence was covered.

Theoretical part of the study was based on literature review on the topics of robotic process automation and artificial intelligence, and empirical part of the study consisted of semi-structured theme interviews. For the empirical part two expert interviews were held. One of which covered the topics related to intelligent automation and the other interview focused on artificial intelligence and its capabilities.

The outcome of this study indicated that artificial intelligence capabilities can complement robotic process automation in many ways when automating business processes. Robotic process automation is a valuable and efficient tool to automate processes but there are certain challenges related to it which limit the possible use cases. These challenges relate muchly to requirements to handle structured data and incapability to make complex decisions and handle processes where there is no clear rules. Artificial intelligence capabilities then often aim to address exactly these challenges. They aim to either extract information from unstructured or from huge amounts of structured input data, or make sense out of the transformed structured data and understand what is happening. However the study also indicated that there are many challenges and considerations related to using artificial intelligence capabilities together with robotic process automation. Most pressing ones were related to data. Use of artificial intelligence requires huge amounts of data and it needs to be of good quality – this isn't always self-evident with the data organizations have. Also critical challenges related to validation of the data, trusting the results given by artificial intelligence, as well as to GDPR and possible future regulations with artificial intelligence.

TIIVISTELMÄ

PAVEL KAARNIJOKI: ÄLYKÄS AUTOMAATIO – TEKOÄLY KYVYKKYYKSIEN POTENTIAALIN ARVIONTI YHDISTETTYNÄ OHJEMISTOROBOTIIKKAAN
Tampereen teknillinen yliopisto
Diplomityö, 71 sivua, 2 liitesivua
Tammikuu 2019
Tietojohtamisen diplomi-insinöörin tutkinto-ohjelma
Pääaine: Tiedon ja osaamisen hallinta
Tarkastaja: Professori Hannu Kärkkäinen, Professori Henri Pirkkalainen

Avainsanat: ohjelmistorobotiikka, älykäs automaatio, tekoäly, robotiikka, älykkyys, automaatio

Tämän tutkimuksen tavoitteena oli ymmärtää miten tekoäly kyvykkyydet voivat tehostaan liiketoimintaprosessien automaatiota yhdessä ohjelmistorobotiikan kanssa. Tutkimuksen tarkoitus oli ymmärtää, mitä ohjelmistorobotiikalla pystytään tekemään ja minkälaisia rajoitteita sen käyttöön liittyy, sekä ymmärtää mitä tekoälyllä tarkoitetaan, minkälaisia kyvykkyyksiä siihen liittyy ja miten nämä kyvykkyydet voivat tukea ohjelmistorobotiikkaa liiketoimintaprosessien entistä tehokkaammassa automaatiossa. Lisäksi tutkimus kattaa tarkastelua tekoälyn hyödyntämiseen liittyvistä haasteista.

Tutkimuksen teoria osuus perustui kirjallisuustutkimukseen jonka aiheina olivat ohjelmistorobotiikka ja tekoäly. Empiirinen osuus tutkimuksesta puolestaan koostui puolistrukturoiduista teema haastatteluista. Empiiristä osuutta varten tutkimuksessa haastateltiin kahta alan asiantuntijaa. Haastatteluista toinen keskittyi teemaltaan älykkääseen automaatioon ja toinen tekoälyyn sekä tekoäly kyvykkyyksiin.

Tutkimuksen lopputulos osoittaa, että tekoäly kyvykkyyksien hyödyntäminen yhdessä ohjelmistorobotiikan kanssa voi monella tapaa tehostaa liiketoimintaprosessien automaatiota. Ohjelmistorobotiikka itsessään on jo tehokas tapa automatisoida prosesseja, johon kuitenkin liittyy tiettyjä rajoitteita ja haasteita, mitkä puolestaan rajoittavat potentiaalisten käyttökohteiden määrää. Nämä haasteet liittyvät usein edellytyksiin käsitellä strukturoitua dataa tai ohjelmistorobotiikan kyvyttömyyteen tehdä kompleksisia päätöksiä ja suorittaa prosesseja, jotka eivät seuraa selkeitä sääntöjä. Tekoäly kyvykkyydet puolestaan pyrkivät usein vastaamaan juuri näihin haasteisiin. Niiden tarkoitus on tuottaa informaatiota strukturoimattomasta tai valtavasta määrästä strukturoitua dataa, tai luoda ymmärrystä muokatusta datasta sekä ymmärtää mitä tapahtuu. Tutkimus kuitenkin osoitti, että tekoälyn hyödyntämiseen yhdessä ohjelmistorobotiikan kanssa liittyy myös monia haasteita sekä tärkeitä huomioon otettavia asioita. Tärkeimmät huomiot liittyvät dataan. Tekoälyn hyödyntäminen vaatii usein ensinnäkin suuren määrän dataa, ja sen pitää olla hyvä laatuista – tämä ei aina ole kuitenkaan itsestään selvää kun tarkastellaan organisaatioilla hallussa olevaa dataa. Lisäksi huomionarvoisia haasteita liittyy datan validointiin, tekoälyn päätöksentekoon luottamiseen ja GDPR-tietosuoja asetukseen sekä mahdolliseen tulevaan tekoälyn sääntelyyn.

PREFACE

I am a master's degree student at Tampere University of Technology studying Information and Knowledge Management, with a major in Knowledge and Competence Management. This thesis was undertaken for Cargotec Oyj where I wrote my thesis along with my day job. The initial idea for the thesis came from the company but the final subject was defined together with my superior Inga Rantanen.

The start of this project and the whole research process started already in fall 2017 but due to some limitations in time most of the process took place in spring 2018 and the final parts were finished during the following summer. The whole process of writing the literature review and conducting the interviews was enjoyable and gave me much more insight and deeper understanding on the subject of artificial intelligence and its possibilities when combined with robotic process automation. The idea for the thesis came from discussions with my superiors at Cargotec, and the final topic was decided by the preferences of the organization and mine combined.

First I want to start by thanking especially my superiors, Inga Rantanen and Tuula Virtanen. They made it possible to write this thesis at Cargotec, and also gave me valuable ideas and insight during the whole process. I also want to give special thanks to Professor Hannu Kärkkäinen who was eager to help and guide me through the process. In addition I want to send my thanks to the interviewees Samu Paajanen and Tuomo Pursiainen who gave valuable insight in the subjects of this thesis. Lastly I want to thank my family and girlfriend who supported and motivated me during the process.

Helsinki, 21.01.2019

Pavel Kaarnijoki

SISÄLLYSLUETTELO

LIST OF SYMBOLS AND ABBREVIATIONS

AI	artificial intelligence
BPM	business process management
CA	cognitive automation
CI	cognitive intelligence
DNN	deep neural network
IA	intelligent automation
ROI	return on investment
RPA	robotic process automation
NLP	natural language processing
NLU	natural language understanding
ML	machine learning

1. INTRODUCTION

This chapter starts with background information on the topic of this thesis. Following the introduction to the topic of this study, the main research problem, as well as the research questions are introduced. In the end of this chapter also the structure of this thesis is explained.

1.1 Research background and motives

In recent centuries and decades many drastic developments in technology have affected how we, as humans, work. One major development areas have been in the automation of work tasks. First automation entered the manufacturing world in the 19th century; increasing efficiency, cost effectiveness and removing bottlenecks in production lines. This increased the productivity in manufacturing significantly and enabled humans to focus on different kind of tasks. (Davenport & Kirby 2016.) In recent decades the same kind of development has been going on in the knowledge intensive lines of work (Frey & Osborne 2017). Now automation is moving to those knowledge intensive back and front office tasks and business processes (Jesuthasan & Boudreau 2017), and as a result companies are slaying off people's jobs – or at least that might be first impression for people about one of the latest innovations in this area, so called Robotic Process Automation (RPA). Name itself describes the technology quite well. It means automating business processes, in other words robotizing knowledge intensive manual tasks on computers that previously were handled by humans (Aguirre & Rodriguez 2017; Asatiani & Penttinen 2016; Anagnoste 2017; Craig et al. 2015; Forrester Consulting 2014; Jesuthasan & Boudreau 2017; Kukreja & Nervaiya 2016; Lacity & Willcocks 2015; Lacity & Willcocks 2016b).

This change and interest towards robotizing business processes have increased tremendously in the past few years (Lacity & Willcocks 2016b). However as most companies are still continuing their day-to-day operations without implementing RPA as part of their operations, then some are already looking forward toward to the next step - what to do after implementing RPA. After all, many of those companies who already have implemented, piloted or just have done deeper investigation into the technology might have had noticed that in spite of RPA's tremendous possibilities, regarding automation of business processes, it is just the first step in the automation journey, and applicable to only certain kind of processes. RPA is highly powerful tool to automate rule based high volume simple processes but as soon as complexity of the process increases, the data becomes more unstructured or some kind of reasoning is needed in decision points, then

pure RPA isn't enough anymore (Anagnoste 2017; Aguirre & Rodriguez 2017; Burnett & Modi 2017; Lacity & Willcocks 2016a).

To achieve a true end-to-end process automation, or to even get closer this stage, other supporting technologies are also needed. To address this need and possibility for end-to-end process automation existing literature shows implications that combining basic RPA solutions with different artificial intelligence (AI) capabilities, depending on the nature of the process, could result in a much higher level of automation where in an ideal state the whole process could be almost completely automated without need for humans to intervene (Aguirre & Rodriquez 2017; Anagnoste 2017; Burnett 2017; Burnett & Modi 2017; Lacity & Willcocks 2016a). This combination of RPA and AI is often referred to as cognitive automation (CA), cognitive intelligence (CI) or intelligent automation (IA) depending on the source material. Here in this study this combination of RPA and AI will be referred to as intelligent automation, as shown in Figure 1.

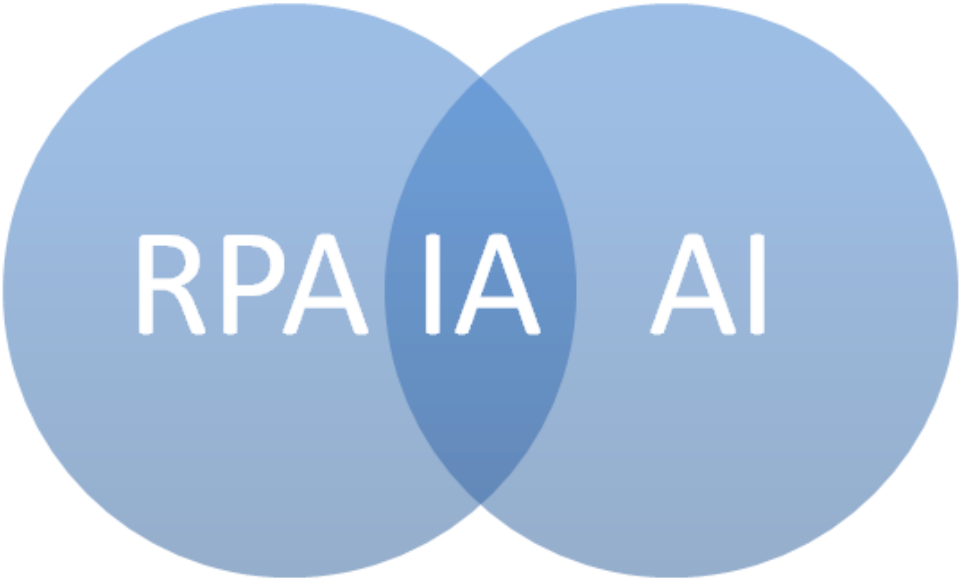

Figure 1. *Intelligent automation as a combination of robotic process automation and artificial intelligence capabilities.*

Thus the idea of intelligent automation is to enhance RPA with the AI capabilities. This in essence could enable much better automation levels and the use cases of RPA could widen up tremendously, as the intelligent automation solutions could then for example also make decisions, and interpret unstructured data (Lacity & Willcocks 2016a).

This in turn points out to a very interesting research area where very little existing research has been done. Research on AI and related technologies have been around for quite some time but in the context of automating business processes and especially combined

use with RPA the research is very limited, and most existing research just briefly mention this possibility without clarifying what AI capabilities could be used and how.

1.2 Research problem, research questions and objective

The purpose of this thesis is to provide insight on how AI capabilities could complement RPA in a big picture. Thus the aim is to understand first what can be done with RPA and what limitations relate to using it. Then the main purpose of this thesis is to provide understanding how AI capabilities could address these challenges and limitations that relate to using RPA. This is why the main research question is: ***How artificial intelligence capabilities can increase the automation level of business processes when combined with robotic process automation?*** To better answer this question there are five sub-research questions:

- *What is robotic process automation?*
- *How robotic process automation can be used to automate processes and what benefits as well as limitations does it have?*
- *What is artificial intelligence?*
- *What are different artificial intelligence capabilities and how they can be used?*
- *What kind of risks are related to using artificial intelligence?*

These sub research question then aim to give vital information to fully answer the main research question. In this sense the sub-research questions main purpose is to provide thorough understanding on all the important aspects of RPA and AI. The sub-research questions handle the business side of RPA and AI and thus the technical aspect to these are not covered in this thesis more than it's necessary to understand the concepts and their capabilities. This is because the focus of this whole thesis is to understand from the business perspective how RPA and AI capabilities can best work together to achieve higher level of automation in business processes.

1.3 Research scope

The scope of this research was defined by the need of the organization, Cargotec, this study was conducted for. The organization has started RPA initiatives during the last couple of years and already within the automated and assessed processes can been seen that RPA is a great technology but it certainly has its limitations. RPA can be used in business processes to automate very specific manual and repetitive tasks, but when it comes to automating more complex processes or processes that have decision points where clear rules can't be defined, then problems start to occur. Thus the problem is that with RPA only certain parts of processes can be automated, and humans are needed to complete the tasks that RPA cannot complete. Similarly if the input data the process gets isn't in a well-defined, predictable and structured format, RPA struggles to perform. These challenges

has raised a question in Cargotec, that could something be done about this, and is there possible a technology or a way that could help RPA with these challenges.

Recent publications and literature already however shows implications that these RPA related problems and tasks, or at least some of them, could be addressed now with AI technologies. However the publications have merely implied about the possibilities without really showing which certain AI capabilities could address these issues and how. Therefore this study focuses on understanding what AI actually mean and what capabilities and technologies there are within this concept. Also an important part on this study is be to gain insight on what challenges using these AI capabilities might bring about. Thus the scope of this thesis is definitely not to go into details on how each AI technology work in a technical manner under the surface, but the aim is to present their capabilities in a general way. So the aim is to understand what can be done with these AI capabilities and what things to consider in the context of using them together with RPA solutions.

1.4 Research structure

Structure of this thesis can be split to two main sections. First section, chapters one to three, covers the theoretical part of this thesis. Following this section, chapters four to six, are based on empiricism. The literature review was first conducted to gain insight about the topic of this thesis. After this the empirical research was conducted, which consisted of interviews that were held based on the findings and implications of the literature review on the theoretical part. This research structure thus follows Figure 2 which is presented in the next page.

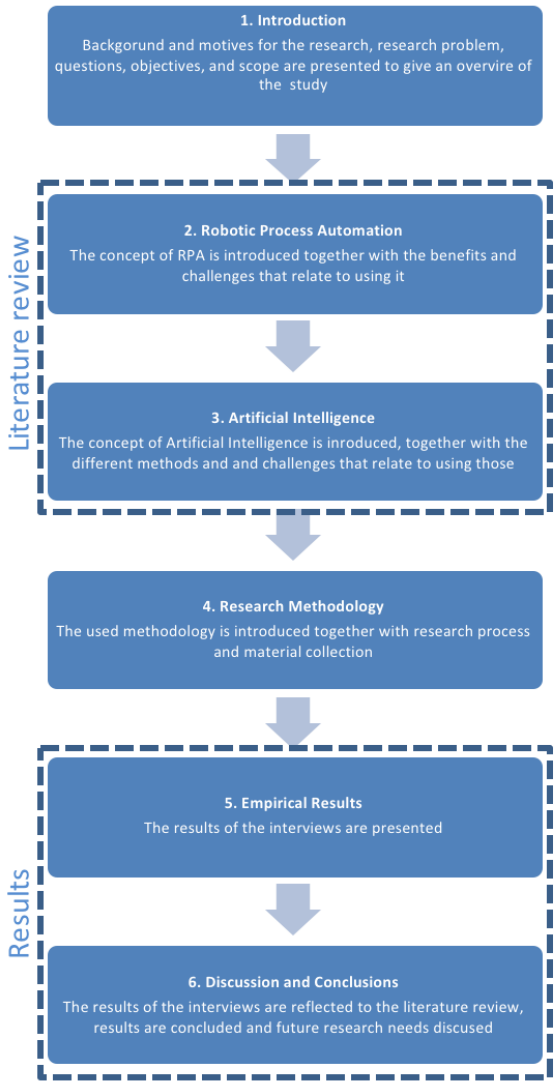

Figure 2. *The structure of the thesis.*

First chapter contains background information including research problems, research questions, scope as well as content of this thesis. Second chapter give theoretical information about the subject of RPA and what this technology is capable of. Third chapter, the last theory based chapter, covers basic information about AI, what it actually means, what different capabilities does it consists of and what kind of challenges relate to using AI. Thus second and third chapter conclude the literature review part of this thesis.

Fourth chapter starts the empiricism part. This chapter is about the research methodology, and so includes information about research methodology, research process and material collection. The following fifth chapter includes the empirical results from the interviews that were held for the thesis. Then in the last sixth chapter, Discussion and Conclusions, the results from the whole research are examined and summarized by reflecting the gained insight and results of the interviews with the literature review. The fifth and the sixth chapter conclude to the results of this thesis.

2. ROBOTIC PROCESS AUTOMATION

This chapter starts with theoretical information of robotic process automation (RPA) in a general level. Therefore this chapter focuses on introducing what RPA is, how and where it can be used and what kind of benefits as well as on the other what kind of limitations there are to utilizing it to automate processes.

This second chapter therefore answers to the following research questions: "*What is robotic process automation?"* and "*How robotic process automation can be used to automate processes and what benefits as well as limitations does it have?"*. The outcome of this chapter is to give a good overview and theoretical understanding on the subject of RPA and to provide insight on what kind of factors and considerations relate to using it. This chapter is fully theory based.

2.1 Robotic process automation concept

In today's fiercely competitive environment companies and organizations are often faced with challenges to staying competitive and still being able to run effective and successful businesses (Forrester Consulting 2014; Richter & Brühl 2016). This in turn has led to a situation where streamlining processes have become a more and more critical success factor (Richter & Brühl 2016). The effect of this can especially be seen in companies' so called back offices, which cover many parts of companies core process areas in functions like procurement, human resources (HR) and finance, where the operational support systems serve as key tools to create, manage and deliver services (Craig et al. 2015). These include systems such as enterprise resource planning (ERP), customer relationship management (CRM) and supply chain management (SCM) to support and ease the processes. (Kukreja & Nervaiya 2016.)

Traditionally companies have widely adopted many ways to improve these kind of back office processes. Solutions have been centralizing the processes to shared service centers, improving the processes by standardization and optimization, improving the processes with technological enablers i.e. ERPs, CRMs and such systems or outsourcing and offshoring to lower cost countries (Craig et al. 2015). Within last few decades especially outsourcing and offshoring these kind of back office operations to BPOs, abbreviation of Business Process Outsourcing, has grown to become increasingly popular as they usually offer the same or better level of services with a fraction of a cost. Despite this being often a great solution for cost reduction, there are still a few challenges that often arise when using outsourcing solutions. For example hidden costs of management, service level agreements are often overly complex and problems within communication sooner or later emerge. (Asatiani & Penttinen 2016.)

Similarly though technology, gradual process improvements and system automation at some level have eased up the required working hours spent with these back office processes but they still conclude to a big workload with lots of repetitive transactions and tasks that have to be executed monthly, weekly or even daily. Thus these processes also require a lot of human resources and the people handling these processes may consume a significant amount of their working hours to handle just these repetitive transactions over and over again (Aguirre & Rodriguez 2017). This in turn means that they sum up to a large and ongoing expense also from the companies' perspective. In addition, when there are humans intervening, there is also a change and even a great probability for errors to happen at some point (Kukreja & Nervaiya 2016; Varghese 2017), after all human beings are fallible.

In a last couple of years a new kind of technological solutions addressing business process related problems and aiming to further improve and automate processes has entered the markets called robotic process automation, or as usually just shortened to RPA. Even though the term "robotic process automation" might first bring to mind a vision of a physical robots handling human tasks in offices, the term actually means different kind of a robot (Asatiani & Penttinen 2016; Craig et al. 2015; Lacity & Willcocks 2015). RPA refers to a software robot that automates knowledge intensive, high-volume and repetitive tasks that were previously performed by humans (Aguirre & Rodriguez 2017; Asatiani & Penttinen 2016; Anagnoste 2017; Craig et al. 2015; Forrester Consulting 2014; Jesuthasan & Boudreau 2017; Kukreja & Nervaiya 2016; Lacity & Willcocks 2015; Lacity & Willcocks 2016b).

In essence RPA aims to replicate and mimic human actions on the computer manipulating and using existing applications, like ERPs, CRMs, spreadsheets and email, working on the presentation layer to perform certain tasks configured to it (Aguirre & Rodriquez 2017; Asatiani & Penttinen 2016; Kukreja & Nervaiya 2016; Lacity & Willcocks 2016b). This said, RPA is a software that is configured to perform processes and replace human factor in the middle. This revolution of automating business processes is now happening inside many companies' business operations and it's already replacing physical labor in many processes (Lacity & Willcocks 2015; Kukreja & Nervaiya 2016).

2.2 Robotic process automation vs. traditional IT development

RPA's great advantage lies in the logic how it works. It takes a very much different approach to software development compared to traditional IT development like business process management (BPM). RPA enables automating manual and repetitive rule-based tasks in a cost effective way as the integration is done without disrupting the underlying logic in legacy systems, as RPA uses applications like humans navigating in the user interface (Aguirre & Rodriguez 2017; Asatiani & Penttinen 2016; Burnett et al. 2017; Fersht & Slaby 2012; Kukreja & Nervaiya 2016) with its own IDs and passwords.

Thus RPA takes more lightweight approach to process development as it operates on the presentation layer whereas traditional IT development like BPM interacts on the business logic and data access layers (Aguirre & Rodriguez 2017; Fersht & Slaby 2012; Lacity & Willcocks 2016b) as shown in Figure 3 below. BPM is also more about finding new solutions for existing business problems as well as creating new opportunities for gaining competitive edge (Aguirre & Rodriguez 2017; Niehaves et al. 2014). This said BPM is about re-engineering and redefining business processes to make them more efficient (Niehaves et al. 2014), whereas RPA is more about creating a more efficient way to execute existing processes by software automation (Anagnoste 2017; Fersht & Slaby 2012; Lacity & Willcocks 2016b; Varghese 2017).

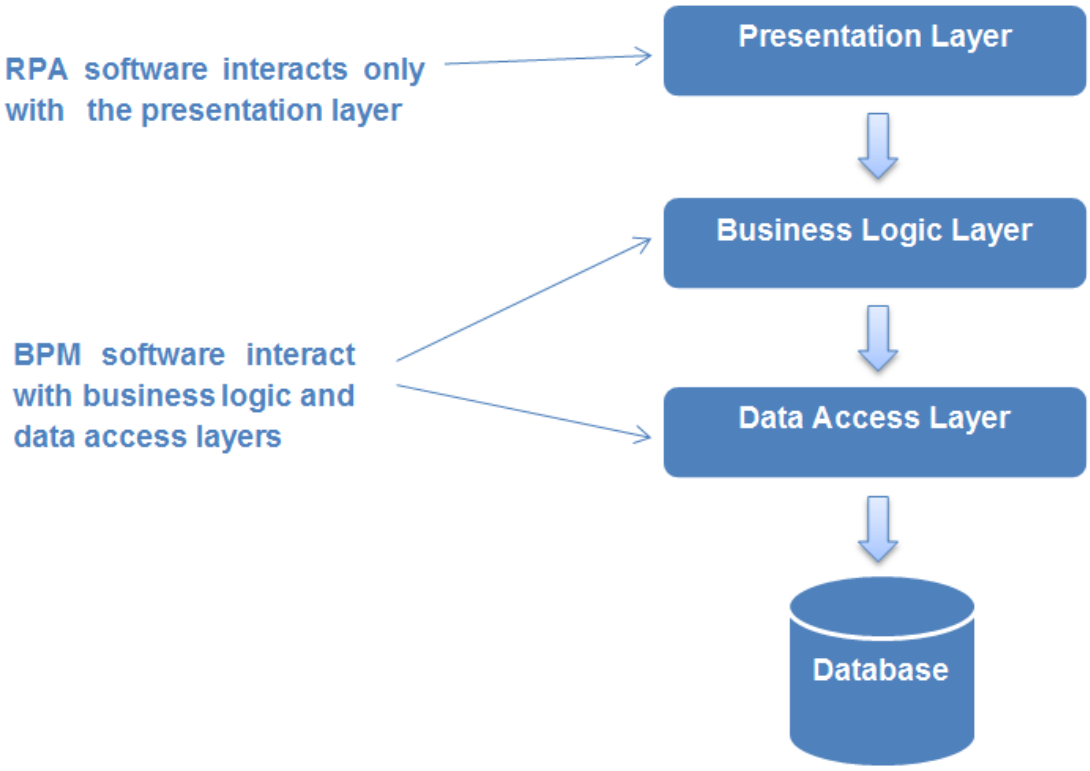

Figure 3. *Representation of the discrepancies between RPA and BPM (Lacity & Willcocks 2016b).*

Thus RPA isn't something that supersedes BPM, but rather complements it. They are both meant for different purposes and are suited for different kind of processes (See Table 1 on the next page). Hence BPM is more suited for process re-engineering purposes and it's best suited for processes requiring high level of IT expertise on expensive and extensive investments like developing ERP or CRM systems (Lacity & Willcocks 2016b).

Another distinguishing fact is that BPM projects are usually owned by the IT department but RPA solutions on the other hand are usually deployed inside business units (Lacity & Willcocks 2016b). This is due to the fact that RPA is relatively easy technology to understand and learn as it executes the processes similarly to how human counterpart would. This in mind, robotizing processes using RPA doesn't necessarily require comprehensive

technical and programming skills (Lacity & Willcocks 2016b; Anagnoste 2017), whereas traditional IT development requires experienced engineers and software developers (Fersht & Slaby 2012) as automation is done by back-end integrations which frequently requires some redesigning in the existing systems (Asatiani & Penttinen 2016). An important difference with RPA compared to traditional IT-solutions is also the reusability of the solutions. With RPA, functions and modules used with one robot can easily be transferred to new robots which also applies to traditional IT-solutions but with them it's considered much more expensive (Fersht & Slaby 2012).

Table 1. *Robotic process automation development vs. traditional IT development (Adapted from: Fersht & Slaby 2012; Forrester Consulting 2014; Lacity & Willcocks 2016b).*

Issue	Robotic Process Automation	IT Development via BPM
Business goal	Automate existing processes	Reengineering processes
Development skills required	Modest; can be done by process modelers and analysts with a few months of training	Extensive; requires software architects and engineers with extensive experience with relevant programming languages and BPM tools
Development methodology	Lightweight; development is done on the presentation layer	Heavyweight; requires business logic and possibly data access -layer integration
Component re-use	High; functions are reusable when developing new robots	High; relatively expensive to develop

As BPM solutions require extensive IT knowledge, with RPA what is more needed is a profound understanding of the business processes targeted for automation (Anagnoste 2017; Lacity & Willcocks 2016b). This makes it possible that business people with process and subject matter expertise can be trained to automate processes within only a couple months and even without any prior experience from programming (Anagnoste 2017; Asatiani & Penttine 2016; Fersht & Slaby 2012; Lacity & Willcocks 2016b). This combined with the fact that RPA doesn't interfere existing systems' underlying logic means that the threshold of business processes worth automating becomes much lower (Asatiani;

Penttinen 2016; Lacity & Willcocks 2016b). This is also shown below in Figure 4 where the "long tail" illustrates the processes suitable for automation with the use of PRA.

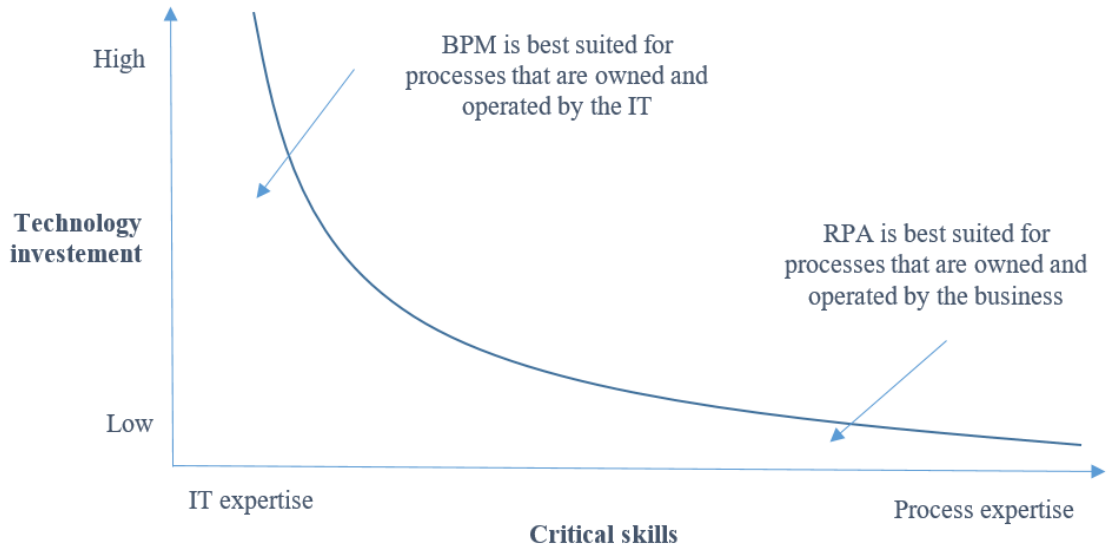

Figure 4. *RPA and BPM application areas (Adapted from: Lacity & Willcocks 2016b).*

The key differentiation here is that organizations often have a few specialized high value processes where some level of automation is already built in. These typically relate to systems like ERP, SCM and CRM and incorporate the kind of processes which are owned by the IT function. In addition to these, companies typically have a long tail of processes that are of lower value and are usually deployed and handled by the operations people. These lower value processes then often are too small to justify the use of scarce IT resources. (Lacity & Willcocks 2016b.) This in mind these processes provide a great deal of possible use cases to implement RPA.

2.3 Adoption of robotic process automation

A successful adoption of RPA into companies' operations and processes requires an in-depth analysis of the targeted processes for robotization. When assessing automation suitability with RPA it should be determined whether the process can be written down step-by-step with all the possible events, errors and outcomes that could occur. This is because the robot can only execute certain steps that are taught to it. (Asatiani & Penttinen 2016.) It doesn't know how to behave if an unexpected error message prompts up or if any pre-taught step doesn't happen exactly the way they was taught to the robot. Thus it's very important to understand that software robots need a much more explicit and unambiguous instructions than humans. (Asatiani & Penttinen 2016; Fersht & Slaby 2012; Fung 2014.) RPA solutions are in this sense "stupid robots" as they are not capable of using judgment when making decisions nor have they common sense. Thus human beings are better equipped at responding to these kind of unexpected situations as they're more adaptable

and creative by nature (Singh et al. 2009). Processes requiring interpretation, adhoc decision making or subjective judgement aren't in this sense suitable for RPA (Fung 2014). This is also why low cognitive requirements and minimal human intervention on the process are critical success factors for RPA implementations (Fersht & Slaby 2012; Fung 2014).

An important factor when assessing the automation potential is also to look at the volumes of the process. Here high-volume processes are often the most potential implementation areas for RPA as high-volume also indicates that there's opportunity to cut costs (Fung 2014; Fersht & Slaby 2012; Lacity & Willcocks 2016b). High volume usually also indicates that processes are rather simple and routine like (Aguirre & Rodriquez 2017). Routine tasks in turn are defined as tasks that follow certain rules (Frey & Osborne 2017), so they are manual tasks that don't necessarily require cognitive capabilities. In this sense the automation potential and implementation possibilities for RPA increases the more high-volume routine like and the more manual rule-based the task is, as illustrated below in Figure 5.

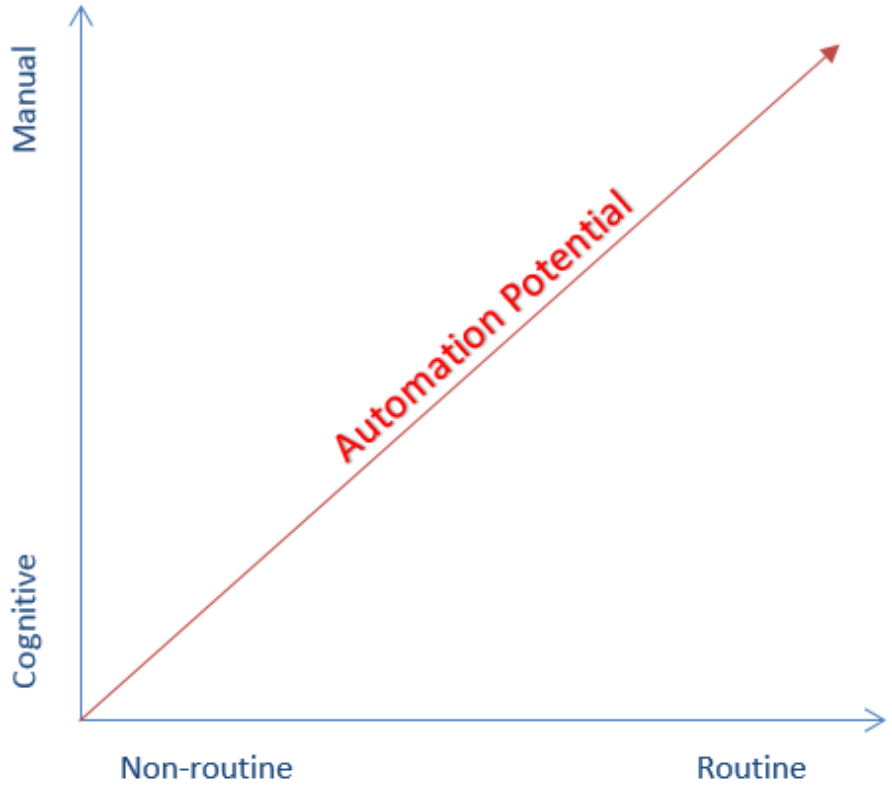

Figure 5. *Robotic process automation potential (Asatiani & Penttinen 2016).*

Another important factor when assessing RPA's potential is standardization of the processes. This is because standardized processes are often measured, already well documented, run on a stable environment and have predictable outcomes (Lacity & Willcocks

2016b). Predictable outcomes here are very important as the robot needs exact and unambiguous rules to work with as every possible step in the process need to be taught to the robot so that the execution of the process can be flawless (Fehrst & Slaby 2012). Similarly, stable environment is important due to the fact that the systems and environments that the robot works with need to stay the same every time the robot executes processes so that there are no unexpected issues that could crash the robot in the middle of performing the tasks assigned to it (Fung 2014). This also minimizes the error handling need which is a critical success factor with RPA as the more there are errors and exception cases, the more delay it causes to the development of the process as well as to the execution of the ready built process.

Standardization on the other hand drives another important benefit for assessing RPA's potential as when the processes are measured and well documented there's also usually good visibility to the different costs that relate to them (Lacity & Willcocks 2016). Visibility to the costs here is important because when companies understand the cost structure of the current process they are also able to estimate the cost difference between the robot and human worker and so calculate the return on investment (ROI) that RPA implementation would generate (Fehrst & Slaby 2012). Without visibility to the costs of the current process, and so to the ROI of RPA implementation, it might be difficult to justify the costs related to robotization.

Good process candidates for RPA also require access to multiple different systems. This is attractive feature as communicating between different, possible otherwise not together configurable, systems is easy to carry out with RPA as it accesses systems simply using the presentation layer and doesn't touch the underlying logic of the systems (Aguirre & Rodriguez 2017; Asatiani & Penttinen 2016; Burnett et al. 2017; Fersht & Slaby 2012; Kukreja & Nervaiya 2016). This way systems that otherwise often couldn't be integrated or the cost of traditional IT development initiatives couldn't be justified for the process, can now be integrated to work together (Lacity & Willcocks 2016b). In conclusion the criteria of using RPA solutions to automate processes are concluded on Table 2 on the next page.

Table 2. *Criteria for robotic process automation.*

Criteria	Description
High volume	Process has high volume of transactions or it's performed frequently
Process uses multiple systems	Process involves accessing multiple systems that otherwise couldn't be easily integrated
Stable environment	Mature systems and environments which remain the same every time process is executed
Process can be broken down to unambiguous	Process can be broken down into exact step-by-step rules which have no room for misinterpretation
Minimum need for exception handling	Highly standardized process with little need for handling exceptions
Manual costs are known	Cost structure of the current process is known and ROI can be calculated for the RPA solution
Low cognitive requirements	Process doesn't require judgement or complex interpretation skills

However it's worth noting that process doesn't need to meet all criteria to be a suitable candidate process for automation with RPA. These criteria work as more of a guideline to finding suitable processes and for building a compelling business case for robotization (Fersht & Slaby 2012).

Thus RPA isn't an all-around solution for every case. There is always a need for process assessment where the process in scope should be thoroughly assessed to get a more holistic overview on all the aspects related to it, and to figure out what kind of solution would give the best and most optimal result. In some cases it might be e.g. better to use BPM methods and reengineer the process, because there might be a more profound reason to change the way process is executed. Thus it's important to carefully assess the process; situation where RPA is needed and find out is there a better solution to solve the problem, and not to just jump to any solution.

2.4 Benefits and limitations of robotic process automation

RPA as a technology solution for automating processes is a very powerful tool and can drive great benefits when implemented properly. RPA's first and foremost benefit is the fact that it can save a lot of time and increase productivity with minimal process change (Aguirre & Rodriguez 2017; Fersht & Slaby 2012; Lacity & Willcocks 2016b; Le Clair et al. 2017; Kukreja & Nervaiya 2016). With RPA processes can be between 30-70% faster than when executed by humans (Kukreja & Nervaiya 2016). In the speed there are

nevertheless some limitations that come from the fact that RPA operates in the user interface navigating through systems like human, meaning the process can't be executed faster than the overall process and systems allow (Lacity et al. 2015). Thus RPA doesn't often exceed BPM or traditional IT development solutions when it comes to maximum speed of handling tasks but neither is this necessary as with RPA the task can still be done significantly faster than the human counterpart could. This in mind robots can work 24/7 without interruptions (Fersht & Slaby 2012; Kukreja & Nevaiya 2016; Willcocks et al. 2016b), but again the existing process might set some limitations here (Willcocks et al. 2016b). In addition when robots are doing tedious, mundane and time consuming manual tasks, humans are able to concentrate on more creative, meaningful and value creating tasks (Kukreja & Nervaiya 2016).

Another great benefit and often the one benefit that companies first start to look with RPA is the cost effectiveness. Robotization of a process with RPA is relatively fast and cost effective way for automating processes as the time needed to develop the process with RPA is usually just a few months (Fersht & Slaby 2012; Lacity & Willcocks 2016). This combined with the fact that robots can work around the clock without disruptions and without additional costs after implementation, other than costs related to running the platform (Asatiani & Penttinen 2016; Kukreja & Nevaiya 2016), makes RPA a very competitive solution for handling and automating processes. However the cost savings of a software robot might vary greatly firm to firm and depending on was the work previously handled in-house or was it maybe outsourced to an offshore location (Burnett et al. 2017; Fersht & Slaby 2012). In addition the actual realized savings might also be difficult to accurately assess as often people aren't let go but are redeployed elsewhere often to more non-routine value creating tasks (Lacity & Willcocks 2016b; Kukreja Nervaiya 2016).

As RPA operates on the presentation layer manipulating applications the same way as humans it's also possible to implement RPA solutions virtually with almost any software that humans use (Aguirre & Rodriguez 2017; Asatiani & Penttinen 2016; Burnett et al. 2017; Fersht & Slaby 2012; Kukreja & Nervaiya 2016), even if the applications aren't open to third party integrations. This very much broadens the possibilities of RPA solutions compared to traditional IT system integrations as many corporate IT systems don't have public application programming interfaces (APIs) which greatly limit the communication possibilities with other systems (Asatiani & Penttinen 2016). Even if the integration can be done, RPA solutions are significantly less expensive as the solution is built on top the existing systems and the underlying business logic and data access layers aren't touched (Aguirre & Rodriguez 2017; Fersht & Slaby 2012; Lacity & Willcocks 2016b).

Another significant benefit that results from the use of RPA is accuracy. RPA itself minimizes errors in the processes as robots don't make mistakes because everything in how they work is backed by the logic taught to the robot. (Kukreja & Nervaiya 2016.) This of course requires that all the necessary rules and exception handling for the process have been described to the robot's logic. If all this is done correctly then robot will do the job

exactly as it's taught and if it counters an error it should relate to transaction that doesn't follow the set up parameters, and then the transaction is put aside for the human to handle (Kukreja & Nervaiya 2016). This in turn points to the fact that RPA is best suited for managing structured processes with structured data (Lacity & Willcocks 2016a), because in these cases a specific correct result can be derived and so defining specific rules for RPA software is possible. Related to the fact that RPA solutions follow exact rules taught to them they also log every step in the process and thus the results are easily auditable which decreases compliance risks (Fersht & Slaby 2012; Lacity & Willcocks 2016b) and gives transparency to the processes. This means that it's also easy to find problems in the process flow and possible bottlenecks can easily be identified (Fersht & Slaby 2012).

Whereas RPA can drive multiple benefits, it does come with its flaws and limitations on where and how it can be utilized. First and most important one is that RPA is a software solution and thus needs digital inputs and the target process must have unambiguous rules where there is no room for interpretation. Software robots can't make decisions by themselves without clear predefined rules on how to move forward on each and every possible scenario (Aguirre & Rodriguez 2017; Fersht & Slaby 2012; Fung 2014; Lacity & Willcocks 2016; Kukreja & Nervaiya 2016). In this sense robots need much more explicit information than humans. This also means that processes requiring some kind of judgement or ad-hoc decision making are directly out of RPA's scope of automatable processes – processes need to have a single correct answer as a result (Lacity & Willcocks 2016c). Therefore, processes that use RPA need to have a deterministic outcome (Aguirre & Rodriguez 2017; Lacity & Willcocks 2017c).

Another limitation for RPA relates to the kind of input data software robot needs. RPA in essence is a software solution and thus requires input data in a digital format. Secondly, and most importantly, the input data need to be in structured format (Aguirre & Rodriguez 2017; Lacity & Willcocks 2016c). As mentioned robots can't make decisions, use judgement or interpret data, and so if the input data is not exactly in the expected format robot will have problems identifying the needed information from it. Therefore, robots are good with input data from systems like ERP, CRM or SCM, as well as handling data from spreadsheets and other digital input that have data in a structured predictable format (Aguirre & Rodriguez 2017; Anagnoste 2017; Asatiani & Penttinen 2016; Lacity & Willcocks 2016a). Issues however arise when robot starts to handle other kind of data sources, be it emails, free-text forms or any other data format that doesn't have strict limitations and predefined rules on how the data is presented. In these cases interpretation of the data isn't so straightforward as the way data is presented might vary from case to case, and so the robot would have to decipher what the data actually means. (Lacity & Willcocks 2016a.) For humans of course these kind of input data usually doesn't cause any problems as humans are by nature more flexible and creative (Singh et al. 2009), and so can adapt to new situations better and interpret also other kind of inputs than those that are in a

strictly defined format. This however isn't the case with RPA. Table 3 below shows a short summary of the benefits and limitations of robotic process automation.

Table 3. *Benefits and limitations of robotic process automation.*

Benefits	Limitations
Can save a lot of time and significantly increase productivity with minimal process changes	RPA can only handle input data that is in a structured format
Cost effective solution for automating processes	Process needs to have unambiguous step-by-step rules
Robots don't make mistakes if process is taught correctly	Can't handle processes where use of judgement or adhoc decision making is needed
Can be applied to virtually any digital process that humans perform	

As can be seen from Table 3 many of RPA's greatest benefits also have their downsides that create actual limitations to the possible implementations and use cases of RPA solutions. This doesn't of course mean that RPA would be insufficient solution for automating business processes, but it means that RPA is suitable for certain manual, well defined and mature processes.

2.5 Robotic process automation vs. Intelligent automation

As a technology RPA is an ideal candidate to automate rule-based processes and activities that transact with structured data. (Anagnoste 2017; Aguirre & Rodriguez 2017; Burnett & Modi 2017; Lacity & Willcocks 2016a). Therefore the outcome of these processes and activities need to be highly predictable and thus deterministic (Aguirre & Rodriguez 2017; Lacity & Willcocks 2017c). This is also the problem of RPA as the prerequisites of the processes limit the scope of possible implementation areas of RPA solutions. However more recently also another kind of automation area has gained increasing interest called intelligent automation or cognitive RPA, which is often referred as the next step in automating business processes after pure RPA (Anagnoste 2017; Burnett & Modi 2017).

Here intelligent automation differs from traditional RPA in a sense that it combines other technologies and capabilities, like machine- and deep learning, with RPA to enable learning and automation of processes that traditionally require human decision making. The key differentiating factor here is that these AI based solutions enable processes to learn and adapt based on experience. (Burnett & Modi 2017.) Whereas purely RPA based processes must be taught to perform specific tasks based on certain logic – they can't adapt to changing situations and changing input data, but need to be taught every varying scenario . This also means that RPA solutions are good only for clearly defined processes which stay the same and use standardized structured data as their input. (Aguirre & Rodriguez 2017; Asatiani & Penttinen 2016; Anagnoste 2017; Kukreja & Nervaiya 2016; Lacity & Willcocks 2016b.) Intelligent automation, which takes advantage of AI solutions, in addition to RPA, use different algorithms and capabilities such as statistical analysis, machine learning, image recognition, natural language processing to also work with unstructured data. (Burnett & Modi 2017.)

As intelligent automation, where RPA is combined with AI capabilities, has the ability to learn, handle varying process scenarios and exceptions means that the processes where it's implemented can have multiple possible outcomes with varying certainty of occurrence, and thus be probabilistic in nature. This means that the outcome of the process is then dependable on how AI handles its parts in the process. (Burnett & Modi 2017.) This is therefore one critical difference with RPA and intelligent automation – former can only handle processes with deterministic outcome and latter can also handle ones with probabilistic outcome.

As RPA can only handle structured and sometimes semi-structured data sources as input data, the other key differentiating factor is intelligent automations ability to work with unstructured data sources (Burnett 2017; Lacity & Willcocks 2016a). This enables extraction of information from various data sources like texts, scanned documents and PDFs that often come in varying layouts. (Burnett 2017.) These are for RPA robot often very cumbersome or even impossible data sources to handle.

RPA and intelligent automation are after all very close to each other when it comes to what they are capable of. RPA here represents the technology and solution that is meant for automating logic and rule based process which use structured data as their input and have deterministic outcome i.e. steps to the outcome can be clearly defined. Intelligent automation on the other hand is muchly an extension for RPA. It combines AI technologies with RPA to enable wider automation potential by enabling handling of unstructured source data and handling of varying process scenarios. AI capabilities therefore fill in the gaps in the processes where RPA can't handle the job. Intelligent automation is therefore a set of AI technologies and capabilities combined with RPA, whereas talking about RPA means RPA technology on its own.

3. ARTIFICIAL INTELLIGENCE

This chapter starts with introducing the concept of artificial intelligence (AI) to give a better understanding on the context of this thesis. Following this comes introduction to different kind of capabilities that relate to AI, as well as considerations on what kind of limitations there are to utilizing AI.

This chapter therefore answers to the following research questions: *"What is artificial intelligence?"*, *"What are different artificial intelligence capabilities and how they can be used?"* and "*What kind of risks are related to using artificial intelligence?"*. The main outcome of this chapter is not go into detail on how different AI capabilities work in a technical manner, but to give a good overall view on what AI actually means and provide understanding on what kind of capabilities there are within the concept of AI as well as provide insight on the constraints and challenges related to them. This chapter is fully theory based.

3.1 Concept of artificial intelligence

Artificial intelligence, or as often just shortened to AI, is a concept that has become somewhat a hype word during the last few years due to technological developments. This can be also seen from the Gartner's hype cycle of emerging technologies where many AI based technologies like Machine Learning, Deep Learning and Artificial General Intelligence are mentioned (See Figure 6 on the next page). However many times these concepts and definitions are used somewhat interchangeable and there don't seem to be consensus on what AI actually even means (Burgess 2017, p. 1; Scherer 2016). Companies that create AI applications stress the potential of it and make grand promises on how it'll change the way people live and how businesses operate whilst any real value creation often drowns in the midst of hyped marketing. Then there is also the technology part to AI which consists of computers, data and mathematic models which often seem to go beyond understanding to people from outside the developer world. (Burgess 2017, p. 1.)

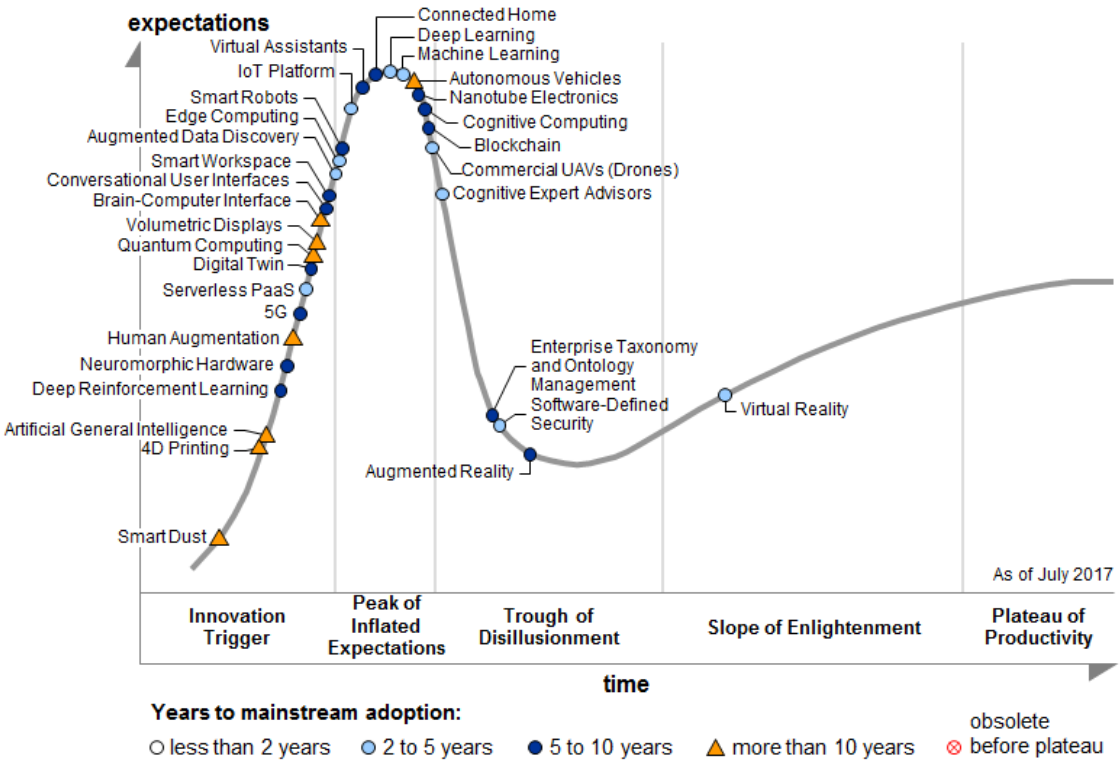

Figure 6. *Gartner's hype cycle for emerging technologies (Gartner 2017).*

All this has led to a situation where the definition and meaning of AI is very much yet to be settled (Burgess 2017, p. 2; Reuner 2017). One definition describes AI as: "capability of a computer program to perform tasks or reasoning processes that we usually associate to intelligence in a human being" (Rossi 2016). Similarly to this are Oxford Reference's (2018) definition: "the theory and development of computer systems able to perform tasks normally requiring human intelligence", Everest Group's definition: "AI is the ability of machines to exhibit human-like intelligence" (Burnett 2017) and Harvard Journal of Law & Technology's definition: "'artificial intelligence' refers to machines that are capable of performing tasks that, if performed by a human, would be said to require intelligence." (Scherer 2016). All of these definitions point to very similar definitions but are however maybe a bit circular as they include the word 'intelligence', and therefore raise a question of what intelligence and being intelligent actually means. This in turn is more of a philosophical question and thus defining intelligence is not addressed in this thesis. Therefore the issue of defining AI doesn't actually relate to the difficulty of defining the concept of 'artificial' but the issue is with the ambiguity of conceptualizing intelligence (Scherer 2016). Therefore definition of AI also relies on what we consider as intelligent behavior of a human being (Rossi 2016).

AI concept and approaches to defining it have thus also shifted through time. Often AI concepts have been tied to some certain ability to perform intelligent like tasks. This in turn has meant that as over time technological advances have increased and expanded computers' capabilities, these tasks have become more 'ordinary' and therefore executing

these tasks don't seem so intelligent anymore. (Burgess 2017, p. 3; Scherer 2016.) Similarly if belief of what is human intelligence changes, also requirement for computer programs performing tasks that are considered AI like, change (Rossi 2016). Thus methods and technologies that previously has been thought as examples of intelligence move to just being ordinary everyday applications.

AI concept itself has existed for over half a century as the term "artificial intelligence" was first mentioned in 1956 by John McCarthy during a scientific conference discussing if machines could be intelligent (Rossi 2016). However for a long time progress in the field was slow and even stagnated a few times due to inflated expectations and withdrawal of funding from AI researches. Only in the last couple decades real improvements and more actual use cases in the field has started to appear which have muchly been made possible due to the rapid technological developments; especially in processing power, growing amounts of data masses, cheapening storages and by ever increasing connectivity. (Burgess 2017, p. 12-19.) For example many everyday applications that are used today in people's everyday life actually take advantage of AI capabilities (Burgess 2017, p. 2; Rossi 2016; Snow 2017). Most noticeably for everyone this can be seen in smartphones. Vast majority of them have sophisticated AI capabilities, whether it's Google Assistant, Siri or Cortana. All of these examples are already pretty mature solutions and they can help in everyday life but they are still rather limited in what they can actually do. All of these rely to turning speech into words and then aim to find meaning for those words. Once the intent has been figured out, the following parts of the process are rather simple automation; be it checking the weather, getting timetables or starting an application. AI as a whole is however much more than just speech recognition and natural language understanding (NLU). (Burgess, p. 2.)

3.2 Artificial intelligence capabilities

To better understand what AI means and what it consists of, the concept of AI can be divided into three main capability groups (see Figure 7 on the next page) based on the objective of what the capabilities are aiming to achieve – what is the goal. These are "capturing information", understanding "what is happening" and understanding "why something is happening". (Burgess 2017, p. 3-4.)

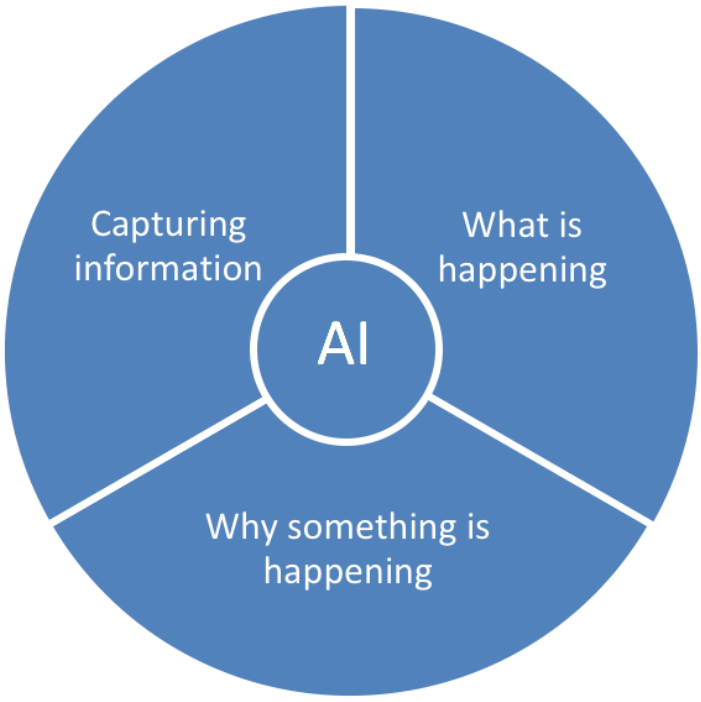

Figure 7. *Three main objectives of artificial intelligence.*

Here then AI can be divided into core capabilities which can then be associated with one or more of the main objectives. These are:

- image recognition,
- speech recognition,
- search,
- clustering,
- natural language understanding (NLU),
- optimization,
- prediction,
- understanding.

In theory all AI technologies and applications fall under some of these capabilities (Burgess 2017, p. 3-4.) Therefore AI is more of an umbrella term for different capabilities and technologies rather than a certain capability in itself.

All these capabilities and technologies that are in the field of AI are however still meant for very specific purposes, in which their capabilities are restricted (Burgess 2017, p. 4; Burnett 2017; Rossi 2016). They might be able to outperform humans already in certain tasks, but then lack in others. Thus these kind of technologies and methods are incorporated with a term Narrow AI, in contrast to general AI which in theory is capable of performing any intellectual task give to it (Burgess 2017, p. 4; Burnett 2017). General AI, or in other words artificial general intelligence or strong AI, in this sense requires machines

to actually understand what is going on and not just simulate thinking or decision making in strictly defined context (Burgess 2017, p. 4; Russell & Norvig 2010, p. 1020). Thus this thesis won't cover the third objective of artificial intelligence "Why something is happening", and "understanding" capability as these fall under the area of general AI, which doesn't really exist yet. AI solutions, no matter how sophisticated, don't yet have the ability to understand what the data means. For current AI systems it doesn't matter what the data actually represents, it could about invoices or ice cream flavors, all it does is correlate data sets against the trained model. (Bergstein 2017.) Thus general AI is still a very theoretical concept (Burnett 2017), and therefore this thesis focuses only on Narrow AI or in other words weak AI capabilities and technologies.

One other way to categorize AI capabilities is to divide them between 'supervised' and 'unsupervised' learning. When AI system is trained using large data masses with labeled data points it uses supervised learning (Burgess 2017, p. 7; Russell & Norvig 2010, p. 695-696). This approach could be used for example to identify object X in a picture, X referring to certain object, out of thousands of pictures where some actually were pictures of object X, and some weren't. Crucial here however is that the pictures would have to be labeled either 'object X' or 'not an object X'. Using supervised learning AI system would learn with large amounts of training data which are inherent characteristics of the object X (Burgess 2017, p. 7; Ng 2016). After teaching the system what object X looks like it could be tested with a new set of similar data but this time the labels would not be revealed. If system was trained well enough it would identify correctly whether the picture has object X on it or not. Supervised learning is usually used when the data is in unstructured or semi-structured format, like pictures, sounds or free form texts. (Burgess 2017, p. 7).

The other approach, unsupervised learning in essence means that the AI system is just fed with very large structured dataset that however means nothing to it. How the AI system then approaches this data is by identifying data clusters. As the system in this approach doesn't know what the data actually represents it's just searching for patterns within the data. (Burgess 2017, p. 7; Russell & Norvig 2010, p. 694-695.) As an example unsupervised learning could be used to define house prices in a particular area. The house price depends on multiple variables: location, room count, age, square footage etc., which'll make prediction of value rather difficult. AI system using unsupervised learning can however find connections and define weighing factors for different variables using statistical analysis. Thus the system learns the weighting factor of each variable on the house value. After training the system with large amounts of data from houses that have all the variables known, the system can then be used to predict the value for a house of which other variables are known but the price is not. (Burgess 2017, p. 7-8.)

As previously pointed AI can be divided to three different main groups based on the objective, and to eight different main capabilities from which all AI applications and technologies can ultimately be derived. However, as a main driver of these AI capabilities

and applications is often machine learning (ML). (Burgess 2017, p. 8.) In the core machine learning is all about computational learning of solving a certain problem where machine writes the 'code' rather than a human – human developer creates the algorithm or algorithms as a base, and the machine creates a solution for this using the data it's fed (Burgess 20117, p. 8; Davenport & Ronanki 2018; Tan et al. 2017, p. 296-298). Problem that machine learning aims to solve can thus be undefined, as in unsupervised learning which aims to find patterns in structured data or defined, as in supervised learning where large data masses, often unstructured, are trained to answer a particular question (Burgess 2017, p. 20).

What is critical with machine learning algorithms is the amount of data – the more data algorithm is fed the more accurate it can be. This can be explained by how one of machine learnings core enablers deep neural network (DNN) works, where the machine creates the 'code' or model to solve the problem by "itself". With DNN AI mimics the human brain's way of processing information – neurons are interconnected and each neuron are first given a random weighing factor that changes as the brain, or in this case machine, learns. How DNNs then in practically work is that they consist of multiple different layers – more difficult and complex the problem is, the more layers there are (see Figure 8 below). Data is then fed to the input layer and the output layer presents answer or answers, because there are as many nodes as there are types of answers. In between these input and output layers are then the hidden layers where all the configuring happens. These hidden layers also have more neurons that input or output layers. Each of these hidden layers look for different features in the data, with increasing complexity the more hidden layers there are. (Burgess 2017, p. 20; Knight 2016; Singh 2017.)

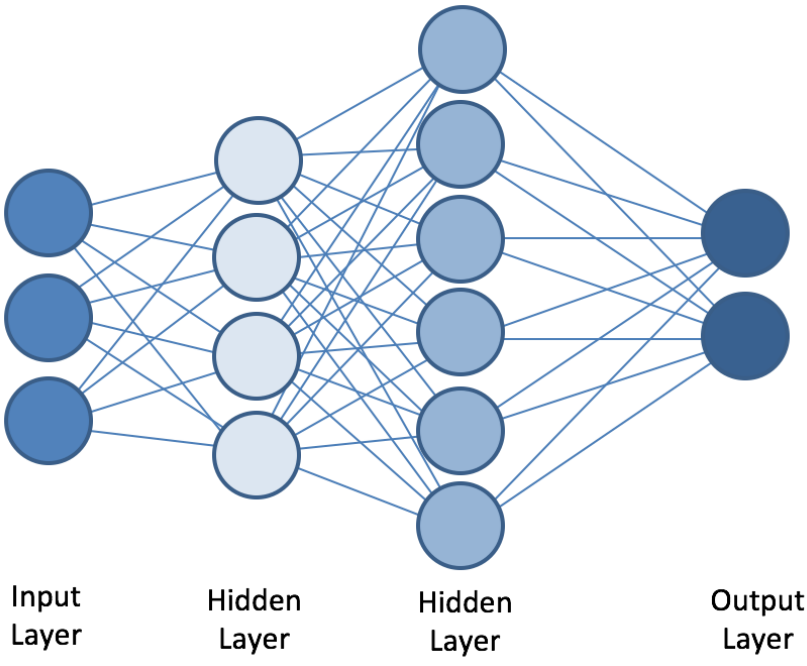

Figure 8. *Representation of neural network (Burgess 2017, p. 21).*

Use case example of such a solution utilizing DNNs could be for example image recognition where different layers look for different things like: shadows, outlines, colors, shapes etc. Each of the neurons have their own weighing which defines how much information is send to the next layer: neurons with strong weighing, which are found through training, imply the 'right' route to the correct answer at the output layer and will propagate more information to the next layer, whereas neurons with low weighing, again found through training, imply route to the 'wrong' answer, and thus won't pass as much information forward. Weights of the neurons are then constantly adjusted the more data the model is fed, until the optimal solution for the model is reached. Therefore the more data the model is fed, the more chances there are to refine the weightings, and so find an optimal and more accurate solution. (Burgess 2017, s. 20-21; Knight 2016; Somers 2017). This means that the models need huge amounts to data so that the neurons in each layer can come up with the right weightings.

Point worth noting is also the fact that when AI is actually used it's most often implemented as a combination of multiple capabilities (Burgess 2017, p. 4; Burnett 2017). Meaning that AI applications and technologies usually consist of more than one of the earlier mentioned capabilities. Therefore once the individual capabilities potential are understood they can be combined like modular parts to create solutions to different kind of problems.

3.2.1 Capturing Information

The first group capturing information consists of methods aiming to get structured, unstructured or big data. Examples of capabilities within capturing information are: image recognition, speech recognition and search capability. Within these capabilities AI aims to capture all the structured, unstructured or big data, and then turn the gathered data into useful information. (Burgess 2017, s. 3.) Thus the goal might be to turn unstructured data into structured format, like recognizing object X from a picture, or make sense out of huge masses of structured data, like finding logical patterns within the data (Burgess 2017, p. 30).

First capability of "capturing information" is image recognition which is based on machine learning and needs vast amounts sample images as training data. Images themselves can be categorized as unstructured data and image recognition has here three main types of applications: identifying images, looking for similar type of images and looking for differences in images. Identifying images is about figuring out if picture contains a certain object or not (as object X in Chapter 3.2). It can be used to find certain type of images or it can be used to group similar images by 'tagging' them. Photo tagging is also perfect example of previously mentioned supervised learning (Chapter 3.2) as with tagging AI system is trained with hundreds of thousands, or millions, of tagged photos. This in essence means that the more available input data, in this cases images, the system can obtain, the more advanced and accurate the system ca be. (Davenport & Dreyer 2018.)

In addition to tagging image recognition can be used to find images that are similar to other images. This on the contrary to tagging is an example of unsupervised learning as the AI system doesn't need to know what is actually in the image, but it just looks for images that have similar characteristics. Last application of image recognition is to spot differences in images (Burgess 2017, s. 32.) Good example of this kind of application is the image recognition use in radiology to spot anomalies in images; like bone fractions or cancerous lesions. Usually the AI system is able to identify these anomalies as accurately as human radiologist, and in some cases even outperform humans. (Davenport & Dreyer 2018.)

Second method of "capturing information" is speech recognition. What this in practice does is encoding of unstructured speech to structured format, text. Speech recognition then uses supervised learning to match the transcripts with tagged training data (Burgess 2017, s. 33-34; Ng 2016). However as speech recognition uses voice as input data, which in essence is very unstructured, the quality can differ drastically depending on how the voice is recorded. Thus the accuracy, which can also be measured by Word Error Rate (WER), of speech recognition system can drop from a high of 7% to over 16% (Burgess 2017, p.34.), compared to human accuracy which is estimated to be around 5% (Brynjolfsson & Mitchell 2017) this is already a radical drop. This obviously puts a great challenge for the system to give consistently good results. Similarly different languages and little things like regional accents bring their own challenge to the mix, and when vocabulary widens also the contextual understanding becomes more difficult for the system, as words can have different meanings in different contexts (Burgess 2017, p. 34; Cambria & White 2014).

Third method in the "capturing information" objective of AI is search. By the term search is referred to cases where structured data is extracted from unstructured text. Search can be categorized as supervised learning and it works with both unstructured data, meaning emails, reports, free text forms and so on, and semi-structured data, which in turn have some consistency between the data points but still enough variation to be troublesome for logic based systems to handle. (Burgess 2017, p. 35.) This in mind a good example of semi-structured data could be invoices as they generally contain similar information, but the information can be presented in different places and in a different ways (Burnett 2017). However eventually once the AI system is trained using different kind of sample invoices it is able to learn with a good percentage all the variables and possible different representations of the same data fields. Thus if the place or the representation of one data variable changes the system can cope to the situation, recognize the variable in another place and change it to a standardized format. Secondly, with unstructured data the AI system can do two things, it can either categorize the text by correlating the patterns of words with the learning data, and it can extract entities from the data; like names or places. (Burgess 2017, p. 35.) This capability can then be very useful for example when combining the categorization and entity extraction to automatically handle customer emails; by

categorizing the context and then extracting the relevant information regarding the sender and the issue.

The last method of "capturing information" is clustering. Clustering differs from the previously mentioned methods as it always handles structured data. This in mind clustering is about finding patterns of similar kind of data, inside the data. As clustering works with structured data it can also learn unsupervised, without tagged data. Basically clustering is thus just about finding similar items and putting them together as well as about separation of dissimilar items. (Burgess 2017, p. 37; Russell & Norvig 2010, p. 694-695; Tan et al. 2017, p. 61.) Usually clustering is then about extracting information from new data based on the alignment to the training data, or it could be used to sport deviations in the new data. Human are of course themselves quite good at identifying patterns in datasets but when it comes to huge data masses where there might be hundreds of different variables then AI system using clustering will prove its worth. (Burgess 2017, p. 37.)

3.2.2 What is happening

The second group which objective is to determine "what is happening", consists of NLU, optimization and prediction methods. All these methods in essence aim to determine, usually based on the just captured information by another AI capability, what is actually happening. (Burgess 2017, p. 3.) The first one of these is NLU which works as a sort of translator between humans and machines, turning unstructured data of sentences into structured intents. Therefore the main purpose of NLU is to figure out the correct structure and meaning of the transcription, based on which a model of the input text is created. The created model itself is probabilistic, and so word meanings can change as the system 'learns'. (Burgess 2017, p. 38-39.) Obviously some great challenges relate to NLU; words can have several meanings, different words can have similar meanings, sentiments might not be clear and context might change the whole meaning of a sentence (Knight 2016; Socher 2018; Srinivasan 2016). In the actual applications NLU is often combined with speech recognition to create a solution where speech recognition transcribes the words from sounds, search function, or often in this context called as natural language processing (NLP), looks for patterns and turn text in to data, and then NLU finally defines the intent (Burgess 2017, p. 31; Srinivasan 2016). The first part using speech recognition is often the more critical part, as if the quality of input sounds are not good then not much of the data can be used. Often speech recognition is then removed from the process by using an application called chatbot, which mitigates risks of speech recognition getting the words wrong, as it uses human's written text as its input (Burgess 2017, p. 39). Chatbot in short thus means a virtual agent which interacts with humans using natural language. (AbuShawar & Atwell 2015; Asbjørn & Brandtzæg 2016). Thus it gives an easily accessible interface for communicating with the AI.

Second method within the AI objective "What is happening" is optimization. This AI capability is in the core of what people often think as an AI, as it gets closest to mimicking

the human way of thinking process. In essence the characteristic of optimization capability is that there is a goal to be achieved, which could be an idea that needs reasoning, a problem that needs a solution or a plan to be made. This goal is then tried to achieve through iteration rounds, where small changes and actions are made that affect the situation and then the situation is reviewed to see whether the made changes affected the goal to be closer or not. If the changes had a positive effect and goal is closer the system carries on the same path making further changes. Respectively if the changes affected negatively for reaching the goal, then the system tries something different. Traditionally these optimization systems have been basic decision trees which needed humans to configure them but after the rise of machine learning much of the models' design can now be done by the AI system itself. (Burgess 2017, p. 41-42). Basically using optimization the AI system aims to find an optimum strategy which to follow.

Last of the AI capabilities is prediction. By prediction is meant AI's capability to match new data points to an identified group based on the historical data (Burgess 2017, .p 46; Jung et al. 2017). This in mind prediction often follows clustering, with which the groups could be identified, or search, with which patterns could be defined that the prediction capability then looks for. From optimization prediction in turn differs in the sense it doesn't aim towards a set goal, prediction just correlates the new data against the historical data. (Burgess 2017, p. 46.) Thus prediction can be particularly useful to help improve decision making, especially when handling large data sets (Agrawal et al. 2017). Similarly to other capabilities within AI, the more features the system has to work with the more it requires training data to figure out the different variables and their weightings correctly. Then when the prediction model is fed with new data it correlates the variables with historical data and gives a prediction with a probability for correct answer (Burgess 2017, p. 47).

3.3 Challenges and limitations

AI capabilities can be useful tools to handle task from capturing information to understanding what is happening. However even AI isn't perfect by any means and there are some serious challenges and limitations that might even prevent the use of AI in certain cases. First one of these relates data. As described already through Chapter 3.2. every AI capability and so application is very data hungry. This means that for the AI systems to properly work they require huge amounts of data for training (Ransbotham et al. 2017). This isn't of course a limitation, AI system can be trained with even little amounts of data, but then the model faces a challenge – is it accurate enough anymore. Thus the more complex the problem the more data it needs to be trained with, so that the AI system can learn all the necessary variables and their weighting for the model correctly.

If the model isn't trained with enough training data the AI system might also face an issue called over-fitting. This essentially means that the model isn't accurate enough and when

correlated with new data against the model results might not represent the reality. Over-fitting then becomes more likely when the amount of variables in the data grow, and respectively decreases the more training data model gets or the amount of variables there are within the data (Russell & Norvig 2010, p. 705.) As an example when using prediction to recommend certain product to consumer in a webshop, if the model is over-fitting then AI system is recommending similar products to everyone no matter what their previous buying behavior implies. (Yeomans 2015.) In this sense when the model is over-fitting it doesn't take into account all the necessary variables to give proper predictions as it hasn't been fed with sufficient amount of data to learn these.

What however really is the true problem of many of today's AI systems it the opaqueness of the decision making. This is due to the fact that validating the decision AI system makes can often be quite troublesome. Many times the systems can be so complicated that even the engineers who designed them in the first place might have serious troubles figuring out the reason for any single action the AI system makes (Knight 2017). This relate especially to unsupervised learning where AI systems use DNN for reasoning. These work in a way that the system is fed with structured data which actually means nothing to the system – the data points are not labeled. What the system then does is that it spots clusters of similar data points within the data. (Burgess 2017, p. 7-8). In these cases it's very hard to understand how the reasoning in these systems happen as they are no coding or algorithms behind in how they work and there's no transparency on how the systems got to the final results. This is also why they are often called "Black Box" systems as they are, in many cases, very difficult for outsiders to examine and even for the creators to understand how they reach certain conclusions. (Knight 2017; Snow 2017; Srinivasan 2016.) Especially in certain, business and medicine related, critical applications where humans are held responsible this can be a problem as the users of AI system would have to be able to trust the AI's reasoning (Rossi 2016; Srinivasan 2016). If this isn't possible some of the most critical use cases might vanish. This in turn might mean that the systems can't be used in many cases where the transparency and validation possibility of the decisions are crucial. Thus for the AI systems to be truly revolutionary for the business, they need to be as transparent as possible (Rossi 2016).

The lack of visibility to the AI's decisions making might also raise another kind challenges. This is the case especially with regulation. Any regulation directly related to the deployment of AI systems still lack behind as the technology is changing so rapidly and so many different applications emerge all the time. Despite the fact that the field still lacks regulation in many areas, there are some issues that are also AI related and that have been considered from the regulation perspective, especially in EU (Rossi 2016). Most noticeably this can be seen in the General Data Protection Regulation (GDPR) which took effect as law across EU in May 2018. There in the GDPR especially the Recital 71 shows implications that restrict the use of AI:

"The data subject should have the right not to be subject to a decision, which may include a measure, evaluating personal aspects relating to him or her which is based solely on automated processing and which produces legal effects concerning him or her or similarly significantly affects him or her, such as automatic refusal of an online credit application or e-recruiting practices without any human intervention. Such processing includes 'profiling' that consists of any form of automated processing of personal data evaluating the personal aspects relating to a natural person, in particular to analyse or predict aspects concerning the data subject's performance at work, economic situation, health, personal preferences or interests, reliability or behaviour, location or movements, where it produces legal effects concerning him or her or similarly significantly affects him or her…." (EUR-Lex 2016.)

Similarly Recital 22 shows implications that affect the use of AI:

"Any processing of personal data in the context of the activities of an establishment of a controller or a processor in the Union should be carried out in accordance with this Regulation, regardless of whether the processing itself takes place within the Union. Establishment implies the effective and real exercise of activity through stable arrangements. The legal form of such arrangements, whether through a branch or a subsidiary with a legal personality, is not the determining factor in that respect." (EUR-Lex 2016.)

In addition the Article 13 Paragraph 2(f) might affect the use of AI:

"Where personal data relating to a data subject are collected from the data subject, the controller shall, at the time when personal data are obtained, provide the data subject with all of the following information: .....

the existence of automated decision-making, including profiling, referred to in Article 22(1) and (4) and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject." (EUR-Lex 2016.)

These extracts therefore show implications that GDPR will eventually restrict automation regarding decision making about individuals. The law will at least effectively create a "right to explanation", as people have now the right to ask explanations on algorithmic based decisions that were made about them (Hume 2018; Rossi 2016). This "right to explanation" is described above in the Article 13 paragraph 2(f) (EUR-Lex 2016). Additionally as stated in Recital 71 decisions related to individuals that are based on fully automated decision making and which might have some legal impacts, are not allowed anymore (EUR-Lex 2016). Thus these extracts already imply that that there might be some serious issues related to using AI in decision making, especially when it relates to individuals. However in addition to these notions about GDPR, the actual impact on AI applications will surely be revealed later when some precedent cases have been dealt with, and thus the boundaries and interpretations of the newly established law have clarified.

As AI starts to automate more and more decisions it's also important to understand how AI actually "thinks" so we are able trust it (Castelvecchi 2016; Rossi 2016; Snow 2017). After all, the credibility of the results that AI systems provide are only as good as the data they're fed on. And it's not only about the actual data that AI is trained with but all too often the problem is that humans who build the algorithms behinds those AIs might end up feeding them their own humanlike biases into the algorithms. (Rossi 2016; Snow 2017). Thus bias can come from polarized data set, where the training doesn't actually represent the whole truth, and so the model might behave in an unexpected way (Brynjolfsson & Mitchell 2017). Then the other way biases might creep into the systems is the above mentioned situation, where biases of the developers who built the algorithm end up in the model. In these situation the person's human biases and their cultural assumptions are transmitted to the choices they make when modelling the algorithms. (Campolo et al. 2017.) Problem here is the fact that often identifying own biases might be extremely difficult, as they are built in to people and so understanding one's own biases isn't always so simple (Soll et al. 2015). These biases might become a huge problem especially in the situations when something goes wrong. Who can be blamed when a machine makes the decision – is it the customer using the system, one who provides the system or the one who designed the algorithm in the first place that the AI system is based on. Thus biases are one of the biggest and also the most troublesome challenges with AI.

In addition to the above mentioned challenges one big challenge also relates to the language. Language itself is a very unique human capability. With AI, especially with Speech Recognition, NLP and NLU, humans are also providing machines these language capabilities (Socher 2018; Srinivasan 2016). These can possibly enable very interesting and highly useful applications but there are also challenges that might occur. These relate to the already mentioned issues in sub-chapter 3.2.2 where NLU capabilities were described. Essentially language is so complex and multidimensional that even the most sophisticated NLU based application might get the meanings of words and sentences wrong (Knight 2016; Socher 2018; Srinivasan 2016). Challenge here doesn't get any easier when more languages and regional accents are added to the mix (Cambria & White 2014). The issue becomes even more apparent with understanding the context of text or spoken language. This is highly critical as it isn't even enough that the AI system understands all the possible variations in the meaning of the words, if it still gets the context wrong. Sentences can often be said in so many ways and same sentence might have a very different meaning depending on the context it's used in. (Knight 2016; Socher 2018.) This might then bring about issues in different contexts where NLUs and NLPs are being used, be it handling emails, documents or conversations with a chatbot.

As a summary, AI technologies despite their huge potential have certain challenges that can seriously affect their usage. These challenges are concluded below on Table 4.

Table 4. *Key challenges and limitation related to using AI.*

Challenge	Description
Requires huge amounts of data	Huge amounts of training data is needed to teach the model all the necessary features within the data
Reasoning behind the models	Many AI applications' method of reaching the decision aren't transparent because they don't operate
Regulations affecting the use cases	Regulations might seriously affect the use cases of AI solutions; like GDPR, which restricts automated decision making related to individuals inside EU
Understanding language	Understanding context and meaning of the words and sentences isn't easy for AI solutions

These challenges of course doesn't relate to every AI application but depend really on the use case. However all these challenges are something that organizations should definitely take into account and assess the potential implications when looking to implement AI solutions into their operations.

4. RESEARCH METHODOLOGY

This chapter introduces the methodologies that were used in this research. Aim of this chapter is therefore to give understanding on the methodological choices and to the philosophical approach that was used in this thesis. First sub-chapter focuses on the used methodology. Idea of this sub-chapter is therefore to understand the reasoning why certain methods were chosen to be used and why. Second sub-chapter then focuses on the material collection related to this thesis. This sub-chapter explains how the materials were collected for the literature review and how the interviews were designed and conducted in the empirical part. Lastly, the third sub-chapter explains the overall research process and how it progressed from start to finish.

4.1 Methodology

Scientific research and the philosophy behind it is steered by multiple factors, and contains assumptions about the way researcher views the world (Saunders et al. 2009, p. 108). Thus when conducting a research it's important that the researcher considers how the research can be scientifically well formed. One way to look at conducting a research is to use the research 'onion' introduced by Saunders et al. (2009, p. 108). This research onion is used to present and reason many important choices that define a scientific research, which are:

- Philosophies,
- Approach,
- Strategy,
- Choices,
- Time horizon,
- Techniques and procedures.

These factors are presented in the research onion as different layers and they represent different scientific issues that the researcher must consider when conducting a research. The idea in the onion is that researcher starts in the top layer and peels the onion layer by layer going into more detailed choices regarding the research. (Saunders et al. 2009, p. 106-108). Thus to begin with the researcher must first consider different scientific philosophies and the way in how to approach the research. After these have been defined researcher can move on to scientific choices, time horizon and techniques and procedures that are used in the research. Below in Figure 9 are presented the choices that define this research.

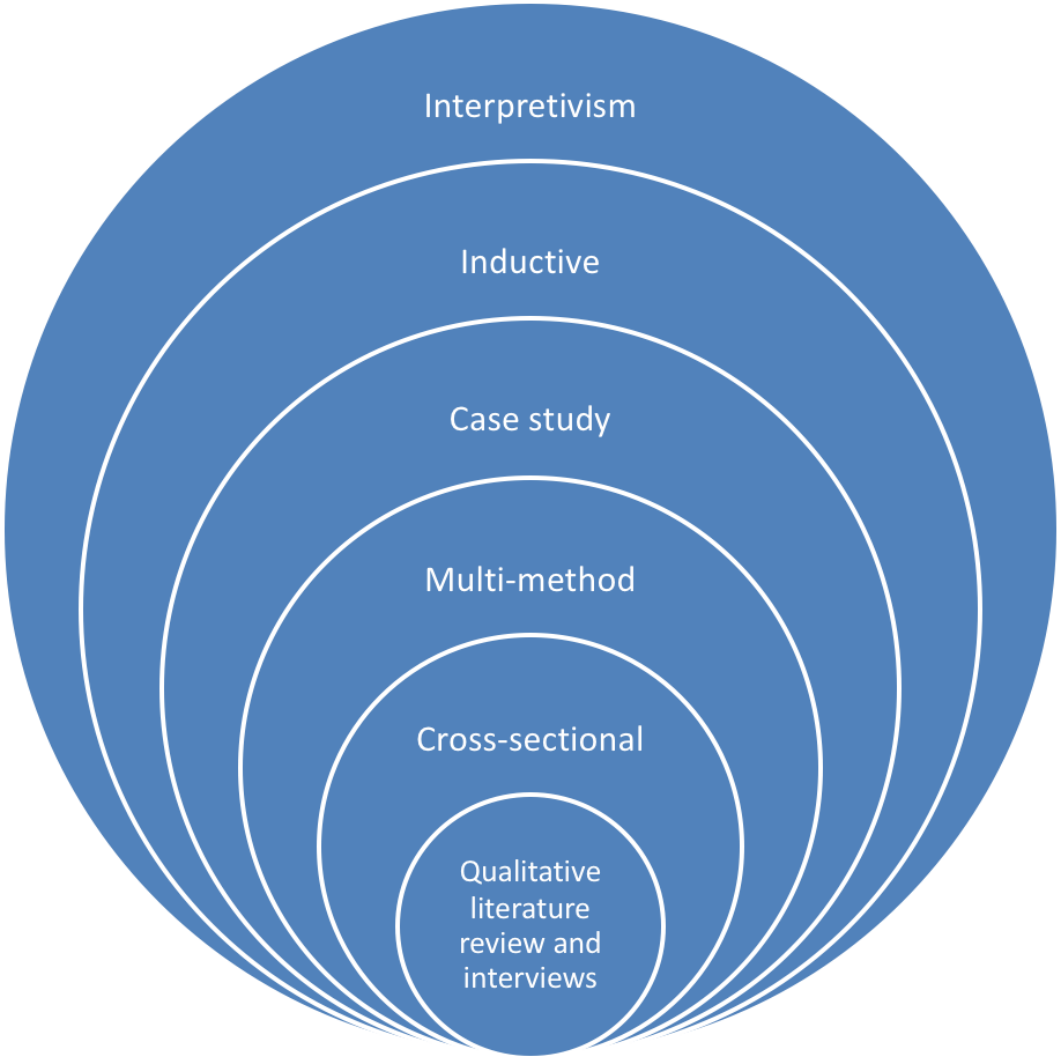

Figure 9. *Research choices based on Saunders et al. (2009) research onion.*

First when starting to conduct a research, the research philosophy should be considered. This is important because the chosen research philosophy shows how the researchers views the world and approaches knowledge in the research. Important notion is that the chosen research philosophy also affects the research strategy and methods which are chosen as part of that strategy. Strategies and methods differ from one another and so will effect on what is viewed important and useful in the research. Thus the chosen research philosophy affects inevitably also to the outcome of the research. Research philosophies then are of four kinds: positivism, realism, pragmatism and interpretivism. Positivism here means that the research is conducted from a very objective viewpoint which can results in almost generalized theories. In positivism focus is on the observable matters that can provide valid facts, without interpretation. Thus the aim with positivism is often to explain phenomena. Second research philosophy, realism, then views the world in a way that the reality is independent of the mind – meaning that observable phenomena is

enough to provide facts. Pragmatism in turn takes a little bit different view. It argues that choosing the right research question is the most important factor for the research, and that using multiple methods, often both qualitative and quantitative techniques, will best enable answering the research questions. Lastly there is interpretivism philosophy to the research. This approach argues that the world of business is more complex than physical science world, and thus what is needed in the research is to take into account more variables, like people and organizations. (Saunders et al. 2009, p. 108-119.) This research takes an interpretivism philosophy as a viewpoint of the world. This is because the research topic relates very closely to organizations. RPA and AI surely have a technical side, which is also very important, but when we are considering the possibilities of combining these two, their perceived value comes from the acceptance of users. Thus this research takes subjective view on the possible benefits and challenges that relate to using AI technologies with RPA.

When starting a research another important factor is how researcher approaches knowledge about the research topic. For this there are two kind of research approaches: deductive and inductive. If the topic is not familiar for the researcher then he or she usually turn to the inductive approach, where data is first collected and the theory is developed as a result of the data analysis. The other approach to conducting research is therefore deductive, where theory and hypothesis are designed in the beginning of the research and this theory is then tested with different hypothesis. (Saunders et al. 2009, s.124; Woo et al. 2017.) This way deductive approach is especially suitable for researches where researcher has already, before the research, obtained understanding on the topic and he or she is familiar with the subject. Inductive approach in turn represents the opposite approach, where theory is designed based on research. This thesis have strong features of inductive approach and thus is the chosen approach for the thesis. There isn't currently almost any theory related to combining RPA with AI technologies, thus this thesis aims to fill this gap by using qualitative data to create new understanding within the topic. This is one of the main purposes of inductive research, whereas deductive research is more about collecting quantitative data and finding relationship within the variables (Saunders et al. 2009, p. 124-127).

After the first two layers on the research onion have been pealed, and so philosophy and approach are decided, the research purpose and strategy can be decided. As one of the main purposes of this study is to gain new insight on how RPA and AI can work together, the purpose of this thesis is exploratory. Exploratory study in essence is about seeking to find new understanding and seeing things in new light. Exploratory studies are especially useful in cases where the problem itself might need clarifying. The three main principals in conducting exploratory study thus are search of literature, interviewing subject experts and conducting focus group interviews. (Saunders et al. 2009, p. 139-140; Woo et al.

2017). Two out of three of these methods, literature review and interviewing subject experts, are used in this thesis. This combined with the aim, finding new insight, shows that purpose of this thesis is clearly exploratory.

Once the purpose of the research has been established researched strategy can be formed. Different kind of research strategies are: experiment, survey, case study, action research grounded theory, ethnography and archival research. Each of these strategies can be used no matter was the chosen purpose for the study exploratory, descriptive or explanatory. (Yin 2003.) This thesis has case study chosen as its strategy. Case study as a strategy means conducting a research which involves empirical investigation of a single or multiple phenomena using multiple sources of evidence, where the context is also present. (Robson 2002, p. 178). What is however relevant is that the phenomena should form some kind of an entity. In this sense case studies can be used in very many ways. Case can be a company, single problem or practically any defined entity or entities. Case study as a strategy enables answering to question like “What?”, “How?” and “Why?”, thus it also suits well for exploratory research (Saunders et al. 2009, p. 146; Yin 2003). In this thesis all research questions also relate to “What?” and “How?”, and in this sense the research strategy of this thesis can be categorized as a case study also from this viewpoint (Yin 2003).

Next in the research onion is choosing the data collection method and time horizon for the research. Data collection methods can be divided into two main groups: mono method and multiple methods. In this thesis the data collection methods were literature review as well as theme interviews. As there were two different data gathering methods, which both are for gathering qualitative data, this thesis is a multi-method research. A multi-method research in essence means that multiple data gathering methods are used but they are both either quantitative or qualitative by nature. (Saunders et al. 2009, p. 151-152.) Time horizon in turn in this thesis is cross-sectional as this is a study conducted of a particular case at a particular time, as the time constraints itself already set the boundaries. (Saunders et al. 2009, p. 155-156.)

4.2 Collection of material

Material collection for this thesis consisted of two parts: materials for the literature review and materials for the empirical research. In the literature review the materials that we used consisted of different kind of academic articles, books, academic journals and other publications. The material collection process for the literature review is explained in the sub-chapter 4.2.1 and following this in the sub-chapter 4.2.2 is explained the material collection for the empirical part which consisted of interviews.

4.2.1 Literature review

Literature review exists for two main purposes. First purpose is to help the researcher to produce better research ideas, and thus ultimately better research questions, about the research topic. The second purpose is to provide researcher a more holistic and wider view about the topic in a big picture. (Sharp et al. 2002.) Literature review also helps to understand what other researches have been conducted related to the topic (Saunders et al. 2009, p. 60) and thus literature review helps to find gaps in the research field. This is also the reason why literature review was conducted for this thesis.

Theoretical part of this thesis that was conducted as literature review consists of two chapters: robotic process automation and artificial intelligence capabilities. This literature review was conducted to gain better understanding on the topics of the thesis. Literature research for these started with using Scopus, which is a database for peer-reviewed literature, to find appropriate articles, books and conference papers for the theoretical part. Other databases like IEEE Xplore, ProQuest and Springer Link were also used to help find appropriate material. Here peer reviewed articles, journals as well conference papers were favored to sustain high scientific quality of the literature, and so results of the research could also be considered more relevant. Material selection for the literature review first started by using specific search terms to limit and steer the search results. Search terms related to the research were:

- “robotic process automation” OR “RPA”
- “Artificial Intelligence” AND “concept”
- “Artificial Intelligence” AND “methods”
- “Artificial Intelligence” AND “technologies”
- “Artificial Intelligence” AND “capabilities”
- “Intelligent Automation”
- “cognitive intelligence”
- “cognitive automation”
- “process automation”

After collecting appropriate material final selection of the literature for the thesis was done by reviewing the abstracts to check how well the literature piece matched the wanted information. The publication year was also used as a selection criteria to find literature pieces that were as new as possible so that they would be relevant for the research. Similarly publishers and citation counts were used as criteria to select literature to guarantee relevance. After good pieces of literature were collected and read, also the citations of these were used to find other possible relevant literature. Same time new search terms were generated based on the new insight and ideas that came up when reading the already collected literature. After new search terms were generated the literature search process

started again. Thus, the literature search part of this thesis was an ongoing iterative process throughout the writing of the literature review part which enabled gaining more and more insight on the topics.

4.2.2 Empirical part

This research method was chosen based on the nature of this thesis. The purpose of this thesis is exploratory and so it combines relatively new literature areas with interviews. The interviews themselves were chosen to be semi-structured theme interviews as they open possibilities to gain deeper insight from the interviews; as semi-structured interviews may consist list of key themes and some questions related to these but the questions may change, depending on the flow of the interview, and additional questions can be asked (Saunders et al. 2009, p. 320). Thus semi-structured interviews doesn't tie the researcher to the question frame but enable adhoc modifications depending on how the interviews proceed. This was the main reason why semi-structured interviews were chosen here.

Therefore the empirical part of this thesis consisted of semi-structured interviews. Interviewees for the empirical part were chosen to be experts in the subject areas of this thesis to get deeper understanding on the subjects as well as new insight. In Table 5 on the next page are presented the interviewees; who they are, from what company, what is their title, area of expertise and what were the central themes that were covered in their interview.

Table 5. *Interviewees' name, title, area of expertise, organization and central themes in the interview.*

Name	Title	Area of expertise (among others)	Organization	Central themes
Tuomo Pursiainen	Senior Consultant, Intelligent Automation	Process automation, intelligent automation and robotic process automation	CGI	Combining RPA with AI solutions; benefits, challenges and use cases
Samu Paajanen	Director, Data Science	Advanced analytics and artificial intelligence technologies	CGI	Use of artificial intelligence technologies, benefits, challenges, and use cases

There were two expert interview that were held for this thesis. They both concerned different topics and had different interview questions. These interview questions and the key themes can be found on Appendixes, as shown later. To shortly summarize, first interview was held with Tuomo Pursiainen who is a Senior Consultant at CGI focusing on process automation, robotic process automation and intelligent automation. This interview focused on intelligent automation; what key challenges of robotic process automation it could address, what kind of challenges there might be with it and how organizations should get started with intelligent automation. The key themes and interview question for this interview can be found on Appendix A.

Second interview was held with Samu Paajanen who is a Director at CGI focusing on advanced analytics and artificial intelligence technologies. The interview with Samu Paajanen then focused more on the use of artificial intelligence technologies; what different technologies there are, what benefits and challenges relate to these and what kind of use cases currently exist. The key themes and interview question for this interview can be found on Appendix B. Both of these interviews took approximately 45 minutes.

4.3 Research process

The research process of this thesis started with getting familiar with the topic related to this thesis. After this started the literature search and the found articles were read. When the topics of the thesis came more familiar and research needs within the topics were

identified the research questions were created. Following this the research strategies and methods were chosen after which more specific literature research was done. After the literature research was ready the literature review for this thesis was written, and based on the findings the process for finding suitable candidates for the interviews started.

After the proper candidates for interviews were found the interview questions as well as frames for the interviews were designed and the dates for the interviews were set. After the interviews were held the interview recordings were transcribed. After this the findings from the interviews were concluded. Lastly the interview results were combined and reflected towards the literature review and the conclusions of the whole thesis was written.

5. RESULTS

This chapter contains the results of the empirical research conducted in this thesis. The chapter is divided into two sub-chapters. First one includes the interview results of the intelligent automation expert interview. This sub-chapter provides deeper understanding to the topics of intelligent automation. Interviews give deeper understanding on the subjects of artificial intelligence and intelligent automation. In addition the interviews aim to shed light on the possible challenges that might relate to using these technologies.

These interviews aim to give additional insight and deeper understanding on the following sub-research questions: "*How robotic process automation can be used to automate processes and what benefits as well as limitations does it have??", "What is artificial intelligence?", "What are different artificial intelligence capabilities and how they can be used?", "What kind of risks are related to using artificial intelligence?"*. These results of the interviews are excluded of the researcher's opinions.

5.1 Intelligent automation expert interview

This interview was held with intelligent automation expert, Tuomo Pursiainen who is Senior Consultant at CGI focusing on process automation, robotic process automation and intelligent automation. This is why the key themes of this interview really focused on intelligent automation. Key themes and finding of this interview can be found on the table 6. First key theme was about "Intelligent automation potential" to better understand what problems with RPA could AI solutions address, and to get insight on what kind applications there already are within the concept of intelligent automation. Second key theme "Challenges related to intelligent automation" then focused on the challenges that AI technologies could bring about when combined with RPA, and how regulation and especially EU's new law GDPR could affect the use cases. Last key theme "How to get started" focused on understanding how organizations should start their intelligent automation journey. These key themes are answering to the following sub research questions: "*How robotic process automation can be used to automate processes and what benefits as well as limitations does it have?"* and *"What is artificial intelligence?"*.

Table 6. *Key themes and findings from the interview with intelligent automation expert.*

Key themes	Intelligent automation potential	challenges related to intelligent automation	How to get started
Key Findings	- AI could be used to address two main issues with RPA; handling unstructured data and coping with complex decisions - Most mature technologies are currently related to NLP and chatbots - In addition to conversational agents chatbots could be used to; handle input data, handle decision points - Machine learning together with RPA could enable interesting applications where predictions could be used as triggers for robot	- There is not enough and of good quality data related to processes to create models - Biases; related to data or to the models - GDPR and its effect on processes that handle data about individuals - Transparency to the decisions and validation of the results	- Look for processes that were previously assessed for RPA but found too difficult - combine chatbot in front of RPA - combine with analytics, e.g. AI to predict IOT data and RPA as a trigger based on the prediction

Intelligent automation potential

The interview was started with a question for interviewee to explain what is artificial intelligence and intelligent automation in order to gain insight on how the interviewee sees these concepts, and so provide understanding on what is their point of view to the topic. Pursiainen started to define artificial intelligence with describing how it can mean very many thing; some are very science fictional, then there is general AI which Pursiainen described as system that is usually able to reason and understand things, and learn faster with ultimate speed. Right after he pointed out his view on general AI stating that the current capabilities aren't yet there and that AI we now have is more of a supporting AI which consists of different narrow AI areas like: natural language processing, natural language generation, image recognition, video analytics, text analytics, machine learning and so on. From here he pointed that intelligent AI means exactly these narrow AI capabilities. To sum up, Pursiainen described AI as set of capabilities that deal with vast amounts of data and the problems of physical world and unstructured data. General AI in

his opinion on the other hand would be a system that could handle with all of these as an input and then give an answer to a variety of different problems.

When asked additional question about the current state of AI. Pursiainen stated that in his opinion most advanced technologies are currently related to natural language processing; from which there are examples from companies like Google, Microsoft and IBM which all have advanced AI solutions that can interpret the context on the level that these system can understand the topic and understand what the humans are intending to say or do. Other developed capabilities have been seen in the field of image processing, which can very well identify certain objects in pictures. Additionally Pursiainen pointed out that chatbots and natural language generation have progressed very much, so that they can understand also different persons and even accents.

When it comes to RPA's challenges and how AI technologies could address these. Pursiainen sees two major challenges. These are handling of unstructured data, and RPA's need for clear rules. Firstly Pursiainen sees unstructured data as a showstopper for automation when the robot should handle input data that isn't in a structured format, like emails from customers – robots just aren't good at handling this kind of unstructured data as input data for the process. It is the robot's lack of being able to understand intent. Robots might be able to extract some simple keywords and understand basic patterns, but to program the robot to understand intent is somewhat yet impossible, and so also RPA's use cases are limited in this sense. How to approach this issue Pursiainen suggests that either texts are analyzed with text analytics or machine learning is used to get the intent out of the input data and so the actual task for the robot. Moreover, Pursiainen suggests that even better solution would be to add chatbot at the beginning of the process to do all the data gathering and discussion with the end user. After this chatbot would translate the unstructured data into structured format and give this translated 'package' for the robot which then would be able to handle and process the structured data from the point forward. This challenge with unstructured input data represents the first major key issue with RPA that AI capabilities could address. The second key challenge that Pursiainen points out is then related to the complex decision rules. In simple use cases robots can make decisions when there are step-by-step logical rules to follow. However, some use cases are so complicated and there are so many factors the robot would need to take into account, that it's not practical or even possible to implement robot to these processes. How Pursiainen then sees that this issue could be solved, is to gather huge amounts of data from decision points where humans are still needed. This would enable getting all the information on those decisions and when enough data is gathered the rules could be derived from the decisions made by human, and then with new use cases machine learning models could be used to make decisions based on the historical data. In this sense the AI models would then give predictions in these previously human held decision points, and RPA could then seamlessly work with other logical parts of the processes.

When talked about actual use cases in combining RPA with AI, Pursiainen first raised chatbots as one of the most common examples, as they are used in processes to guide the user on the process or on a specific task. Other examples that Pursiainen raised were the use of machine learning models together with IOT solutions to predict that something is going to happen after which the RPA is started based on this input trigger. In this sense RPA isn't any more started in traditional way, that is manually or scheduled, but the process is started based on the prediction that something is going to happen. Another common use case mentioned was the use of machine learning model to sort and classify emails and tickets coming to systems. In some use cases these classified and categorized tickets and emails were then automatically handled by RPA, if the confidence on decision making was adequate.

When asked about machine learning and should we, and can we, trust the decision making by it Pursiainen clearly stated that there wasn't a definitive answer to this question but that it really depends on the use case. In mission critical cases like health diagnostics or decisions to hire someone, Pursiainen sees that the AI systems aren't yet confident enough, and even if they were he wouldn't trust them 100%. Therefore Pursiainen suggested that machine learning and AI could rather be used to gather all the data and give the prediction based on it but he would still keep the final decision making on humans' shoulders. Then in some uses which aren't so mission critical, Pursiainen suggested that AI decision could definitely be implemented on full scale. These kind of use cases would thus be for example ticket categorization, basic requests like changing passwords, and other processes where even if the AI makes mistakes there wouldn't be such critical and far-reaching consequences.

Pursiainen also sees that the technical side, for implementing AI solutions with RPA in business processes, is getting easier and easier all the time. He states that most RPA providers offer good services and APIs which enable connection with major AI platforms from major players in the market, as they provide solutions to be deployed through APIs. This way they could be used like building block to create services inside RPA processes. Pursiainen also notes that some RPA providers like UiPath have already started to provide AI capabilities within their own RPA solutions which makes the first step of starting with AI even easier. Thus Pursiainen sees the business cases as more of an issue when starting AI implementations with RPA. The processes need to have good quality data and there needs to be a lot of it, so that the built AI models can be accurate enough.

However requirement for data quality and quantity are difficult to define, as it's quite impossible to predict the needs for building the model beforehand and the requirements also differ case by case. In some cases just tens of thousands of items might be enough to train the model but usually much more is needed, as there needs to be data for both building the model and then for testing it. Regarding quality Pursiainen pointed out that the importance of good quality becomes more important when there is only small amount of data to train with, as this might affect the model. However in cases where there are huge

amounts of data to train the model with, then also little amount of bad quality data is acceptable.

Challenges related to intelligent automation

Pursiainen stated that the biggest issues of implementing AI technologies with RPA relate to data quality and missing good data needed for model creation. Another mentioned challenge was Finnish language, which come to picture when chatbots, machine learning, natural language processing and natural language generation are used. Capabilities regarding understanding Finnish language are progressing but they lack significantly behind of AI capability to understand form example English. Another great challenge that Pursiainen brought up is AI bias. This means that the training data can be biased by some feature in the data set. Then when the model is trained with biased data the model inherits those same biases. This in turn might cause a lot of extra work if the data set has to be cleaned of the biased data. Identifying the biases can however be rather difficult especially when talking about human bias. Then can just be hoped that in the model something correlates in an unexpected way so that the data affecting this can be identified. Thus Pursiainen highlights that the models are really as good as the data they are fed on.

When talking about the new GDPR, Pursiainen raised two effects on how the regulation will likely affect. First is the automated decision making restriction, which will in Pursiainen's view effect AI's use in cases where data related to people are exploited. This is because in many cases AI's decision making is somewhat a 'black box' and there is no transparency on how the model got to the decision. Of course this can be sometimes hard even for people to explain the reasoning behind the decisions but for these black boxes using neural networks it's almost impossible. Then the second part that Pursiainen raised as an effect how GDPR will affect relates to retaining the needed data for training the machine learning model, as the due to GDPR there are new limitations on the retention of data. Additionally all the regulation rules related to data anonymization and removing of identifiable information are important, but by Pursiainen's view also problematic. Even if the model can be trained without using any of this data there could surface issues related to the datasets that need to be kept. This is because when the data is cleaned from the identifiable data, also other important factors of the data could be lost which then directly affect on the results that the trained models give based on the cleaned input data. Executing those processes where AI models are somewhat black boxes can then be very problematic.

Then when talking about the black box issue as well as the possible validation issues with AI decision making Pursiainen states that there might be ways to go around the validation. One approach here to validate the results of a system, that isn't transparent, could be to test the model after removing the labels from the data and to then test the model again against earlier results, if they are available, and see if the model behaves the same way and gives expected results. However black box situation might not even be an issue in all

occasions. For example if the AI is only used to give the prediction and then humans are doing the final decision, as humans can then check the predictions for irregularities.

How to get started

Pursiainen suggests that for starting assessing the intelligent automation possibilities organizations should have a look on the processes that were previously already assessed for RPA but were too complicated; for too complex decision making or rules. These could be the first points to find possible use case, of course requiring that some kind assessment have already been conducted. As a second option Pursiainen suggests that business intelligence and analytics department should be taken under consideration for intelligent automation implementations. There automated reports could be created based on the predictions of AI models and the RPA would be the operator that make things happen. Use cases of these kind of application come especially form predictive maintenance where AI models analyze the IOT data coming from plants, this predicts something, and then automated processes are used to react as wanted when predictions trigger them e.g. to shut down the plant.

5.2 Data Science expert interview

This interview was held with data science and artificial intelligence expert, Samu Paajanen who is Director at CGI focusing on advanced analytics and artificial intelligence technologies. Key themes in this interview focused around artificial intelligence. Key themes and finding of this interview can be found on the next page on Table 7. First key theme was about "Artificial Intelligence" to better understand what artificial intelligence actually means, what it consists of and what are the most mature capabilities currently. Second key theme "Data" then focused on the use of data with AI. Third key theme "Challenges" focused on understanding what kind of challenges there might arise when AI solutions are used as part of business operations. Last key theme "Prerequisites and requirements" focused on how organizations should prepare to using AI solutions and how they should start the journey. These key themes are answering to the following sub research questions: *"What is artificial intelligence?", "What are different artificial intelligence capabilities and how they can be used?", "What kind of risks are related to using artificial intelligence?".*

Table 7. *Key themes and finding from the interview with data science expert.*

Key themes	Artificial intelligence	Data	Challenges	Prerequisites and requirement
Key Findings	- In essence AI is all about very advanced analytics and automation - AI works with probabilities which might limit the use cases, but if some level of uncertainty can be accepted use cases of AI are manifold	- AI requires lots of data, but data doesn't necessarily have to be perfect - More features there are in the data, the more data is needed to train the model	- Data quality and biases - "Black box" issue and transparency to the decision making - How to trust the decisions made by AI; it can be nearly as accurate as humans but there is also a psychological problem to trusting AI	- adequate amount of data; which often isn't anymore a problem especially with mid-sized and big companies - people in the organization should have understanding on the possibilities

Artificial intelligence

Especially the subject of artificial intelligence is quite challenging as the concept is so wide and there isn't yet even a common understanding on what it actually consists of. Thus the interview with Samu Paajanen was started with asking him to describe what he understands as artificial intelligence and what it consists of. Paajanen started by explaining that artificial intelligence is very hyped concept nowadays and people have different understandings about it; for some it's just statistical math but in Paajanen's opinion it's about analytics, very advanced analytics combined with automation. Paajanen also stated that artificial intelligence is not about one technology, it's more like set of technologies that enable artificial intelligence. One of the main types is machine learning which enables the kind of advanced analytics that is not possible to implement in other ways. However machine learning is just one enabler and there are variety of different methods that can be used in advanced analytics. Essentially these all however have the same requirements for successful implementation. First is that they need lots of data, something to process. Secondly they need to interpret the input data they're fed. Like in a case of a self-driving car; they use lots of data to gather information about the surroundings, so plenty of data is needed. Then analytics engine processes this information from the surroundings; be it visual information, what other cars look like, what humans look like and how

to identify where the road goes. Then the analytics engine driven by machine learning combines information from these and makes analysis of the situation and the surroundings. Lastly driving the car is handled by basic automation, based on the analysis and predictions given by the analytics engine. So in the end even a self-driving car is nothing but advanced analytics and automation.

When talking about the most mature AI technologies Paajanen pointed out that the most mature ones are probably those that use unsupervised learning, like Googles DeepMind. These work in a way that they can be given a task but there is no need to define right answers for the system in advance. Thus the systems learns by itself.

In the context of business processes Paajanen pointed out that he doesn't think that there are really areas where AI couldn't be helpful. Often the tasks are however about automating simple things where AI isn't even needed. When AI comes into the picture is in those tasks that require human like decision making or human like conclusions. AI implemented to the process could for example make decisions to approve insurance claim or not. With AI could then be said with a certain probability, let's say 80% probability that the insurance claim is valid. AI system will never however get to definitive results but in the same way rarely do humans. Human beings can say that the something looks valid and that it's very likely to be correct, but who can say that for sure. In this sense if certain amount of uncertainty is accepted in the decision making, then AI solutions can be applied to almost all business processes. Thus AI really plays with probabilities and even the most sophisticated AI systems come down to just that. This is also what really differentiates AI and advanced analytics from traditional programming. Traditional programming is all about zeros and ones – the solution is always true or false. There is no grey area. With AI the approach is quite the opposite, decision are almost always in the grey area and very rarely anything can be said with a 100% certainty.

Data

Data is a very important aspect in the context of AI. AI systems tend to require lots and lots of data. Paajanen although points out that there are differences depending on the use case. It all comes down to the use case and what kind of features there are within the data. The more feature rich the data is, the more data model requires. In the end the results are only as good as the data is, if data is of bad quality, so will be the results. Underlining this Paajanen however states that the data doesn't have to be perfect, because AI is essentially about probabilities. Thus the model doesn't need to give definite results, even 80% correct is often a good result.

How to be sure that the data is of good quality and that there is enough of it can however be troublesome with AI systems. Often you can't be sure of it in advance, and what is then needed is just testing and checking the data sources for anomalies. Problems might also occur when different data sources are used, as then the results driven from the models

might not be comparable between data sources. However usually some surprises emerge and there's no definite way to assess these things about data in advance.

Another problem then with AI is the possible validation and trust issues regarding the results AI system produces. Paajanen, pointed out that again the issues relate muchly to the data. There might not be enough data or the quality isn't good. Another view Paajanen brought up was possible issues with the GDPR. Due to the GDPR data that contains person specific information needs to be stripped out of this identifying information. This however might lead to a situation where useful data is lost in the process. Although, Paajanen points out the fact that this was often a requirement even before the GDPR. Now people just need to be more careful about it.

Challenges

There are of course other risks and challenges, than just amount of data and data quality. One challenging thing here is possible biasness of data. Problem here is essentially humans. Humans are biased creatures and so if you teach an AI system with data that's generated by humans the system inherits their biases. This might be a true risk for example if AI solution is used to make decisions related to who should be hired or not. Thus the AI system isn't objective anymore but it's biased by the views of the people who generated the data the system is taught with. This could in worst cases be a risk for companies and even to whole societies, as Paajanen stated.

The trust with AI is also a big question mark nowadays. The systems are still not mature and developed enough that all decision could be assigned for AI systems to make, as we don't have enough experience. Although Paajanen highlighted that even when the AI systems could possibly do all the decisions he wouldn't trust 100% on fully automated AI decision making. As Paajanen sees this, it's always good to have a human occasionally checking that the system performs as supposed to and that results are meaningful. Trusting the AI systems is however also a bit of a philosophical question because different things are expected from AI than from humans. For example if an autonomous vehicle crashes it always ends up in the news but if humans crash their cars it doesn't necessarily gather any attention. So even if on average AI system drives better than humans, it's still not accepted for AI to make any mistakes.

Paajanen also brought his opinion on the effect of GDPR to AI. He sees that it will definitely limit the scope where AI applications can be used when it comes to processes concerning person specific data. In the big picture he however sees that GDPR is more of a good thing as the privacy concerns are completely valid. In addition AI solutions should always be built in a way that they are good for companies, good for people and don't risk any confidential or private information leaks.

Prerequisites and requirements

There are many prerequisites for using AI solutions and to start a successful AI journey in an organization. First and foremost Paajanen highlights again the importance of data. First and foremost organizations must have adequate amounts of data at their disposal. This isn't usually a problem as most mid-sized and big companies have lots of data that they are however not utilizing. In addition there must be people inside the organization who are willing to invest their time to work together with AI experts. In addition to these an important thing is to have understanding of AI capabilities within the company, so that the potential is actually understood.

6. DISCUSSION AND CONCLUSIONS

This last chapter of this thesis combines the findings from the literature review together with the results of the interviews. First subchapter focuses on interpreting the findings from both theoretical and empirical part, and putting the findings together. This subchapter is divided into two smaller subchapters where the topics are discussed with research questions in mind. Then second subchapter is a brief summary of the findings related to the research questions. Third subchapter focuses on critical evaluation of the whole research process, and finally, in the last subchapter is discussed the possible future research needs in the surroundings of this topic.

6.1 Discussion

This chapter consists of two subchapters that are "Artificial intelligence in business processes" and "Intelligent Automation". In order to properly answer to the main research question: "*How artificial intelligence capabilities can increase the automation level of business processes when combined with robotic process automation?*" deeper understanding to the themes of the two subchapters are needed. First subchapter unravels the concept of artificial intelligence and focuses on different artificial intelligence capabilities. Artificial intelligence is wide and multifaceted concept, therefore the focus is kept on understanding the different artificial intelligence capabilities and their possibilities as well as challenges in the context of business processes.

Second subchapter focuses on context of intelligent automation. Here RPA's challenges are discussed and different artificial intelligence capabilities potential to complement these challenges are argued in more detail. This chapter includes findings that the researcher found most interesting based on the conducted literature review and expert interviews.

6.1.1 Artificial intelligence in business processes

This subchapter is answering mostly to the following sub-research questions: *"What is artificial intelligence?", "What are different artificial intelligence capabilities and how they can be used?", "What kind of risks are related to using artificial intelligence?".* First finding in this theme was related to the concept of AI. It's a wide and in many ways even hard concept to grasp. This is because the field of AI is constantly changing as technologies, capabilities and algorithms evolve and new ones arise. Another problem in the concept itself is in the definition as it consists word 'intelligence'. Thus defining AI would mean a need to define what intelligence and being intelligent means. This philosophical

question however isn't focus of this thesis so the concept of artificial is examined through AI's different capabilities.

As a first key finding, literature review showed that AI as a concept can be broken down by the objectives it's used for. These objectives relate to either capturing information or to understanding what is happening. Then within these objectives AI can be divided into different methods that are used. This is showed in Figure 10 below.

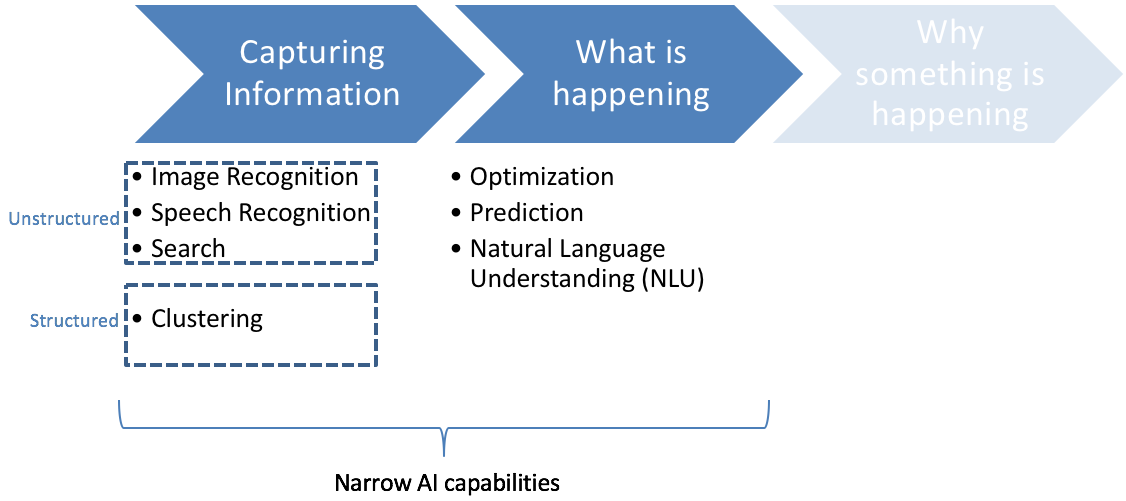

Figure 10. *Objectives and capabilities of artificial intelligence.*

AI's objective capturing information is all about extracting information from the context of certain problems. These problems can then be about handling either unstructured or structured information. After the information has been extracted and converted to a more understandable format then capabilities within what is happening start to work out what the data actually means. Important factor to understand is also that AI capabilities are rarely used as single units, but that most AI solutions combine different AI capabilities to being able to do both task; capture information and then make sense of it. Good examples of these kind of solutions are for example chatbot's, which were discussed in the interview with Pursiainen and also briefly introduced in the literature review. Chatbots usually combine Search capability, or often in this context called as NLP, with NLU. There NLP first processes the unstructured text, captures information, to find pattern and turns the text into data. Then finally NLU tries to define the intent of all this data. Then on the other hand Search capability could be used for example to handle semi-structured invoices, as explained in chapter 3.2.1. Thus the capabilities can be used in very different use cases, and when capabilities are combined together possible number of solutions are even greater.

Another point worth noting was that all these capabilities represent narrow AI. This in essence means that they are meant for only certain specific tasks. In contrast to narrow AI another point of view to AI is the general AI, or as sometimes called general artificial intelligence. This however isn't yet possible with the current capabilities as general AI

would require a system that could handle any types of input and then process the task, whatever it might be. This is thus still much of a "science fiction" version of AI. General AI would also be the one that answers to the AI's third objective "Why something is happening", which simply isn't possible with current capabilities as they don't actually understand what the data represents, but merely correlate the fed data, whatever it might be, against the model. Current capabilities therefore are only meant for specific purposes, be it; capturing information about images, speech, text, data or about understanding the data by extracting the intent out of it. To the same kind of conclusions the interviewees also ended up, when asked about AI. They stressed that current AI solutions are very much meant for certain purposes and specific tasks. The fact that the current AI solutions are meant for specific purposes also means that they are very much like adhoc solutions and so the same AI models probably can't be used elsewhere than to where they were originally designed. This is due to the fact that AI solutions are most often trained with very specific data about certain problem that want to be solved. Thus the trained model can only be used to solve that specific business problem. If the same base model wanted to be used somewhere else then the AI system would need to be trained again with data related to the new problem. In addition if the context has been changed a lot, where the model is implemented, then also the underlying algorithms might need to be changed. Thus in addition to different methods being meant for different purposes also almost every AI solution is target and problem specific, and not applicable elsewhere.

One important finding based both on the literature review and the interviews was that the key to practically all AI solutions is machine learning in some form. This is the very core of most AI application, and thus many challenges and limitations that relate to it are also related to every AI application. This was noticed through the interviews and literature review when the same challenges circulated with every capability. Thus essentially all AI application have certain common requirements. First is that AI solutions require huge amounts of data, this came very clear already after conducting the literature review, where the importance of data was highlighted throughout the literature. Data is the base for all AI applications as AI models usually work in a way that they are fed with huge amounts of data, and the data is used to teach the algorithms on how to behave. The more data algorithm is fed, usually the better as it then has more possibilities to learn all the features within the data, and thus become more accurate.

Business processes are also a valid target for implementing AI solutions. Of course there are plenty of simple processes and tasks where simple automation, like RPA, can do the trick and AI is not needed. When AI however steps into the picture is in those cases where humanlike judgement and decision making is needed. One problem that relates to using AI in business process is the fact AI solutions give a probabilistic outcome. This means that there are multiple possible outcomes to the situation and thus AI system gives a probability. Therefore the AI system will never results in certain results. However an interesting fact that came up in the interview with Paajanen was that this doesn't actually differ

so much from human decision making, because rarely even human decision making can reach hundred percent certainty. After all many times human's decisions aren't either anything but an enlightened guess at the best. Problem here is thus more psychological and philosophical than actually technological. We have the capabilities, the technology is there and most often so is the data. Thus the problem is more or less: can we trust the trust the probability if it's produced by AI. This question is therefore one of the most important and aching issues, as the usefulness of AI isn't only about technological capabilities but it's also related to people's attitudes towards AI decision making. If however some level of uncertainty is accepted the possible applications for AI are numerous. Although as found out in the interviews it maybe isn't wise to fully rely on AI to make decisions, even if it would be possible, as it's always good to have someone to at least oversee the process and make sure that the AI system gives consistent results. How interviewees then would divide this decision making responsibility with humans and AI would be to look at the processes. If the process or task is very mission critical, let's say health care diagnostics or decisions related hiring or letting people go, then AI should only be used as a supportive measure to give additional arguments for people to make the actual decision, as was stressed by the interviewees. However AI could be used in end-to-end solutions where single decisions don't have so far-reaching consequences, like categorizing tickets and other types of processes where volumes are high and few errors in the end results don't result in big problems. Thus the AI decision making can be divided to two types; human assisted decision making, where AI systems give probability driven predictions which humans can then use as additional criteria for final decision making, and to automated decision making where the decision responsibility is fully assigned to AI.

As data is the number one requirement for the use of AI solutions, many of the challenges and considerations when using AI also relate to it. As already mentioned, amount of data is a hugely important factor when using AI solutions. However as Paajanen pointed out in the interview this is also really dependent on the case where the AI solution is implemented. As found out in the literature review AI solutions' underlying models are taught based on the data they are fed with. If the use cases contain data that is rich in features then more data is needed to teach the model all those necessary features. However as brought up in the interview with Paajanen, the results don't have to be perfect, and as mentioned earlier AI can't even result in certain answers. Usually the models are performing accurately enough for business purposes if they give a probability of eighty percent.

When it comes to data, also data quality is an important factor with AI solutions. Thus if the quality of the data is bad, then so will be the results. Quality then can be affected by many factors; data can include abnormalities that doesn't belong there, and so data can be biased. This is one important factor that was highlighted in both literature review and interviews. This is thus a true problem, as AI solutions don't understand what the data is about and so the biasness can't be identified by the system. Humans however would

maybe be able to pick up these data points that don't belong, but AI system can't. Then on the other hand biasness is also related to people. People are biased by their experience, culture and world views. This in turn makes it difficult for even people to be objective, as own biases are usually difficult to identify. People bias can then drive the outcomes of AI solutions in two ways; used data can be biased by peoples' views if the data is generated by humans, and the models can be biased by their creators. This biasness of models is also a tricky question, as when developer builds the algorithm the model eventually inherits the biases of the developer. Thus both of these ways of AI being biased leads to a situation where the AI system isn't any more necessarily objective, but it's guided by the subjective views of the people. Biasness can however be somewhat reduced by trying to clean the data out of this biased data, as Pursiainen stated in the interview. This can be done in a way that the results are checked to see if something correlates against the model in an unexpected way. Then these variations in the data can possibly be identified and stripped away. Here the risk of course relates to the fact how the actual issues and so biased points in the data can be distinguished from the variations that belong to the data, due to just differences and not actual biases, as explained in the literature review. Secondly when cleaning the data also other important factors might be lost from the data set which then might result in different conclusions, than the original dataset would give.

Lastly is the validation issue of the results. Many times AI decisions are driven by complex machine learning algorithms. Issue here often is that there is no transparency on how the AI system ended up in a certain outcome, as the algorithms learn based on the data they are fed and not by certain step-by-step rules. This means that the steps in how the AI system gets to results are somewhat opaque. Thus validating the results AI system gives might be rather difficult. This opaqueness can however be fought against, at least to some extent, as Pursiainen stated in the interview. One approach here could be to test the model after removing the identifying labels from the data. After this the model could be tested against earlier results, from those before implementing AI solutions, and check if the model gives still expected results. This would thus show using unsupervised learning if the algorithm gets to the same results, and so works in a proper. This doesn't of course still show the exact way how the AI gets to the outcomes but at least the results could be then trusted more.

6.1.2 Intelligent automation

This subchapter is answering mostly to the following sub-research questions: *"What is robotic process automation?", "How robotic process automation can be used to automate processes and what benefits as well as limitations does it have?"* but additionally gives insight in the context of intelligent automation also to the following: *"What are different artificial intelligence capabilities and how they can be used?", "What kind of risks are related to using artificial intelligence?"*. First finding based on the literature review was that RPA is a technology meant for specific tasks i.e. automating business

processes, using existing applications and working there on the presentation layer. It has great benefits within its capabilities, but from the context of these benefits can also be derived its greatest limitations and challenges. RPA is useful for automating processes that have clearly defined unambiguous rules, and don't have the need for complex decision making. Due to this it's also very cost effective way to automate processes and increase productivity. In addition as RPA operates on the presentation layer automation can be done with minimal process changes and they can be implemented to virtually any digitalized process that humans perform, as long as other criteria are met. As the robot uses clear defined rules, it also means that the robot can only work by these rules. Thus robots don't make mistakes if the process is taught correctly from the beginning. In addition to these RPA can only work with structured data, as it's doesn't have the ability to interpret.

As from the benefits of RPA can be seen, most of these also have their downsides. Firstly, RPA can only handle structured data and therefore working with unstructured data is nearly impossible. This is because unstructured data contains information that isn't in a readable format for the robot, these would be for example emails, like Pursiainen stated. Thus for the robot to being able to understand this kind of free form unstructured information, that doesn't follow certain representation guidelines, the information would have to be translated to a structured and thus more 'machine readable' format for the robot. The second key issue then with robotics is the fact that it needs clear rules to work with. For this reason processes that require humanlike judgement or decision making are out of the scope of RPA, or at least the process need to be split to pieces where robots and humans handle different parts.

How intelligent automation, meaning combination of AI capabilities with RPA, then could address these issues is an interesting subject. As was stated by the interviewees RPA is a great tool but it has two major problems; handling of unstructured data as well as semi-structured input data and making complex decisions. However when looking at the capabilities within AI (see Figure 10), it can be quickly seen that AI's main objectives also address the main challenges with RPA. "Capturing information" is muchly about transforming unstructured data into structured format, and "what is happening" relates to the making sense from different kinds of structured data, using capabilities like prediction, optimization and NLU. This thus directly addresses the decision making problem with RPA. The same views was also brought up and very much highlighted in the interview with Pursiainen. This however leads to the fact that processes using intelligent automation aren't anymore deterministic, with definitive results, but they are probabilistic in nature, and thus results are always probabilities. This however isn't necessarily a problem, as human decision making is also only based on good predictions as Paajanen highlighted in the interview. This is truly an interesting fact and one that should be kept in mind when utilizing AI with RPA. Together they don't result in perfect answers but good guesses based on the information they're given and the experience they have gained – just like humans.

Thus RPA is like a human without brains that can process well manual and repetitive work. It's highly efficient in specific easy tasks but it can't learn anything new. However when combined with AI capabilities RPA get brains. With these capabilities combined to RPA the robot can first interpret the surroundings to capture information out of it by seeing and hearing, and then converting the unstructured input data into structured format. After data is captured and converted AI capabilities mimic the other unique human capability, understanding and learning, to make sense of what is happening – and so to predict, understand or optimize what the process should do next. Thus in essence RPA and AI complement each other in many ways and when exploited together the best of both technology areas truly emerge. RPA focuses heavily on manual repetitive tasks whereas AI is most useful in situations where complex rules and unambiguous decision making makes the logic based RPA fail to succeed. As intelligent automation then is all about combining RPA with AI capabilities, the differences between these two can be concluded as shown below in Figure 11.

	Robotic process automation	Intelligent automation
Data	• Structured	• Unstructured
Tasks	• Rule-based	• Inference-based
Results	• Deterministic	• Probabilistic

Figure 11. *Discrepancies between robotic process automation and intelligent automation.*

Regarding the actual use of intelligent automation a couple specific use cases to combine RPA with AI were also brought up in the interview with Pursiainen. First one of these concerned the use of chatbots in bit of a different manner, than what the literature review implied. Chatbots were also brought up in the interviews as one of the most used solutions together RPA. Chatbots in essence mean conversational agents that interact with humans, usually in a text based interface but also via spoken language. Thus the main possible use cases that the literature review raised were the use of chatbots merely as a conversationalist from which users could then ask questions within certain topics and so ease their everyday work when information gathering eases. The use case that Pursiainen however brought up was related to using chatbots in the beginning of RPA processes to do data gathering and discussing with the user, and to eventually trigger RPA process for which the chatbot would 'translate' input data into a correct structured format. This interestingly could enable many possibilities. When chatbot would be used in the beginning to gather information, it could for example also gather unstructured inputs i.e. free form text requests from users or fetching data and then combining other AI capabilities like search to extract the data out of that information. In essence here chatbot would then work as an interface for users and also in the role of "capturing information", as the first objective of AI. Thus implementing chatbot in the front would enable multiple possibilities for robots

to handle different kinds of input data. This essentially was one of the key finding from the interviews. Chatbots or overall AI capabilities in the beginning of processes could enable and widen the scope of processes suitable for RPA in many ways, as the format if input data wouldn't any more be a restriction.

The second key use case of AI combined with RPA that interviewees brought up was related to the use of machine learning within processes' complex decision points where humans are currently still needed. As mentioned, current RPA solutions can only handle unambiguous rules and so humans are needed in those steps of the processes' where complex decision making is needed. This means that RPA is often implemented only to some parts of the processes, creating chains of tasks where some are executed by a human and some by a robot. This situation is depicted below in Figure 12. Hence, true end-to-end automation solutions with the use of RPA are still quite rare.

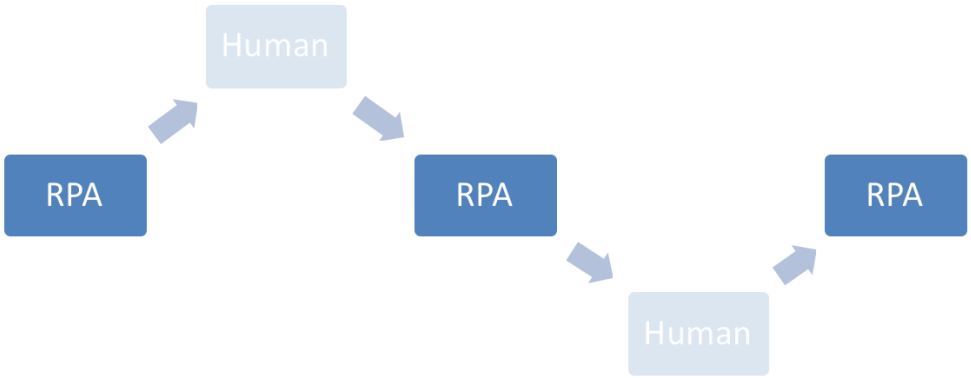

Figure 12. *Robotized process parts with human intervening in the middle.*

This issue of needing humans in the middle of processes to handle decision making steps could however be often replaced with AI. This transition from part-human-part-robot processes could be started by gathering all the available data and information about those decision points where human are still needed. This means that all the decisions, and related data to those decisions, would be logged and gathered for further use. After gathering huge amounts of data from these decision points, this data could be used to train a machine learning algorithm based on those previous decisions, as Pursiainen stated. Thus supervised learning would be used in these steps as the results of the decision are known for the training data, and the machine learning algorithm would need to learn the parameters based on which the results are derived. The more data from these decision points could be gathered the more possibilities the AI would have to learn the right parameters and weightings in the data based on which to make the correct decisions. After this the taught model could be used to make new decisions based on the historical data, and so the 'weak' spots in the RPA process could be filled. Additionally as those decisions by AI are always probabilistic, thus predictions, also some thresholds could be used in a sense that decisions by AI would only be forwarded to the robot if the accuracy would

exceed some certain level. This would partly answer to the trust issue related to AI decision making which was heavily emphasized in both interviews. In this sense using RPA together with AI would also address some issues related to AI. This prediction possibility with AI could also be a solutions for wider range of problems. Essentially what Pursiainen suggested, was to use prediction capability as a possible trigger for starting processes. Thus the use cases could relate to situations where AI triggers the RPA robot in case it predicts that something is going to happen. In this sense implementing AI solutions with RPA also means that robots aren't necessarily started anymore in a traditional way, manually or as scheduled, but that the AI solution triggers them based on the set parameters and tresholds.

The actual combining of RPA and AI has also become increasingly easier as most major RPA solutions have APIs which enable connecting them directly with major AI platforms. This is an important factor to consider with intelligent automation as it opens many possibilities to directly implement AI solutions to RPA processes. Interestingly Pursiainen also mentioned that few RPA providers are already deploying AI capabilities directly to their services and thus enabling even better possibilities for intelligent automation. Thus the interviewees saw the technical side of combining RPA with AI rather easy but the actual problems might then arise when looking for possible actual use cases. This is due to the basic requirements as described in chapter 6.1.1; process need to have lots of data, from where to teach the algorithms, and that data needs to be of good quality. Additionally to these requirements, also validity of the data comes to a greater meaning when used in a context of intelligent automation. Problem often is that the AI models and ways they reach the decisions are so opaque that there is no transparency on how the decisions were actually made. This means that validating of the results might be very hard or even impossible. This in turn can in cases be a showstopper as sometimes the reasons behind those outcomes are as necessary part of the outcome as are the results.

There are of course other challenges related to intelligent automation which might cause some headache, again depending on the use. One of the major ones is the EU's new law concerning data privacy, the GDPR. As found out in the literature review there are many aspect that need to be taken into account when using automation solutions now in the age of GDPR. Firstly it restricts the automated decision making related individual persons. This thus means that intelligent automation solutions can't be used to make decisions, for example about hiring or letting people go. They can be used to gain additional insight for these decision but the actual decisions can't be purely based on automation. Second key point relating to GDPR is the requirement for being able to show how the decisions were made. This then restricts the use of many machine learning based AI solutions as they are often very opaque in how they work and in how they reach the decisions. Additionally the data should also be handled without identifying information about people which might increase the risk that also some valuable data is lost in the process, and the models aren't so accurate anymore, as interviewees stated.

6.2 Summary and conclusions

The objective of this thesis was to understand how artificial intelligence capabilities could complement robotic process automation, what kind of possibilities there are and what kind of challenges relate to adding artificial intelligence to the mix. To better understand the context of this thesis and in order to get answers for the research questions a literature review was conducted and interviews were held with the professionals on the field. In this chapter the results from literature review and interviews are combined to answer first to the sub-research questions and lastly to answer the main research question.

First, sub-research questions, which are derived from the main research question are answered to gain more insight on the bigger picture. Thus the sub-research questions help to understand the topic as a whole, and also identify different aspects to it. The first and the second sub-research questions were derived from the main research question to better understand the context of RPA. The third, the fourth and the fifth sub-research questions were derived from the main research question in order to better understand the context of artificial intelligence and it what it consists of.

The first sub-research question is important as it gives background on the concept of RPA. The second sub-research question then is important as it gives insight on what can be done with RPA, and what kind of limitations there are to its use. The third sub-research gives understanding on what is understood as AI, as the literature review showed defining this concept can be rather difficult. The fourth sub-research question continues from the third one and gives more insight on what kind of different capabilities there are within AI, and how these capabilities can be used in practice. The fifth sub-research question then answers to the possible risks and limitations that might relate to using AI. Lastly the main research question is answered to conclude the findings.

What is robotic process automation?

RPA is a rather new technology focusing on automating processes. It differs from traditional process automation methods and technologies in the sense that it uses existing applications, and more precisely uses them the same way as humans. This is made possible by the fact that RPA uses the presentation layer of applications, meaning it uses the user interface of applications the same way as humans. This means that RPA also needs the same kind of accesses and rights as humans. Therefore automating processes using RPA is done by mimicking human actions on the computer manipulating applications by clicking and typing in ERPs, CRMs, spreadsheets, emails and so on. This said RPA is a solution that's configured to automate certain specific tasks. It needs very specific unambiguous rules together with structured input data in order to work properly. Thus RPA is a solution that's meant to ease the workload of people in mundane, routine like high volume tasks which doesn't require any kind of judgement or interpreting.

How robotic process automation can be used to automate processes and what benefits as well as limitations does it have?

RPA can be utilized in many kind of processes. When certain criteria for RPA are understood the possible implementation areas are also easier to find. Firstly, utilization of RPA requires that the processes have very specific and clear rules that the robot can follow. This is because this kind of software robots are "stupid" in a sense that they can't make any kind of humanlike decisions nor can they interpret information. Thus everything from input data to process steps need to be very structured and clear. This shows then another important criteria for RPA, structured inputs. This said it can be clearly identified that there are certain limitations that limit the scope of automatable processes with RPA. These limitations also indicate the criteria for successfully choosing processes for RPA. As the limitations indicate that processes need to have clear rules, meaning they are also manual processes. This therefore usually implies that the processes should have high volume of transactions or they should be performed frequently to justify the implementation costs. Second key thing derives from the fact that RPA takes advantage of the user interface of applications. This then means that RPA processes can use multiple systems, and thus maybe enable new automation possibilities to applications that previously couldn't have been integrated. Lastly as RPA needs specific rules to follow it also requires, stable systems which remain the same each time processes is executed. Otherwise the process would need to be reconfigured each time. Similarly another thing that relates to this is the requirement for minimal exception handling – the more there are exceptions in the process, the more complicated designing it will be, and thus more costly.

What is artificial intelligence?

AI is a concept that has become increasingly hyped in recent years. This has been muchly due to the rapid development in technologies, and so many new methods and applications in the field has emerged. Many definitions also relate to AI but most of them circle around a certain topic; computers ability to perform tasks, that require human like intelligence. What this human like intelligence then means is a bit problematic subject, as it would require defining what 'intelligence' and being intelligent actually means. However there are certain aspect and objectives that relate to human decision making and so also to AI, through which the concept can be examined. These are capturing information, understanding what is happening and understanding why something is happening. Then inside these objectives AI can be divided still to different capabilities that are used to achieve these objectives. Thus the concept of AI can be approached from the perspective of its capabilities, meaning what kind of tasks AI can be utilized for. Therefore in essence AI is a set of capabilities and technologies that each aim to address a certain kind of problems, and so the current AI capabilities aren't meant for general purposes to answer any given problems but they are so called narrow AI capabilities that can give answers to certain specific problems and issues.

What are different artificial intelligence capabilities and how they can be used?

As mentioned there are three objectives to AI; capturing information, understanding what is happening and understanding why something is happening. Here capturing information relates to extracting information from the subject of interest, be it images, text, sounds, data or some other kind of information source. Ultimately here the types of information can then be still divided into two main categories, to structured information and unstructured information. As there are different kinds of information to be captured there are also different capabilities for capturing information with AI. These are image recognition, speech recognition, search and clustering.

The second part after capturing information is then understanding "what is happening". This part relates to making sense of the captured information, which is often done also utilizing AI. Here the capabilities are NLU, optimization and prediction. The third and final objective of AI would be to understand "why something is happening". This is then very problematic area of AI as this would really require AI to understand what the data is actually about. To AI it however doesn't make a difference what the data represents because it only looks for similarities within the data, and thus aims to correlate data of interest with the taught model that's based on the historical data. Hence this, yet much of a theoretical part of AI is called general AI which is still not in the sphere of AI capabilities. These before mentioned capabilities within "capturing information" and understanding "what is happening" are therefore so called narrow AI capabilities, as they focus on solving specific tasks using certain kind of inputs.

Most of these capabilities are based on capability called machine learning. This is the base of most of the AI applications that actually aim learn as they are used. Of course there are also other methods but ultimately the very core of understanding AI is about understanding how machine learning works and what it requires. Machine learning essentially then is all about data. It need huge amounts of data, as the base of it is a kind of a learning algorithm. This data is then used to teach the model all the features within the data. More features there are, the more data is required to teach the model all those features. Thus the more data the model is fed, the more accurate it can be.

What kind of risks are related to using artificial intelligence?

AI can be useful in very many cases but there also some challenges that should be understood and considered before starting to implement any AI solutions. The most consequential requirement here is related to data. There needs to be lots of it and it needs to be of good quality. This is because if there isn't enough data the learning of the AI model might be limited and thus when used with new data it might not give results that reflect the

actual situation. This then makes the whole model unusable. Data quality then is important because it also makes it possible for the model to be accurate. Here there are many factors that could affect the data quality. Firstly data can be of bad quality in a sense that it doesn't represent the needed information, it can be somehow insufficient or imperfect. Data quality can be also bad if it's biased. This can be caused by two main reasons. First, the data can be biased if it's generated by humans. This is a problem as humans are biased by their experience and cultural background among other things. Here the real problem however is that it's also hard to identify and even acknowledge one's own biases completely. In this case the AI model then inherits those biases form the biased data, which is generated by humans. Secondly, the data can be biased by the developers who built the algorithm. Similarly the base algorithm for the AI model then inherits those biases of the people who created it. Thus both of these ways of AI being biased might lead to a situation where the AI system isn't objective.

How artificial intelligence capabilities can increase the automation level of business processes when combined with robotic process automation?

The goal of combing RPA with AI capabilities is an interesting subject. RPA is solution that offers possibilities to automate simple, rule-based processes with little costs and with minimal process changes. However as earlier mentioned there are certain challenges related to it. Processes need clear rules, there can't be need for human like judgement in the process, and RPA can only handle structured data. These criteria limit the implementation possibilities of RPA to certain kind of processes, and still often only sub-processes can be completely automated. Thus the overall process usually still needs humans to fill the complex decision making points.

In this sense there are three major areas where RPA lacks in capabilities. Handling unstructured data, handling complex decision making and handling processes where there are no clear rules. Then when looking at the AI's objectives "capturing information" and understanding "what is happening" and the capabilities within these objectives, it can be quickly seen that they aim to address similar kind of problems. Objective "capturing information" is muchly about handling unstructured data, and changing it to structured format. Therefore this objective addresses the handling of unstructured data problem with RPA. Then understanding "what is happening" objective addresses the issues with human like decision making and handling of complex decision rules. Thus RPA is like a human without brains. It can handle well explained, manual and repetitive tasks but it can't use judgement. However when AI capabilities are brought to the mix, RPA gets brains. Through different AI capabilities robots can understand different kinds of input data, make sense on what is actually happening and make decisions based on given information and gained experience – like humans. This way AI capabilities augment RPA in human like tasks and together they make intelligent automation.

Even though AI can bring a valuable addition to automating business processes in combination with RPA, there are also certain challenges that it brings along. Firstly, similar challenges that pure AI solutions have with data amount and data quality apply here. This means that processes need lots of data to teach the AI algorithms in order to get consistently good results out of them. Also the quality of data is even more critical when AI is used, as answers to the last sub-research question show. Thus these requirements might limit the use of AI very much if there is now historical data or of good quality data to work with. In addition to these 'basic' requirements for data there are also other considerations when implementing AI with RPA, and especially when implementing to business processes. For example, validity of the data comes to a greater meaning when used in a context of intelligent automation. Problem here however often is that the AI models and ways they reach the decisions are so opaque that there is no transparency on how the decisions were actually made – they are as said 'black boxes'. This means that validation of the results might be very hard or even impossible when the logic in how AI model reaches its decisions isn't transparent. This in turn can be a showstopper as sometimes the reasons behind those outcomes are as necessary part of the outcome as are the results. In addition to this, trusting AI decision making can become hard for people as results from AI systems aren't deterministic, but they are probabilistic. This means that there is always some level of uncertainty with the results – as they are merely probabilities. This combined with the fact that the decision making process is opaque means that people often have hard time to trust AI with decisions. Thus AI decision making is not perfect but nor is human's. When humans make similar kind of decisions they are also just good guesses based on the gained knowledge and experience, and not definite answers. However if it's a machine or a software humans often expect it to perform perfectly, flawlessly without mistakes and always the same way. This has been the common expectation and understanding for a long time but AI works in a different manner so the perspectives and expectations should be adjusted accordingly.

Additionally latest headache with AI solutions utilization is the EU's new law concerning data privacy, called GDPR. This restricts AI usage to some level, or at least requires extra work to make the processes comply with it. Often this means that full automation might not be possible when information about individuals are the core data source for processes. In these cases AI and RPA combination can be used but not to the ultimate level of reaching end-to-end automation.

6.3 Critical evaluation of the research

This thesis has couple issues that affect the results and so to the conclusions. Firstly there is an issue related to sample size of the interviews. For this thesis only two interviews were held. This can be considered as small sample size which will affect the conclusions that are drawn from the interview results. Additionally the two interviewees were both from the same company, so the sample group was homogeneous. This can be seen as a

disadvantage as the opinions might be affected by similar background and experiences. For this reason there might be similar biases for these interviewees affecting the results. Generalizability is also an issue that should be considered because of the above mentioned reasons. However the interview results showed similar implications as could be drawn from the literature review. Therefore the validity of the results can be considered credible because no big difference between theoretical and empiricism parts were found (Saunders et al. 2009, p. 156). Also the fact that interviewed people were experienced professionals with years of experience in their subject areas could be considered as an asset counter-weighing the small sample size, combined with the fact that the interviewees had years of experience from delivering these kind of solutions to customers in different sectors.

The literature review conducted for this thesis handled two different theories, from which especially theory about AI is a rather big entity to write about. The literature about AI is also very scattered and the definitions and capabilities were often interchangeable used which made it rather difficult to have a good grasp of the whole entity. In addition many AI capabilities and technologies are very much overlapping in the literature so defining the boundaries between matters was challenging. This in mind the literature review doesn't cover all the aspect of the theory, especially regarding AI. This however was a known choice and thus the scope was limited to examining AI capabilities and their challenges in generalized way, as explained in sub-chapter 1.3. Nevertheless the theory part consisted of all the most relevant parts related to the subject of this thesis. Sources used in the literature review were also chosen from known publishers and using databases containing only peer-reviewed material. Thus most of the results are also such that they would probably show in the similar kind of researches as well. Therefore this thesis can be considered reliable.

To sum up, the results of this study can be considered credible and valid as the purpose of this thesis was to form new insight that isn't available in the current literature, although the above issues should still be considered. The main results were derived by combining the results from literature review with the results from the interviews.

6.4 Future research

Many future research areas can be derived from the subject of this thesis. This thesis took a very broad and generic perspective to the topic due to organization's and researcher's preferences. This therefore led to situation where the focus of this thesis was on AI capabilities and challenges in the context of combining with RPA. To answer more specifically to sub-research question about different AI capabilities and their use, an interesting research area would be in case studies combining certain AI capabilities with RPA. This would enable taking a deeper look into specific use cases, and would show implications in more detail how certain capabilities can be used, the kind of benefits they can actually yield and what challenges as well as possible limitations might arise.

To answer more specifically to sub research question about challenges and risks related to AI, one further research area would be to focus on these in the context of AI and especially AI's use in business processes and decision making. This is also very philosophical subject as it relates muchly to technology behind AI solutions but more importantly many challenges relate to people's attitudes, assumptions and ethical considerations towards the use of AI, especially in decision making. This is an interesting field of research and case studies in the area would definitely prove useful as the use of AI in any field is a growing trend.

Additionally, also related to challenges, an interesting research area would be to focus on regulation related to AI's use. This of course is still a bit immature field as the use and possibilities of AI are changing very rapidly and the regulation always lags behind. This is however an increasingly important aspect to AI and thus an important field to study, as it will definitely affect the use of AI in the future in many ways – there is no doubt in that. First big implications of regulatory effects were seen in May 2018 when the new EU law, GDPR, took place. However the effects of GDPR aren't still fully understood in the context of using AI, so interesting research area would also be to study the effects of GDPR in more detail.

REFERENCIES

AbuShawar, B., Atwell, E. (2015). Alice Chatbot: Trials and Outputs. Computación y Sistemas, Vol. 19(4), pp. 625–632.

Agrawal, A., Gans, J., Goldfarb, A. (2017). How AI Will Change the Way We Make Decisions. Harvard Business Review. 26.7.2017. Available (Accessed 10.12.2017): https://hbr.org/2017/07/how-ai-will-change-the-way-we-make-decisions.

Asbjørn, F., Brandtzæg, P., B. (2016). Chatbots and the New World of HCI. Interactions, Vol. 24(4), pp. 38-42.

Aguirre, S. & Rodriguez, A. (2017). Automation of a Business Process Using Robotic Process Automation (RPA): A Case Study. Communications in Computer and Information Science, Vol. 742, pp. 65-71.

Anagnoste, S. (2017). Robotic Automation Process - The next major revolution in terms of back office operations improvement. Proceedings of the 11th International Conference on Business Excellence.

Asatiani, A., Penttinen, E. (2016). Turning robotic process automation into commercial success – Case OpusCapita. Journal of Information Technology Teaching Cases, 1 November 2016, Vol 6(2), pp. 67-74.

Asatiani, E., Kasslin, H., Penttinen, E. (2018) How to Choose Between Robotic Process Automation and Back-end System Automation. Twenty-Sixth Conference on Information Systems (ECIS2018). Portsmouth, UK.

Anifowose, F.A. (2013). Artificial intelligence application in reservoir characterization and modeling: Whitening the black box. Society of Petroleum Engineers - Saudi Arabia Section Young Professionals Technical Symposium 2011, pp. 45-57.

Bergstein, B. (2017). The Great AI Paradox. MIT Technology Review. 15.12.2017. Available (Accessed 20.02.2018): https://www.technologyreview.com/s/609318/the-great-ai-paradox/.

Brynjolfsson, E., Mitchell, T. (2017). What can machine learning do? Workforce implications: Profound change is coming, but roles for humans remain. Science, Vol. 358(6370), pp. 1530-1534.

Burgess, A. (2017). The executive guide to artificial intelligence: how to identify and implement applications for AI in your organization. Springer International Publishing AG

Burnett, S., Jain, A., Modi, A. (2017). Rise of Automation in P&C Insurance. Everest Group. 8.2.2017.

Burnet, S., Modi, A. (2017). RPA Futures – Accelerated and Intelligent automation. Everest Group. 8.12.2017.

Burnett, S. (2017). Automating Content-Centric Processes with Artificial Intelligence (AI). Everest Group. 8.12.2017.

Castelvecchi, D. (2016). Can we open the black box of AI? Nature, Vol. 538(7623), 5 October 2016, pp. 20-23.

Cambria, E., White, B. (2014). Jumping NLP Curves: A Review of Natural Language Processing Research. IEEE Computational Intelligence Magazine, May 2014 Issue.

Campolo, A., Sanfilippo, M., Whittaker, M., Crawford, K. (2017). AI Now 2017 Report. Available (Accessed 20.4.2018): https://ainowinstitute.org/AI_Now_2017_Report.pdf.

Craig, A., Lacity. M., Willcocks, L. (2015). Robotic Process Automation at Xchanging. The Outsourcing Unit Working Research Paper Series. 15/03.

Davenport, T. H., Dreyer, K. J. (2018). AI Will Change Radiology, but It Won't Replace Radiologists. Harvard Business Review. 27.3.2018. Available (Accessed 30.4.2018): https://hbr.org/2018/03/ai-will-change-radiology-but-it-wont-replace-radiologists.

Davenport, T. H., Kirby, J. (2015). Beyond Automation. Harvard Business Review . June 2015 Issue.

Davenport, T. H., Ronanki, R. (2018). Artificial Intelligence for the Real World. Harvard Business Review. Jan-Feb 2018 Issue. Available (Accessed 20.2.2018): https://hbr.org/2018/01/artificial-intelligence-for-the-real-world.

EUR-Lex. (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). Available (Accesses 27.12.2017): http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32016R0679.

Fersht, P., Slaby, J. R. (2012). Robotic automation emerges as a threat to traditional low cost outsourcing. HfS Research. October 2012.

Forrester Consulting. (2014). Building A Center Of Expertise To Support Robotic Process Automation – Preparing For The Lifecycle Of Business Change. February 2014.

Frey, C. B., Osborne, M. A. (2017). The Future of Employment: How Susceptible Are Jobs To Computerisation?. Technological Forecasting and Social Change, Vol. 114(C), pp. 254-280.

Fung, H.P. (2014). Criteria, Use Cases and Effects of Information Technology Process Automation (ITPA). Advances in Robotic and Automation. Vol. 3(3), pp. 1–11.

Gartner. (2017). Gartner Identifies Three Megatrends That Will Drive Digital Business Into the Next Decade. Gartner Available (Accessed 28.10.2017): https://www.gartner.com/newsroom/id/3784363.

Hume, Kathryn. (2018). When Is It Important for an Algorithm to Explain Itself?. Harvard Business Review. 6.7.2018. Available (Accessed 20.7.2018): https://hbr.org/2018/07/when-is-it-important-for-an-algorithm-to-explain-itself.

Jesuthasan, R., Boudreau, J. (2017). Thinking Through How Automation Will Affect Your Workforce. Harvard Business Review. April 2017 Issue.

Jung, J., Concannon, C., Shroff, R., Goel, S., Goldstein, D., G. (2017). Creating Simple Rules for Complex Decisions. Harvard Business Review. 19.4.2017. Available (Accessed 20.5.2018): https://hbr.org/2017/04/creating-simple-rules-for-complex-decisions.

Knight, W. (2016). AI's Language Problem. MIT Technology Review. 9.8.2016. Available (Accessed 20.12.2017): https://www.technologyreview.com/s/602094/ais-language-problem/.

Knight, W. (2017). The Dark Secret at the Heart of AI. MIT Technology Review. 11.4.2017. Available (Viewed): https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/.

Kukreja, M. & Nervaiya, A. (2016). Study of Robotic Process Automation (RPA). International Journal on Recent and Innovation Trends in Computing and Communication. Vol. 4(6), pp. 434 – 437.

Lacity, M., Willcocks, L. (2015). What Knowledge Workers stand to gain from Automation. Harvard Business Review. June 19, 2015.

Lacity, M., Willcocks, L. (2016a). A New Approach to Automating Services. MIT Sloan Management, Vol. 58(1), pp. 41 – 49.

Lacity, M., Willcocks, L. (2016b). Robotic Process Automation at Telefonica O2. MIS Quarterly Executive. Vol. 15(1), pp. 21 – 35.

Lacity, M., Willcocks, L. (2016c). Robotic Process Automation: The Next Transformation Lever for Shared Services. The Outsourcing Unit Working Research Paper Series, London School of Economics and Political Science, January 2016. p. 35.

Lacity, M., Willcocks, L., Craig, A. (2015). The IT Function and Robotic Process Automation. The Outsourcing Unit Working Research Paper Series, London School of Economics and Political Science, October 2015. p. 38.

Le Clair, C., Cullen A., King M. (2017). The Forrester Wave™: Robotic Process Automation, Q1 2017, Available (Accessed 20.5.2018): https://www.edgeverve.com/wp-content/uploads/2017/02/forrester-wave-robotic-process-automation.pdf.

Liang, H., Dickerson, C.E., Champion, D., Battersby, D. (2016). A Framework for Technology Assessment from Different Scientific Viewpoints: Preliminary Report. 11th System of Systems Engineering Conference (SoSE), Kongsberg, pp. 1-5.

Ng, A. (2016). What Artificial Intelligence Can and Can't do Right Now. Harvard Business Review. 09.10.2016. Available (Accessed 24.2.2018): https://hbr.org/2016/11/what-artificial-intelligence-can-and-cant-do-right-now.

Niehaves, B., Poeppelbuss, J., Plattfaut, R., Becker, J. (2014) BPM Capability development – a matter of contingencies. Business Process Management Journal. Vol. 20(1), pp. 90 – 106.

Oxford Reference. (2018). Artificial Intelligence. Available (Accessed 20.4.2018): http://www.oxfordreference.com/view/10.1093/oi/authority.20110803095426960#.

Pavaloiu, A. (2016). The Impact of Artificial Intelligence on Global Trends. Journal of Multidisciplinary Developments. Vol. 1(1), pp. 21 – 37.

Ransbotham, S., Kiron, D., Gerbert, P., Reeves, M. (2017). Reshaping Business With Artificial Intelligence – Closing the Gap Between Ambition and Action. MIT Sloan Management Review. Fall 2017.

Reuner, T. (2017). Navigating the AI Hype With a Clearer Segmentation. 28.11.2017. HfS Research.

Richter, P.C. & Brühl, R. (2016). Shared service center research: A review of the past, present and future. European Management journal. Vol. 35, pp. 26-38.

Robson, C. (2002) Real World Research, 2nd ed. Oxford, Blackwell.

Rossi, F. (2016). Artificial Intelligence: Potential Benefits and Ethical Considerations. European Parliament, Policy Department C: Citizens' Rights and Constitutional Affairs. 12.10.2016. Available (Accessed 25.10.2017): http://www.europarl.europa.eu/thinktank/en/document.html?reference=IPOL_BRI(2016)571380.

Russell, S. & Norvig, P. (2010). Artificial Intelligence: A Modern Approach. Prentice Hall, Upper Saddle River, NJ.

Saunders, M., Lewis, P., Thornhill, A. (2009). Research Methods for Business Students, 5th ed. Harlow, England, Pearson.

Scherer, M. U. (2016). Regulating Artificial Intelligence Systems: Risks, Challenges, Competences, and Strategies. Harvard Journal of Law & Technology. Vol. 29(2), pp. 354-400.

Sharp, J., Peters, J. and Howard, K. (2002) The Management of a Student Research Project, 3rd ed. Aldershot, Gower.

Singh, A. (2017). Deep Learning Will Radically Change the Ways We Interact with Technology. Harvard Business Review. Available (Accessed 12.2.2018): https://hbr.org/2017/01/deep-learning-will-radically-change-the-ways-we-interact-with-technology.

Singh, A. L., Tiwari, T., Singh, I. L. (2009). Effects of Automation Reliability and Training on Automation-Induced Complacency and Perceived Mental Workload. Journal of the Indian Academy of Applied Psychology. October 2009, Vol. 35, Special Issue, pp. 9-22.

Snow, J. (2017). New Research Aims to Solve the Problem of AI Bias in "Black Box" Algorithm. MIT Technology Review. 7.11.2017. Available (Accessed 12.12.2017): https://www.technologyreview.com/s/609338/new-research-aims-to-solve-the-problem-of-ai-bias-in-black-box-algorithms/.

Socher, R. (2018). AI's Next Great Challenge: Understanding the Nuances of Language. Harvard Business Review. 25.7.2018. Available (Accesses 28.7.2018): https://hbr.org/2018/07/ais-next-great-challenge-understanding-the-nuances-of-language.

Soll, J. B., Milkman, K. L., Payne, J. W. (2015). Outsmart Your Own Biases. Harvard Business Review. May 2015 Issue.

Srinivasan, V. (2016). Context, Language, and Reasoning in AI: Three Key Challenges. MIT Technology Review. 14.10.2016. Available (Accessed 12.2.2018): https://www.technologyreview.com/s/602658/context-language-and-reasoning-in-ai-three-key-challenges/.

Woo, S., Boyle, E., Spector, P. (2017). Best practices in developing, conducting, and evaluating inductive research. Human Research Management Review, Vol. 27(2), June 2017, pp. 255-264.

Tan, Y., Takagi, H., Shi, Y. (2017). Data Mining and Big Data: Second International Conference, DMBD 2017, Fukuoka, Japan, July 27 – August 1, 2017, Proceedings. Springer International Publishing.

Varghese, F.C. (2017). The Impact of Automation in IT Industry: Evidences from India. International Journal of Engineering Science and Computing. Vol. 7(3), pp. 5000-5004.

Yeomans, M. (2015). What Every Manager Should Know About Machine Learning. Harvard Business Review. 7.7.2015. Available (Accessed 20.5.2018): https://hbr.org/2015/07/what-every-manager-should-know-about-machine-learning.

Yin, R.K. (2003) Case Study Research: Design and Method, 3rd ed. London, Sage.

APPENDIX A: INTELLIGENT AUTOMATION EXPERT INTERVIEW

Intelligent automation potential

- How would you describe the meaning of artificial intelligence and intelligent automation?

- What do you think are the most critical problems with robotic process automation where artificial intelligence could help and how?

- What kind of use cases have you seen or experienced in combining artificial intelligence technologies with robotic process automation?

- What kind of prerequisites are there to start using artificial intelligence technologies with robotic process automation?

Challenges related to intelligent automation

- What kind of challenges and problems might arise when implementing artificial intelligence technologies with robotic process automation?

- How do you think the new General Data Protection Regulation is going to effect the use of artificial intelligence technologies?

How to get started

- What do you think would be a good starting point for organizations to start utilizing artificial intelligence technologies with robotic process automation?

APPENDIX B: DATA SCIENCE EXPERT INTERVIEW

Artificial intelligence

- How would you describe what artificial intelligence is and what it consists of?
- Which in your experience are the most mature artificial intelligence technologies?

Data

- What are the requirements for data when it comes to using artificial intelligence technologies?
- How can we ensure and can we know that we have enough data and of good quality?
- How do we know we can trust the decision making with artificial intelligence technologies?

Challenges

- What are challenges and risks when using artificial intelligence technologies?
- Are there typical challenges related to certain artificial intelligence technologies?
- How do you think artificial intelligence and built in biases either in the models or data sets will affect?
- How do you think General Data Protection Regulation (GDPR) will affect the use of AI? Will it limit the use cases?

Prerequisites and requirements

- What kind of perquisites are there to start using artificial intelligence technologies?
- What are the most critical factors when implementing AI technologies to business processes?
- How to get started with implementing artificial intelligence technologies to business processes?