UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Artificial Intelligence within Financial Services

-In Relation to Data Privacy Regulation

Master Thesis Project in Innovation and Industrial Management
Spring 2018

Johanna Moberg & Alexis Olevall
Supervisor: Rick Middel
Graduate School
Innovation & Industrial Management

Artificial Intelligence within Financial Services
In relation to Data Privacy Regulation

By Johanna Moberg & Alexis Olevall

School of Business, Economics and Law, University of Gothenburg,
Vasagatan 1, P.O. Box 600, SE 40530 Gothenburg, Sweden

Acknowledgement

Johanna Moberg
johanna.e.moberg@gmail.com

Alexis Olevall
alexis.h.olevall@gmail.com

ABSTRACT

Background: The data that is processed about individuals is increasing rapidly, which is one contributing factor to the increased usefulness of Artificial Intelligence (AI) within today's businesses. However, this extensive processing of personal information has become heavily debated, and is an area that the General Data Protection Regulation (GDPR) aims to regulate. At the same time, it has been argued that the formulation of the GDPR is infeasible with AI technology. One industry where an extensive amount of data about customers is processed, including automated processing based on AI technology, is financial services.

Purpose and Research Question: The purpose of this research is to examine what impact the GDPR has on AI applications within financial services, and thereby the research question stated is: *What is the potential impact of the GDPR on Artificial Intelligence applications within the financial services industry?*

Methodology: To fulfil the purpose of this research, a qualitative research strategy was applied, including semi-structured interviews with experts within the different fields of examination: law, AI technology and financial services. The findings were analysed through performing a thematic analysis, where coding was conducted in two steps.

Findings: AI has many useful applications within financial services, which currently mainly are of the basic form of AI, so-called rule-based systems. However, the more complicated machine learning systems are used in some areas. Based on these findings, the impact of the GDPR on AI applications is assessed by examining different characteristics of the regulation. The GDPR initially imposes both an administrative and compliance burden on organisations within this industry, and is particularly severe when machine learning is used. These burdens foremost stem from the general restriction of processing personal data and the data erasure requirement. However, in the long term, these burdens instead contribute to a positive impact on machine learning. The timeframe until enforcement contributes to a somewhat negative impact in the short term, which is also true for the uncertainty around interpretations of the GDPR requirements. Yet, the GDPR provides flexibility in *how* to become compliant, which is favourable for AI applications. Finally, GDPR compliance can increase company value, and thereby incentivise investments into AI models of higher transparency.

Conclusion: The impact of the GDPR is quite insignificant for the basic forms of AI applications, which are currently most common within financial services. However, for the more complicated applications that are used, the GDPR is found to have a more severe negative impact in the short term, while it instead has a positive impact in the long term.

Contribution: This research makes a theoretical contribution to the field of research about the feasibility of the GDPR with technology, by examining how this regulation will impact one specific technology, that is, Artificial Intelligence. This study also makes a practical contribution by reducing the ambiguities for companies about how the GDPR will impact AI applications.

Keywords: *Artificial Intelligence, Machine learning, Rule-based Artificial Intelligence, Regulation, General Data Protection Regulation, Innovation, Financial Services.*

Definitions

Artificial Intelligence (AI) - AI is described to be different technologies that enable machines to perform tasks that historically have required human intelligence (Tecuci, 2012).

Rule-based systems - AI systems where humans determine and program the rules (Kingston, 2017).

Machine learning - AI systems where the machine has the ability to learn from data without predetermined rules from humans (Mittelstadt, Allo, Taddeo, Wachter & Floridi, 2016).

Artificial Neural Networks (ANNs) - A type of machine learning that has been developed with the human brain as inspiration that has a complex structure with many interconnected layers (Lake, Ullman, Tenenbaum, & Gershman, 2017).

Statistical machine learning - A type of machine learning that is based on statistics and probabilistic reasoning (Ghahramani, 2015).

Natural Language Processing (NLP) - The ability to communicate in natural language, which is the language used by humans when communicating (Lake et al., 2017).

General Data Protection Regulation (GDPR) - An EU regulation that regulates data protection and privacy for all individuals residing within the European Union (EU GDPR, 2018a).

Data subject - A living individual whose data is processed. Any information related to a natural person or 'Data Subject' that can be used to directly or indirectly identify the person (EU GDPR, 2018a).

Data controller - A controller is the entity that determines the purposes, conditions and means of the processing of personal data (EU GDPR, 2018a).

Unstructured data - Data without a predetermined structure, for example text or pictures (Datainspektionen, 2018a).

Structured data - Data with a predetermined structure, such as data registries and databases (Datainspektionen, 2018a).

Innovation - The introduction of new ideas into the market, that are translated into commercial or technological outcomes that are socially desirable through the usage of new processes, products, or services (Ranchordás, 2015).

Regulation - A legislative act that is binding (EU GDPR, 2018a).

Financial services - An industry that encompasses a range of institutes, including banks, insurance companies, securities brokers, investment companies (Hämmerli, 2012).

Table of Contents

List of Figures

List of Tables

1. INTRODUCTION

The initiating chapter begins with a description of the background to the research topic, which is followed by the purpose and research question. After that, the delimitations and contribution of this study are described. The chapter then ends by outlining the disposition of the study.

1.1 Background

In today's world, companies collect greater amounts of information than ever before (Villaronga, Kieseberg, & Li, 2017). With the increased volume of data that is being produced about individuals, in combination with technical advancements, it is possible to make more in-depth analyses and gain more insights about collected data (Oliver Wyman, 2017). Such developments create opportunities to decrease costs and develop new business models (ibid.). At the same time, how companies manage and process data about their customers has become one of the most discussed topics of this decade (ibid.). One technology that is increasingly applied for processing, and to generate better insights about data is Artificial Intelligence (AI) (De Laat, 2017). This technology is becoming increasingly important within human society (Villaronga et al., 2017). The significant increase in computing power and storage capacity, along with the extensive amount of available data, have contributed to major advancements within the field of AI (Kaplan, 2016; Villaronga et al., 2017).

AI is described to be different technologies based on algorithms that enable computers to automatically perform tasks that historically have required human intelligence (Van de Gevel & Noussair, 2012; Kaplan, 2016). While the technology has its roots in the 1950's, it is first in recent years that applications of AI have become more relevant and useful (Lake, Ullman, Tenenbaum, & Gershman, 2017; Tecuci, 2012; Van de Gevel & Noussair, 2012). AI technologies can perform a wide range of tasks both faster and at a lower cost than humans, but also more ambitious tasks than what humans can carry out by themselves (Kaplan, 2016). The field of AI has taken different directions over the years (Lake et al., 2017), and can be divided into two broad approaches; Artificial General Intelligence (AGI) and Narrow AI. The area of AGI aims to fully replicate human-level general intelligence in machines or computers (Goertzel, 2014; Van de Gevel & Noussair, 2012), whereas Narrow AI only can solve a narrow set of specific tasks (Goertzel, 2014). While AGI has not been achieved yet, and some doubt that it ever will, Narrow AI has achieved remarkable success (Bostrom, 2014; Goertzel, 2014). AI technology has been identified to drive innovation for both products and services (McKinsey, 2017). As of today, AI is successfully applied across a wide range of industries, such as for providing buying recommendations based on previous behaviours, at border crossings for face recognition, in autonomous vehicles (Bostrom, 2014), and as decision-support for credit evaluation (Bahrammirzaee, 2010).

In fact, AI algorithms are today present in our everyday life, and automated decision-making is becoming increasingly common (Art. WP 29; Mittelstadt, Allo, Taddeo, Wachter & Floridi, 2016). In this kind of algorithm-driven society, Malgieri and Commandé (2017) point out that it is crucial that the decision-making of AI algorithms is transparent and comprehensible for

individuals to understand how companies use their information. However, the decision-making process of many AI models is often complicated and difficult to understand (Bohanec, Robnik-Šikonja & Kljajić Borštnar, 2017; Mittelstadt et al., 2016). At the same time, consumers demonstrate an increased awareness of privacy, and are becoming more restricted in sharing their data (Kieselmann, Kopal, & Wacker, 2016; Van Otterlo, 2014).

A response to the increased automated processing of data is the introduction of the General Data Privacy Regulation (GDPR), enforced in May 2018 (Kingston, 2017). The GDPR aims to strengthen the rights for individuals by imposing stricter privacy and safety requirements on organisations, such as increased transparency of automated data processing based on AI technology (Art. WP 29; Wachter, Mittelstadt & Floridi, 2017a). Nonetheless, in attempting to protect citizens, it has been argued that the GDPR could have a negative impact on current technologies, as it is stated that the requirements are not feasible with current technologies and are difficult to comply with when AI is used (Kieselmann et al., 2016; Kingston, 2017; Malgieri & Commandé, 2017; Villaronga et al., 2017; Wachter et al., 2017a). These requirements mainly refer to an alleged right for individuals to receive an explanation to how automated decisions have been taken (Malgieri & Commandé, 2017; Wachter et al., 2017a) and a right to request that one's personal data is erased (Villaronga et al., 2017). The aspect of explaining decisions is argued to be problematic since some AI models are difficult to understand due to their complex structure (Kingston, 2017). To that, the erasure requirement has been criticised for being formulated with respect to how humans think and forget, without accounting for how machines function (Villaronga et al., 2017). However, it is widely debated how these requirements should be interpreted and what the practical implications actually will be (Malgieri & Commandé, 2017; Villaronga et al., 2017; Wachter et al., 2017a).

What can be said though is that the impact of the GDPR will become more severe the more personal data a company collects and processes (Oliver Wyman, 2017). One industry that processes an extensive amount of data about their customers is the Financial services industry (ibid.), including actors that provide services within banking, insurance, security brokerage and investments (Hämmerli, 2012). Such extensive processing is needed since the services provided require access to customer data and frequent interaction with the customers (Oliver Wyman, 2017). To that, banking and insurance are identified to be industries where AI-based automated decisions about customers are conducted on a more regular basis (Art. WP 29; PWC, 2017). For example, AI is used as decision-support for credit evaluation (Bahrammirzaee, 2010), as well as to analyse risk and price premiums within insurance (Rouse & Spohrer, 2018).

It is concluded that there are uncertainties about what the GDPR requirements mean for businesses using AI. It has been argued that GDPR could have a negative impact on current technologies, there among AI. Nonetheless, regulations have a multifaceted impact on innovation and technologies (Ashford, Ayers, & Stone, 1985; Blind, 2012; Pelkmans & Renda, 2014; Ranchordás, 2015). The impact of regulations can be both positive and negative, which depends on the characteristics of the regulation (ibid.). Thereby, since AI is one technology that is increasingly used to automate data processing within businesses (Art.

WP 29), there is a need to examine the different aspects of the GDPR in greater detail to determine the impact that this regulation will have on AI applications.

1.2 Purpose and Research Question

Automated decision-making based on AI is becoming increasingly common (Art. WP 29; Mittelstadt et al., 2016). However, the GDPR raises questions about the extent that it will be possible for companies to continue to use AI in data processing. The Financial services industry processes extensive amount of information about customers and has therefore been identified to be significantly affected by the GDPR (Oliver Wyman, 2017). Accordingly, the purpose of this thesis is to examine whether the GDPR could impact AI applications within financial services. Hence, the following research question is formulated:

- *What is the potential impact of the GDPR on Artificial Intelligence applications within the financial services industry?*

With Artificial Intelligence applications, it is meant how AI technology actually is used in different ways, such as for predictive purposes or to automate organisational processes. Stating this research question means that both the potentially positive and negative effects that the GDPR could have on AI applications will be examined. The research question will be answered by gathering information from three different groups of respondents with different expertise: Legal experts, AI experts and Industry actors within financial services. The Legal experts are interviewed to explain the content of GDPR and the AI experts to assess the technical aspects of AI. Finally, Industry actors within financial services are interviewed to assess the current state of AI applications and how the development is likely to be in future years, as well as how these organisations perceive the GDPR. These different areas will then be connected, along with the findings from the literature review, in the analysis section of the report.

1.3 Contribution of the Research

To derive from section 1.1, AI technology demonstrates great potential to create value within financial services, but it is uncertain what impact the GDPR will have on the usage of AI. First, it is argued that there exists a gap between the formulation in GDPR and what is feasible with current technologies (Kieselmann et al., 2016). Secondly, it is stated to be unclear how the requirements of the GDPR should be interpreted, and what the practical implications will be (Malgieri & Commandé, 2017; Villaronga et al., 2017; Wachter et al., 2017a). Hence, this research will make a practical contribution by reducing the ambiguity for financial services practitioners by presenting an overview of expert opinions about AI characteristics and how the regulation should be interpreted, in relation to the industry conditions. At the same time, this study will make a theoretical contribution to the field of research about the feasibility of the GDPR with technology, by examining how this regulation will impact one specific technology, that is, Artificial Intelligence.

1.4 Delimitations

This thesis is focused on the overall financial services industry rather than individual organisations, and also takes a provider perspective of the industry as opposed to consumer considerations. Moreover, the financial services industry is defined to include companies within banking, insurance, security brokerage and investment (Hämmerli, 2012). Due to time constraints and the scope of this study, the focus is on actors that provide services within banking, investment, and insurance. Thereby security broker firms are excluded, and within insurance, the focus is mainly on life and pensions insurance. Hence, the results may not be representative for the whole financial services industry.

Moreover, the focus of this thesis is on Sweden even though the GDPR is an EU regulation. This decision was made since the GDPR is a regulation from the EU, meaning that country-specific regulatory bodies will enforce the regulation in each country (EU GDPR, 2018a), and therefore the precise enforcement may be somewhat different between the EU-member states. Thereof, respondents in the research are limited to people working at companies located in Sweden.

Furthermore, since AI is a broad field that includes multiple and diverse technologies, the focus of this study is limited to Narrow AI. This kind of AI refers to intelligent systems that can perform a narrow range of tasks, as opposed to the Artificial General Intelligence that aims to fully replicate human intelligence (Goertzel, 2014). Hence, when the term "AI" is used hereafter in this report, it refers to Narrow AI. Moreover, the GDPR includes 99 different articles (EU 2016/679), and this study focuses on the parts that are described to have the greatest impact on AI applications; Articles 13-15, 17, and 22 (Villaronga et al., 2017; Wachter et al., 2017a). These articles are further complemented with information from the two initial chapters of the regulation that specify the regulations' general provisions and principles of processing personal data. Hence, this thesis is not a guide for how organisations should become GDPR compliant, instead the focus is on the potential impact the regulation will have on AI applications.

Thereto, this thesis examines perceptions about the future. Hence, the collected information from literature and interviews about the impact that the GDPR will have are speculations, and may not become a reality. Neither are other factors than the GDPR that could affect AI applications taken into consideration, such as other regulations.

1.5 Disposition of the Report
The report will follow the disposition shown in Figure 1 below.

1. Introduction — • Describes the background as well as purpose and research question of the study. Thereto, the contribution and delimitations of the study are presented.

2. Literature Review — • Presents the literature within the field of AI, regulatory impact on innovation, as well as the general data protection regulation (GDPR) in relation to AI.

3. Methodology — • Presents the research strategy, research design, and research method for this study, as well as a discussion about the quality and ethical consideration of this research.

4. Empirical Findings — • Presents the findings of the conducted interviews with AI experts, Legal experts, and Industry actors within Financial services.

5. Analysis — • Connects the findings from the literature review with the empirical findings.

6. Conclusion — • Answers the research question by presenting the main findings, as well as provides suggestions for future research.

Figure 1. Disposition of the report.

2. LITERATURE REVIEW

This chapter presents the literature relevant to this research and begins by introducing the concept of Artificial Intelligence, after which the regulatory impact on innovation is discussed, followed by a description of how the GDPR relates to Artificial Intelligence. The chapter then finishes with some concluding remarks of the literature review.

The literature relevant for answering the research question of this study includes the fields of AI technology and regulatory impact on innovation, as well as how these two fields relate to each other. Hence, the literature review will begin with an introduction to the concept of AI and thereafter presents current literature about regulatory impact on innovation. These two fields will then be discussed in relation to each other, which is visualised in Figure 2 below.
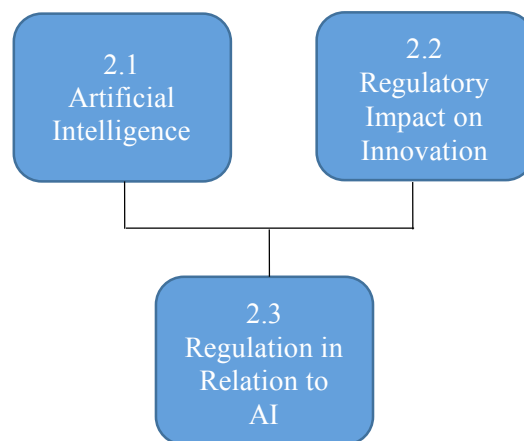
Figure 2. Structure of the literature review.

2.1 Artificial Intelligence

The concept of AI is a broad field and includes several subfields, which is visualised in Figure 3 below and will be discussed in this section. Firstly, there are different forms of AI, which in a broad sense can be divided into rule-based systems and machine learning. Subsequently, there are many different kinds of machine learning, which are classified according to the kind of model that the system is based on, that is Artificial Neural Networks (ANNs) or statistical machine learning. Furthermore, independent of the kind of machine learning model, different learning techniques can be applied for the learning process, where the main ones are supervised, unsupervised and reinforcement learning.
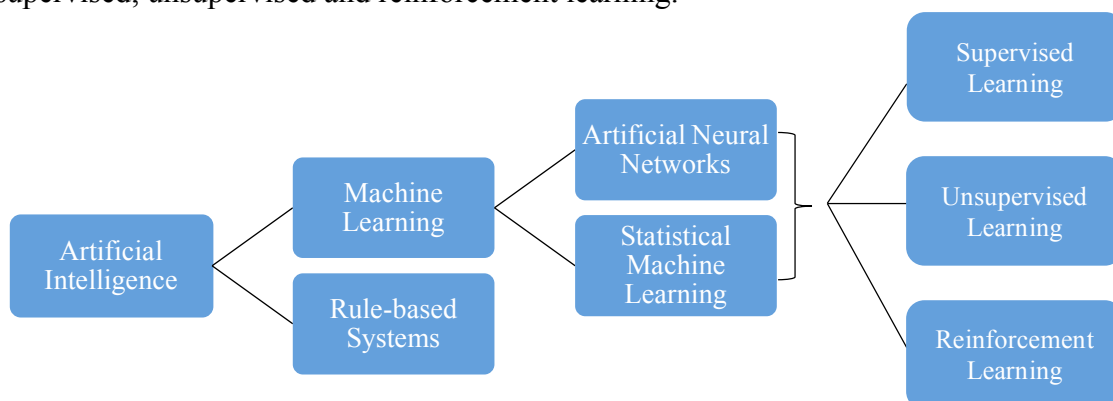
Figure 3. Overview of the concept of Artificial Intelligence.

2.1.1 Introducing the Concept of Artificial Intelligence

Artificial Intelligence (AI) is described as different technologies that enable computers to perform tasks that historically have required human intelligence (Van de Gevel & Noussair, 2012; Kaplan, 2016). To be able to perform such tasks, AI systems inhabit several different capabilities, where some of the important ones are the ability to acquire knowledge and learn, communicate in natural language, have visual abilities, as well as being able to take action such as answering questions or solving problems (Tecuci, 2012). Indeed, AI is a broad field that consists of numerous subfields, including computing, mathematics, linguistics, psychology, neuroscience, statistics, and economics (ibid.). The concept of AI has existed for a long time, but it is explained that it is first in recent years that applications of AI have become more relevant and useful (Lake et al., 2017; Tecuci, 2012; Van de Gevel & Noussair, 2012). The advancements in the field are attributed to a significant increase in computing power and storage capacity, as well as the extensive amounts of data that is available (Kaplan, 2016). It is foremost the AI systems that are capable of performing a narrow set of tasks that have been successfully applied, the approach called Narrow AI (Goertzel, 2014), which hereafter is meant when reference is made to "AI" in this report.

AI is useful to apply in a broad range of industries (Bostrom, 2014), among which financial services is one (Goertzel, 2014). For example, AI systems have surpassed human intelligence within trading in analysing large and complex quantities of transactions in a short time frame (ibid.). Thereto, complex AI systems have been used as decision support in credit evaluation, within asset portfolio management as well as to predict the behaviour of investors (Bahrammirzaee, 2010). Another area where AI has proven to be successfully applied is for fraud detection within banking systems (Gómez, Arévalo, Paredes & Nin, 2017). To that, AI is used within insurance to predict risks and price premiums, where especially machine learning is suitable due to the massive datasets that are analysed (Rouse & Spohrer, 2018).

Nonetheless, despite the significant variety of different AI models, a distinction can be made between more basic forms of AI, which are referred to as rule-based systems, and more advanced systems that inhabit an element of self-learning, which goes under the term "machine learning" (Kingston, 2017). However, these two forms are not mutually exclusive, instead, components of both forms of AI can be combined with each other into hybrid systems (Kluegl, Toepfer, Beck, Fette, & Puppe, 2016). A hybrid of rule-based and machine learning has for example been successfully used to forecast the price movement on the stock market (Chiang, Enke, Wu, & Wang., 2016).

2.1.2 Rule-based Systems

In the traditional approach to AI, humans predetermine and program the rules for what decisions the AI system should take in different situations (Mittelstadt et al., 2016), that is so-called rule-based systems (Kingston, 2017). Rule-based systems are still used in a wide range of applications (ibid.), such as for information extraction of unstructured and textual data (Kluegl et al., 2016). Furthermore, these systems have also been applied to assess mortgage applications to determine the risk that an applicant will default on a loan (Kingston, 2017).

Even though the rules programmed into the system could be based on policy or regulation documents, it is most commonly humans that have experience in the area that determine the rules (Kingston, 2017). Thereby, there is a high demand for highly educated employees to ensure high quality of the applications, and to that, the process of writing and defining the rules is very time-consuming (Kluegl et al., 2016). Hence, this requirement of qualified engineers results in an expensive process of developing rule-based systems, and therefore there exists an economic interest to develop systems that are cheaper and faster to develop (ibid.). On the other hand, rule-based systems have some advantages over machine learning systems, in that rule-based systems in some situations are more suitable to apply than self-learning systems due to the limited availability of example data to train these models (ibid.). Thereto, these systems are easy to understand, and thereby it is also possible to trace how rule-based systems have made decisions (Kingston, 2017; Kluegl et al., 2016). Nonetheless, rather than being based on predetermined rules, AI systems are increasingly coming to rely on machine learning (Mittelstadt et al., 2016).

2.1.3 Machine Learning
Machine learning models are AI systems that can improve performance through its self-learning capability (Mittelstadt et al., 2016). Machine learning methods and techniques can use data to find new patterns and knowledge, and subsequently create models that can be used to make predictions about analysed data (ibid.). The algorithms in these artificially intelligent systems have the capability to autonomously define or modify decision-making rules (ibid.). Hence, the main difference between machine learning and rule-based systems is that machine learning systems can learn on its own independent of the human designer (Kluegl et al., 2016). Consequently, this also means that it is not necessary for the human developer to understand how the algorithm operates and takes decisions (Mittelstadt et al., 2016). In turn, this means that learning algorithms include some level of uncertainty about how and why decisions are made (ibid.). Machine learning algorithms have been applied in a broad spectrum of situations, ranging from identifying objects in images, transcribing speech into text, and matching new products with users' interests (LeCun, Bengio & Hinton, 2015), and machine learning algorithms have shown notably higher performance than more predictive and simpler models (Bohanec et al., 2017). Nonetheless, many different machine learning systems exist, both in regards to the kind of model that underlies the system as well as the technique for how the system learn, which will be described in greater detail in the following two sections.

2.1.3.1 Learning Models
There are several kinds of machine learning models, but two of these have especially seen recent advances and contributed to the rapid progress within the field of AI; Artificial Neural Networks (ANNs) and learning models that are based on statistics (Ghahramani, 2015; Lake et al., 2017), hereafter referred to as statistical machine learning.

2.1.3.1.1 Artificial Neural Networks (ANNs)
Artificial neural networks (ANNs) are adaptive information processing systems (Bahrammirzaee 2010), and have been developed with the human brain and biological neural

networks as inspiration (Lake et al., 2017). ANNs consists of processing units with many interconnected layers (ibid.). In these models, little engineering by hand is required, and thereby the machines can make use of the increasingly available data and computational power that exists today, and therefore it is predicted that models based on ANNs will continue to be increasingly used within AI (LeCun et al., 2015). For financial applications, ANNs have proven to be superior to those of traditional methods, such as regression analysis, and are advantageous for solving complicated nonlinear problems (Bahrammirzaee, 2010). Consequently, these models are also useful to apply to unstructured data (ibid.), such as text analytics (Kluegl et al., 2016). However, one significant drawback of ANNs is that they are highly difficult to understand (Ghahramani, 2015).

2.1.3.1.2 Statistical Machine Learning

An alternative to ANNs is to build machine learning models that are statistically based, which for example includes probabilistic models that account for risk (Ghahramani, 2015; Kluegl et al., 2016). Hence, such models enable aspects of uncertainty to be included and are therefore advantageous to apply to problems where uncertainty is an essential element, such as in forecasting or when data is limited (Ghahramani, 2015). In contrast to ANNs, statistical based machine learning models can learn from fewer examples of data (Lake, Salakhutdinov, & Tenenbaum, 2015), as well as being conceptually simpler, and thereby often easier to understand the models' behaviour (Ghahramani, 2015).

2.1.3.2 Learning Techniques

For machine learning models, including both ANNs and statistical machine learning, different techniques can be used for the model to learn (Lake et al., 2017). These techniques can be classified into supervised, unsupervised and reinforcement learning (Sathya & Abraham, 2013).

The most common learning technique in machine learning is supervised learning (LeCun et al., 2015), where a supervisor feeds the learning model with a set of data that has been labelled and assigned correct classifications by humans (Sathya & Abraham, 2013). Supervised learning can be applied to solve both linear and non-linear problems, and is an efficient tool to use in for example forecasting and predictions (ibid.). Although, a prerequisite for using supervised learning is that there are example data available to train the model (Littman, 2015).

However, it is also possible for machine learning models to learn without labelled data, a learning technique called unsupervised learning (Mittelstadt et al., 2016). When unsupervised learning is used, the network within the AI system organises information and searches for patterns by itself, without instructions from a human supervisor (Sathya & Abraham, 2013). It does so by defining models that fit the identified patterns the best (Mittelstadt et al., 2016). This approach is advantageous to use since it enables relationships that have not been considered beforehand to be identified (Sathya & Abraham, 2013). Thereto, this method of learning is more similar to how humans learn and is a more natural representation of neurobiological behaviour (ibid.). Unsupervised learning has therefore been useful in many

real-world applications, such as speech recognition and texts analytics (ibid.). However, since the input data is not labelled and thereby does not include any information about what the data represents, unsupervised learning is argued to be the most difficult learning technique to use (Jones, 2014).

A third approach to learning is called reinforcement learning, which is a continuous process of trial and error between the machine and its environment (Sathya & Abraham, 2013). Reinforcement learning uses feedback loops and evaluative feedback to receive information whether its decision was correct or not, and this information is then used to make adjustments and improvements in future decisions (Littman, 2015).

2.1.4 Concluding Remarks about AI

AI is described as various technologies that enable computers to perform tasks that historically have required human intelligence. This technology has been applied to various tasks, across a wide range of industries, including financial services. In a broad sense, AI systems can be divided into rule-based and machine learning systems. While the rule-based systems are straightforward with predetermined rules for decisions, machine learning systems are more complicated. However, the degree of complexity varies between different models, where ANNs are almost impossible to understand, and the models based on statistics often are possible to understand, at least to some extent. In addition to different models of machine learning, there are also different learning techniques, where the most common technique is supervised learning, in which the model is trained with labelled data. However, unsupervised and reinforcement learning techniques have certain advantages over supervised learning because these models learn without labelled and pre-classified data, which enables relationships to be found that humans have not considered.

At the same time as AI technology is becoming increasingly applied, Wachter, Mittelstadt, and Floridi (2017b) describe that this technology has become a regulatory priority within several governments during recent years, that including the EU.

2.2 Regulatory Impact on Innovation

Ranchordás (2015) states that innovation is of great importance since it stimulates long-term growth and creates competitiveness. Innovation can be defined as "*the ability to introduce new ideas into the market, translating them into socially desirable commercial or technological outcomes by using new processes, products, or services*" (Ranchordás, 2015, p. 208). Government's view towards innovation and how it should be regulated has changed during recent years, and it is now a priority by the majority of governments to stimulate innovation and economic growth (ibid.). However, regulators face increasingly complicated innovations within various technologies that "*challenge existing regulatory paradigms*" (Ranchordás, 2015, p. 201).

Even though innovation brings opportunities, it also includes both uncertainty and complexities. In turn, innovations become difficult to predict, such as how it will develop (Ranchordás, 2015). Another difficulty for regulators is that there often exists an information

asymmetry between regulators and innovators about complex technologies, and regulators also lack knowledge about the potential impact that the technology could have (ibid.). To that, it is challenging for regulators to keep up with the high pace of technological development since the regulatory process is prolonged, which results in that regulations often lag behind innovation and decrease the rate of innovation (ibid.). For example, while regulators wish to stimulate innovation, they also have the outset to minimise potentially negative effects and control risks, which could cause innovations to become deferred (ibid.).

According to Ranchordás (2015), regulations can have multifaceted effects on innovation, and can both enhance and diminish the incentive to innovate, as well as affect what point in time the innovation is launched. Furthermore, it is described that a regulation's impact on innovation is dependent on the balance between innovation-inducing and innovation-constraining elements of that specific regulation (Ashford et al., 1985; Blind, 2012; Pelkmans & Renda, 2014). Innovation-constraining elements are for example compliance costs, while innovation-inducing elements create incentives for innovation (Blind, 2012). Five characteristics of a regulation can be identified that determine the impact a specific regulation has on innovation, which are administrative burden, compliance burden, timing, flexibility, and uncertainty (Ashford et al., 1985; Pelkmans & Renda, 2014).

2.2.1 Administrative Burden
Administrative burden refers to the extent that the regulation takes time and resources away from entrepreneurial activities, and is a direct result of the information requirements imposed by regulations (Pelkmans & Renda, 2014; Poel, Marneffe, Bielen, Van Aarle, & Vereeck, 2014). In turn, such an administrative burden is disadvantageous for innovation. Thereby, it has become a policy priority to reduce administrative burdens of regulations within the EU since there exists empirical evidence that decreased administrative burden stimulates economic growth (Poel et al., 2014).

2.2.2 Compliance Burden
Compliance burden, or stringency, is the difficulty and cost companies face in conforming to new regulations with the technologies and business models the companies currently have (Ashford et al., 1985; Pelkmans & Renda, 2014). A regulation is stringent if organisations have to make notable changes to their behaviour or develop new technologies to comply with the regulation, and thereby compliance burden leads to considerable compliance cost (ibid.). This characteristic is stated to be the one that has the greatest impact on technological innovations, and may require that companies change technologies or behaviour (ibid.). Costs of complying to a regulation can have an adverse effect on competitiveness and therefore also abilities to innovate, as well as decrease the resources that can be spent for research and development (Blind, 2012). However, compliance burden can also trigger innovation since it could also enhance the incentives to invest in innovation activities or research (ibid.). For example, stringent environmental regulations have been found to trigger investments in more environmentally friendly products (ibid.). This finding is also corroborated by Pelkmans and Renda (2014), who state that very stringent rules can have a positive impact on innovation if the changes that are required to be made by the stakeholders are not too significant.

Furthermore, Blind (2012) points out that when analysing the impact that regulation has on innovation, it is crucial to differentiate between the long-term and short-term effects of regulations. In this regard, the compliance burden can initially hinder innovation, while the long-term effect becomes more diverse, and dependent on the type of regulation as well as the business environment (ibid.).

2.2.3 Timing
The characteristic timing is the timeframe that organisations are given to comply with the regulation (Ashford et al., 1985; Pelkmans & Renda, 2014). Too little time could potentially have a negative impact on innovation since the workload becomes too extensive for companies, while too much time could create too low pressure to meet set requirements and therefore decreased innovative efforts (Pelkmans & Renda, 2014). The most optimal time that is given to become compliant dependents on the specific case, but it is crucial that regulators consider the timing characteristic when developing the regulation (Ashford et al., 1985; Pelkmans & Renda, 2014).

2.2.4 Flexibility
The next characteristic is flexibility, and the more flexible a regulation is, the more it will spur innovation (Ashford et al., 1985; Pelkmans & Renda, 2014; Ranchordás, 2015). Ranchordás (2015) explains that, since innovation is characterised by uncertainty and constant changes it is not well-suited with rigid rules, and therefore the focus should be on flexibility to stimulate innovation and allow for new developments. It is further described that so-called outcome-based regulations are more flexible and stimulate more innovation as compared to prescriptive regulations (Pelkmans & Renda, 2014; Ranchordás, 2015). In contrast to outcome-based regulations, prescriptive regulations specify specific technology or material aspects that have to be fulfilled, which results in fewer opportunities to find innovative methods to comply with the regulation (ibid.).

2.2.5 Uncertainty
Finally, the level of uncertainty is also one characteristic of regulations that affect innovation, which refers to ambiguities in how to comply with a regulation (Ashford et al., 1985; Pelkmans & Renda, 2014). It is described that in some situations uncertainty can be favourable, which is the case if firms explore and test different alternatives in attempts to avoid the negative effects of regulations (ibid.). At the same time, Ranchordás (2015) explains that incentives to invest in technologies affected by the regulation could decrease if there are uncertainties in how the regulation will be enforced. Indeed, both Pelkmans and Renda (2014) and Ranchordás (2015) point out that the innovative process can be negatively affected when there exists uncertainty about a regulation, and substantial investments are required for developing innovations.

2.3 Regulation in Relation to Artificial Intelligence
The progress that has been made in analysing extensive sets of data, as well as within AI technology has contributed to increased automated decision-making (Art. WP 29). However, many of these automated systems are inscrutable and lack accountability, which has resulted

in that the ethical and social impact of such systems, including AI, has become an important issue within governments (Wachter et al., 2017a).

2.3.1 Regulating Automated Decision-making

One potential problem with automated processes is that they could include bias since conclusions often are made about individuals based on studies of a large group of people, and therefore decisions may not be representative for the individual (Goodman & Flexman, 2017; Van Otterlo, 2014). Thereby, automated decision-making could, for example, discriminate against marginalised groups in society (Mittelstadt et al., 2016). Furthermore, it could cause incorrect predictions to be made, which could lead to inaccurate evaluations about, for example, an individual's credit or insurance risk (Art. WP 29). In this regard, it becomes particularly problematic when machine learning models are used since these models are only as reliable and neutral as the data that is used to train the algorithm (Goodman & Flaxman, 2017; Mittelstadt et al., 2016). Hence, if the input data is inaccurate or includes bias, this will be reflected in the decisions (ibid.). Additionally, as the algorithm itself defines the rules for how new inputs are processed, humans have less control over the processing and therefore uncertainty arises about how and why decisions are made (Mittelstadt et al., 2016).

The new EU general data protection regulation, GDPR, is to some extent an attempt by governmental bodies to increase transparency and accountability of AI models and other automated systems (Kingston, 2017). This regulation was enforced 25th of May 2018 (Kingston, 2017), and replaces PUL, the previous Swedish regulation for data privacy from 1998 (Datainspektionen, 2018b). The GDPR aims to *"protect all EU citizens from privacy and data breaches in an increasingly data-driven world"* (EU GDPR, 2018b). The regulation contains rules for processing personal data and stipulates individuals' right to protection of their personal data (Art. 1, Art. 2, EU 2016/679). Such processing concerns manual, wholly as well as partly automated processes (Art. 1 and Art. 2, EU 2016/679). Personal data is data that can be directly or indirectly connected to a specific individual, for example names, photos, banking details, and email addresses (Art. 4, EU 2016/679; EU GDPR, 2018a).

The regulation further restricts the possibilities for organisations to process personal data by specifying certain conditions that have to exist for the processing to be lawful, that is, specific legal grounds (Art. 6, EU 2016/679). Some of the legal grounds specified in Article 6 of the GDPR are that the data subject, an individual whose data is processed, has given his or her consent to the processing, that it is necessary to process the data to comply with other legal obligations, or that the processing is required to fulfil a contract between the company and data subject (Art. 6, EU 2016/679). Furthermore, GDPR applies to all companies that process personal data about EU members, and therefore applies to businesses located inside as well as outside EU (EU GDPR, 2018a). If organisations do not comply with the regulation, they risk a penalty up to 4 % of their annual global turnover, or 20 million Euro, depending on which is the higher amount (Art. 83, EU 2016/679).

Some parts of the GDPR specifically concerns the use of AI within organisations (Villaronga et al., 2017; Wachter et al., 2017a), which is the focus of this report. One of these parts

regards information to be provided to data subjects, a right to access personal data and automated decision-making, which is specified in Article 13-15 and 22 (Wachter et al., 2017). Together, these articles are commonly argued to form a "right to explanation" (ibid.). The second part of the GDPR that concerns AI is the "Right to erasure" in Article 17 (Villaronga et al., 2017).

Table 1 below presents a short description of the articles that are of focus in this report.

Table 1. The GDPR articles that are of focus in this report.

	Article	Title	Description
Right to explanation	13	Information to be provided where personal data is collected from the data subject.	A right for data subjects to receive information about the personal data that organisations have collected about them. This includes automated processing referred to in Article 22, in which case information also has to be provided about the logic of the process.
	14	Information to be provided where personal data has not been obtained from the data subject.	Similar to Article 13, with the difference that the personal data has been collected from other sources than the data subject itself.
	15	Right of access by the data subject.	The data subject has the right to get access to their personal data that is being processed by an organisation.
	22	Automated individual decision-making, including profiling.	The data subject has a right to not be subject to a decision based solely on automated processing. There are some exceptions to this rule, for example consent from the individual.
Right to erasure	17	Right to erasure ('right to be forgotten').	The data subject has the right to have their personal data erased from the organisation, when certain conditions apply, such as that the purpose for which it was collected is no longer viable.

2.3.1.1 The Alleged Right to Explanation of Automated Decisions

Article 13-15 and 22 have received significant attention as there is a common understanding that these articles infer a right for data subjects to receive an explanation from organisations about decisions taken by fully automated means, including processes based on AI (Wachter et al., 2017a). It is described that this right is an attempt by governments to increase accountability and transparency in automated systems such as AI (Wachter et al., 2017a). However, many concerns have been raised in regards to this right since there are ambiguities about how the requirements should be interpreted, and thereto restrictive formulation in the articles (ibid.). It is even questioned what protection the GDPR actually will provide to individuals since it is argued that *"the GDPR lacks precise language as well as explicit and well-defined rights and safeguards against automated decision-making, and therefore runs*

the risk of being toothless" (Wachter et al., 2017a, p. 1). One part of the concerns stems from the formulation in the Articles 13-15, and then additional questions are raised about the formulation of Article 22.

Regarding Articles 13-15, these stipulate that data subjects have a right to receive certain information when personal data relating to them are collected by organisations. Such information is, for example, whether their personal data is processed, the purpose of the processing, and thereto they have a right to access that data (Art. 13-15, EU 2016/679). More specifically when automated decision-making as referred to in Article 22 takes place, the data controller (the organisation that processes the data) has to provide the data subject with additional information (Art. 13-15, EU 2016/679). It is specified that in the occurrence of automated decision-making, the data controller has to provide the data subject with "*meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subjec*t" (Art. 13 (2f), Art. 14 (2g), Art. 15 (1h), EU 2016/679). In turn, it is clarified in Article 22 (EU 2016/679) that such automated decision-making refers to situations when decisions are based on *solely* automated processing.

It is argued that explaining automated decisions for individuals could be difficult technically when AI models are involved (Wachter et al., 2017a). According to Kingston (2017), the information requirement is difficult to comply with if machine learning models are used since the complex structure makes its decision process difficult for humans to understand. Thereto, machine learning models have limited capability to itself provide information about its reasoning (ibid.). On the contrary, with rule-based AI systems it is significantly easier to fulfil the requirements since humans know what the rules are and accompanying consequences can be derived, and thereby information about what determined the outcome of these decisions can easily be described (Bohanec et al., 2017; Kingston, 2017).

However, concerns have been raised about how this explanatory requirement in the GDPR should be interpreted in how extensive information that has to be provided, as well as what point in time the information should be provided (Kingston, 2017; Malgieri and Commandé, 2017; Wachter et al. 2017a). In this regard, Wachter et al. (2017a) argue that to comply with Article 13 to 15 it is only mandated to provide meaningful but limited information, and that it is only explicitly required to inform about the functionality of the solely automated decision-making before a decision is made, and thereby not about specific decisions (ibid.). In contrast, Malgieri and Commandé (2017) argue that it does exist a legal obligation to provide an explanation about the rationale behind particular decisions after a solely automated decision has been made. In addition to these different standing points, there are parts of the GDPR, called recitals, that provides clarifications of how the articles should be interpreted (EU 2016/679). Specifically for Article 22, Recital 71 specifies that the data subject in the occurrence of *solely* automated decision-making indeed has a right *"to obtain an explanation of the decision reached after such assessment"* (Recital 71, EU 2016/679). At the same time, Wachter et al. (2017a) argue that recitals are not legally binding and that this is one reason to question that there exists a right to explanation since Recital 71 is the only part of GDPR that

explicitly mentions this right. Anyhow, Kingston (2017) means that what can be derived from the content of the GDPR is that data subjects should be provided with enough information to be able to contest the decision.

Furthermore, in addition to the ambiguities about the interpretation of Articles 13-15, further concerns are raised about the formulation in Article 22 (Wachter et al., 2017a). This Article states that the data subject should not "*be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her*" (Art. 22 (1), EU 2016/679). This wording results in three different ambiguities: what is considered to be a "solely" automated process, what situations that are considered to produce "legal or similarly significant effects", and whether the content of this Article should be interpreted as prohibition or a right for individuals to object. This will be discussed in the following paragraphs.

Firstly, in regards to the interpretation of a "solely automated process", Wachter et al. (2017a) argue that only a low level of human involvement will suffice for the processing *not* to be considered *solely* automated. With this interpretation, the requirement to provide an explanation about decisions would only apply to a very limited range of decisions (ibid.). On the contrary, others mean that the term "solely automated processing" should be interpreted more extensively, an argument that is based on the guidelines provided by the Article 29 Working Party (Malgieri & Commandé, 2017). These authors argue that the human involvement must have an actual effect and that a human must conduct meaningful oversight for the process *not* to be considered as solely automated (Malgieri & Commandé, 2017).

Nevertheless, regardless of what is considered a solely automated process, there are exceptions to the right in Article 22, and therefore some situations when organisations are allowed to conduct solely automated decisions. For example, solely automated decision-making is allowed if the decision "*is necessary for entering into, or performance of, a contract between the data subject and a data controller or is based on the data subject's explicit consent*" (Art. 22 (2), EU 2016/679). Yet, in these situations there are still some requirements for companies to fulfil, and certain actions have to be taken since it is stated that "*the data controller shall implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision*" (Art. 22 (3), EU 2016/679).

Secondly, concerns are raised about the interpretation of "legal or similar significant effects", since it is not explicitly defined what is meant with "significant" (Wachter et al., 2017a). Nonetheless, a few examples are provided in the recital that is related to Article 22, such as credit applications and e-recruiting practices without human intervention (Recital 71, EU 2016/679). Wachter et al. (2017a) further describe that, depending on how "significant" is interpreted, the burden might come to fall on the data subject to prove that the processing of their data affects them *significantly*.

Thirdly, it is argued that it is ambiguous whether the content of Article 22 should be interpreted as a "prohibition of *solely* automated decision-making" or as "a right for individuals to object" (Wachter et al., 2017a). This is of importance since if Article 22 is interpreted as a prohibition organisations have to establish a legal ground to be allowed to conduct solely automated decision-making. On the other hand, if Article 22 instead is interpreted as a right to object, solely automated decision-making will only be restricted if the subject actually objects (ibid.). Consequently, if it is interpreted as a prohibition, greater protection will be provided to data subjects, but if it is interpreted as a right to object it instead places a burden on the data subject since they have object (ibid.).

2.3.1.2 The Right to Erasure

It has been argued that the requirement of data erasure, which infers a right for individuals to be "forgotten", in Article 17 is inconsistent with AI technology, and that it does not reflect the complexity of this technology (Villaronga et al., 2017). Article 17 paragraph 1 in the GDPR specifies that *"The data subject shall have the right to obtain from the controller the erasure of personal data concerning him or her without undue delay and the controller shall have the obligation to erase personal data without undue delay"* (Art. 17(1), EU 2016/679). This means that organisations are required to erase personal data if it is requested by the individual (Kieselmann et al., 2016). Such a requirement can come to be enforced when for example the subjects withdraw their consent or if the original reason to why the data was processed no longer applies (Art. 17, EU 2016/679). Villaronga et al. (2017) claim that this right is problematic in relation to AI and argue that the regulation has been formulated based on an understanding of how humans process and remember information, and not considered how memory and "forgetting" function in machines. It is explained that only because information is deleted, it does not necessarily mean that the machine forgets the information, as a human mind would when this information is no longer available to access (ibid.). The reason for this is that when data is deleted, it is not deleted instantly, rather, it is only when the deleted space is reused again that the old data is destroyed, which might not be until after some time (ibid.) Villaronga et al. (2017) further describe that deletion is particularly problematic within machine learning due to the large datasets that are used in the training process, and that data is continuously allocated and deleted (ibid.).

Furthermore, Villaronga et al. (2017) describe that the problem is that "deletion" can have several different meanings in AI systems, such as overwriting in file systems, erasing from backups or as extensive as deletion from all internal mechanisms. With this wording of Article 17, it is not explicitly defined what kind of erasure that would be sufficient to comply with the regulation, and it is questioned if organisations using AI systems even will be able to comply to GDPR's erasure requirement (ibid.). The authors conclude that the problem with the right to erasure in relation to AI can be summarised as *"Humans forget, but machines remember"* (Villaronga et al., 2017, p. 19).

2.3.2 Summary and Implications of the Alleged Right to Explanation and The Right to Erasure

Taken together, it is argued that there exists a gap between legislation and the current technical possibilities (Kieselmann et al., 2016), and a greater understanding between lawyers and computer scientists is sought after to ensure that GDPR is compatible with new technologies (Villaronga et al., 2017). This is of importance in an economic perspective as well since a regulation that has low real-life applicability may impact innovativeness, and therefore also affect competitive advantages between the EU and other countries with less restrictive data protection regulations (ibid.). At the same time, Goodman and Flaxman (2017) point out that the GDPR also creates an opportunity for computer scientists to design algorithms that are more transparent, comprehensible, and less discriminating, even though it also could lead to significant challenges for AI.

Table 2 below is an extension of Table 1 that was introduced in section 2.3.1 as an overview of the GDPR Articles that are discussed in this report. In Table 2, a summary for why these articles raises questions in regards to AI is provided. It was described in section 2.3.1.1 that Articles 13 to 15 and 22 in combination form a right to explanation, and are therefore presented together.

Table 2. Questions raised about how the GDPR affects AI.

Article	Description	Questions Raised in regards to Artificial Intelligence
13-15 and 22	Information to be provided where personal data is stored and processed as well as when automated decision making takes place.	How thorough and detailed information is required in solely automated decisions; details about particular decisions after the decision has been made or only about the functionality of the automated decision-making process? How should significant effects of automated processes be interpreted? How extensive human involvement is needed for an automated decision not to be considered solely automated? Should Article 22 be interpreted as a prohibition of automated decision-making or as a right to object?
17	The right to request that personal data is erased from organisations.	How will the term "erasure" be enforced technically? Will the erasure request be possible to comply with when AI systems are used?

Right to explanation — Articles 13-15 and 22

Right to erasure — Article 17

2.4 Concluding Remarks of the Literature Review

Previous research has found that regulations can have both positive and negative impacts on innovation and the development of new technologies. In regards to the negative aspects, the problems are that the regulatory process is slower than the technological development, and that regulators often lack knowledge about complex innovations (Ranchordás, 2015). Although, if formulated in the right way, a regulation could instead stimulate innovation.

Two parts of the GDPR are identified to be particularly problematic for AI applications, which are the right for individuals to receive an explanation to solely automated decisions and to request that their data is erased (Villaronga et al., 2017; Wachter et al., 2017a). The explanatory requirement is particularly problematic when machine learning models are used since it is difficult to explain the reasoning behind decisions compared to rule-based systems (Mittelstadt et al., 2016; Kingston, 2017). However, it is debated if there actually exists a right to explanation or not, and the extent of such a requirement (Malgieri & Commandé, 2017; Wachter et al., 2017a). The erasure obligation is problematic since it is complicated to delete data from machine learning models, and thereto it is undefined what level of deletion that is required (Villaronga et al., 2017). If the harsh version of erasure is required, it is questioned if it even will be possible for companies using machine learning to comply with the regulation (ibid.). Hence, according to the literature, the erasure requirement makes it difficult to use machine learning models in the presence of the GDPR.

Nevertheless, innovation and technologies can be affected by regulation. The impact of a specific regulation depends on the interplay between five different characteristics of the regulation; administrative burden, compliance burden, timing, flexibility and uncertainty (Ashford et al., 1985; Blind, 2012; Pelkmans & Renda, 2014; Ranchordás, 2015). These characteristics have a diverse impact on innovation, which is visualised in Table 3 below.

Table 3. Regulation characteristics.

Regulation Characteristics	Impact on Innovation
1. Administrative Burden	• High administrative burden has a negative impact on innovation.
2. Compliance Burden	• High compliance burden negatively affects innovation. • High compliance burden could also trigger incentives to invest in innovation.
3. Timing	• Both too short and too long time has a negative impact on innovation, and what the optimal time is varies.
4. Flexibility	• Outcome-based regulations (high flexibility) have a positive impact on innovation. • Prescriptive regulations (low flexibility) results in fewer opportunities to find innovative methods to comply with a regulation.
5. Uncertainty	• High uncertainty can have a negative impact on innovation. • High uncertainty can also stimulate innovation if firms explore different alternatives in attempts to avoid the negative effects of a regulation.

3. METHODOLOGY

This chapter presents the methodological approach that has been applied in this research. Initially, the research strategy is described, which is followed by a presentation of the research design and research method. Finally, the chapter closes with a discussion about quality aspects of the report, as well as ethical considerations.

3.1 Research Strategy

In this study, a qualitative research strategy has been applied. One of the differences between qualitative and quantitative research is the connection between theory and method (Bryman & Bell, 2011). In qualitative research, an inductive approach is most often used, where the emphasis is on theory generation, meaning that theory is generated based on the data, rather than testing a theory by formulating a hypothesis (ibid.). As this thesis aims to examine the potential impact of the GDPR on AI applications, which is a novel area concerning a future state, this thesis has an explorative approach. To fulfil the purpose of the study, a qualitative strategy is therefore appropriate. The thesis mainly takes an inductive approach, as the research question is of explorative character and the focus is on generating theory rather than testing a theory. Thereto, an iterative approach was used where literature review, data collection and analysis of empirical data is a parallel process, allowing adaption to each part in accordance with findings in the other. However, the research has some elements of a deductive approach since it takes a stance in existing theory. Hence, the research could be said to have an abductive view of the relationship between theory and research.

Due to the novelty and relatively complicated nature of the researched subject, it is desired to attain informative and in-depth answers, why an emphasis on words and explanations is preferred over quantifiable data, which a qualitative strategy allows (Bryman & Bell, 2011). Thereto, the aim is to capture the respondents' views and opinions about the researched topic (ibid.). Therefore, it was beneficial to use a qualitative approach which allows unexpected answers from respondents, who also are allowed to develop their reasoning (ibid.). Additionally, as this is a highly technical subject that focuses on one industry, it is expected that it would be difficult to reach a large number of respondents with the appropriate level of expertise within the area. Therefore, a qualitative research strategy of interviews is motivated to use instead of a quantitative approach where more respondents would have been required (Bryman & Bell, 2011). However, one disadvantage of a qualitative approach is the lack of generalisability of its results (ibid.), which will be further discussed in section 3.5, "Quality of the Findings".

3.2 Research Design

The research design is the overall framework for the process of collection and analysis of data (Bryman & Bell, 2011). The research design of this thesis is exploratory since it examines and attempts to answer the research question about the potential impact that the GDPR will have on AI applications within financial services. This is a novel area that examines a future state since the regulation was enforced at the end of writing this report. Information was at a first phase gathered through pilot interviews with technology experts to attain a more in-depth

knowledge about AI technology and its applications. These interviews were conducted with three practitioners within the field of AI in relation to accounting, self-driving cars, and data structure. After that, secondary data was collected by conducting a systematic literature review to derive information and background about previous research within the researched topic. Subsequently, primary data collection was conducted through semi-structured interviews with legal experts and AI technology experts as well as industry actors at one concentrated period. These three groups of respondents were interviewed to get a more in-depth view of the expertise within each area of the research topic, and be able to apply the findings to the specific industry. In the analysis, a connection was made between the theory and the empirical data, which was a continuous process during the research.

3.3 Research Method

Research method refers to the technique of data collection (Bryman & Bell, 2011). In this research, a systematic literature review was conducted for the secondary data collection as well as semi-structured interviews for the primary data collection, which are described in the following section.

3.3.1 Secondary Data Collection

In the process of examining the current literature relevant to the research, a systematic literature review was performed, which is described by Bryman and Bell (2011) to be a transparent and replicable process. This approach was chosen to minimise author bias and achieve a thorough description of the literature (ibid.). First, an initial review of the research field was conducted to identify a research topic of interest, thereafter search words, inclusion and exclusion criteria were determined based on the research question and delimitations of the study, which was used in the screening process to identify articles within the scope of the research. In a second step, a more extensive systematic literature review was conducted based on specific keywords, which are presented in Appendix 1. In conducting the systematic literature review, electronic databases were used to search for relevant literature, and the databases used were Google Scholar as well as the electronic databases available on the University of Gothenburg's portal.

3.3.1.1 Critical Review of Sources

During the systematic literature review, the identified keywords were used to search through electronic databases, and the first 25 articles found were screened by their title and abstract to determine if they were within the scope of the research and followed the determined inclusion and exclusion criteria. In the literature review, most of the articles used are peer-reviewed, where one exception is an article ordered on governmental incentive by the DG Research and Innovation of the European Commission. Furthermore, law text and accompanying guiding documents about the GDPR were included, as well as books about AI to gain a deeper insight of the technology. Additionally, for the introductory section management consultancy reports were included to present an up to date overview of the research area.

In regards to the inclusion criteria, this thesis focuses on the regulation GDPR and its impact on AI applications. Therefore, the inclusion criteria used in the systematic literature review

were set to include articles and books that had one of the following topics as their primary focus; "Artificial Intelligence", "regulation and innovation" as well as "the GDPR in relation to AI". Thereto, law texts and complementary guiding documents about the GDPR were included. Only articles in English and Swedish were included. In regards to exclusion criteria, this thesis does not examine how the GDPR could affect other technologies, nor how AI is affected by other regulations. Therefore, articles about how regulations in general impact financial services were excluded since GDPR is not a regulation specifically for the financial services industry, and academic articles that discussed general aspects of the GDPR and not explicitly focused on the GDPR in relation to AI were excluded, as well as how regulations have been found to affect specific industries.

3.3.1.2 Snowballing

As the research topic is a novel area for the authors, there was a risk that relevant keywords were missed in the systematic literature review. Nonetheless, by reviewing citations within the articles found in the systematic literature review, a more thorough search for literature that is relevant to the research topic was conducted (Bryman & Bell, 2011). The snowballing method was a continuous process along with the systematic literature review to increase the likelihood that the most relevant literature was included.

3.3.2 Primary Data Collection

To collect empirical data, several interviews of semi-structured character were conducted, meaning that a list of specified themes as guidance was used, while the respondent still had a great deal of freedom to answer how he or she wished (Bryman & Bell, 2011). This structure was beneficial for this research since it allowed for flexibility during the interviews and thereby enabled interesting themes that emerge to be followed up, as opposed to a structured interview that follows a predetermined sequence (ibid.). Flexibility is of importance in this study since the research question is about a future state, and therefore it is crucial to capture the opinions of the respondents, to enable useful findings. Thereto, the researchers' limited knowledge about the topic constrained their possibility to be aware of all relevant aspects beforehand. Moreover, applying a semi-structured approach in the interviews was preferred over an entirely unstructured approach since this research has a somewhat clear focus with some predetermined aspects that were desired to be discussed (Bryman & Bell, 2011).

3.3.2.1 Selection of Respondents

Interviews were conducted with lawyers, people involved with research or education of AI, and actors within the financial services industry, to collect data about the potential impact that the GDPR is likely to have on AI applications. One criterion used in the selection process was that all of the respondents were active in Sweden since this study is focused on Sweden. When selecting the respondents, the criteria were specifically determined for each of the groups of respondents. For AI experts, the criteria were that they work with or conduct research within AI. For the legal experts, the criteria were that they currently work with the GDPR and have a law degree. For the Industry actors within financial services the criteria used was that they worked within our classification of financial services, and were involved either in AI projects or GDPR projects within the company. Due to a variation in background

and position within the company, the interviewed Industry actors have different knowledge and therefore also different ability to answer each question, which in turn caused a slight variation in the comprehensiveness of the answers.

In the selection of respondents, a convenience sampling technique was used where potential respondents within the authors' network were contacted, which were approached and selected based on their expertise and willingness to partake in an interview. Additionally, a specific sampling technique of snowball sampling was applied by asking interviewees to recommend other potential respondents, which increased the likelihood of reaching knowledgeable respondents. The respondents were contacted through email. Additionally, potential respondents were also searched for through web searches and platforms such as LinkedIn by using appropriate keywords. These keywords, in Swedish and English, were: "Artificial Intelligence", "Machine learning", "Financial services", "GDPR", "Credit", "Insurance", "Bank" and "Management consulting financial services". Furthermore, to increase the likelihood that respondents would participate, an email was sent to them which explained the purpose of the study and why the findings were of interest for them as well, in line with what is recommended by Bryman and Bell (2011). If the request for an interview was accepted, the interview guide with the central themes was sent to the respondent the day prior to the interview session for them to familiarise with the topics, but not given too much time nor too much information to have formulated the answers in advance (ibid.).

Since some respondents wished to be anonymous, the decision was made to present all of the respondents anonymously in the report. To get an overview of legibility of each respondent, some information is provided in Appendix 2 about the respondents as well as the details of the interview process. Thereto, a distinction is made in regards to the age of the financial services organisations, which are divided into two groups; those established within the past two years, and those established before that time. Two years was chosen since that was when the GDPR was approved. A comprised description of respondents is presented in Table 4 below, whereas the complete table of respondents is found in Appendix 2.

Table 4. Overview of the interviewed respondents.

Respondents	Description
AI experts	Three academic experts who are researcher or professor within AI, all working at Swedish technical universities. Four business experts working with AI and have expertise in the field.
Legal experts	One legal counsel and two legal associates who currently work with GDPR.
Industry actors	Ten financial services actors that work at companies either within investment, credit evaluation, banking and insurance. Three management consultants working within financial services.

3.3.2.2 The Interview Guide

The findings from the pilot interviews together with conducted literature review formed the basis for the construction of the interview guides in the thesis. The questions in the interview guides were mainly of open character to allow respondents to answer freely and provide rich answers, which is aligned with the exploratory nature of this study (Bryman & Bell, 2011). Open questions are also beneficial when studying a subject that the researcher is not highly familiar with (ibid.), which is the case for this research. Since it is of importance to access the interviewees' individual opinions, open questions are further beneficial since respondents' answers are not lead in a particular direction (ibid.). Thereto, at the end of the interview, a "catch-all" question was asked to enable the interviewee to add information that had not yet been covered (ibid.).

Three different interview guides were used to be able to adapt questions depending on the expertise of the respondents, one for the Legal experts, one for the AI experts, and one for the financial services actors. Within each group, all respondents were asked the same questions for the results to be comparable, although the sequence varied (Bryman & Bell, 2011). The interview guides were divided into five themes; *General aspects of Artificial Intelligence, General aspects of the GDPR, Explaining automated decisions, Erasing data and the main consequences of the GDPR from an AI perspective.* However, due to limited knowledge, "general aspects of AI" was excluded from interviews with the Legal experts and "general aspects of GDPR" was excluded from interviews with the AI experts. The structure and themes of the interview guide were consistent throughout the interview process. Thereto, before the interview guide was finalised, it was reviewed by the authors' supervisor and the assisting supervisor before the first interview was conducted, to ensure that the questions were clearly formulated and easy to understand. Although initial interview guides were established and remained intact, some questions were added along the interview process that arose from interesting aspects discussed in previous interviews. The complete interview guides can be seen in Appendix 3.

3.3.2.3 The Interview Process

The interviews were performed face-to-face to the extent possible to enable observation of body language, which can inform about potential confusion or anxiety that a respondent has about a question. Interviews took place at respondents' offices for their convenience, and otherwise telephone interviews were conducted. It was also desired to perform interviews in a calm environment for the respondent to feel comfortable to express their opinions (Bryman & Bell, 2011). Both researchers were present at all interviews, where one focused on note-taking and the other on asking questions. In this way, it was possible to capture and follow up on interesting themes at the same time as relevant topics were documented and thereby eased the process of coding in the analysis. In case face-to-face interviews were not possible, due to geographical distance or other reasons, interviews were conducted over the telephone.

All interviews were recorded, which avoided the risk of note-taking disrupting the researcher's attention (Bryman & Bell, 2011), as well as avoided the risk of misinterpretations or information loss. The recording also enabled the respondents' own words to be captured

and ensured that important phrases or expressions were documented (ibid.). Despite the extensive time requirements, all interviews were also transcribed to prevail the respondent's phrasing (ibid.). The transcriptions were then also sent to the respondents for confirmation and assurance that accurate interpretations had been made.

During three of the conducted interviews, two respondents were interviewed at the same time. In such situations, there is a potential risk that the respondents' answers are affected by the other person. However, actions were taken to minimise this impact by ensuring that both respondents were involved in the discussion and were allowed time to speak, thereby enabling both to express their individual opinions. Thereto, since the interviews were transcribed, the answers provided by each respondent were analysed separately. The respondents that were interviewed at the same time were Business AI expert 2 and 3, Financial Services Actor 6 and 7, as well as Financial Services Consultants 1 and 2 (see Appendix 2). Thereto, Financial Services Actors 2 and 3 were interviewed separately, but they work at the same company. The decision was made to interview both of them since they have different roles within the company, and therefore can provide answers from different perspectives. The difference between Academic AI experts and Business AI experts is that the Academic AI experts conduct research or teach at universities about AI, while Business AI experts work with AI within businesses.

3.3.2.4 Language
When the interviews were conducted in Swedish, it was necessary to translate the respondents' answers to English. The process of translating respondents' answers is a form of interaction with the data, and therefore there is also a risk that the information presented does not fully reflect the respondent's precise words (Bryman & Bell, 2011). The translations made are to some extent dependent on the researcher and his or her knowledge and personal background (ibid.). Translations can also miss cultural and national differences, and in the interviews respondents might use words or phrases where there is no counterpart in English (ibid.). To minimise the impact of such language discrepancy, the two authors were both involved in the translation of each interview and thereby the translation was not only dependent on one single person's knowledge and experience.

3.4 Data Analysis
As mentioned above, the analysis was an iterative process where the data gathering, literature review and analysis were conducted in parallel, meaning that an initial literature review was conducted through the systematic literature review, but was revised along the process and adapted to the information that was acquired. A thematic analysis was used to analyse the data from the primary data collection, where data was coded in two steps, into concepts and subsequently into categories. Thematic analysis was advantageous to use since it is flexible to use and easy to understand compared to other analysis techniques (Nowell, Norris, White, & Moules, 2017). A thematic analysis also enables identification of central themes and comparisons between theoretical aspects and empirical data (Nowell et al., 2017).

The interviews were transcribed word for word based on the recordings from the interviews. These transcriptions were then used as a basis for the coding process. In a first step, the empirical findings were colour-coded into concepts which followed the themes of the interview guides. The transcriptions for Legal experts, AI experts and Industry actors were coded separately from each other. Following the recommendations of Bryman and Bell (2011), the coding was conducted continuously during the interview process to increase the understanding of the data. In a second step, the concepts were further coded into categories by finding connections between the data within the concepts. Thereby, the specific categories came to vary between the different groups of respondents since the focus during the interviews was adapted in accordance with the respondents' knowledge. Below in Table 5, it is exemplified how the coding process was conducted.

Table 5. Example of coding.

Empirical Findings	Concept	Related Literature
LE 1 explains that the fact that a right to an explanation is only mentioned in a recital does not matter, it still applies, and the recital only amplifies what it stated in GDPR's articles.	Explaining Automated Decisions.	Whether a right to explanation exists in the GDPR or not & Compliance burden. (Malgieri & Commandé, 2017; Wachter et al., 2017a; Pelkmans & Renda, 2014).
C1 explains that another reason that companies within financial services will not have to erase a lot of data is that banks and insurance companies are subject to several regulations that will triumph the right to erasure and therefore require that specific information is stored.	Erasing Data.	Artificial intelligence and the right to erasure & Administrative burden. (Villaronga et al., 2017; Pelkmans & Renda, 2014).

Using a thematic analysis also has its limitations. Firstly, since thematic analysis is conducted in a flexible manner, the process of identifying themes can become inconsistent and incoherent (Nowell et al., 2017). Even though the aim was to conduct the coding as consistent as possible, it was sometimes difficult since the answers from the respondents had some level of variation. To mitigate this difficulty, the transcriptions were read through by both authors and discussed before the coding process was initiated. In this way, it was ensured that both had an overview of the findings so that the coding would be as consistently performed as possible. To that, during the coding process, the authors kept a continuous communication to remain focused on the themes that had been identified. Secondly, in the coding process there is a risk that the context could be lost when data is coded and categorised together (Bryman & Bell, 2011). Nonetheless, since both authors were involved and the analysis was conducted in synthesis, this risk was mitigated as the findings were constantly questioned and an iterative approach was taken, as well as constantly reviewing the content of the empirical findings.

3.5 Quality of the Findings
There are several criteria for evaluating business and management research, among which some of the most commonly used are reliability and validity (Bryman & Bell, 2011). However, there is a discussion about the appropriateness of applying these criteria on qualitative studies as they are argued to be grounded in quantitative research ideas, and therefore it has been suggested that alternative criteria, which are adapted to qualitative research characteristics, should be used instead (ibid.). However, there exists no consensus about which criteria are most appropriate, and a lot of qualitative research is evaluated based on criteria associated with reliability and validity (ibid.). Although, some versions of reliability and validity have been adapted to better suit qualitative research (ibid.), which therefore are used to evaluate the findings of this research.

3.5.1 Reliability
Reliability is the extent that the research results can be repeated, and thereby concerns consistency in the measurement of concepts (Bryman & Bell, 2011). For qualitative research, reliability refers to the extent that proper procedures are followed in the research process (ibid.). Throughout this research process, the authors have documented all phases in great detail, which have been shared with two peers who critically reviewed the procedures and provided feedback at four different occasions, after which revisions were made in accordance to the feedback. This kind of "auditing" approach is something that enables the reliability of the findings to be improved (ibid.). The Reliability criteria can be further divided into the consistency of interpretations among the researchers, referred to as internal reliability, and the possibility of replicating the study, referred to as external reliability, which will be discussed in greater detail below.

3.5.1.1 Internal Reliability
Internal reliability refers to consistency in interpretations when two or more researchers are involved in the same study, and such consistency is essential when subjective judgments are made (Bryman & Bell, 2011). The research design of this study involved thematic analysis, in which subjective interpretations are made (Nowell et al., 2017). To ensure that the researchers were consistent in their interpretations, both participated in all interviews. The answers from the interviews were then discussed by the authors to address potential responses that were difficult to understand, or inconsistent interpretations between the researchers before transcription began. To the extent possible, each interview was transcribed continuously during the interview process to facilitate that non-verbal cues such as gestures and body-language was remembered. Coding in the analysis of the empirical data was thereto discussed continuously to establish common agreement about the categories that were developed.

3.5.1.2 External Reliability
External reliability refers to the extent that the research can be replicated, and can be improved by describing the procedures of the research method in great detail (Bryman & Bell, 2011). However, it is seldom possible to fully replicate a qualitative study due to its commonly unstructured approach, and that the researchers' preferences and characteristics influence the research. In turn, this implies that there are no standard procedures to follow

within qualitative research (ibid.). Moreover, the research question of this thesis concerns a future state, which therefore includes some level of uncertainty. As a result, the findings of this research is to a large extent speculative and based on the respondents' knowledge and opinions, implying that it would be even more difficult to replicate the findings.

3.5.2 Validity
Validity is associated with the integrity of findings in a research and can regarding qualitative research be classified into two different subgroups; Internal validity and External validity (Bryman & Bell, 2011).

3.5.2.1 Internal Validity
In qualitative research, the criteria internal reliability refers to the extent that there is a congruence between the empirical data and the theoretical ideas developed, as well as the extent that the findings have credibility, meaning whether they are believable and will be accepted by others (Bryman & Bell, 2011). To improve the internal validity of this study, all interviews were recorded and the transcriptions were made word for word. To that, the two researchers have continuously discussed the collected empirical data and the concepts derived in relation to the literature, to ensure that these were congruent.

3.5.2.2 External Validity
External validity refers to the generalisability of the study's results, whether it is possible to generalise the findings to other contexts than the specific research (Bryman & Bell, 2011). In general, the scope of the findings from qualitative studies are often restricted, and difficult to generalise to other settings beyond the research conducted (ibid.), which is also true for this research. To that, as convenience sampling was applied in this research, the generalisability of the results is limited (ibid.). However, in regards to the generalisation of qualitative findings, evaluation should be based on "*the quality of theoretical inferences that are made out of qualitative data*" (Bryman & Bell, 2011, p. 409), which means to generalise to theory rather than to populations (ibid.). Therefore, the findings of this report can be said to be somewhat generalizable to how regulation affects innovation and usage of technology, rather than generalise to a population. Nonetheless, this generalisation is limited to Sweden since this geographical delimitation was made in the research.

3.6 Ethical Considerations
When business and management research is conducted, several ethical considerations have to be made (Bryman & Bell, 2011). To ensure transparency of the research, the authors presented themselves as students from the School of business economics and law in Gothenburg when contacting respondents, and thereto described the background and purpose of this thesis. In this way, respondents were given enough information to give their informed consent to participate in the study, while avoiding the risk of disclosing too much information to influence upcoming answers during interviews (ibid.). Furthermore, after the interview the respondents were offered to receive the transcripts of the interviews with them, to allow them to confirm that their opinions had been interpreted correctly. Since some respondents wished to be anonymous, careful considerations were taken so that their identity was not revealed.

4. EMPIRICAL FINDINGS

This chapter presents the findings derived from the primary data collection, which are divided into three sections based on the different groups of respondents that were interviewed; Legal experts, AI experts and Financial services actors (Industry actors).

The three different sections are structured slightly differently, but follows the themes that were used in the interview guides. An overview of the division used are presented in Table 6 below.

Table 6. Structure of the empirical findings.

AI experts	Legal expert	Industry actors
General aspects of AI	*Not applicable*	General aspects of Artificial Intelligence applications within Financial Services
Not applicable	General aspects of the GDPR	General aspects of the GDPR within Financial Services
Explaining Automated Decisions	Explaining Automated Decisions	Explaining Automated Decisions
Erasing Data	Erasing Data	Erasing Data
The main consequences of the GDPR from an AI perspective	The main consequences of the GDPR from an AI perspective	The main consequences of the GDPR from an AI perspective

4.1 AI Experts

The AI experts have been divided into Academic experts that conduct research or teach at universities about AI, and Business experts that work with AI within businesses. The Academic experts are named AE 1, AE 2 and AE 3, and the Business experts BE 1, BE 2, BE 3 and BE 4.

4.1.1 General Aspects of Artificial Intelligence

AE 1 and AE 3 explain that Artificial *General* Intelligence (AGI) has not been achieved yet and that all AI applications that exist today are Narrow AI, which are machines that can perform a narrow set of tasks, as opposed to fully replicating human-level intelligence. Nonetheless, AE 1 clarifies that there are many different strands of AI and that the narrow kind of AI can still solve highly complicated problems. AE 1, AE 2 and AE 3 mention that one recent success within this field is a computer that managed to develop a playing strategy for the Chinese game "GO" that was new to humans and thereto performed better than humans. The game "GO" has traditionally been considered too difficult for computers since it includes heuristics and undefined rules. AE 3 expresses;

"With these advancements, people realised that AI could be applied to many things that we did not think was possible."

In regards to the different forms of AI that are used, BE 2 describes that rule-based systems can be classified within the field of AI even though the rules are handwritten by humans, although, AE 1 points out that rule-based systems are used less and less. At the same time, BE 2 mentions that both rule-based AI and machine learning applications are extensively used within for example trading. Thereto, BE 4 describes that rule-based AI and machine learning is used within companies, there among financial services, for threat intelligence and to improve IT security. BE 4 explains that financial services are particularly exposed to IT threats since it handles money, and therefore cybersecurity is a high priority for the industry.

Regarding the current state of machine learning in general, it is described that ANNs receive a lot of attention (AE 3, BE 2, and BE 3). Models based on ANNs are explained to be beneficial to use when there is extensive data at hand and when the data is complicated, such as sound, images and text (BE 2 and BE 3). However, BE 2 describes that many companies today do not use ANNs, but rather simpler models based on mathematical statistics. One difficulty of using ANNs is that they require extensive sets of data compared to simpler models (BE 3), and it is pointed out by all AI experts that these models have low transparency, and thereto are challenging to understand due to its complex structure. According to AE 1, another disadvantage of ANNs is that these models do not take uncertainty into consideration;
> *"It is crucial to have control over uncertainty in machine learning since there exist uncertainties in the real world, as well as in the data gathering process."*

Hence, it is explained that it therefore often is beneficial to use more statistically-based models, for example, probabilistic models that do account for uncertainty and are easier to understand (AE 1, BE 2 and BE 3). However, BE 2 mentions that statistical machine learning models are complicated mathematically, which has resulted in a reluctance to apply it. Nonetheless, BE 4 describes that both statistical machine learning models and ANNs are used in several parts of the process of detecting data breaches and prevent cyber attacks, for example to identify potentially fraudulent IP-addresses. At the same time, BE 2 explains that statistical machine learning models most often yield better results than ANNs within the financial services industry since there is a lot of structured table data that contains features such as salary, or historical data of who has repaid loans and who has not.

What concerns different learning techniques, AE 1 means that supervised learning has been the most successful and commonly applied machine learning technique as of today, and BE 1 points out that this is also the easiest learning technique within machine learning. AE 1 and BE 1 express that there should exist great application areas for supervised learning within the financial services industry since there are large datasets and time series data available. Similarly, BE 4 explains that supervised learning is the primary method used in their fraud detection processes, but that some unsupervised learning also can be used for preliminary examination of datasets to identify categories of breaches that they have not previously considered. However, supervised learning also requires extensive training data that is labelled, which in turn requires that there are vast amounts of example data available (AE 3). To that, it is a tedious task to label all data (AE 1 and BE 4).

Due to the difficulties of supervised learning, AE 1 describes that there are movements towards using unsupervised learning, where the machine is trained with raw data that has not been labelled by a human. AE 3 expresses that another advantage of unsupervised learning is that it requires less training data than supervised learning. However, the major problem with unsupervised learning is to extract useful information from the data since it is often difficult to know what you should be looking for (AE 1). Nevertheless, AE 1 believes unsupervised learning will be the next big paradigm of AI;

"The big holy grail of machine learning is that you need to go into systems where you can learn without a huge amount of supervision."

Moreover, AE 1 describes that there is an additional learning method called reinforcement learning, where the system has a feedback loop that instructs if the result of the performance was correct or not, meaning that there is some level of supervision but much more restricted than in supervised learning. BE 1 clarifies that in reinforcement learning the machine makes a judgment and improves with each decision it makes. Thereto, AE 1 describes that the system can learn by itself, that it does not have to be a human that informs the machine whether the right or wrong result was achieved. AE 1 exemplifies that reinforcement learning was the learning technique used in the Chinese game "Go" mentioned above. However, AE 1 further points out that reinforcement learning is difficult to use within areas where wrong decisions have severe effects, such as autonomous driving. In such situations, you do not want the adverse outcome to occur for it to be known whether it was the right or wrong action.

4.1.2 Explaining Automated Decisions

BE 2 explains that it is not given any clear answers of how the requirements in the GDPR should be met. In accordance, AE 1 and AE 3 mean that there currently is a gap between the formulation of GDPR and existing technologies, and AE 1 expresses;

"There seems to be significant confusion in the area of GDPR, particularly the right to erasure and the right to explanation. There is a legal aspect, and there are technical aspects, and there is a big disconnection between the two at the moment."

According to AE 1, there exists a spectrum of AI models with different level of interpretability. As mentioned, all AI experts agree that models based on ANNs are complicated to understand and that it is nearly impossible to explain how decisions have been made. In contrast, statistical machine learning models, such as probabilistic models, are easier to understand (BE 2, BE 3 and AE 1). Just by the nature of statistical machine learning models, it can be seen how decisions have been made, and also more clearly see which factors were essential for the model's output (AE 1, BE 2), meaning that it can be explained how the model works in a general sense (AE 2). However, it is pointed out that it is not possible to get precise information about how the system reasons (BE 2, BE 4), nor understand details (AE 2). Although, BE 2 explains that in statistical models it is possible to test how different factors affect the outcome since it is predefined in these models how data points are connected. AE 1 clarifies that it can be explicitly described what has happened with probabilistic models, which is not the case with ANNs;

"ANNs is more like a black box; you do not know what happens inside, just what the input and the output is."

It is pointed out that if something goes wrong when machine learning is used, it could have more substantial consequences within financial services than other industries since people's' money is at stake (AE 3 and BE 1). Additionally, AE 3 points out that an accountability problem could arise if, for example, an investment by an AI machine turns out badly, and if it is not possible to understand how decisions were made it is also difficult to determine who is responsible. In this regard, it is pointed out that a major research area today is to develop AI models that are easier to understand and that have a simpler structure to enable explanation of decisions, referred to as "explainable AI" (AE 1, AE 3 and BE 2). Thereto, AE 1 and BE 2 explain that there is a trade-off between the ability to explain decisions and performance, meaning that when attempts are made to make algorithms easier to understand, it has a negative effect on accuracy and speed. BE 2 concludes;

"The question is if you want to have a model where you can understand exactly how decisions are made and where the performance suffers, or a model where you cannot understand everything but the model performs better."

4.1.3 Erasing Data

AE 1 perceives the right to erasure in the GDPR to be highly unclear and means that computers cannot forget in the same way as humans and that erasure could have different meanings in computers. AE 1 further describes that the standard delete operation is to erase the link to the data, which means that the data is not completely erased since it is possible to reconstruct the link. AE 1 and AE 3 explain that to erase the data completely, it would require much more effort and cause a significant change in technology, and this would also affect the speed of the systems negatively. Hence, the difficulty and effort to erase data depends on what the practical meaning of "erasure" in the GDPR is (AE 1).

According to AE 3, erasure is complicated due to the difficulty of keeping track of the data that is stored, and therefore it is also difficult to ensure that everything has been erased. Both AE 3 and BE 2 express that back-ups are constantly made for databases and if it would be necessary to erase data from these back-ups extensive work would be required. AE 3 points out that another difficulty is that it is unclear if data is considered erased by GDPR's standards if the data can be reconstructed later. If for example a personal number is split in two and stored in different locations, it could potentially be reconstructed later, and it is diffuse if this counts as personal information according to the GDPR.

In this regard, AE 3 mentions that it is easier to be compliant to the GDPR if systems are built from scratch, and the software is designed knowing the requirements that have to be complied with, taking the so-called "privacy by design" concept. AE 3 states that to abide by the GDPR's requirements companies should take privacy and security into consideration from the very beginning, which is not standard practice today. AE 3 also expresses that the importance of accounting for security and privacy from the start of designing systems could make it easier for new companies to comply with the GDPR.

Furthermore, concerning machine learning AE 1 explains that after the machine has been trained the training data is not needed anymore. However, AE 2 points out that one reason to save the training data could be that the original model did not perform well, and therefore it can be advantageous to have the data saved to be able to retrain the system with another method. Thereto, AE 1 makes a connection to the right to an explanation and says that if you have to explain how a specific decision has been made, then you also have to go back and look at the training data. Besides, BE 3 explains that when a model has been trained with specific data, and this data is erased from the database, the behaviour that is based on the data will remain in the trained system, and to remove a behaviour you have to retrain the system with new data. However, BE 3 describes that retraining models could become problematic if it has to be done too often;

"It would be manageable to retrain a system weekly, but if the system has to be retrained instantly, it would be costly and unsustainable."

In regards to the consequences of erasing data from machine learning system, BE 2 and BE 3 express that the more data companies have access to, the better it is since it in general results in more accurate outcomes.

AE 1 concludes that there is a difference between deletion in a legal and technical sense, and in accordance, AE 3 points out that a more concrete specification is needed to comply to the erasure part of GDPR. AE 3 expresses that even though the objective of the law is not to be precise but to state principles that are open for interpretation, it becomes problematic when it affects technologies since technicians are used to the concrete definitions used within computer science. Thereto, AE 3 points out that since GDPR regulates privacy and applies to many domains, the regulation cannot be adjusted to the current technology since any technology change would require a new regulation to be developed. Moreover, since it takes years of writing and revising to develop a regulation, it makes sense to construct the regulation away from technology (AE 3). Nonetheless, AE 3 expresses that;

"I think many companies are anxious about being compliant, but the thing is that everyone has the same problem, and it will be challenging to penalise all companies."

4.1.4 The Main Consequences of the GDPR from an AI Perspective
BE 2 explains that it is costly to deal with the GDPR, and that it takes time to think about since its requirements are complicated. To that, AE 3 describes that for many of the technical problems of enforcing GDPR there are no generic solutions, but specific solutions can be found given the problem. Moreover, BE 1 points out that first of all companies need to know where data is stored, and since this is often not the case BE 1 perceives this problem to be the most significant challenge of the GDPR. Even though reconstructing data systems to become compliant with GDPR could become an extensive process, lacking insight about the content of data is also costly, and therefore reconstructing data systems could lead to benefits beyond GDPR compliance in that it saves money in a longer perspective (BE 1).

BE 1 and BE 2 believes that rule-based AI could be beneficial to use to keep track of where data is stored, for example through using a system that searches and creates alerts for data with a specific structure, for example, personal identification numbers or names. BE 1 and BE 2 express that these systems that keep track of data are particularly beneficial for large companies that often have many different databases. Also, BE 2 and AE 1 express that machine learning can be beneficial in flagging information that potentially is illegal to store according to GDPR. AE 3 mentions that the AI component Natural Language Processing (NLP) that processes language of both voices and text, is currently used for analysis of legal contracts by using rule-based NLP, but it is also stated that there are many kinds of NLP models which can be based on either rule-based models or machine learning. Moreover, BE 3 points out that since the GDPR does not differentiate between structured and unstructured data, AI and machine learning could be useful to apply for unstructured data, compared to a database where the data is structured, and it is easy to find and delete data.

According to AE 1, how the GDPR will come to affect the usage of AI will depend on how the GDPR is enforced. If harsh interpretations are made there will be extensive constraints on the usage of AI, and it would also negatively affect performance. BE 1 states that the GDPR has caused a lot of developing projects of AI to be put on hold since companies feel that they have to get their heads around the GDPR before they can continue. Thereto, BE 1 explains that the development will be explicitly hindered for AI models that have access to personal information, projects will be put on hold until companies know what they will be able to do with the GDPR in force. However, BE 1 highlights that this postponement of projects will be short-lived, and it will only persist until EU court has processed the first GDPR cases, and it thereby becomes more clear what companies are allowed to do and not. Although, in a broader sense, both BE 1 and BE 3 believe that the research within machine learning will continue, and thereby the development will persist, even though some applications may become delayed by some years. Thereto, BE 3 and AE 3 state that less data collected and stored may result in fewer resources to develop machine learning. BE 1, BE 3 and AE 3 express that now with the GDPR companies will have to be more careful and restrictive in how they collect data.

"Now you have to be careful about what you collect, before you just collected as much as you could, and you tried to decide later what to do with the data" (AE 3).

4.2 Legal Experts

The interviewed Legal experts are named LE 1, LE 2, and LE 3.

4.2.1 General Aspects of the GDPR

LE 1 describes that in contrast to what many believe, a lot of the content in the GDPR already exist in the previous personal data legislation "PUL", such as the right for individuals to be informed about automated decision-making and to request that their personal data is erased. LE 1 means that the great attention that the GDPR has received rather can be attributed to the new sanction possibilities:

> *"It is commonly expressed that the GDPR is like PUL with teeth."*

4.2.1.1 Administrative Burden and Compliance Burden

All Legal experts explain that since many of the requirements in the GDPR already exist in the current regulation for personal data (PUL), the administrative burden that the GDPR imposes depends on the extent companies have been compliant with PUL previously. According to both LE 2 and LE 3's knowledge, there has been limited compliance with PUL among Swedish companies, and therefore many companies have to build systems from scratch to comply with the GDPR, which is a significant burden. LE 3 believes that one reason to the previously lacking compliance to PUL is that organisations have not fully understood the importance of personal data protection, and the importance EU places on personal data integrity. In line with this, LE 1 explains that companies previously have collected and stored "good-to-have data", without considering that it actually is personal data.

In accordance, LE 2 states that one problem often is that there are many kinds of data that people are not aware of actually is personal data, and therefore unknowingly store personal data in their systems, such as IP-addresses. Indeed, LE 2 and LE 3 point out that it is a lot of information that count as personal data, such as a name, a picture, employment number, or encrypted information that you can access with a key. Thereto, voices are most often classified as personal data, especially if it is possible to identify the person if information such as timestamps, location, or other similar factors is available (LE 1). Nonetheless, LE 1 means that;

> *"The greatest challenge of the GDPR is to combine law with every other area of the business since the regulation requires that everyone in the organisation conform to the regulation."*

Furthermore, even though many similarities exist, LE 3 explains that there are some changes in the GDPR compared to PUL. The main changes are improved rights for individuals and increased focus on security, such as the principle to incorporate privacy already in the system design (privacy by design), as well as that unstructured data is included in the regulation. LE 2 and LE 3 explain that Sweden excluded unstructured data from the previous legislation PUL, but this exception is not made in the GDPR, and thereby information in for example emails, pictures and videos will no longer be excluded from the definition of personal data. LE 3 expresses that many companies have years of email saved in inboxes, which now have to be screened for personal data. To that, LE 3 clarifies that information that on its own is not personal data could still be classified as personal data if it can be used to identify a specific

individual when it is combined with other pieces of information, such as age and street address. However, LE 3 points out that it is not defined in the GDPR how many pieces of information that would have to be combined for it to be considered personal data. Similarly, LE 1 mentions that anonymising data can be one way to prevail privacy of personal data, but that it is possible to de-anonymize information if the right tools are at hand. Therefore, it is unclear whether "Datainspektionen" would consider this to be a sufficient security precaution (LE 1).

In regards to the necessary changes to become GDPR compliant, LE 1 explains that the GDPR places stringent requirements on companies to ensure privacy standards, meaning that companies may need to conduct a significant transformation if they do not already have routines in place for personal data treatment, since for example, new IT solutions may have to be implemented. Thereto, LE 1 continues to describe that to be able to comply with the regulation, organisations have to keep track of the personal data they have stored, and to that, it is beneficial to document how long the information is expected to be used. According to LE 3, the critical part is to have control over the personal data stored, and to ensure that the data is accurate and that data that should no longer be stored is screened out in time.

All Legal experts agree that the time that personal data is allowed to be stored depends on the purpose, that personal data can be stored as long as necessary for the given purpose, and after that it should be erased. Thereto, the same information could be stored for different purposes and thereby also for different lengths of time (LE 2). One example is that data has to be stored for seven years to comply with the Swedish accounting law "Bokföringslagen", but if there does not exist any other purpose for storing the data during this period, companies cannot use the data for other purposes, such as direct marketing (LE 2). In this context, LE 1 also points out that the evidence burden according to the GDPR always is on the organisation, meaning that as an initial rule it is not allowed to store personal data, and then as a second rule companies are allowed to store personal data if they have a legal ground for it, such as consent from the data subject. LE 3 adds that other legal grounds to process data exists, such as a necessity for the performance of a contract with the data subject or to fulfil another legal obligation. Additionally, in regards of automated decision-making, LE 3 points out that companies still have to inform individuals that automated processing takes place and that they have a right to object according to Article 13, even though there is a legitimate legal ground.

4.2.1.2 Timeframe for Implementing the GDPR

All Legal experts agree that there has been enough time for organisations to become GDPR compliant until its enforcement since the proposition was approved as early as 2016. If companies do not become compliant in time, the reason is explained to be that they have not started the process soon enough (all Legal experts). However, LE 1 point out that it could be challenging to become compliant in time due to uncertainty about how the regulatory oversight body, "Datainspektionen", will enforce the different requirements, meaning that there will be limited time to know what specific changes that have to be made.

4.2.1.3 Flexibility

LE 2 explains that EU-regulations, in general, are unclear and target-focused. LE 2 further states that GDPR strives to be neutral in technologies, and also to be industry neutral. Since the GDPR covers such a wide range of different sectors and organisations its' formulation has to be general, which also enables authorities to make judgments about individual cases and industries (LE 1). However, LE 3 explains that the legal part of GDPR has low flexibility, in that, for example, specific documentation and registries have to be in place, although the exact details of how the registries should be constructed are not specified. LE 3 clarifies;

> *"Regarding the technical aspects, such as privacy by design, there are several ways to achieve the appropriate level of security, and how you build your organisation to have as strong security standards as possible."*

LE 2 and 3 clarify that the regulation is quite open to interpretation in regards to safety measures and how an appropriate level of protection can be achieved. LE 2 also states that if companies are not capable of making technical safety solutions, they could take organisational security measures, such as implementing a routine for associates about how they are allowed to use for example Outlook. LE 1 clarifies that the overall requirements of the GDPR are that companies should take reasonable safety measures, and exemplifies that one such action could be to anonymise personal data, although it has to be ensured that enough measures have been taken to obstruct de-anonymization of the data. LE 2 further describes that nothing is explicitly targeted towards financial services since the GDPR is industry neutral, but expresses that financial services is heavily regulated from other laws, and therefore companies are restricted in what they can do.

4.2.1.4 Uncertainty

As described above, LE 1 and LE 3 describe that it is still unclear how some parts of the GDPR should be interpreted, such as what is considered to be sufficient effort to inform an individual about the processing of their data (LE 1). In turn, this lack of clarity causes some ambiguity since there is no established practice of the newly added contents in comparison to PUL (LE 1 and LE 3). Nonetheless, LE 1 explains that the knowledge about the GDPR continually increases, and LE 2 mentions that a board will be instituted at the enforcement date which will provide further guidance, but that it is currently not known what the content of these guidelines will be. However, all Legal experts point out that the Article 29 Working Party has released guidelines about how the content of the GDPR should be interpreted. LE 2 explains that Article 22 about automated decision-making is one of the areas where guidance is provided, and clarifies what level of human involvement that is required to *not* be considered a *solely* automated decision;

> *"It is not enough that a human oversees the automated process; instead this person must consider the whole picture of the case, as well as being capable of changing the decision taken, and also have authority to do so."*

Furthermore, LE 1 adds that the word "solely" in the formulation of Article 22 is of importance since many companies use automated processes, such as credit applications and online recruitment, and by including "solely", companies are offered an opportunity to, at

least in a first step, use automated processes. Thereto, regarding another wording in Article 22, LE 1 exemplifies that "significant effect" could be a situation where a fully automated process is used to decline applications within e-recruiting, whereas a situation that does *not* significantly affect an individual could be an automated process to screen for incomplete applications.

Moreover, all Legal experts agree that Article 22 should be interpreted as a "prohibition", that individuals do *not* have to object for Article 22 to apply. LE 2 explains that it even is specified in the guidelines by the Article 29 Working Party that the subject does not have to claim their right and make an objection, instead it is the company that has to ensure that they have legal ground for the processing. Although, LE 3 emphasise that the individual still has a right to object to the automated processing. Furthermore, LE 1 means that it can also be derived from the content of the regulation that the burden is on the company;

> *"If you read the content of Article 22 together with the information obligation, it can be derived that it is the companies that have to act first and inform the individual that they are doing an automated treatment, for which they must have a legal ground to do."*

4.2.2 Explaining Automated Decisions

According to LE 1, there have been many questions in regards to the right to receive information and automated data processing. In regards of *when* individuals should be informed in the occurrence of *solely* automated decision-making, LE 2 means that the information that such processing takes place should be provided in close proximity to when the personal data is collected. However, LE 2 highlights that this would not count as consent, but instead merely a proof that the company has fulfilled this information requirement. Thereby, companies also need to ensure that a legal ground is established for the processing, which for example could be an agreement specifying that the company has to conduct such processing if they are to be able to grant a loan to the applicant (LE 2). LE 1 describes that the information should be provided before the processing, and also during the process if any changes occur. LE 1 explains that this is important as individuals have to know in what part of the process that the automated decision has been made to determine whether this affected the decision that was taken or not, and thereafter be able to contest the decision. At the same time, LE 3 means that information also has to be provided after a decision has been made for the individual to be able to contest the decision.

Regarding the extent of information and explanation that companies have to provide, LE 1 means that the information requirement for *solely* automated decisions is very extensive. To fulfil the purpose of enabling individuals to contest decisions, LE 1 emphasises that detailed information about a specific decision has to be provided. Thereto, LE 1 further explains that the information requirement does not become less extensive simply because a right to explanation only is mentioned in a recital, and that this applies anyway. In accordance, LE 2 and LE 3 clarifies that recitals in the GDPR can be seen as binding since their purpose is to add further explanations to the articles and how the content is intended to be interpreted. However, this extensive information requirement could also bring some positive aspects for companies since it could decrease the risk of errors in the automated systems (LE 1).

Nevertheless, LE 2 points out that there is some level of reasonableness to the information requirement and that it is not required to disclose how the algorithm works in great detail:

"The individual has a right to understand how the decision is made and what aspects that are taken into consideration, but you do not have to write five pages about it."

4.2.3 Erasing Data

In regards to the right for individuals to request that personal data is erased, LE 1 once again points out that this right has existed previously, but explains that many individuals may not have been aware of this right. Thereby, it is likely that customers will want to test the system and companies might see an increase of erasure requests initially, but this will most probably level out with time (LE 1). However, LE 2 explains that the right to erasure is quite hollow and that it does not apply unconditionally, and to that companies can often find legal grounds for not erasing the data. Hence, LE 2 concludes that the right to erasure probably will not have a significant impact after all. LE 3 clarifies that the banking and insurance industry in particular is heavily regulated and subject to many other regulations that triumph the GDPR, such as regulations for money laundering, which require organisations to document and store specific information.

4.2.4 The Main Consequences of the GDPR from an AI Perspective

LE 1 believes that the GDPR will create increased awareness among companies, and also that there will be increased inquiries from the public regarding the content of the GDPR, which thereby imposes a higher workload on many businesses. Thereto, LE 3 describes that there is a trend in society where people are becoming increasingly concerned with how their personal information is used, and LE 3 means that the GDPR reinforces such perceptions. In accordance, LE 2 perceives that it lately has been an interesting development for companies with a strong connection to the Swedish market;

"Swedes have always been very integrity-extrovert and have not cared if data is collected about them, but today we see a quite strong regression of this attitude, in that people are becoming increasingly aware of one's rights, and what companies can do and not do."

Moreover, in regards to AI, LE 3 means that in one way it can be said that the GDPR could impede the development on AI, but also states that on the other hand all technologies are regulated in hindsight. LE 3 concludes that it takes a long time to develop regulations and that it additionally is challenging to know what a technology will look like in the future. EU-regulators have attempted to be forward-looking with the GDPR, in that for example automated decision-making is believed to increase in the future. However, the Internet had recently been established in 1995 when the old directive was enforced, and PUL was not adapted to the extensive processing of personal data online, and thereby it is possible that the regulators have been aiming in the wrong direction once again with the GDPR (LE 3).

4.3 Industry Actors

The majority of the financial services actors are named FS 1 to FS 10 respectively, whereas the management consultants have the abbreviations C 1, C 2 and C 3.

4.3.1 General Aspects of Artificial Intelligence Applications within Financial Services

Talking to the respondents, it is soon established that there is no clear or shared understanding of how AI should be defined. Nonetheless, it is described that rule-based AI is widely used within financial services, both within banks and insurance companies (C1), and it is pointed out by the banks that they have many automated processes based on rule-based AI that have replaced human administrators (FS 4, FS 5, FS 8, FS 9). In regards to machine learning, several respondents express that the use of machine learning within financial services is limited, and mean that there is more talk than action when it comes to machine learning applications (FS 3, FS 6, FS 9 and C2). Although, FS 10 expresses that companies within the industry are currently working on implementing more machine learning solutions. FS 3 states that there is excellent potential for machine learning where there is a lot of data, and for this reason, FS 1, FS 3 and FS 6 mean that machine learning is beneficial within the banking and insurance industry. FS 6 expresses;

> *"The insurance industry is ideally suited for machine learning solutions since there is a lot of historical data and many manual tasks that could be automated."*

FS 10 states that most companies use rule-based AI systems since they are easier to implement. However, some machine learning applications are in place, but all respondents agree that AI models based on ANNs are rare in this industry. However, FS 10 points out that ANNs are becoming increasingly popular, mainly to analyse large datasets. The different areas where AI applications are used within the financial services industry are further discussed below.

4.3.1.1 Credit Evaluation

One common application where companies within financial services apply AI is credit approval processes. It is described that AI is used to calculate credit-worthiness of individuals (FS 2, FS 3, FS 5, FS 10 and C2), and that one bank is looking into machine learning applications in anti-fraud processes to for example identify fraudulent online applications for loans (FS 5). FS 8 expresses that AI can be advantageous to use for credit evaluations since it looks at data objectively and minimises bias. Thereto, FS 2 and FS 8 state that it is crucial that credit assessments are based on clear parameters about the creditworthiness of a person, rather than just relying on human judgement and gut feeling.

At the bank FS 5 works at, statistical models are used to estimate the probability of repaying a loan, and machine learning is a part of the systems used for making predictions. The company FS 2 and FS 3 work at uses machine learning to define credit risk by estimating the likelihood that customers will default on loans based on historical data. The data analysed include the characteristics of people who previously have paid back loans, including individual factors such as salary and address. FS 2 explains that it is beneficial to use machine learning since it allows analysis of larger datasets and enables companies to make better decisions. Although,

FS 2 and C3 state that there is still a lot of manual processing and only a limited part of the credit approval process that is automated currently, and mean that AI is only used as decision-support to specify the risk and what the recommended credit is. However, C3 adds that some banks use fully automated processes for smaller loan applications.

In regards to different kinds of machine learning models, FS 2 and FS 5 explain that in what they do, they have better use of simpler statistical machine learning models that are easier to understand than ANNs, which is vital within for example credit evaluation. FS 2 and FS 10 express that they have tested to use machine learning models with ANNs, but it did not yield better results than the current statistical based models. FS 3 mentions that they do use ANNs in some parts related to other services than credit decisions, such as categorisation of bank transactions. FS 2 further describes that they are working with supervised learning, which is more suitable to use than unsupervised learning in their applications since they work with structured table data.

4.3.1.2 Investments

Another area where AI is applied is within trading and investments (FS 8, C1, C2 and C3). Algorithmic trading is common for large companies, but these are not self-learning algorithms (C1 and C3). FS 8 expresses that it is possible to use machine learning in trading, even though it is quite complicated. The company FS 1 works at uses a robo-advisor to invest in index funds based on a questionnaire where the customer has answered questions about for example their economic situation and risk perception. Everything in the system is automated and pre-programmed dependent on the answers to the questionnaire, although there is no self-learning element in this system.

4.3.1.3 Customer Service

Several respondents point out that companies in this industry increasingly apply AI within customer services, that is in the development of so-called virtual assistants that are applied to have some form of automated dialogue with the customer (FS 4, FS 6, FS 8, FS 9 and C3). These virtual assistants are used in areas such as investment counselling and in responding to general customer inquiries. As of today, it is only a few of these virtual assistants that can understand voices, speak and take calls, which is enabled by applying NLP based on machine learning (FS 4, FS 5, FS 6 and FS 9). Even though most of the virtual assistants only operates in text form, progress is being made toward voice recognition and to increase the usage of machine learning models (FS 4, FS 5, FS 6 and FS 9). FS 9 is particularly optimistic in this regard;

"Our future vision is that the virtual assistant will be the first-line customer service, where humans take over the communication when inquiries become too complicated, or there is a desire to talk to the customer for various reasons."

FS 8 and C3 explain that the companies' virtual assistant based on machine learning becomes better the more people talk to them, and these learning models have been trained through supervised learning. Although, FS 8 points out that there are also parts of the virtual assistant that is not machine learning, and instead are based on specified rules of what it should do.

However, there are some limitations to extend the usage of virtual assistants, where one problem lies in the old data architecture of banks which limits the virtual assistant from finding all the services that the customer asks for (C3).

4.3.1.4 Insurance

C1, C2, and FS 6 describe that AI can be useful to apply within the insurance industry. FS 6 points out that there are already many applications that contain rule-based systems, such as an automated business system for administration of pensions, insurances, and long-term savings based on algorithmic described rules. However, FS 6 points out that there is still extensive human involvement in current insurance processes. Regarding machine learning, it is explained that self-learning solutions have not been applied yet to a great extent, but it is stated that for example pricing of insurance premiums would be a perfect job for a machine learning system since large datasets are analysed (FS 6 and C1). It is further expressed that the statistical models that are currently used to calculate risks and premiums are believed to be quite close to machine learning (FS 6 and C1). FS 6 and C2 further describe that AI has been more widely used in property and casualty insurance than life insurance. For example, AI has been used in car insurance premiums by analysing driving behaviour. FS 6 explains that one reason to this is that less sensitive information is processed within property insurance than life insurance.

4.3.1.5 Payment

FS 5, FS 8 and C2 express that AI has been successfully applied to analyse payment card transactions to detect fraud, and the company FS 10 works at has implemented machine learning models for detecting IT-frauds of payments related to credit purchases. In the latter case, extensive data is analysed, and data points that potentially are fraudulent are flagged, after which each case is manually analysed. FS 5 points out that it is difficult to automate this process entirely. This process uses supervised learning, and it is expressed that it would be challenging to use reinforcement learning. In reinforcement learning, the model is not told which data points are fraud and which are not, which is needed for the model to operate accurately (FS 10).

4.3.1.6 Other AI applications within Financial Services

It is mentioned that virtual assistants also are used internally at one bank, for simple processes like unlocking accounts when passwords are forgotten (C3) and enabling associates to get answers to a range of questions (FS 8). C1 mentions that some companies in financial services use AI for marketing purposes, by registering different behaviours of the customer, and thereby they can identify customer preferences. In this process, data is automatically collected according to specific rules, for example from recording payment card activity. Thereto, FS 10 describes that AI at some banks is used to identify customers who are about to leave the company, which enables the company to contact them and attempt to persuade them to stay, and this is something that has given positive results.

Many of the respondents see great potential for AI within financial services (FS 3, FS 4, C2 and C3). C3 describes that there is excellent potential for both rule-based systems and machine learning within credit approval processes. C3 clarifies that because credit policies include numerous rules and requirements that have to be fulfilled, rule-based solutions are suitable. At the same time, machine learning can also be beneficial since people are different and sometimes one customer deviates from the "normal", and therefore the system has to be able to adapt rather than to always follow predetermined rules (C3). In accordance, FS 2 further explains the advantages of applying machine learning for credit evaluation;

> *"It is no longer possible to explain that a loan has been granted based on gut feeling, the decisions instead have to be based on facts and statistics, which is possible when using machine learning. At the same time, there has been an explosion in the amount of data, and therefore something more advanced than traditional models is needed to be able to analyse such extensive datasets."*

Furthermore, FS 5, FS 9 and C3 express that they see the highest potential of AI within customer service. Thereto, FS 6 perceives that rule-based systems and machine learning have great potential within insurance. Especially for Life and Pension insurance, FS 6 believes that the highest potential is within health insurance since it would make it possible to register a large amount of individual health data in real-time. However, it may not be a straightforward process since it is doubtful if customers would be willing to share their health information with insurance companies (FS 6).

However, even though there is great potential for using AI within financial services, some challenges exist. FS 8 expresses that customers still want to have a human they can discuss a decision with, and it is therefore unlikely that banks will implement a fully automated model for decisions such as credits. Thereto, one difficulty of implementing AI within financial services is that these organisations are finance companies and not technology companies, and AI is a particular area of technical knowledge (FS 2). It is a substantial change to use technologies like AI to a greater extent, and companies have to consider many aspects before applying AI (FS 3). This transition becomes further complicated with the old systems that many of the larger banks have (FS 3). Thereto, FS 9 expresses;

> *"Machine learning requires so much data, and therefore extensive server- and computing power is needed, which we do not have, and I think few banks and companies, in general, have this level of computer power in-house."*

FS 9 explains that the requirement of computer power is problematic in this industry since there is a reluctance to use external tools such as cloud solutions due to the highly sensitive business data that companies process.

4.3.2 General Aspects of the GDPR within Financial Services
FS 7 states that the financial services industry is under burdensome regulation and has a lot of administrative burden from other regulations and that GDPR further adds to this burden. C3 expresses;

"It is an enormous part of the companies' IT portfolio that is allocated only to ensure that they are compliant with all regulations."

C1 states that the GDPR is a tremendous administrative burden for companies, and it is a massive project that involves many employees. FS 7 further explains that the GDPR requires a new way of thinking, which could be quite challenging. This statement is confirmed by several respondents who express that there has been a lot of work with the GDPR and that it has taken a lot of time (FS 4 and FS 5). FS 3 and C2 express that one problem with enforcing GDPR is that the definition of personal data includes unstructured data, and it could be challenging to find this kind of data in all systems. Nonetheless, FS 7 and C1 point out that compliance is more straightforward for companies with more modern and scalable systems since an agile work method is necessary to insert technical changes from a regulation. Indeed, respondents FS 1, FS 2 and FS 3 work at companies founded within the last two years, and they mean that it has not cost their companies anything to become compliant with GDPR since the systems have been built from scratch knowing that GDPR would become enforced. However, FS 6 and FS 7 point out that the difficulty of what the regulation means in practice remains, and FS 6 expresses;

"As an engineer you want everything to be clearly formulated, to be true or false, which is not the case with law."

In accordance, both C1 and C2 state that the formulation of the regulation is explicit, but they perceive the practical implementations to be complicated. FS 7 expresses that the information from "Datainspektionen" has been insufficient, and four respondents (FS 5, FS 6, FS 10 and C1) mean that uncertainty still exists since there is no best practice, and this uncertainty will persist until after the enforcement (FS 6, FS 10 and C1). C2 means that companies could decrease the risk coupled to the uncertainty in how GDPR will be enforced through documenting how they have reasoned in regards to diffuse terms since this could make it more difficult for regulatory bodies to have objections. C2 believes that the GDPR includes diffuse terms since it focuses on what should be achieved rather than how since it applies to all industries processing personal data.

Some respondents (FS 4, FS 5 and FS 9) believe that there has been enough time to become compliant with the GDPR, and C1 adds that companies have known about the regulation for a long time. However, the general perception seems to be that it has been stressful to become compliant with the GDPR (FS 9 and FS 10). Thereto, C2 points out that since the proposition specifying what the actual law would look like was presented first in December 2017, there has been limited time to prepare. Nonetheless, FS 6 thinks that the time pressure of GDPR's enforcement date has benefited the industry since the time pressure together with the high fees has resulted in an evident change. Thereto, FS 9 points out that in comparison to some other regulations, the GDPR can be turned to something that benefits customers, and therefore companies have incentives to become compliant beyond regulatory demands.

4.3.3 Explaining Automated Decisions

It is expressed by FS 6, FS 9, and C3 that the possibility to understand and explain decisions made by machines is crucial within the finance industry since it is under strict surveillance. C3 clarifies;

"Even if the company trusts the decision taken by a machine, you also have to understand the reasoning of the machine since in this industry the supervisory body, "Finansinspektionen" requires an explanation to how decisions have been made, which becomes more difficult if you have a self-learning machine."

C3 continues to describe that "Finansinspektionen" has previously mainly been interested in the rules and policies that companies have in place, but have now also become increasingly interested in how systems are constructed. In line with this, FS 9 believes that the more automated processes become, the more "Finansinspektionen" will look into details of the systems companies use. However, FS 5 points out that the requirement to explain decisions varies among applications within financial services, and there are stricter requirements for credit approvals compared to for example fraud detection. Nonetheless, C2 and FS 6 state that regulatory bodies are in a disadvantage since it is difficult to understand AI and how it functions, and FS 6 believes this will also be true for "Datainspektionen" when checking for compliance to GDPR in AI applications. All respondents agree that the possibility to explain how AI systems work depends on the kind of model used, where rule-based systems can be understood quite easily compared to machine learning models. According to FS 6, purely rule-based systems do not seem to conflict with the GDPR since it is easy to show the factors that gave a specific result. In regards to machine learning, FS 3 describes that many companies use statistical models which are quite easy to interpret, for example within credit scoring it can easily be derived what variables that would need to change to get a better credit score, such as a higher salary.

FS 2 further describes that in the models they use for credit assessment it is relatively easy to understand the overall functionality of AI models, but means that it is more challenging to explain details about specific decisions. Other respondents also share this perception (FS 5, and FS 9). Although, FS 3 explains that the models used produce statistics and graphs for how decisions have been made, so it is still possible to get a lot of information about the process. It can also be visualised which factors had the most significant effect, such age, which makes it possible to have control over how the model operates (FS 3). Additionally, the respondents working with machine learning models for credit evaluation purposes and fraud detection of credit payment point out that it is possible to increase the understanding of specific decisions by running tests where values or composition of variables are changed, and then observe what happens to the result (FS 3 and FS 10). However, FS 2 and FS 8 describe that there exist complicated models that could improve the performance of AI systems, such as ANNs, but that these models are so complicated that it is not possible to understand how the system derived at a specific decision.

According to FS 3, FS 8 and FS 10 it is easier to understand decisions of machine learning models if supervised learning has been used, as compared to unsupervised and reinforcement

learning. The reason is described to be that in reinforcement learning the machine continuously learns on its own, whereas in supervised learning the machine is only trained in the beginning and then this internal state stays the same. FS 3 and FS 8 point out that it is problematic to use unsupervised learning because of the "black box syndrome", meaning that it is difficult to understand how the machine derived to the decision. FS 8 further explains that current machine learning technologies cannot provide an answer to why it knows what it knows. FS 3 summarises the problematic of using unsupervised learning in the financial industry;

"Today people still want to have an understanding of why a decision has been made, but unsupervised learning is often highly abstract, and is a lot like a black-box. In other words, you do not know why decisions are made."

FS 8 emphasises that these characteristics of unsupervised learning cause a problem in lacking accountability: if an explanation cannot be provided how a decision has been made, it will also be problematic to determine who is responsible for the mistake. Moreover, C1 and C2 describe that the large banks have a lot of inaccurate data, and FS 2 states that machine learning will never question if the data it receives is correct, and will just analyse the data and try to find patterns. Thereto, FS 8 and C1 express that there also are ethical considerations to why some actors are reluctant to adopt machine learning that has not been trained with labelled data. FS 8 describes that although AI can improve consistency in decision-making, there is a risk that decisions become biased if the AI machine gets to operate freely without supervision. Therefore, it is neither possible to control what factors the machine includes in decisions, and thereby it cannot be established that the decision is not discriminating (FS 8). FS 8 describes that there currently does not exist any good solution to exclude bias in machine learning models and is therefore sceptical that unsupervised learning will be used in the near future:

"In industries similar to ours we will not trust unsupervised learning for quite some time, partly due to regulatory reasons, but also due to ethical reasons, we want to act in the best interest of our customers, and we cannot guarantee this if a machine takes a decision, which we cannot explain it."

4.3.4 Erasing Data
FS 8 explains that one of the main challenges of GDPR for companies is the requirement of erasing data, in terms of being able to locate and delete data. However, FS 8 points out that the level of difficulty depends on how good internal structure and control over data the company has had historically. Thereto, FS 7 and C2 state that removing data is not something that has been done previously within financial services. In this regard, FS 7 explicitly points out;

"Companies already have many standards in place for complying with the GDPR, but it has not been best practice in the industry to have a screening process for continuously erasing data that should not be stored."

C2 further describes that one problem is that privacy has not been considered to a great extent by companies, and therefore a lot of information about customers have been stored over the

years. FS 1, FS 3, and FS 5 do not see any difficulties with erasing data. FS 5 means that the problem instead is to erase only specific information about an individual instead of just erasing everything that is stored about that individual. Another issue highlighted by C2 is to remove data from back-ups. C1 and C2 further express that it would be extremely time-consuming to erase a person from all back-ups since companies have conducted extensive documentation over so many years within financial services.

FS 1 and FS 3 describe that they currently do not receive many requests for erasing personal data. However, it is expressed that there are always people who want to test these regulations, and therefore the requests may increase after the GDPR enforcement (FS 1, FS 4 and FS 10). Thereto, respondents express that they do not believe that banks and insurance companies will be the primary target for data erasure, this will instead be actors like the social media platform Facebook where the information can be made public and sold (FS 7, FS 8 and C1). In contrast, banks have rules preventing them from sharing data about their customers due to bank secrecy (FS 7, FS 8, and C1). C1 explains that another reason that companies within financial services will not have to erase a lot of data is that banks and insurance companies are subject to several regulations that triumph the right to erasure and therefore require that specific information is stored.

However, C2, FS 4 and FS 7 describe that the erasure obligation becomes relevant when customers have left the company. FS 7 continues to describe that they now have an automated process that erases this kind of data when it is time, a system that FS 6 explains is based on a quite clear algorithmic business rule. Although, FS 7 points out that it has taken considerable effort to build this data structure since the entire insurance engagement and associated transactions are personal data, and therefore it is an extensive amount of information that has to be erased.

Nevertheless, C1 means that for companies to become compliant many larger banks have to rebuild their systems, which is also true for the company FS 10 works at who describes that they have needed to reorganise their systems to be able to delete data and comply with the erasure request of GDPR. Extensive work is needed for larger bank and insurance companies to understand the data they have, and since information related to one individual has been stored in several different systems, this could make it challenging to locate all personal information about an individual and therefore it is also challenging to comply with the erasure request (FS 4, FS 5, FS 7, FS 8, FS 10 and C1).

FS 10 describes that to comply with the erasure request companies can either erase data altogether or anonymise the data. FS 10 states that anonymization of data is quite easy. However, one difficulty of anonymising data is that different variables that are not directly personal information potentially could be connected and together identify a person (FS 2 and FS 3). FS 10 adds that even when data is anonymised the data could be combined to identify a person, and if this is possible it is not enough safety precautions to have the data anonymised. C1 expresses that especially insurance companies have a lot of anonymised information stored about historical claims, which potentially together could identify an individual.

However, the question is how many of the anonymous data points that have to be deleted (C1, FS 2 and FS 3). C2 describes that there is a reluctance towards these erasure programs since companies are so unused to erasing information, and many actors therefore instead wish to anonymise the data.

4.3.5 The Main Consequences of the GDPR from an AI Perspective

C1 describes that most industry actors agree that the intentions of the GDPR are positive and that it is appreciated. Several respondents (FS 1, FS 2, FS 3 and C2) express that it is positive that the processing of personal information becomes regulated, and that this will be beneficial for customers. Most of the Industry actors do not believe that GDPR will have a significant impact on AI. One of the reasons is that companies often are not interested in personal information specifically; companies want to analyse risks and find driving attributes, and therefore the data can be made anonymous without negatively affecting the performance of AI models (FS 2, FS 3, FS 5, FS 6, C1 and C2). FS 4, FS 6 and C2 believe that the future focus will be on how the application of AI can become better rather than have a negative impact or hindering the development of AI and machine learning since the usage of AI enables to make processes more efficient and autonomous. C2 expresses;

> "Everything you have done before you will also be able to do after 25th May, we may just need to have more knowledge about what we actually do. However, I do not believe that you will not be able to use machine learning or AI."

C2 further points out that the GDPR imposes a significant workload, but that GDPR also could lead to a greater understanding and insight of data, which today is lacking in many companies, especially for the large actors. C2 believes that this could be an enabler for AI since companies hopefully will get a better structure and understanding of data with the push from GDPR, which means that they would have better conditions to apply machine learning. Furthermore, FS 1 means that the GDPR is positive since its requirements for transparent decisions will lead to more meticulously designed automated systems.

However, FS 2 and FS 10 believe that, at least in the short term, GDPR could hinder some projects to develop AI since companies will be more careful and it will take more time. Thereby, companies may miss some projects in the short-term. FS 2 clarifies that both rule-based AI and machine learning will become negatively affected by the GDPR since companies will become more restrictive in how their data is used, for example in automated processes based on AI, until companies have ensured GDPR compliance. FS 10 also mentions that companies might want to train the AI machine on data several years back, and companies might not be allowed to save some of the data now with GDPR, and this might negatively affect some AI learning. C3 and FS 9 believe that since GDPR could make it more challenging to apply AI since companies may be less willing to apply AI and instead to a greater extent use models that have greater transparency. Thereto, FS 6 points out another potentially negative impact of the GDPR in regards to machine learning;

> "The prerequisite for machine learning is large amounts of data, and with the GDPR the control of personal data is given back to consumers for how the information can be used, which could impede machine learning applications."

5. ANALYSIS

This chapter presents the analysis of the study, where the literature is connected to the empirical findings. Initially, an overview of the framework for the analysis is provided, which is followed by a discussion of the different forms of AI applications within the financial services industry. Thereafter, the different characteristics of GDPR are evaluated in relation to AI. Finally, the analysis finishes with a summary and discussion of the implications of the findings.

To be able to fulfil the purpose of this research, the analysis is divided into two parts; an initial section that discusses the forms of AI that are currently used within financial services, which also assesses the future potential of AI applications within the financial services industry. The findings in this section are then used to evaluate what impact the GDPR has on AI applications within financial services. This second section is structured according to the five characteristics that were identified in the literature to determine how a regulation impacts innovation. Figure 4 below visualises the framework for the analysis. In the end, to make the findings clearer the analysis ends by summarising what has been derived as well as discusses the implications of the findings.
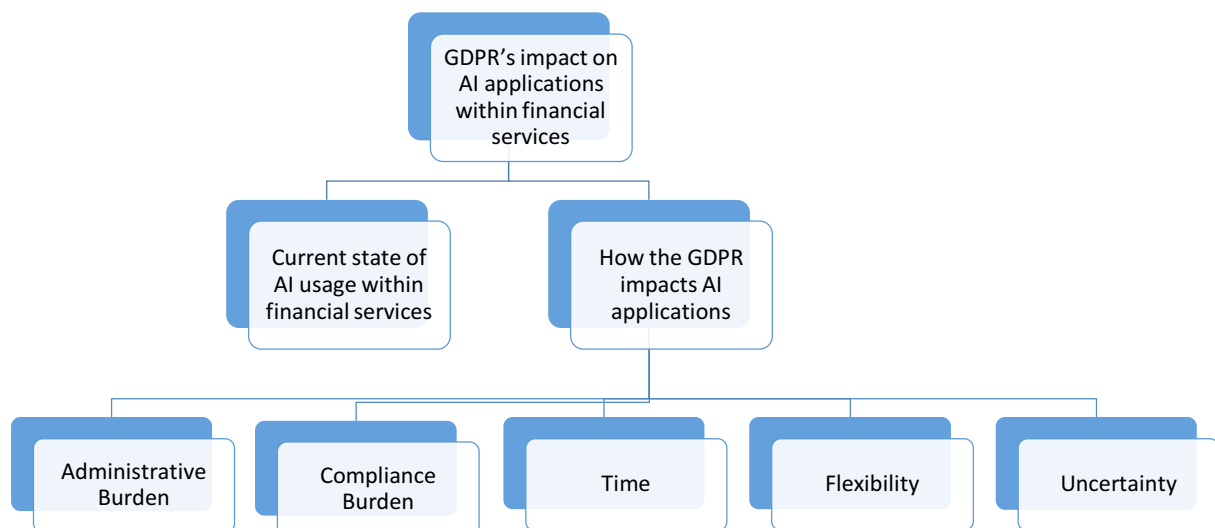
Figure 4. The framework for the analysis.

5.1 Artificial Intelligence Applications in the Financial Services Industry

In line with Tecuci (2012), the AI experts explain that AI is a broad field and that there is no commonly agreed upon definition. In turn, this could explain why Industry actors have different opinions about what the term "AI" includes. Regardless, the literature (Bostrom 2014) and AI experts agree that the Artificial General Intelligence (AGI) does not yet exist and express that all applications so far have been and are Narrow AI. According to previous research, AI has been successfully applied to many different areas of financial services. These findings are also confirmed by the Industry actors who describe that AI, for example, is applied as decision-support within credit approval processes, to improve security for

consumers and companies by identifying IT threats and detect frauds, to automate investment processes, as well as to assist customers more efficiently within customer service.

5.1.1 Current Forms of Artificial Intelligence Applications within Financial Services

Despite the perceived confusion about the definition of "AI", there seems to be an agreement between literature and empirical findings that a distinction can be made between two different forms of AI: systems based on rules that have been pre-programmed by humans, that is rule-based systems, and AI that has an element of self-learning, called machine learning. Although, while the AI experts mean that rule-based systems are used less and less today, the Industry actors describe that this form of AI is more commonly used than machine learning within financial services. Nonetheless, machine learning models have been successfully applied within some areas of the industry, and are described to become increasingly applied.

5.1.1.1 Machine Learning Models

The Industry actors describe that the machine learning models that currently are most common within financial services are statistical based models, while ANNs are quite rare. Although, ANNs are becoming increasingly common for analysing large data sets (Industry actors), which is aligned with the benefits of ANNs described by LeCun et al. (2015). Nonetheless, statistical machine learning models are described to perform better than ANNs for the majority of tasks within this industry due to the vast amount of structured data that is processed. Additionally, ANNs are problematic to use for companies within this industry since ANNs are difficult to understand and at the same time, industry-specific regulations require transparency and accountability of decisions (AI experts and Industry actors). The GDPR imposes even higher pressure on these requirements since customers have the right to receive information about data processing and how automated decisions are made. Hence, it seems unlikely that ANNs will become widely adopted within this industry in the nearest future. Although, the industry actors point out that the level of explanatory requirement varies within the industry. In turn, these differences could explain why ANNs for example are used within fraud detection and for parts that are not explicitly related to financial decisions, where there is a lower need to understand the reasoning of models.

Despite the differences, the AI expert describes that both statistical based models and models based on ANNs require technical expertise to develop and implement. In accordance, lack of technical knowledge is described by one Industry actor to be a difficulty of implementing AI within financial services. Consequently, this could explain why the simpler rule-based models are more common than machine learning in the industry. Furthermore, this requirement of technical expertise could also explain the intra-industry differences indicated in the empirical findings: that newer actors with a more technical knowledge-base seem to use AI to a greater extent compared to the old-established actors. Nonetheless, since technical knowledge is an asset that can be acquired, machine learning solutions may become more extensively adopted in upcoming years, when the old-established companies have been able to restructure and adapt to the increasingly technology-driven industry environment.

5.1.1.2 Machine Learning Techniques

In regards to learning techniques used for machine learning models, both LeCun et al. (2015) and the AI experts state that supervised learning is most common, which is confirmed by the Industry actors to also be true at the current state in the financial services industry. Nonetheless, the AI experts describe that supervised learning requires a significant amount of training data, and labelling this training data is also a tedious task. To that, the AI experts, in accordance with Kluegl et al. (2016), describe that the available training data is at times insufficient. These limitations indicate that there could be a lot to gain for companies in this industry to instead use unsupervised learning, which the AI experts explain require both less training data and human work. However, the AI experts mention that specifically for the financial services industry, there is historical data available, such as of who has repaid loans and not, which indicates that for at least some tasks within this industry there are already labelled data, and therefore less work effort is required to label the data. Yet, both the AI experts and Sathya and Abraham (2013) point out that unsupervised learning can identify patterns that humans may not have considered. Consequently, it could be beneficial to use unsupervised learning in some tasks since there often are extensive amounts of data to be analysed, and therefore challenging for humans to find all correlations that exist in the dataset.

However, some Industry actors explain that with unsupervised learning it is difficult to remain in control over what the machine learns and also challenging to understand why decisions have been made. In turn, the Industry actors mean that it becomes problematic to use unsupervised learning in this industry due to regulatory and ethical reasons. In accordance, Jones (2014) points out that unsupervised learning is the most difficult learning technique since the input data does not include any information about what the data represents. Hence, even though the AI experts believe that unsupervised learning will be the next big paradigm of AI, this is not likely to be true for the financial services industry within a close future. Moreover, deriving from the Industry actors, neither reinforcement learning seems to be a preferable choice of learning technique in this industry as it currently is nearly non-existing. One explanation to this could be that reinforcement learning is difficult to use when "wrong" decisions result in severe effects (AI experts), which commonly can be said to be the case within the financial services industry where many high-stakes decisions are made.

5.1.2 Future Potential of Artificial Intelligence Applications within Financial Services

The AI experts describe that many advancements within AI technology have been made during recent years and that the development is moving forward rapidly, why it is likely that AI applications will have even higher potential in upcoming years. More specifically, within this "rule-based AI dominated" industry there seems to be benefits to gain from extending the use of machine learning systems. Such systems have demonstrated higher performance (Bohanec et al., 2017), at the same time as it is time-consuming for humans to write the rules for rule-based systems (Kluegl et al., 2016). Indeed, it is recognised that machine learning models are becoming increasingly common (Mittelstadt et al., 2016; Industry actors). The Industry actors see a high potential for using AI technology, and particularly machine learning, within several areas of financial services. For example, AI is pointed out to be advantageous for making credit approval processes more objective and consistent, fully

automating the customer service experience as well as pricing insurance premiums (Industry actors). Hence, it is likely that the use of machine learning will expand in upcoming years, which are likely to be statistical machine learning models using supervised learning to derive from section 5.1.1 above.

However, extending the use of machine learning may not be entirely straightforward for companies within this industry. For example, the Industry actors describe that it is difficult to use machine learning due to insufficient computer power in-house, and that the highly sensitive business data makes it is too risky to use external cloud solutions. Furthermore, it is described that customers still want some human contact when interacting with financial services providers, and thereto there is a need for specific technical expertise to use machine learning, which many organisations lack (Industry actors). Thereby, these factors are some challenges that this industry faces in adopting machine learning models.

Besides these identified challenges, regulation could also impact innovation (Ashford et al., 1985; Blind, 2012; Pelkmans & Renda, 2014; Ranchordás, 2015), such as applications of AI. Indeed, regulation is pointed out by the Industry actors to be one significant restricting factor for using AI, and it has been argued that the new regulation GDPR could be burdensome to comply to with AI technology (Malgieri & Commandé, 2017; Villaronga et al., 2017; Wachter et al., 2017a). This will be discussed in the following section by taking a stance in the current state and future potential of AI applications within financial services.

5.2 The Impact of the GDPR on Artificial Intelligence Applications

Lake et al. (2017) and Tecuci (2012) describe that AI applications have become more relevant and useful in recent years, which is confirmed by the AI experts who point out that AI technology has become increasingly adopted during the past years. As described in section 5.1 above, this development can also be said to be true for the financial services industry where many innovative applications of AI technology are found and also identified to have great future potential. However, Ranchordás (2015) describes that regulations can decrease the rate of innovations. Indeed, the literature also suggested that AI could be negatively affected by the new data privacy regulation, GDPR (Malgieri & Commandé, 2017; Villaronga et al., 2017; Wachter et al., 2017a). In this regard, concerns were raised about some of the GDPR's articles, which are presented in Table 7 below, along with the Legal and AI experts' opinions about these issues. This table is a continuation of Table 2 in section 2.3.2 in the literature review; "Summary and implications of the alleged right to explanation, and the right to erasure".

Table 7. Findings of how the GDPR affect AI.

	Article	Questions Raised in Regards to Artificial Intelligence	Findings
Right to explanation	13-15	**1.** How thorough and detailed information is required in solely automated decisions; details about particular decisions after the decision has been made or only about the functionality of the automated decision-making process?	**1.** Extensive information requirement exists about how solely automated decisions have been made, but not as detailed information as how the algorithm works. Thereto, information should be provided to the extent that the individual has the possibility to contest the decision, which can be both before, during and after the decision.
	22	**2.** How should "significant effects of automated processes" be interpreted? **3.** How extensive human involvement is needed for an automated decision not to be considered solely automated?	**2.** Significant effects are exemplified to be denials on online applications, such as in recruitment and for credits. To get a more comprehensive understanding of this term, best practice has to be established. **3.** For it *not* to be considered solely automated, the human must oversee the process as well as have authority to change the decision.
Right to erasure	17	**4.** What is the technical meaning of "erasure"? **5.** Will the erasure request be possible to comply with when AI systems are used?	**4.** It is still unclear what kind of erasure that is meant and what actions that have to be taken to be compliant. **5.** It requires more effort, but it is likely to be possible.

Table 7 describes how some aspects of the GDPR relates to AI. However, to derive a more holistic perspective of the impact that the GDPR has on AI applications, a more in-depth examination must be conducted, which also includes the findings in Table 7 above. It is described that the impact of a regulation depends on the balance between innovation-inducing and innovation-constraining elements (Ashford et al., 1985; Blind, 2012; Pelkmans & Renda, 2014; Ranchordás, 2015). According to Ashford et al. (1985) and Pelkmans & Renda (2014), the impact that a specific regulation has on innovation is determined by five characteristics; administrative as well as compliance burden imposed by the regulation, timeframe until enforcement, the degree of flexibility, and the level of uncertainty in the regulation. Hence, to fulfil the purpose of this research, these attributes of the GDPR will be evaluated in relation to AI technology, and how this technology is applied within the financial services industry. This evaluation will be conducted by taking a stance in the GDPR's primary requirement of restricting the processing of personal data, as well as focusing on the articles that are presented in Table 7 above, which are related to automated decision-making.

5.2.1 Administrative Burden

According to Pelkmans and Renda (2014), administrative burden refers to the time and resources that companies allocate to become compliant with a specific regulation. In turn, this decreases the time that companies can spend on entrepreneurial activities, and is thereby disadvantageous for innovation. In this regard, there are several aspects of the GDPR that needs to be considered, which mainly can be attributed to the GDPR's general restriction of processing personal data as well as the right to erasure.

5.2.1.1 The GDPR's General Restriction of Processing of Personal Data

Firstly, the GDPR imposes an obligation on companies to establish a legal ground to be allowed to process personal data (Art. 6, EU 2016/679). The Legal experts recurrently mention the importance of having a legal ground to process data. The need to establish a legal ground becomes particularly severe within financial services since this is an industry where companies handle a lot of personal data, and therefore the procedure of establishing a legal ground is likely to require a quite significant effort. This is particularly true in the short-term until routines have been implemented. To that, it is described by the Industry actors that companies also have to document the reasoning behind decisions to later on be able to motivate for both regulators and individuals, why a specific legal ground applies. Hence, such documentation processes are likely to increase the administrative burden that the GDPR imposes on organisations. In turn, fewer resources will be available for entrepreneurial activities and thereby have a negative impact on AI applications.

Secondly, in contrast to the previous data privacy regulation in Sweden, unstructured data such as emails and voice recordings will also be included in the definition of personal data with the GDPR in force (Legal experts). Therefore, companies now have to screen systems for this kind of data and delete the data that they do not have a legal ground to store, which is likely to require a quite significant effort. To that, the Legal experts point out that many companies have stored a lot of data without considering that it could be personal data, meaning that they have considerable amounts of data in storage. Additionally, one AI expert points out that a prerequisite for companies to become compliant with GDPR is to know where their data is stored, which many companies do not know. Consequently, the obligation to screen through systems is likely to impose an administrative burden on many companies initially as it will require a comprehensive work effort to go through all systems. This burden can be said to be especially significant for old-established firms since it is stated by the Industry actors that these companies have old internal systems where data is dispersed, at the same time as they have collected and stored extensive amounts of data over the years. Thereby, to become compliant, systems are likely to have to be restructured to gain better control and knowledge about the data. In fact, one of the AI experts explains that it is easier for companies that have been established more recently to comply with the GDPR. The reason for this is explained to be that it is possible to build systems that fulfil the GDPR requirements if security and privacy aspects are taken into consideration from the beginning of the system design. Indeed, the newer companies interviewed do not see any problems with becoming GDPR compliant. At the same time, there seem to be some positive aspects that could come out of this administrative burden for the old-established companies since one AI

expert points out that it is also costly to lack insight about the content of one's data. Therefore, a reconstruction enables companies to save money in the future.

Finally, in regards to AI applications, the Legal experts explain that personal data can only be used for the purpose that it was collected. This restriction matters for AI applications in terms of machine learning since organisations who wish to use personal data to train machine learning models will have to ask for consent from the data subjects, or establish another legal ground. It is no longer possible for companies to collect and use data as they wish, indicating that time and resources will have to be allocated to establish a legal ground for the specific purpose of being used to train machine learning models. This burden could become quite significant considering the extensive amount of data that is required to train machine learning models. At the same time, deriving from the above paragraph, complying with the GDPR enables many companies to keep track of and gain insight about their data, which is pointed out by the Industry actors to create favourable conditions for machine learning. If companies know what data they already have stored, they can better identify what data that can be used for training machine learning models, although a legal ground has to be established first. Hence, the initially imposed administrative burden could have a positive impact on AI applications in regards to machine learning models. Although, this will most likely be in a longer time perspective since it takes some time to develop and implement new applications. This finding, that an administrative burden can have a positive impact on innovation, is not mentioned in the current literature, but could be an additional aspect of importance to consider when evaluating the impact that regulations have on innovation.

5.2.1.2 Erasing Data

The Legal experts describe that the erasure requirement also has existed in the previous data privacy regulation PUL, but yet the Industry actors perceive this requirement as one of the main challenges of the GDPR. It is stated that a screening process for continuously erasing data has not been best practice within the industry. Hence, this indicates that there is some truth in the Legal experts' statement that the compliance with PUL has been limited. In turn, this could explain why companies perceive the GDPR as a burden that requires significant time and resources to become compliant. Thereto, there are indications that companies might receive more erasure requests than they have previously. The reason for such a claim is that the Legal experts explain that many individuals previously have not been aware that there actually exists a right to erasure. Therefore, consumers are likely to exploit this right to a greater extent due to the great attention of GDPR and increased obligations for companies to inform individuals about their rights. Indeed, both the Industry actors and Legal experts believe that more individuals will want to test the system with the introduction of the GDPR. To that, the Legal experts also point out that Swedish citizens are becoming increasingly concerned about how companies process their personal data, which also could contribute to an increase in erasure requests. In turn, handling a significant number of erasure requests is undoubtedly something that could take considerable time and resources for companies to process.

However, the Industry actors express that their businesses are not consumers' primary targets for requesting that personal data is erased. Although, following the reasoning of the previous paragraph about individuals' increased concerns about how personal data is processed and the attention that the GDPR has received, there is a possibility that companies in the industry have underestimated the number of requests they will receive. On the other hand, it is explained by Industry actors and Legal experts that the financial services industry is subject to other regulations that triumph GDPR, and therefore there are many situations when companies will not have to fulfil erasure requests that they receive from customers. As a result, the administrative burden will probably not be as extensive as was suggested at a first assessment.

Nevertheless, companies must be prepared and capable of erasing data since the regulatory exceptions do not apply to all cases. In fact, the Industry actors describe that there are situations when companies in this industry are obliged to fulfil erasure requests, which often is the case when previous customers request that their data is erased. In such situations, the Industry actors point out that it becomes time-consuming to fulfil the erasure requests since there is such extensive amount of data that have to be erased, often meaning the entire customer engagement. Additionally, considering that it was established in the paragraphs above that many companies have old and dispersed internal systems, it becomes difficult to locate the data that has to be erased. Therefore, some companies will have to dedicate effort to restructure systems in preparatory purpose to be capable of meeting erasure requests that have to be fulfilled. To this, both the AI experts and Industry actors describe that an extensive work would be needed if it will be required to delete data from backups to comply with the GDPR's erasure requirement. In turn, a significant administrative burden would be imposed on many organisations considering the extensive amount of data that has been collected in this industry over the years, and thereby also would have to be erased. Although, it is at the current state unclear to what extent backups will be included in the GDPR requirement, and therefore further regulatory guidelines or court cases have to established to clarify this question.

5.2.2 Compliance Burden
The second characteristic, compliance burden refers to the cost and difficulty of complying to a regulation with a company's existing technologies and business models (Ashford et al., 1985; Pelkmans & Renda, 2014). Several aspects are identified for this characteristic of the GDPR, which appeared to all stem from the GDPR requirements of explaining automated decisions, and erasure of personal data. Hence, this section is divided into these two different subsections.

5.2.2.1 Explaining Decisions
Firstly, according to Article 13, 14 and 15 in the GDPR, the informative requirement becomes more extensive in the occurrence of automated decision-making since information in these cases has to be provided about the logic involved in the process (EU 2016/679). In this regard, Wachter et al. (2017a) argue that it could be challenging for companies to explain how decisions have been made when automated processes such as AI takes the decisions. The AI experts agree with this statement in the sense that it is problematic to provide details about

specific decisions taken by machine learning models, as these often lack transparency and are difficult to understand. Hence, it could be difficult for companies to fulfil the informative requirement with current technologies if machine learning is used in automated decision-making. In turn, this results in a compliance burden for companies who use or wish to implement machine learning solutions to automate decision-making. To that, if it would be impossible to provide sufficient information about decisions, these models may even have to be replaced with new ones that are easier to understand, which would become costly. Thereto, the compliance burden could become further significant if the machine learning models are based on ANNs since the AI experts explain that there is a trade-off between performance and ability to explain decisions made by such models. Thereby, increasing the level of understanding of ANN-models could become a compliance burden in terms of lower quality in the decisions that the machines take.

Nonetheless, it is debated whether there truly exist "a right to explanation" in the GDPR, that is, to what extent companies have to provide information about decisions stemming from automated processing (Malgieri & Commandé, 2017; Wachter et al., 2017a). In this regard, Wachter et al. (2017a) argue that it is enough for companies to explain the overall functionality of the automated system *before* a decision is made. Such level of explanation is described by the AI experts to be quite easy to fulfil with most AI models, besides the highly complicated models based on ANNs. Thereby, if such an overall explanation of functionality would be the correct interpretation, the compliance burden of the GDPR for companies within this industry would be quite low since it was established in section 5.1 that ANNs are rarely used. On the other hand, Malgieri and Commandé (2017) argue that information has to be provided about a particular decision after the decision has been made. This interpretation would indicate a more severe compliance burden since the AI experts explain that such information often is difficult to derive from machine learning models, including statistically based models, but especially ANNs.

According to the Legal experts, the information requirement about automated decisions should be extensively interpreted, and it is repeatedly mentioned that the purpose is that individuals should be able to contest a decision. Hence, it indicates that the crucial thing is that enough information is provided about the decision-making process for the individual to be able to contest a decision, which is also in line with Kingston's (2017) standing point. For it to be possible to contest a decision, it is explained by the Legal experts that information has to be provided both before and after a decision has been made. Thereby the Legal experts' standing point is closer aligned with Malgieri and Commandé (2017) than Wachter et al. (2017a). This line of reasoning is further amplified by the Legal experts' explanation that recitals should be considered to be part of the regulation and viewed as legally binding. Therefore, the argument made by Wachter et al. (2017a) that a right to explanation does not exist because it is only mentioned in a recital becomes ineffective. In turn, this also indicates that there is more protection in the regulation than suggested by the same authors (ibid.). Consequently, it is concluded that extensive information about automated decisions has to be provided, but it could be difficult to provide complete explanations when machine learning is

involved in the decision-making process. Therefore, a quite significant compliance arises from the informative requirement for some companies that use machine learning.

However, deriving from Article 22 it becomes clear that such extensive information requirement only applies when the decision is based on *solely* automated processing. One Legal expert points out that the word "solely" is of importance since this enables companies to use automated processes in some parts of the decision-making process. Additionally, Article 22 further specifies that the decision taken by *solely* automated means, must *significantly* affect the individual for the informative requirement to apply (EU 2016/679). One example of such a decision is exemplified to be credit applications (EU 2016/679), meaning that the requirement of Article 22 often would become applicable within the financial services industry if decisions were to be taken by *solely* automated means. Nevertheless, deriving from the Industry actors, it seems that nearly no *solely* automated decision-making takes place within this industry. Instead, automated data processing is mainly used to support human decision-making. In turn, this means that the informative requirement of Article 22 many times become irrelevant. Consequently, since there are few *solely* automated decisions, it does not matter that many decisions within this industry were established to have "significant effect". Hence, companies within this industry using machine learning models will not be imposed with the quite significantly compliance burden suggested in the previous paragraph.

Nonetheless, it is described that a few *solely* automated decisions take place within some banks for smaller loan applications (Industry actor). Thereby, there may be further automated decision-making processes than what has been found in this study, why there may become somewhat of a compliance burden for some companies within the industry if these models are difficult to understand. Additionally, since the Industry actors emphasise that high value can be achieved by making use of AI technology, there may be a desire to increase automation with these businesses, and to a broader extent conduct fully automated decisions, which would then come to fall under the GDPR's definition of *solely* automated decision-making. Anyhow, in that situation it is not likely that there will be a significant compliance burden for companies since it was established in section 5.1 that the most commonly used AI models in this industry are of the simple form and easy to understand, that is, rule-based systems. In turn, this means that the extensive informative requirement would often be fulfilled by simple means if similar models are continued to be used. This line of reasoning is also in accordance with the Industry actors who do not see any conflict with GDPR when rule-based AI is used.

Although, there are some machine learning models in use within the industry which are not as simple to understand as rule-based AI. The majority of these machine learning models are statistically based, instead of ANNs. From such statistical based models, it is described that it is often possible to derive quite extensive information about decisions, such as which factors that affected the outcome (Industry actors). Thereto, supervised learning is used rather than unsupervised learning, which according to the Industry actors makes it possible to understand why decisions have been made. However, following the AI experts' opinion described above, some industry actors that use machine learning models point out that it is difficult to explain

specific details about individual decisions. Nonetheless, it is likely that the information that can be provided will be sufficient to fulfil the requirements, considering that the Legal experts explain that it is not required to disclose information as detailed as algorithmic specificities. In turn, since it is not necessary to increase the level of understanding in the machine learning models in use, the suggested trade-off in performance and explanatory power within complicated machine learning models do not appear to become relevant. Thereby, it can be concluded that there neither would be a significant compliance burden stemming from the informative requirement in future years even if *solely* automated decision-making would increase, assuming that similar AI applications as currently are in place are continued to be used.

On the other hand, in regards to machine learning models, there is potential value to gain from ANNs and unsupervised learning techniques (AI experts and Industry actors). Therefore, companies could become more motivated to invest in new machine learning models that have a higher level of explanatory power, since that would be required for companies to use such models and at the same time comply with GDPR's requirements. The AI experts point out that there is already ongoing research into such explanatory AI, and the GDPR's requirement of transparency indicates increased importance of such research. Thereby, the compliance burden stemming from the difficulty of explaining decisions made by highly complicated machine learning is likely to have a positive impact on AI applications in the longer-term, by inducing the development of more informative and transparent models. This reasoning is in line with Blind (2012) who points out that a compliance burden can trigger innovations by enhancing incentives to invest in innovation activities or research.

5.2.2.2 Erasing Data
In addition to the above-identified administrative burden that the right to erasure gives rise to, there are some aspects of the erasure requirement that also could impose a compliance burden in terms of difficulty to fulfil the requirements with current technologies.

The AI experts express that the erasure requirement is problematic since erasure within computer systems can be conducted on different levels, but it is not specified in the GDPR what kind of erasure that is sufficient. In this regard, the AI experts explain that it will be difficult for companies to comply with the regulation if an extensive level of erasure is required since it would cause a significant change in technology. In accordance, Villaronga et al. (2017) state that it is difficult to erase data technically from systems, and primarily from machine learning models. It is even argued that it may be impossible to comply with the erasure requirement if machine learning systems are used since its complicated structure makes the deletion process difficult, and it often takes time for these systems to forget the deleted information (AI experts; Villaronga et al., 2017). Thereby, if companies would no longer be able to continue to use machine learning models with the GDPR in force, it would impose a significant compliance burden due to the difficulty of complying with existing technology. Nonetheless, the Legal experts explain that if companies are not capable of implementing technical safety solutions, other security measures could be taken to fulfil the GDPR's requirements. The legal experts clarify that what matters is that companies have

taken reasonable safety measures. Hence, it is likely that the same will be true for handling erasure requests, which thereby indicates that machine learning models most likely will be possible to use. In turn, there would not be a compliance burden following from the restrictive wording of the right to erasure.

Yet another potential difficulty expressed by the AI experts when discussing erasure in regards to machine learning models is that deleted data remains in the system, regardless of the kind of deletion that is conducted, since it exists in the behaviour of the system. Hence, the AI experts mean that it could become challenging to comply with the requirements if the GDPR requires that even behaviours in a machine learning model are deleted when erasure requests are received. According to the AI experts, a machine learning model has to be retrained on new data to be able to remove a specific behaviour, which becomes burdensome if it has to be done too often. However, an extensive amount of data is used to train machine learning models, why it seems unlikely that a specific behaviour could be connected to a particular individual. Thereby, the GDPR will probably not require that systems are retrained, at least not for a low number of requests, meaning that this is not likely to result in a compliance burden for companies.

Furthermore, even though it is unlikely that it will be required to delete the existing behaviour of models, erasure requests could still affect new models that are to be implemented since there will be less data available to train machine learning models if many erasure requests have to be fulfilled. The AI experts explain that less data could have a negative impact on the performance of machine learning models since more training data often yields higher accuracy. Consequently, the GDPR could hinder companies to exploit the benefits of machine learning applications, which would result in a cost concerning foregone gains of using a technology. To that, erasure requests could impose a further compliance burden since the AI experts point out that the training data often is beneficial to save after machine learning models have been trained. This could be of importance to, for example, be able to correct flaws in the model or explain why a particular decision was taken (AI experts). Consequently, this is another reason to why erasing data could lead to lower performance of machine learning models, or difficulty of fulfilling regulatory requirements of explaining decisions with current technologies. Although, the significance of this impact will depend on how many erasure requests that companies receive, and how many of these that have to be fulfilled, which will not be known until some time after the GDPR has been enforced.

At the same time, the Legal experts explain that companies can store personal information if it is anonymized, and the Industry actors describe that machine learning models can learn from anonymized information. Thereby, companies could find ways to avoid that machine learning models are negatively affected by the potentially lower amount of training data due to erasure requests. However, companies must take enough safety precautions to obstruct de-anonymization, but it is difficult to know what level will be considered sufficient by regulatory bodies, a perception that is shared by both Legal experts and the Industry actors. Thereto, the Industry actors describe that they do not necessarily need personal data to train machine learning models. In turn, another solution to increase the available training data

could be to only erase the personal data that is directly related to an individual in a dataset and save other parts. However, it is mentioned that only deleting some parts of a dataset is technically difficult (Industry actor). Additionally, both the Legal experts and Industry actors emphasise that it is unclear how many data points that are not on its own personal data, but relates to the same individual, that companies can store before it is considered to be personal data. Hence, the GDPR will still impose a compliance burden in that it will become a struggle to ensure that enough effort has been dedicated to unable de-anonymization or identifying individuals by combining different data points.

It can be concluded that there are many potential difficulties identified for machine learning models in regards to the right to erasure that could have a severe negative impact if strict interpretations are made by supervisory bodies, which is unclear at the current state even though it seems as if the strictest interpretations are not likely. Furthermore, it seems to be ways for companies to continue to use machine learning despite the erasure requirement, and thereby avoid this potential negative impact. Anyhow, there are some aspects of the erasure requirement that do impose a compliance burden on companies that use machine learning models or wish to do so, due to the lower amount of training data or the effort to obstruct de-anonymization of personal data. Thereby, this compliance burden will have a somewhat negative impact on machine learning models since the difficulty of using such models will increase. Nonetheless, this burden will be particularly significant initially until it has been clarified what measures that will be considered sufficient to ensure that an individual cannot be identified.

5.2.3 Timing

The time frame that companies have available to become compliant to a regulation is described as a double-edged sword for innovation incentives, in that too much time decrease innovation incentives, and too little time discourage innovation due to the extensive workload to become compliant (Ashford et al., 1985; Pelkmans & Renda, 2014).

In the case of the GDPR, the Legal experts agree that there has been enough time for companies to become compliant. However, some Industry actors describe that it has been stressful to become compliant with the GDPR for many companies within this industry. Indeed, the Industry actors point out that the GDPR has been a large project that has taken considerable time, although the companies with newer and more flexible systems have experienced a smoother transition. Hence, since the GDPR is not considerably different from the previous data privacy regulation PUL (Legal experts), but Industry actors perceive the GDPR as a large project, it confirms the Legal experts' statement that companies have not been fully compliant with PUL. However, some Industry actors perceive that there has been sufficient time to become compliant, but that there has been short of time to make adjustments in accordance with the regulation since it was first in December 2017 that the final version of GDPR was released. This is also recognised by one Legal expert to be a challenging aspect of the GDPR. Hence, when even legal professionals are uncertain what is required, it is no wonder that the industry practitioners are confused. In turn, this indicates that in the case of

GDPR it actually has been too short of time, rather than too low effort within companies to conduct necessary changes.

Regarding the impact that the timing of a regulation can have on innovation, it is explained that both too much time and too little time can be innovation-constraining (Ashford et al., 1985; Pelkmans & Renda, 2014), which makes it quite difficult to determine whether the timeframe of the GDPR has a positive or negative impact on innovation, and more specifically on AI applications. What is derived though is that there seems to have been enough time to become compliant considering that companies have known about GDPR for two years. However, it was described that the initial approved version of the regulation was perceived to be too unclear to start the adjustment process, which thereby could have restricted the possibilities of developing new solutions since it was not known what was required. Moreover, later on when the final version was released, companies perceived a time pressure, which could indicate that it was too little time to develop new solutions. This reasoning is also in line with Pelkmans and Renda's (2014), who describe that too short of time results in a too extensive workload and thereby constrain innovation. Thereby, these different time aspects seem to have decreased incentives to invest in new solutions, such as AI, rather than enhanced innovation. Nonetheless, in the case of the GDPR, it does not seem that it is a too long timeframe that is the reason for decreased incentives to invest in the development of new solutions, but rather inadequate information about what the final content would look like. In turn, it is indicated that there are further aspects to consider in regards to the timeframe than what previous literature has identified, that it is not merely the number of days from when a regulation becomes approved.

5.2.4 Flexibility

Concerning flexibility, previous research describes that the higher flexibility a regulation has, the more innovation-enhancing it is (Ashford et al., 1985; Pelkmans & Renda, 2014; Ranchordás, 2015). According to Ranchordás (2015), innovations are uncertain and continuously change, and therefore it is essential for regulations to be flexible to better fit with current technologies.

The flexibility characteristic is found to be more straightforward to assess for the GDPR as compared to the previously discussed characteristics. Foremost, the Legal experts and Industry actors describe the GDPR to be an "industry neutral" regulation which thereto is not adapted to specific technical characteristics. This description of the GDPR is in line with what is described as an outcome-based regulation, which has higher flexibility compared to prescriptive regulations that specify detailed requirements (Pelkmans & Renda, 2014; Ranchordás, 2015). Supporting this argument is the Legal experts' statement that the details of how companies should comply with the GDPR are not specified, and one AI expert also states that it does not exist any generic solutions to many of the technical problems of enforcing GDPR. This means that companies themselves to some extent can choose the solution they want to adopt to become compliant. For example, in Article 17 the right to erasure, it is specified that data should be erased upon request (with some exceptions), but

there is no further information about how this should be done. Hence, it is up to organisations to find the most suitable method for them to comply with this requirement.

Furthermore, it is stated that high flexibility in regulations could lead to more innovative methods to comply with the regulation (Ashford et al., 1985; Pelkmans & Renda, 2014; Ranchordás, 2015), which to some extent is indicated to be true for the GDPR. The AI experts mention that rule-based systems are suitable for, and is currently used, to ensure compliance with regulations. The AI experts further point out that rule-based systems and machine learning are beneficial to use to go through large sets of structured and unstructured data, which according to the legal experts both are included in the definition of personal data. Even though few such solutions seem to have been implemented at the current state, one example is the described administrative system within insurance that is based on algorithmic rules, meaning a rule-based system, which was implemented to continuously erase data. That few such solutions are identified could be explained by what was derived about the timeframe for the GDPR enforcement, that there has been too little time from when the final version was released and it became clearer what needed to be done to comply with the regulation. However, in a longer perspective, the GDPR may lead to new applications of rule-based and machine learning systems for complying with the GDPR, and could thereby have a positive impact on the future usage of AI technology.

5.2.5 Uncertainty
Ashford et al. (1985) and Pelkmans and Renda (2014) describe the uncertainty characteristic to be ambiguities in how to comply with a regulation, and that uncertainty can be both innovation-inducing and constraining. In this regard, both the Legal experts and Industry actors perceive the formulation of the GDPR to be clear and easy to understand but point out that the practical meaning of the requirements is somewhat unclear. Indeed, this seems to be a common perception considering the expressed confusion about how specific terms in the GDPR should be interpreted and what it means for AI applications (AI experts; Industry actors; Villaronga et al., 2017; Wachter et al., 2017a).

One such diffuse term is stated to be "solely automated processing" in Article 22 where it is questioned what level of human involvement that is required for the automated processing to be considered *not* solely. Wachter et al. (2017a) mean that only a low level of human involvement is sufficient for it to be considered *not* solely, while Malgieri and Commandé (2017) argue that the processing has to include *meaningful* human involvement for it not to be considered solely automated. In this regard, the Legal experts agree with the latter reasoning of Malgieri and Commandé (2017). In turn, this means that AI applications are negatively affected by this extensive interpretation of human involvement as it becomes more difficult to use fully automated processes, including both rule-based systems and machine learning. Although, it is noteworthy that there are many exceptions to this rule, and solely automated processing will still be possible to use if for example explicit consent is received from customers, or that it is necessary for fulfilling a contract (Art. 22, EU 2016/679). Considering the services that these financial companies provide to their customers, it is likely that companies often will be able to claim such a legal ground due to that contracts usually have to

be entered for the company to be able to provide the service. Thereby, the negative impact on AI applications is likely to be less severe. Additionally, since the Industry actors state that there are nearly no solely automated decision-making processes in use at the current state, the impact will be even less severe. Accordingly, it is neither likely that the second identified diffuse term, "legal or similar significant effects" in Article 22 (Wachter et al., 2017a), will have a significant impact on AI applications within this industry since it only becomes relevant if the decision-making process is entirely automated.

Furthermore, a third diffuse term is pointed out by Wachter et al. (2017a) to be whether Article 22 should be interpreted as a "prohibition" or as "a right for data subjects to object". In this regard, the Legal experts all agreed that it should be interpreted as a prohibition. Consequently, this will increase the administrative burden for companies because they will have to establish a legal ground to conduct automated processing, rather than being allowed to do it at all times besides when data subjects explicitly object to the processing. Since the Legal experts see it as a prohibition, this interpretation thereby creates a greater struggle to implement automated processes, which thereby has a negative impact on AI applications. Nevertheless, even though the Legal experts clarify how this term should be interpreted, there is still a possibility that the supervisory body's interpretation will be different. Hence, the impact could become more or less significant than what is believed, which is also true for the other mentioned diffuse terms in the GDPR.

Finally, another confusion is expressed by the AI experts and Industry actors to be what the technical meaning of "erasure" is. This kind of uncertainty has a negative impact on AI applications since it is difficult to develop solutions when it is unknown what actions are sufficient to be compliant. Thereby, even though flexibility in alternatives how to fulfil the erasure requirements was identified to be positive for AI applications, it also results in a negative impact.

Deriving from the above discussion, it can be established that there are several ambiguities about how the GDPR will be enforced, which according to both Pelkmans and Renda (2014) as well as Ranchordás (2015), reduces the incentive to invest in innovations. Hence this characteristic of the GDPR could constrain innovation, and thereby is likely to have a negative impact on AI applications. Indeed, the AI experts and Industry actors explain that some projects to develop AI have been postponed due to the uncertainty about the interpretation of GDPR's requirements. Thereby, AI applications will be negatively affected in the short term until best practice is established. Although, this high uncertainty in the GDPR can be explained by the Legal experts' statement that the regulation aims to be industry neutral and that more detailed specifications of the requirements would make it challenging for GDPR to fit all industries that process personal data. Thereto, it is expressed by the Legal experts that complementary guidelines on how interpretations should be made will be provided by supervisory bodies, why it is likely that the uncertainty will decrease when such guidelines are released, or best practice has been established. Hence, in the longer term when the uncertainties have become resolved, GDPR will not have a significantly innovation-constraining impact anymore.

It is stated that uncertainty around regulatory compliance can have a positive effect on innovations if companies explore different compliance alternatives to avoid that future regulations affect the firm negatively (Ashford et al., 1985; Pelkmans & Renda, 2014). However, in this study there are no indications that the uncertainty of how GDPR's requirement should be interpreted has resulted in increased efforts to elaborate on different ways to become compliant. Hence, the uncertainty characteristic can instead be said to be innovation-constraining, foremost because the uncertainties surrounding the GDPR have caused AI projects to become postponed. Although, this impact is most prominent in the short term until there are established court cases. Thereto, Ranchordás (2015) describes that when high investments are required into the development and implementation of technologies, and there also exists regulatory uncertainty, the innovation-constraining effect becomes more severe. In this regard, it is described by the industry actors that it is a substantial change to use technologies like AI to a greater extent for many companies within this industry. To that, many of these financially oriented companies have to acquire more technical expertise to be able to extend the usage of AI. Thereby, it is indicated that developing and implementing AI applications requires extensive investments, which further reinforces the negative impact on AI applications within this industry.

5.2.6 Utility
Besides the five characteristics that have been identified in the previous literature to determine how different regulations impact innovation, it has in this study been derived that an additional, sixth, characteristic also seems to contribute to the impact that the GDPR regulation has on AI applications: "utility". This utility characteristic refers to a potential for value-creation within businesses by complying with the GDPR requirements. It is a recurring topic mentioned by the Industry actors that the GDPR benefits customers. Indeed, since the GDPR imposes strict requirements on companies to establish safety standards in how they process personal data, and a trend exists where consumers are increasingly concerned with personal integrity (Legal experts), complying with this regulation is likely to increase customer value. In turn, by not just taking the necessary actions to become compliant with the GDPR, but to also go the extra mile to implement industry-leading privacy and security standards companies could establish a competitive advantage. This potential value creation could in turn increase incentives for companies to invest more resources into different machine learning solutions with a higher explanatory power to increase transparency in automated decision-making. In accordance, Goodman and Flaxman (2017) point out that the GDPR creates incentives for companies to develop machine learning models that have greater transparency, are easier to understand and include less bias.

Furthermore, the Industry actors point out that the GDPR's requirements of transparency in decision-making will lead to more meticulously designed automated systems, which could be both rule-based and machine learning. Hence, another potential source of an increased customer-value following from the GDPR requirements is increased accuracy of decisions that affect customers. In line with this reasoning, the Legal experts point out that the extensive informative requirement in the GDPR could decrease the risk of errors in the automated

system. In turn, this means that the GDPR will increase the pressure for investments in improving the quality of AI systems.

Considering the increased incentives for companies to invest in better AI models, particularly within machine learning, the utility characteristic of the GDPR can be argued to have a positive impact on AI applications in the longer term when investments yield returns. In accordance, the Industry actors perceive many benefits of AI, and therefore believe that the focus with the GDPR in force will be on how to make AI applications better rather than impeding its development and application.

5.3 Summary and Implications of the Findings

Evaluating the different characteristics of the GDPR in relation to AI applications within financial services, various findings are identified, which are visualised in Table 8 below.

Table 8. How the different GDPR characteristics impact AI applications within financial services.

GDPR Characteristics	Findings	Impact on AI Applications within the Financial Services Industry			
		Rule-based Systems Short-term	Long-term	Machine Learning Short-term	Long-term
1. Administrative Burden	Initially Significant	Somewhat negative	No impact	Negative	Positive
2. Compliance Burden	Moderate	No impact	No impact	Somewhat negative	Positive
3. Timing	Too short	Somewhat negative	No impact	Somewhat negative	No impact
4. Flexibility	High	No impact	Positive	No impact	Positive
5. Uncertainty	High	Negative	No impact	Negative	No impact
6. Utility	High	No impact	Somewhat positive	No impact	Positive

Firstly, it is derived that the GDPR initially imposes a quite significant administrative burden on organisations within the financial services industry since companies have to make organisational adjustments to become compliant. Thereby, this burden to some extent constrains applications of AI and its development in the short term, which becomes even more significant if machine learning applications are in place, or companies wish to implement it. Poel et al. (2014) point out that it is a priority within the EU to decrease the administrative burden imposed by regulations, but the result of this study indicates that EU has not achieved this objective in the case of GDPR. However, one reason for the high administrative burden could be that many companies have not been compliant with the previous data privacy regulation PUL, and therefore companies may perceive the burden

more significant than it would have been otherwise. Nevertheless, even though companies are initially negatively affected by the administrative burden of the GDPR, the changes that are made to comply with the regulation are found to create favourable conditions for machine learning applications. Thereby, the administrative burden in the case of the GDPR, can have a positive impact on machine learning solutions in the longer term.

Secondly, the compliance burden of the GDPR in an aggregated assessment is concluded to not have any significant impact on the rule-based form of AI applications. Nonetheless, machine learning solutions are somewhat negatively affected in the short term due to erasure requirements but positively affected in a long-term perspective due to higher pressure for explanatory models. This finding is in line with Blind's (2012) statement that the compliance burden can initially hinder innovation, while the effect becomes more diverse in a longer perspective. However, in contrast to what Ashford et al. (1985) state is the typical situation for regulatory impact on innovation, the compliance burden is not identified to be the GDPR characteristic that has the most significant impact on AI technology, at least not within the financial services industry. Noteworthy though is that the results might have been different if another industry had been examined.

In regards to the timing characteristic of the GDPR, it is derived that the timeframe until enforcement has been too short, and has thereby caused innovation to suffer due to the extensive resources allocated to ensure compliance with the regulation. In turn, at least in the short term, the timeframe has somewhat negative effects on AI applications, including both rule-based systems and machine learning. Furthermore, it is concluded that the GDPR has a high level of flexibility and therefore is likely to have a positive impact on both rule-based and machine learning applications in the longer term since it encourages companies to develop new technical solutions for compliance to the regulation. The findings from the assessment of the fifth characteristic, uncertainty, is established to have a negative impact on both rule-based systems and machine learning, at least in the short term since the many uncertainties surrounding the GDPR causes AI projects to become postponed. In addition to these characteristics, "utility" was identified to be a sixth aspect of the GDPR that positively impacts AI applications in the longer term. This positive effect mainly concerns machine learning applications due to the potential of establishing a competitive advantage, which in turn increases the incentives to invest in better performing models.

As demonstrated in Table 8 above, the different GDPR-characteristics have a very diverse impact on AI applications. In an aggregated sense, the impact that the GDPR has on AI applications within the financial services industry is summarised in Figure 5 below. In assessing the aggregated impact, a distinction can be made both between the different forms of AI, as well as the time perspective.
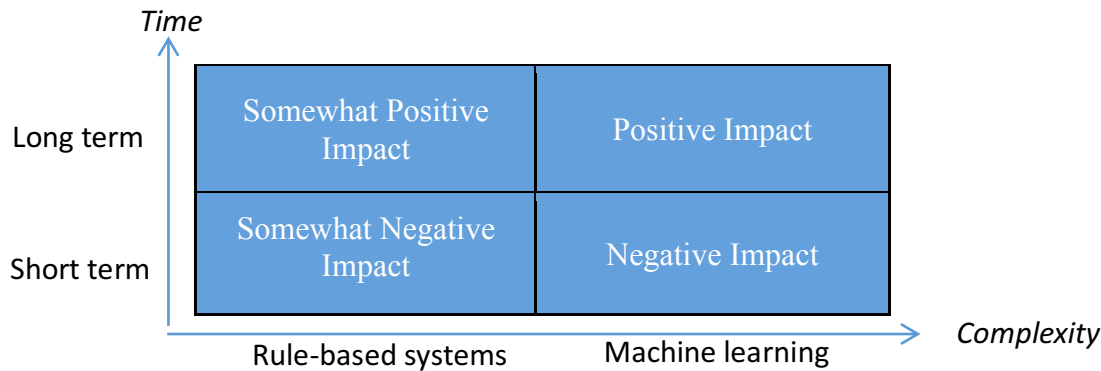
Figure 5. Long- and short-term impact of the GDPR on AI applications within financial services.

As is visualised in Figure 5, the GDPR is likely to have a somewhat negative impact on AI applications in the short term, and most significantly on machine learning. When considering a longer time perspective, the GDPR instead enhances incentives to invest in further applications of AI technology. This negative short-term impact could to some extent be explained by Ranchordás's (2015) argument that a strong regulatory focus on mitigating risks can cause innovations to become delayed. In regards to the GDPR, regulators seem to have focused on the risks connected to the processing of personal data and on protecting citizens, rather than the economic consequences of the regulation. Indeed, considering that the respondents mention that AI projects have been delayed, there are indications that some favourable business opportunities could be missed. Furthermore, the negative impact could also be attributed to the fact that AI technology develops in a more rapid pace than the regulatory process, and that the GDPR is not adapted to the recent developments in technologies (Kieselmann et al., 2016, AI experts and Legal experts). This line of reasoning is further supported by the arguments of Villaronga et al. (2017), who state that the GDPR could have an adverse effect on innovation if the regulation is incompatible with new technologies. On the other hand, Ranchordás (2015) states that this kind of discrepancy between regulation and technology is a common situation, indicating that this is a larger regulatory issue than merely in the specific case of the GDPR.

Finally, it has been established that the formulation of the right to erasure has caused confusions among technicians since it is not clear what kind of "erasure" that is referred to in technical terms, and thereby it seems as the GDPR to some extent is infeasible with AI technology. Hence, it is indicated that EU regulators have limited knowledge about AI technology, which is in line with Ranchordás (2015) and Villaronga et al.'s (2017) statement that there often exists an information asymmetry between regulators and innovators about complicated innovations. However, rather than a lack of knowledge, it could be that regulators have not been able to account for specific technical aspects since the GDPR applies to a wide range of industries. Either way, this confirms Villaronga et al.'s (2017) argumentation that there is a need for a greater understanding between law and technology, which in turn indicates increased importance of collaboration between the two parts in the future.

6. CONCLUSION

In this closing chapter of the report, the stated research question is answered by presenting the main findings of the study, and then closes with recommendations for future research.

6.1 Answering the Research Question

Recently, there has been considerable attention around the new data privacy regulation, referred to as the GDPR. This regulation aims to strengthen the rights of individuals and restricts the processing of personal data, that including automated processing. Despite the GDPR's good intentions, it has also been argued to constrain the usage of AI technology, a critique that mainly refers to the increased informative requirement in the occurrence of automated decision-making, as well as the data erasure requirement. However, there are different opinions how these requirements should be interpreted and what the practical implications will be for AI. One industry where AI applications have become increasingly adopted, and where companies also process extensive amounts of personal data is the financial services industry. Hence, this study aimed to answer the following research question:

"What is the potential impact of the GDPR on Artificial Intelligence applications within the financial services industry?"

AI technology is demonstrated to have many useful applications within various financial services, such as for credit evaluation, investments, customer service and fraud detection. In examining the impact that the GDPR has on AI applications within the financial services industry, it is derived that a distinction can be made between a short- and long-term perspective, as well as between the different forms of AI, which vary in level of complexity; rule-based systems and machine learning. The results are visualised in Figure 6 below.
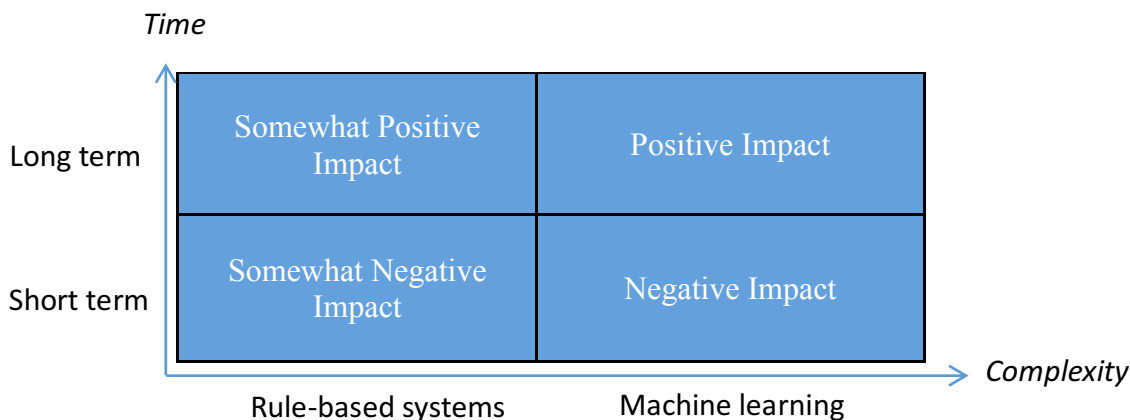
Figure 6. Long- and short-term impact of the GDPR on AI applications within financial services.

These findings were derived by conducting an initial assessment of the impact of six different characteristics of the GDPR in relation to AI applications. Thereafter, the impact of each characteristic was combined to assess the aggregated impact of the regulation, by taking a stance in the parts of the GDPR that have been argued to be problematic for AI. Firstly, the

GDPR initially imposes a significant administrative burden on organisations that takes both time and resources away from entrepreneurial activities. This administrative burden stems from the requirement of finding a legal ground to process personal data, and the need to restructure internal systems to be able to comply with the data erasure requirement. In turn, this additional workload has a negative impact on AI applications, particularly machine learning applications. Such applications require extensive amounts of data, and with the GDPR in force, it will require more effort by companies to collect and be allowed to use personal data for the specific purpose of training machine learning models. At the same time, these actions result in greater control of data within companies which is likely to have a positive impact on machine learning solutions in the longer term.

Secondly, the compliance burden that the GDPR imposes on companies within financial services is quite low. The reason for this is that there are nearly no *solely* automated decision-making processes in place, instead it is mainly used as decision-support and therefore companies are not affected by the extensive informative requirement of explaining automated decisions. Although, noteworthy is that even if *solely* automated decision-making would be conducted, it would not be a significant burden anyhow since it is quite easy to explain outcomes from the kind of AI models that are used in this industry. Regarding the rule-based systems, which is the most common form of AI in this industry, these are particularly easy to understand, but it also seems possible to derive sufficient information about decisions from the kind of machine learning models that are used. However, somewhat of a compliance burden is imposed on organisations using machine learning since it becomes more challenging to fulfil erasure requests, as well as a potentially decreased performance of machine learning models due to lower access to training data. While complicated machine learning models currently are not the primary form of AI within this industry, there are many benefits to gain from using such models. In turn, the GDPR is likely to incentivise companies to invest in research and development of more explanatory AI models. Thereby the compliance burden of the GDPR could have a positive impact on machine learning models when considering the long-term impact.

Thirdly, the timeframe until enforcement of the GDPR is concluded to have a somewhat negative impact on both rule-based and machine learning applications since the initial release of the regulation was too unclear to start adjustment processes, and there was too short of time when the final content was established. This limited period of time means that extensive effort is allocated to ensure compliance with the regulation and thereby innovation suffers. In regards to the fourth characteristic, flexibility, the GDPR is found to include a quite high level of flexibility, by specifying *what* should be achieved by organisations, rather than stating detailed prescriptions of *how* compliance should be achieved. In turn, this flexibility provides incentives to invest in new technical solutions to become compliant, such as AI, which is a technology pointed out to be advantageous to apply for achieving GDPR compliance.

Furthermore, in regards to the fifth characteristic uncertainty, the GDPR is concluded to have quite high uncertainty about what the practical implications of the requirements are, which in

turn causes projects of AI to be postponed. However, this effect will only be short-lived until court cases have clarified the ambiguities. Finally, besides the five characteristics of a regulation that are identified in the previous literature to impact innovation, an additional sixth characteristic is identified in this study to also contribute to the impact that GDPR has on AI applications, which is "utility". By going the extra mile to implement industry-leading privacy standards and automated decision-making processes with high accuracy and transparency, companies could establish a competitive advantage since these are aspects that customers increasingly value. This opportunity is likely to motivate companies to increase investments into better solutions, and the development of more transparent machine learning models with higher explanatory power.

In conclusion, as visualised in Figure 6 above, the aggregated impact of the GDPR on AI applications within financial services turned out to be dependent on the level of complexity of the AI models and thereby what form of AI that is used, as well as the time perspective. In the short term, the GDPR has a negative impact on AI, although the impact is less severe for the simpler applications based on rule-based systems. However, the GDPR's negative impact on AI is not as severe as suggested in the literature, at least not within the financial services industry. This finding is foremost explained by that this industry is not in the forefront of AI technology and mainly uses the simpler forms of AI, meaning that the industry is not affected by the more severe impact that GDPR has on highly complicated models. At the same time, companies struggle with the interpretation of several ambiguous aspects of the GDPR, especially concerning the use of technology. Hence, the insights of this research help companies to comprehend what the GDPR means for the usage of AI within their businesses. To that, the ambiguities surrounding technical interpretations also indicates that there is a knowledge gap between law and technology, which highlights a need for closer collaboration between regulators and technicians in the future.

On the other hand, in a longer time perspective, the negative impact on AI applications within financial services is likely to be near insignificant since companies then have had time to adapt to the new requirements, and the uncertainties surrounding GDPR have been resolved. This study showed that even for the companies within this industry that do use more complicated AI models, there seem to be ways to continue to use such models. Some methods identified are to anonymise personal data, or only use the data related to an individual that is not personal data. In turn, these findings show that the GDPR does not seem to be as unfeasible with technology as suggested in the literature. This finding is at least true for the kind of AI technology that is used within the financial services industry, and when a long-term perspective is considered. To that, this research found that there are many benefits to gain from applying AI within businesses, which makes it likely that companies will focus on developing better AI models rather than allowing the GDPR to impede its usage. In fact, the GDPR is concluded to have a positive impact on AI applications within financial services in a longer perspective. This positive impact is particularly significant for applications based on the more complicated machine learning AI models, which are extensively data driven. The reason for this is that companies achieve a greater knowledge of their data, and thereto the

regulation creates an increased pressure on financial services actors to use transparent AI models in automated decision-making processes. In turn, such pressure enables the development of AI technology to be directed towards models with an increased level of understanding, and thereby consumers will be ensured insight into how their data is processed and how decisions are made about them.

6.2 Future Research

In this report, a recurring topic was that complicated AI models are difficult to understand, and that bias could be included in these decisions. It is pointed out that there currently is no solution for how to ensure that bias is excluded from AI models. Therefore, it would be interesting to conduct research within this area of AI technology, that is, how to ensure that machine learning models have an objective representation of reality, and thereby include less bias.

Moreover, the findings of this study confirmed a previously identified gap between law and technology regarding information asymmetry as well as a discrepancy in the pace of regulatory and technology development. This finding indicates that this is a field of research that needs to be examined in greater detail. In this regard, it would be of interest to examine how legal and technical aspects can become more integrated.

Furthermore, this report has taken an industry perspective in how the GDPR could affect financial services, but it is possible that the impact varies between different organisations. Hence, it would be interesting to conduct a more in-depth case study of how the GDPR impacts AI applications on an organisational level. A case study would allow for a greater understanding how GDPR affects the usage of AI technology within a specific organisation. For example, in this study few fully automated decision-making processes were identified, but examining organisations in greater detail could reveal more extensive use of such processes.

Finally, since a sixth characteristic of the GDPR was identified in this study to contribute to how this regulation impacts AI technology, it would be interesting to examine if this finding can be generalised to other regulations and technologies as well. Similarly, it would also be of interest to examine if there are further regulation-specific characteristics that also matter for what impact a regulation has on technologies and innovation.

REFERENCES

Law text

Article 29 Data Protection Working Party Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679, WP 251, adopted on 3 October 2017 (cit. Art. WP 29).

The Regulation 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 (GDPR). (cit. EU 2016/679).

Literature

Ashford, N., Ayers, C., & Stone, R. (1985). Using Regulation to Change the Market for Innovation. *The Harvard Environmental Law Review: HELR, 9*(2), 419-466.

Bahrammirzaee, A. (2010). A comparative survey of artificial intelligence applications in finance: Artificial neural networks, expert system and hybrid intelligent systems. *Neural Computing and Applications, 19*(8), 1165-1195. doi:10.1007/s00521-010-0362-z

Blind, K. (2012). The influence of regulations on innovation: A quantitative assessment for OECD countries. *Research Policy*, *41*(2), 391-400. doi:10.1016/j.respol.2011.08.008

Bohanec, M., Robnik-Šikonja, M. & Kljajić Borštnar, M. (2017). Decision-making framework with double-loop learning through interpretable black-box machine learning models. *Industrial Management & Data Systems,* 117(7), 1389-1406. doi:10.1108/IMDS-09-2016-0409

Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press, UK.

Bryman, A., & Bell, E. (2011). *Business research methods*. Oxford University Press, USA.

Chiang, Enke, Wu, & Wang. (2016). An adaptive stock index trading decision support system. *Expert Systems With Applications*, 59, 195-207. doi:10.1016/j.eswa.2016.04.025

De Laat, P. (2017). Algorithmic Decision-Making Based on Machine Learning from Big Data: Can Transparency Restore Accountability? Philosophy & Technology, 1-17. doi:10.1007/s13347-017-0293-z

Ghahramani, Z. (2015). Probabilistic machine learning and artificial intelligence. *Nature*, *521*(7553), 452-459. doi:10.1038/nature14541

Goertzel, T. (2014). The path to more general artificial intelligence. *Journal of Experimental & Theoretical Artificial Intelligence, 26*(3), 343-354. doi:10.1080/0952813X.2014.895106

Gómez, J.A., Arévalo, J., Paredes, R. & Nin, J. (2017). End-to-end neural network architecture for fraud scoring in card payments. *Pattern Recognition Letters, 105,* 175-181. doi:10.1016/j.patrec.2017.08.024

Goodman, B., & Flaxman, S. (2017). European Union Regulations on Algorithmic Decision Making and a "Right to Explanation". *AI Magazine, 38*(3), 50-57.

Hämmerli, B. (2012). Financial services industry. In *Critical Infrastructure Protection. Springer, Berlin, Heidelberg,7130,* 301-329. doi:10.1007/978-3-642-28920-0_13

Jones, N. (2014). The learning machines: Using massive amounts of data to recognize photos and speech, deep-learning computers are taking a big step towards true artificial intelligence. *Nature, 505*(7482), 146-148. doi:10.1038/505146a

Kaplan, J. (2016). Artificial intelligence: think again. *Communications of the ACM*, *60*(1), 36-38. doi:10.1145/2950039

Kieselmann, O., Kopal, N., & Wacker, A. (2016). A novel approach to data revocation on the internet. *Lecture Notes in Computer Science (including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 9481*, 134-149. doi:10.1007/978-3-319-29883-2_9

Kingston, J. (2017). Using artificial intelligence to support compliance with the general data protection regulation. *Artificial Intelligence and Law*, *25*(4), 429-443. doi:10.1007/s10506-017-9206-9

Kluegl, P., Toepfer, M., Beck, P. D., Fette, G., & Puppe, F. (2016). UIMA Ruta: Rapid development of rule-based information extraction applications. *Natural Language Engineering*, *22*(1), 1-40. doi:10.1017/S1351324914000114

Lake, M.B., Salakhutdinov, R., & Tenenbaum, J. (2015). Human-level concept learning through probabilistic program induction. *Science (New York, N.Y.), 350*(6266), 1332-8. doi:10.1126/science.aab3050

Lake, M.B., Ullman, T.D., Tenenbaum, J.B & Gershman, S.J. (2017). Building machines that learn and think like people. *Behavioral and Brain Sciences 40,* 1-72. doi:10.1017/S0140525X16001837, e253

LeCun, Y., Bengio, Y. & Hinton, G. (2015). Deep Learning. *Nature. 521*(7553), 436-444. doi:10.1038/nature14539

Littman, M. (2015). Reinforcement learning improves behaviour from evaluative feedback. *Nature, 521*(7553), 445-451. doi:10.1038/nature14540

Malgieri, G. & Comandé, G. (2017). Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation. *International Data Privacy Law, 7*(4), 243-265. doi:10.1093/idpl/ipx019

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S. & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society, 3*(2), 1-21. doi:10.1177/2053951716679679

Nowell, L., Norris, J., White, D., & Moules, N. (2017). Thematic Analysis. *International Journal of Qualitative Methods, 16*(1), 1-13. doi:10.1177/1609406917733847

Pelkmans, J., & A. Renda. 2014. Does EU Regulation Hinder or Stimulate Innovation? Available at: https://www.ceps.eu/system/files/ No 96 EU Legislation and Innovation.pdf

Poel, K., Marneffe, W., Bielen, S., Van Aarle, B., & Vereeck, L. (2014). Administrative simplification and economic growth: A cross country empirical study. *Journal of Business Administration Research*, *3*(1), 45-58. doi:10.5430/jbar.v3n1p45

Ranchordás, S. (2015). Innovation-friendly regulation: The sunset of regulation, the sunrise of innovation. *Jurimetrics Journal of Law, Science and Technology, 55*(2), 201-224.

Rouse, W., & Spohrer, J. (2018). Automating versus augmenting intelligence. *Journal of Enterprise Transformation,* 1-21. doi:10.1080/19488289.2018.1424059

Sathya, R., & Abraham, A. (2013). Comparison of Supervised and Unsupervised Learning Algorithms for Pattern Classification. *International Journal of Advanced Research in Artificial Intelligence, 2*(2), 34-38. doi:10.14569/issn.2165-4069

Tecuci, G. (2012). Artificial Intelligence. *Wiley Interdisciplinary Reviews: Computational Statistics*, *4*(2), 168-180. doi:10.1002/wics.200

Van de Gevel, A. J., & Noussair, C. N. (2012). *The Nexus between Artificial Intelligence and Economics*. Heidelberg: Springer.

Van Otterlo, M. (2014). Automated experimentation in Walden 3.0: The next step in profiling, predicting, control and surveillance. *Surveillance & Society, 12*(2), 255-272.

Villaronga, E. F., Kieseberg, P., & Li, T. (2017). Humans forget, machines remember: Artificial intelligence and the right to be forgotten. *Computer Law & Security Review, 34*(2), 304-313. doi:10.1016/j.clsr.2017.08.007

Wachter, S., Mittelstadt, B., & Floridi, L. (2017a). Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation. *International Data Privacy Law, 7*(2), 76-99. doi:10.1093/idpl/ipx005

Wachter, S., Mittelstadt, B., & Floridi, L. (2017b). Transparent, explainable, and accountable AI for robotics. *Science Robotics, 2*(6). doi:10.1126 scirobotics.aan6080

Websites
Datainspektionen, 2018a
https://www.datainspektionen.se/lagar-och-regler/personuppgiftslagen/strukturerat-eller-ostrukturerat/
[Accessed 2018-05-08]

Datainspektionen, 2018b
https://www.datainspektionen.se/lagar-och-regler/personuppgiftslagen/
[Accessed 2018-05-18]

EU GDPR, 2018a
https://www.eugdpr.org/gdpr-faqs.html
[Accessed 2018-02-12]

EU GDPR, 2018b
https://www.eugdpr.org/key-changes.html
[Accessed 2018-02-03]

McKinsey, 2017
https://www.mckinsey.com/~/media/McKinsey/Industries/Advanced%20Electronics/Our%20Insights/How%20artificial%20intelligence%20can%20deliver%20real%20value%20to%20companies/MGI-Artificial-Intelligence-Discussion-paper.ashx
[Accessed 2018-04-28]

Oliver Wyman, 2017
http://www.oliverwyman.com/content/dam/oliver-wyman/v2/publications/2017/apr/future-proofing-privacy.pdf
[Accessed 2018-05-07]

PWC, 2017
https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf
[Accessed 2018-01-28]

APPENDIX

Appendix 1 - Keywords Used in the Systematic Literature Review
When conducting the Systematic Literature Review, the following keywords were used. The columns with "second" and "third" key words are included to show the searches that included more than one keyword.

Table 9. Keywords used in the systematic literature review.

First keyword	Second keyword	Third keyword
Artificial intelligence		
Machine Learning		
Rule-based learning		
Supervised learning		
Unsupervised learning		
Reinforcement learning		
Financial service		
Artificial Intelligence	Risk*	
Artificial Intelligence	Challenge	
Artificial Intelligence	Regulation	
Artificial Intelligence	Risk*	Financial service
Artificial Intelligence	Challenge	Financial service
Artificial Intelligence	Regulation	Financial service
Artificial Intelligence	Insurance	
Artificial Intelligence	Bank*	
Artificial Intelligence	Loan*	
Artificial Intelligence	Credit*	
Artificial Intelligence	Invest*	
Regulation	Innovation	
Regulation	Technology	
GDPR	Artificial Intelligence	
GDPR	Machine Learning	
GDPR	Financial services	

Appendix 2 - List of Respondents and Interview Details

Table 10. List of respondents and interview details.

Title	Company	Date	Duration	Language	Channel
AI Experts					
Academic AI Expert 1. *Professor in computer science and conducts research about AI.*	Swedish Technical University.	2018-03-05	55 minutes	English	Face-to-face
Academic AI Expert 2. *Associate professor within AI.*	Swedish Technical University.	2018-03-05	30 minutes	Swedish	Face-to-face
Academic AI Expert 3. *Professor in computer science and conducts research about AI.*	Swedish Technical University.	2018-03-08	55 minutes	English	Face-to-face
Business AI expert 1. *Assistant manager data analytics.*	Consulting firm active within the Swedish market.	2018-03-07	1 hour 15 minutes	Swedish	Telephone
Business AI Expert 2. *CTO - Machine learning developer.* Business AI Expert 3. *Data Scientist - Machine learning developer.*	Swedish company focusing on Machine Learning.	2018-03-08		Swedish	Face-to-face
Business AI expert 4. *Software engineer.*	Threat Intelligence company using AI and machine learning.	2018-03-08	30 minutes	Swedish	Face-to-face

Legal Experts					
Legal expert 1. *Legal counsel working with GDPR.*	Swedish industrial company.	2018-03-01	1 hour 10 minutes	Swedish	Telephone
Legal expert 2. *Legal associate working with GDPR.*	Swedish law firm.	2018-03-15	50 minutes	Swedish	Telephone
Legal expert 3. *Legal associate working with GDPR.*	Swedish law firm.	2018-03-22	60 minutes	Swedish	Face-to-face
Financial Services Industry Actors					
Financial Services actor 1. *Client relationship manager.*	Swedish Investment company. Founded within the past two years.	2018-03-08	40 minutes	Swedish	Telephone
Financial Services Actor 2. *Chief Executive Officer.*	Swedish company providing AI-based credit ratings. Founded within the past two years.	2018-03-09	50 minutes	Swedish	Face-to-face
Financial Services Actor 3. *Chief Technical Officer.*	Swedish company providing AI-based credit ratings. Founded within the past two years.	2018-03-09	30 minutes	Swedish	Face-to-face
Financial Services Actor 4. *Release manager.*	Swedish bank. Founded more than two years ago.	2018-03-14	30 minutes	Swedish	Telephone
Financial Services Actor 5. *Data Scientist.*	Swedish bank. Founded more than two years ago.	2018-03-15	45 minutes	Swedish	Face-to-face

Financial Services Actor 6. *Head of Digital Solutions.* Financial Services Actor 7. *Senior Product Analyst - Responsible for GDPR.*	Swedish Fintech company developing business systems and digital solutions for the pension and life insurance industry. Founded more than two years ago.	2018-03-20	1 hour 30 minutes	Swedish	Face-to-face
Financial Services Actor 8. *Head of Information, Strategy and Architecture.*	Swedish bank. Founded more than two years ago.	2018-03-20	40 minutes	Swedish	Face-to-face
Financial Services Actor 9. *Chief Product Owner.*	Swedish bank. Founded more than two years ago.	2018-03-21	30 minutes	Swedish	Telephone
Financial Services Actor 10. *Machine Learning Specialist.*	A Swedish company within financing solutions. Founded more than two years ago.	2018-03-20	30 minutes	Swedish	Face-to-face
Consultant 1 within Financial Services. *Director Financial Services Advisory.* Consultant 2 within Financial Services. *Manager Advisory Services. Involved in GDPR projects.*	A Management Consulting Firm active in Sweden, where Financial Services is one of their business areas. Founded more than two years ago.	2018-03-19	1 hours 30 minutes	Swedish	Face-to-face

Consultant 3 within Financial Services. *Senior manager in the business area financial services.*	A Management Consulting Firm active in Sweden, where Financial Services is one of their business areas. Founded more than two years ago.	2018-03-19	30 minutes	Swedish	Face-to-face

Interview Guide- AI Experts

Background information
1. What is your position within the company you currently work at?
2. What is your experience of Artificial Intelligence?

General information about AI
3. How would you describe what Artificial Intelligence is?
4. How would you describe the difference between Artificial General Intelligence (AGI) and Narrow AI?
5. What kinds of AI systems are most commonly applied today (e.g. rule-based or machine learning)?
   a. What kind of approach to AI do you perceive to be the most important for the future development of AI?
6. In brief, what are the main methods that are applied for machines to learn (for example supervised learning)?
   a. What are the key differences between these different learning methods?
   b. To your knowledge, what are the possibilities of creating AI models that learns from new data without storing it?
7. What are the most commonly applied learning algorithms within Artificial Intelligence today?
   a. In what kind of applications within financial services have Artificial Intelligence been adopted?

Erasing data
8. What are the possibilities of erasing data from different Artificial Intelligence algorithms?
   a. If data is deleted, what are the consequences for the machine's performance?
9. The new data privacy regulation GDPR contains an article that gives individual data owners the right to have their personal data erased given certain conditions. How can a balance be attained between this legal requirement and an Artificial Intelligence model's need to remember information that has been used to train it?

Explaining automated decisions
10. How would you describe the level of transparency in Artificial Intelligence models?
    a. How do you believe that increased transparency in Artificial Intelligence would affect performance of machine learning applications?
11. How would you describe the possibility for humans to understand how decisions are made within Artificial Intelligence?
    a. And also how the degree of interpretability differs between different learning techniques?

12. The GDPR seem to create an increased awareness of individual data privacy rights, and therefore imposes an increased pressure on organizations to handle personal data more restrictive, how do you believe that this will affect Artificial Intelligence applications?

    a. And how do you think the GDPR will affect the future development of Artificial Intelligence?

13. Do you have anything else to add that you consider to be relevant for this topic?

Interview Guide- Legal Experts

Background information

1. What is your role within the company you currently work at?
2. How are you working with the GDPR?

General aspects of the GDPR

3. What would you say counts as personal data in the GDPR?
   a) What is the time frame in regards to how long personal data can be stored legally?
4. What do organizations in general have to do to become GDPR compliant?
   a) Do you know anything of specific importance for the financial services industry, such as bank and insurance?
5. What articles pose the most significant challenges for organisations to become compliant?
   a) To your knowledge, how can the potential challenges be resolved?
6. How would you describe the administrative burden in terms of time and resources that GDPR imposes on organisations?
   a) What is the perceived cost of complying with the GDPR? (do companies for example have to make changes to their internal systems or develop new technology to comply with the regulation).
   b) Considering the scope of GDPR, what is the perception about the time-frame?
7. In your perception, how much flexibility do organisations have in the approaches they apply to fulfil the requirements of GDPR?

Explaining automated decisions

8. What is your opinion about the degree of explicitness of the definitions in the GDPR?
   a) In regards to the right to be informed about *automated* decision-making, stated in paragraphs in Articles 13-15, how would you say that the term *"meaningful information about the logic involved"* should be interpreted?
      I. Would you say that subjects have a right to get an explanation about how their data is being processed in the existence of automated decision-making?
         • *If yes, at what point in time would the subject have the right to get an explanation in relation to a certain decision (before, after or both)?*
         • *If yes, how extensive explanation would be required? (That is, merely overall system functionality or details of how a specific decision has been made).*
9. What are the potential consequences of the recitals in the GDPR?
   a) What do you consider the effects to be of the guidelines on *Automated individual decision-making and Profiling for the purposes of Regulation 2016/679* provided by the Article 29 Working Party?

      b) More specifically, what effect do you think it will have that a right to *explanation* in regards of automated decision-making is only explicitly mentioned in Recital 71?

      c) How would you say that the content in Article 22 should be interpreted, as a prohibition or a right to object (i.e. is it prohibited for the data provider to make an automated decision when certain criteria are apparent, or do the individual data owner have to make an actual objection for it to apply).

10. In Article 22 it is expressed that it only applies to decisions that are based *solely* on automated decision-making. How should "solely" be interpreted in regards to the level of human intervention that is allowed for this article to apply?

      a) In the same Article (22), it is also stated that this right only applies if it "*significantly* affects him or her". How is it determined what is "significant"? (I.e. do the subject has to prove that it affects him or her significantly?)

Erasing data

11. What do you think the consequences of Article 17, the "Right of Erasure", will be in practice?

      a) In your opinion, to what extent do you believe that individuals will request to have their data erased?

The main consequences of the GDPR from an AI perspective

12. In your opinion, what are the implications that the regulation has several vaguely defined expressions?

      a) What possibilities do Sweden have to add additional legal requirements that complements GDPR?

13. What do you think will be the main effect of GDPR enforcement in the short term?

      a) And the effects in the long term?

14. How would you say that the GDPR requirements fit with today's increasing use of automated systems?

      a) How do you think that the GDPR will affect the usage of automated processes?

15. Do you have anything else to add that you consider to be relevant for this topic?

Interview Guide - Industry Actors

Background information

1. Could you briefly describe what services the company you work at offers?
2. What is your role within your company?

General information about Artificial Intelligence applications in the financial services industry

3. To your knowledge, in which applications are automated processes that are based on Artificial Intelligence used within financial services?
4. What kind of Artificial Intelligence systems do you perceive are the most commonly applied (e.g. machine learning, rule-based systems, neural networks)?
   a) What potential do you see for Artificial Intelligence within the financial services industry in the future?

Explaining automated decisions

5. To your knowledge, in what extension are personal data included in the automated processes used within financial services?
6. In the situations where automated processes are used, to what extent would you say these involve human influence?
7. In your perception, what are the possibilities for humans to understand the overall functionality of the systems in the automated processes that are based on Artificial Intelligence?
   a) How would you describe the possibility to understand how specific decisions have been made (i.e. to find out exactly what factors that have been considered in the specific decision)?

Erasing data

8. What do you consider to be the main changes that organizations have to implement to become compliant with the new EU General Data Privacy Regulation (GDPR) that is coming into force this May?
   a) In your perception, what parts of the GDPR requirements poses the greatest challenges for organizations in your industry?
      I. *How do you perceive that these challenges could be resolved?*
   b) How do you perceive the administrative burden in terms of time and resources that the GDPR imposes on organizations within the financial services industry?
   c) How do you perceive the time-frame for becoming compliant with the GDPR requirements until its enforcement date?
   d) How much flexibility do you perceive that companies have in the different approaches that they can choose among to become GDPR compliant?
9. In your experience, to what extent do your organization receive requests from external individuals that want you to erase their personal information that your organization is storing?
10. The GDPR seem to create an increased awareness of individual data privacy rights, and therefore imposes an increased pressure on organizations to handle personal data

more restrictive. How do you believe that this will affect the future usage of Artificial Intelligence solutions within the financial services industry?

11. Do you have anything else to add that you consider to be relevant for this topic?