MASTER THESIS

Course code:
BE323E

Candidate name:
Anthony Christopher Caffrey

The impact of machine learning on the Norwegian legal industry

Date: May 29th 2020

Total number of pages: 96

NORD University

www.nord.no

Abstract

Rapid advances in computational capacity combined with ever more advanced statistical models has over the course of only a few years revolutionised the way many tasks are performed. The subset of Artificial Intelligence known as *machine learning* is solving all manner of tasks such as diagnosing cancer, predicting the future value of stocks, and replacing human call-centres with chatbots. The algorithms behind machine learning even allow for the extraction of semantics from written text, paving the way for "legal tech" performing tasks that until recently required lawyers. This has resulted in various predictions of loss of jobs in the legal industry, including predictions of industry disruption.

This MBA Thesis explores how Norwegian law-firms are responding to machine learning, with an emphasis on the digital transformation of the industry, what the future of the industry may look like, and how law-firms are moving towards an ever more digital future.

The most important findings are that Norwegian law-firms are still in the early stages of utilising the potential of Information Technology. Though some previously manual tasks have been simplified through partial automation, there is limited evidence of lawyers utilising machine learning software. The industry is still in early days of adopting technology that would simplify the performance of repetitive tasks. However, in the narrow field of document review and electronic discovery, machine learning is making a significant impact. There are also indications of potential industry fragmentation in the low end of legal complexity.

Foreword

This MBA Thesis at Nord University / Nord Business School was submitted on May 29th 2020. The Master of Business Administration programme at Nord University is a 1,5 year post graduate education that includes the following subjects; Applied method, social economics, leadership, project management, strategic management, economic analysis, marketing and entrepreneurship, organisational management, accounting and analysis, corporate finance and an MBA Thesis.

I am a jurist with fourteen years of experience in the practical application of the law in both the Private and Government sectors. When I first heard of the breakthroughs in machine learning, I dismissed the technology as irrelevant to my profession. I wish to thank my relentless cousin Carl Eidsgård (formerly Oracle, now Microsoft) for telling me that I was wrong, and my friend Peter Flem, MsC, whom I met at the Ukrainian 2018 Lviv IT Arena, for explaining the technology in layman's terms.

I wish to thank the attorneys who graciously lent me their insight and time:

Partner Kyrre W Kielland with Advokatfirmaet Ræder AS

Partner Eivind Grimsø Moe with Advokatfirmaet Hjort DA

Partner Nicolai Halbo and Chief Digital Officer Peter van Dam with Advokatfirmaet Simonsen Vogt Wiig AS, and

Specialist counsel Hanna Beyer Olaussen and senior attorney Emily Mary Weitzenboeck, PhD, with Advokatfirmaet Wikborg Rein AS

Science and academia have always eluded me, and I would probably have ended up writing an opinionated op-ed if not for the questions, advice and guidance from my supervisor, professor Bjørn Willy Åmo. Thank you for nudging me in the right direction.

And finally, thank you Kristin and Olivia, for understanding that dad needed to spend evenings, weekends and a few holidays reading and typing.

Anthony Christopher Caffrey

May 29h 2020

Index

1.0 The subject of and approach to research

1.1 The research problem and research questions

The research problem is the exploration into the impact of *machine learning* on the Norwegian legal industry.

Machine learning is a programming method where computer software is designed to solve problems based on having learned the rules for solving them. This differs from traditional programming where computer software solves problems based on rules given by human programmers. Machine learning isn't a new technology as such, but the massive increases in computational power over the last decade or so has provided machine learning with a renaissance. As long as there is enough data to learn from, or an environment to interact with, machine learning provides super-human analytical prowess and/or automation of repetitive tasks well enough as to replace humans completely. The stunning performance of machine learning software has led to consultancy firms and experts predicting that machine learning will replace humans in a great many tasks, causing significant unemployment and industry disruption – even in the legal industry.

However, it is one thing to task a computer with comparing pictures of a tumour in order to predict if the tumour looks survivable or not, and quite another to understand ethics and the deeper semantic meaning of words. Since the days of Cicero, applying the law has been a highly creative task (see subchapter 3.6.3.1), understanding and planning the possible approaches to solving a legal problem, manoeuvring within the constraints and possibilities of the law, and then ultimately arguing the matter in front of decision-making humans so that they may favour a specific outcome. At first glance it doesn't seem likely that a brick of silicone and plastic could be of any other use to the legal industry, than as a somewhat glorified typewriter that allows for the storage of letters on the C:\ drive rather than in a traditional filing cabinet. And yet, a significant loss of jobs and possible industry disruption has been predicted in the legal industry.

The legal industry itself is especially diverse, ranging from one-man firms dealing with a broad range of legal areas, to very large firms specialising in very large and complex matters within narrow fields; the clients ranging from single individuals to multi-billion Pound corporations and Governments. Some firms are still getting used to rudimentary computer software, while other firms have a highly specialised IT strategy. This poses the problem of how to research this relatively new phenomenon within the scope of an MBA

thesis in such a way that the results have generality, i.e. that there is a functional relationship between the individual respondents / informants, and the industry in general. This thesis assumes that approaching very large law-firms are most likely to result in results that have generality. An initial exploration into the technology showed that developing sufficiently complex machine learning software as to be of any use in the legal industry, is a very expensive process that commonly requires large volumes of data to learn from. Large firms draw their experience and financial clout from hundreds of employed lawyers and clerks, giving them a superior ability to invest in expensive projects. In addition, large firms generally work internationally, in English, removing the language barrier that might otherwise curb the introduction of English language machine learning software into the Norwegian legal industry. These considerations and more indicated that the largest firms were most likely to have relevant information about the impact of machine learning, and that these findings would have generality to the entire Norwegian legal industry.

A qualitative approach was selected because there are so many unknown factors. A quantitative approach might tell us more about what many firms think, but the questions would be limited to what one was able to divulge from studying other sources. Within the scope of an MBA thesis, a qualitative approach seemed favourable because it would reveal more first-hand information about what is going on in the industry.

Against this backdrop the overall problem of this thesis can be generalised into the question of what impact machine learning is having on the legal industry.

This overall problem can be further divided into three specific research question:

1. How are law-firms responding to the machine learning phenomenon?
2. What do law-firms believe the future role of machine learning in the legal industry to be?
3. What are law-firms doing to be a part of that future?

1.2 What we know thus far about the research questions

Machine learning and business strategy are well researched academic fields taught in universities and other accredited institutions around the World. There are countless publications, textbooks, peer-reviewed articles etc, and the challenge is not finding relevant material, but condensing it within the scope of an MBA thesis. In short: we know that business strategy is the planning for long term profitability/sustainability, and we know that

machine learning works, how it works and why it in many areas provides super-human analytical and predictive prowess.

There is limited research into the Norwegian legal industry, on how lawyers perform their tasks, what their tasks comprise of, what skills they need to perform those tasks, and the tools and techniques they use. There is equally limited research on "legal tech", i.e. a buzzword that generally comprises a broad range of software used by the legal industry and possibly other industries. Research on the legal industry focuses mainly on the composition of its employees and owners, profitability, invoicing practices etc, and this is of limited value; it tells us that the industry mainly bills by the hour, that the hourly rate does not reflect an objective quality standard, and that the need for legal aid is not sufficiently met by the industry [i.e. that there is an untapped demand for legal services in Norway]. There seems to be a potential for growth in the industry, especially in relation to consumer law, small claims etc.

To summarise; Machine learning works, but differences in approach and understanding are, to some degree, fuelling both hype and resignation that may represent an uncertainty in relation to business strategy. Machine learning thus needs to be explored. Business strategy is well researched and is also covered in a later chapter. The limited relevant research on the legal industry is explored, cross referencing it with other research/theory in order to explain what lawyers do, why machines currently are far from performing the creative and intellectual tasks required for jurisprudence, why machines are incapable of statistically replicating the work of lawyers, and which explorative or repetitive time-consuming legal tasks machine learning is well apt to perform.

1.3 The reasons for doing this research and target group

The most important reason for doing this research, is that there is little research into the legal industry and the potential impact of machine learning. This thesis aims to fill a gap in the research and to provide a basis for future research into the industry.

The target groups for this thesis are Norwegian lawyers and the legal-tech industry, i.e. computer software developers attempting to exploit machine learning in software intended for the legal industry. In addition, this thesis may be of interest to future students, the Norwegian bar-association, and possibly the media.

1.4 An introduction to the relevant research method

The **object** is to find out what law-firms are doing to exploit the opportunities and avoid the threats of Machine Learning legal-tech, and to explore this technology in such a detail that a valid analysis can be performed.

The **design** is phenomenological, i.e. an exploratory qualitative study. Parts of the study relate to well-established academic fields, while other pars are of an explorative nature, and the data is triangulated to improve understanding.

The **data sources** are interviews with six informants from four of the largest law-firms in Norway.

The **studies reviewed** are a combination of articles into the legal profession and industry, general and narrow articles and textbooks dealing with Artificial Intelligence, and articles and textbooks about business strategy.

The **main results** are that the findings indicate that:

- law-firms aren't adapting competitive strategies to exploit the possible opportunities presented by machine learning. This confirms the theories of Clayton Christensen and Jay Barney.
- the industry is still heavily invested in a pricing strategy of billing by the hour.
- the speed and extent of future development will decide if the technology is disruptive, as the industry seems to be walking into the trap Christensen calls "the innovators dilemma".
- the current impact of machine learning is limited, but that there are many tools that utilise machine learning models.
- machine learning is having a potentially significant impact on electronic discovery and document review.

As a **conclusion:**

A strong case has been made that machine learning tools are impacting the legal industry and will continue to do so, over time reducing the current requirements of manually performing the various routine tasks in a law-firm. Machine learning tools are already assisting lawyers in discovering errors and potential risks within narrow areas with a higher degree of precision/quality than human lawyers are capable of on their own, and certainly

speeding up the processes. However, the complexity of the legal industry, its environment and the creativity required to perform the tasks of lawyers is an important limiting factor.

This thesis provides a preliminary exploration into various areas that could be researched in the future.

The degree of impact rests with the ability to develop software that becomes more broadly applicable, beyond the narrow tasks currently served. Future research is recommended to deal with this problem.

If the software being developed becomes broadly applicable, spanning many or all of the tasks currently performed by lawyers and augmenting the way lawyers perform those tasks, the industry isn't adapting its strategies to cope. According to the theories of Christensen, within the framework of Porter, and the theories of Barney, there certainly is a potential for industry disruption. Advantages in machine learning software will not primarily be gained by the legal industry, but by software developers, rival industries and new entrants. Future research is recommended to deal with this problem, because the data isn't sufficient to draw definite conclusions.

1.5 Structure of the thesis

Chapter 1 presents an overview of the research problem and method.

Chapter 2 explores the nature of the legal industry and applying the law, the predictions of impending disruption, the statutory requirements for practicing law in Norway, and finally segmenting the work of lawyers into specific tasks.

Chapter 3 deals with machine learning and how it works, how we may understand the levels of technology in terms of Artificial Intelligence, definitions of some of the terms encountered, forces possibly limiting and forcing the introduction of machine learning into the legal industry, and attempts to apply this understanding to the specific tasks of lawyers explored in the previous chapter.

Chapter 4 explores current legal tech, both reference works, various tools for automation, advanced tools for performing complex tasks, and software that directly competes with lawyers.

Chapter 5 presents a business strategy framework for understanding the industry forces that may be affected by machine learning, and then explores what strategy theory predicts that the legal industry will do when faced with potentially disruptive innovation.

Chapter 6 deals with the research method and why confidence may be placed in the results.

Chapter 7 deals with the findings.

Chapter 8 analyses the findings.

Chapter 9 presents conclusions to the research questions.

1.6 Significant challenges and how they were overcome

The main challenge was understanding machine learning, not only on a conceptual level, but actually understanding how predictions could be made by training models by applying algorithms to large amounts of data, and then exploring the limitations. I spent several months simply reading about machine learning.

Another challenge was to find a relevant approach to researching machine learning and the legal industry within the scope of an MBA thesis. After dismissing exploring the specific strategies of selected law-firms, I decided on performing a broader exploration into the industry structure and the forces likely affecting how law-firms view and implement machine learning, and the risks involved.

Covid19 presented further challenges regarding finding informants and performing the interviews, but a sufficient number of informants were interviewed as to give the findings a sufficiently perceived generality.

However, further research is recommended in order to follow up this preliminary research into an under-researched field.

2.0 The legal industry

2.1 Introduction

Dramatic predictions into the future of the legal industry have arisen because of a recent renaissance in machine learning, a computer technology that doesn't follow rules made by human programmers, but rather learns its rules from experience. In order to explore whether or not the predictions have merit, we first need to explore which skills a lawyer require in order to perform his/her tasks, and subsequently which specific tasks the lawyer commonly performs.

2.2 Predictions of impending disruption of the legal industry

If you were a weaver in the late 18th century, working your profession on a traditional vertical loom, and someone told you that a steam-driven power loom was under construction and that it would weave patterns based on punch-cards, you might disregard it as pure fiction. How could a complex task such as weaving be replaced by machines? And then, one day, you'd wake up, realising that your entire industry had been replaced by machines. Is the same thing happening to the legal industry?

In 2014 the Israeli American company LawGeex began developing an Artificial Intelligence (AI) for spotting legal issues in non-disclosure agreements (NDA's). They started by using unsupervised learning algorithms to teach the AI- system the core legal language of NDA's. Then a convolutional neural net was used to train and fine-tune the system through supervised learning, and finally a unique augmentation algorithm was applied to boost the learning of the system, according to Professor Yonatan Aumann at the Bar Ilan University[1]. In February 2018 the AI was pitted against 20 very experienced lawyers, and the speed and accuracy of discovering possible legal issues in 5 NDAs was compared. The LawGeex AI used **26 seconds and achieved a 94% accuracy rate**. The 20 very experienced lawyers spent an average of **92 minutes and achieved an 85% accuracy rate**. In other words, the LawGeex AI apparently outperformed expert lawyers with not only a massive increase in speed, but also with a significant boost in quality.

In May of that same year, professor of corporate- and finance law at Tilburg University in The Netherlands, Eric Vermeulen, presented the LawGeex demonstration to Norwegian lawyers at the annual Norwegian Bar-Association conference "Fagdager". According to Vermeulen "*the legal industry will change completely*" and "*this scares the wits out of the lawyers*" [2]. According to Vermeulen this causes some lawyers to resign themselves to the disruptive nature of machine learning AI. Other lawyers seem to believe that they are

technologically relevant by knowing how to use "*Outlook, Word and maybe Excel*", implying that they are, in fact, already irrelevant.

In a 2017 Harvard Business Review cover-story, MIT-researchers Erik Brynjolfsson and Andrew McAfee hailed machine learning as *the* most important general-purpose technology of our age and wrote: "*The status quo of dividing up work between minds and machines is falling apart very quickly. Companies that stick with it are going to find themselves at an ever-greater competitive disadvantage compared with rivals who are willing and able to put [machine learning] to use in all the places where it is appropriate […]. [AI] can achieve superhuman performance […] and their impact will be profound."* [3].

At the same time, two Facebook "*AI-robots developed their own language*" creating much hype, with "*droves of articles predicting an oncoming robotic revolution*" [4].

> *"I am putting myself to the fullest possible use, which is all I think that any conscious entity can ever hope to do."*
>
> - *HAL 9000*

A few years prior, in 2016, Peter Nussey, Vice President with a company called ThoughtRiver in the UK, wrote an interesting Linked-in article on recent predictions into machine learning replacement of lawyers[5]. The article deals with a 2015 study by McKinsey[6] and a 2015 study[7] by Dana Ramus with the University of North Carolina School of Law, and Frank S. Levy with Massachusetts Institute of Technology asking, "Can robots be lawyers?".

McKinsey & Co suggests that 69% of paralegal/clerical work in the legal industry and 23% of lawyer work in the legal industry, is automatable. Levy and Remus suggests that the figure for lawyers lies between 13 and 23%. Correspondingly, in 2016 Deloitte released a study[8] which predicted that 39% of jobs in the legal industry would become obsolete as a result of implementing machine learning, a specific type of Artificial Intelligence, in that industry. It seems, as Peter Nussey writes, that the main difference is that Levy and Remus believe that "*several tasks are too opaque and complex or require a level of emotional intelligence that means they won't be automated anytime soon*". McKinsey takes another approach, "*believing that the speed with which advances in artificial intelligence and machine learning, challenge our assumptions about what is automatable"*. Levy and Remus point out that the main challenge with the assumptions is that there is a failure to engage with technical details that appreciate the capacities and limits of software, and that there isn't enough data on how lawyers divide their time amongst various tasks.

The Norwegian Association of Lawyers (Juristforbundet), an organisation who represents around 20,000 jurists employed by the government, the various municipalities, private organisations including law-firms, and by the courts as judges, performed a November

2019 study into the current level of AI in the workspace. The findings are reported in a November 15th 2019 article in the magazine "Juristen"[9]. According to the Juristforbundet around 1 in 10 jurists outside law-firms state that some level of AI has been applied in the workspace, and another 1 in 10 state that the employer is considering implementing AI. Within law-firms a total of 3 in 10 state that they are currently implementing or consider implementing AI. And around 3 in 10 state that Robotic Process Automation, RPA, is in use in their workspace. RPA is the automation of certain processes that usually don't require decision-making. However, 3,9% state that RPA is performing advanced tasks. It is worth noting that 39% of respondents state that their knowledge of technology is limited, 13% state that they have no knowledge of technology, and 37% state that they have some knowledge of the technology; the 3,9% that stated that RPA is performing advanced tasks may not be qualified to differentiate between advanced tasks and pure automation of tasks that do not require decision making. Private sector lawyers evaluate their technical expertise the highest, while law-students report a poor level of AI understanding. Around 20% of all respondents believe that AI will have a significant impact within the next five years, according to the study.

If machine learning does have such an impact on the legal industry, it may be disruptive, i.e "*ability to create new markets and/or replace leading actors*" [10]. This ought to concern the industry. And because the predictions claimed that these changes would occur during the next ten years, we're half-way there; we should be able to witness machines making a serious impact on the legal industry.

In order to answer the questions that these demonstrations and studies beg, we need to examine what is done in the legal industry, what makes jurisprudence and legal advice unique, which component subtasks are required to give legal advice, and the degree to which machines can perform parts of those tasks. This includes exploring current legal-tech, i.e. computer software designed in order to assist lawyers in their work, and some of the limitations of the technology.

2.3 Statutory requirements for practicing law in Norway

The legal industry comprises offices of legal practitioners known in Norway as "*advokater*", i.e. a lawyers, attorneys or counsellors-at-law primarily engaged in the practice of law – that is, giving legal advice to clients who are not their customers (as opposed to

giving legal advice as a result of an employment contract). The legal industry does not include lawyers/jurists employed in the public sector, special interest organisations etc.

In order to practice law in Norway one is required, amongst other things, to be educated in both the body of laws and academic method for solving legal questions. According to the Courts Act (domstolsloven) § 220 first section there are two degrees that are considered sufficient: the current five-year *Legum Magister/Magistra* master's programme, and the recently discontinued six-year professional degree in law *Candidatus/Candidata Juris*. The degrees are comparable to the post-graduate professional doctorate *Juris Doctor* degree in the United States.

The three Norwegian universities that are authorised to award the degrees state that the master's degree is a degree in *Rettsvitenskap.* The English translation of the Norwegian word *Rettsvitenskap* is "legal science", or, in a more precise translation, "jurisprudence"; the philosophy, science and study of law, according to Spaak (2003)[11] and Blacks Law Dictionary[12].

According to the Code of Conduct laid down in the Justice Department regulations given in accordance with the Courts Act (domstolsloven) chapter 11, also known as "Advokatforskriften" chapter 12, a lawyer is a hands-on practitioner, an ambassador for whomever his or her client (or employer) may happen to be, applying the law to whatever questions the client may have, and assisting the client in reaching whatever goals are possible within the scope of the law. Correspondingly, the Norwegian legal education is not an education in philosophy and ethics, though the education includes a few days of compulsory lectures on ethics, but rather structured as a practical approach to solving legal problems. An example is the curriculum at the largest Norwegian legal faculty, at the University of Oslo[13], where the five-year education focuses on learning the valid academic method for solving legal questions and the various relevant legislation in the various fields of law that are considered necessary for someone to be skilled in the law.

The difference between theoretical and practical law is a matter of some debate, raising valid questions about why such a practical profession should be approached as an academic abstraction. An excellent 1933 article by Jerome Frank in the University of Pennsylvania Law Review[14] explores this problem. Frank writes: "*no sane person will deny that a knowledge of the rules and principles, of how to 'distinguish' cases, and of how to make an argument as to the true ratio decidendi of an opinion [but] the tasks of the lawyer do*

not pivot around these rules and principles." Frank goes on to write that that the two tasks of the practicing lawyer are:

(1) attempting to predict the outcome of a future enforceable court decision, and
(2) attempting to induce the desired outcome for his or her client.

This description is very similar to the Norwegian Code of Conduct, and these two tasks are as relevant today as in the day of Marcus Tullius Cicero; a client approaches the lawyer with some grievance or concern, commonly caused by a layman, a public institution or some legal requirement, and the lawyer explores the options that may be open according to the law, where prediction of possible enforceable outcomes lie at the core. In layman's terms, if this goes to court, what will the outcome be, and what may we do to make sure that the outcome is the outcome the client wants? Advocacy, from which the Norwegian term for a practicing lawyer is derived, is a broad term that, in essence, means to support someone; a lawyer could be considered a professional supporter, someone who knows what will happen if the case goes to court, offers counsel [i.e. legal advice] and pleads for someone [i.e. argues the case], see Blacks Law Dictionary.

2.4 Legal subsumtion and legal method

As explained in the previous subchapter, one of the main tasks of a lawyer is to attempt to predict the possible verdict of the courts, and advice his or her client accordingly. This is done by examining the relevant sources of the law, and applying them to the case at hand through a process called *legal subsumption*; the letter of the law (i.e. the relevant legislation) is applied to the case, and compared to the intention of the law as described in the many documents and minutes from the various government departments and the legislative assemblies, prior court rulings, possible treaties and other direct or analogous sources, private regulations, legal theory etc. These documents are called *the sources of the law* and provide insight into how the law should be applied in different situations. The authority of statute law, regulations, circulars et.c, commonly referred to as *legislation*, is organised in steps depending on the proximity of the legislation to the Constitution (the *lex superior*), the level of detail or vagueness (*lex specialis*) and the recency of the source (*lex posterior*). The process of applying the various sources of the law to a specific case, is called *legal subsumption*. When performing legal subsumption, a valid legal argument is reached by applying weight to the various sources of the law; If a highly detailed rule is given in recent degree, higher weight may be attributed to the specifics of that decree, giving that rule

primacy over vague and possibly antiquated rules in statute law or the Constitution. Understanding the basics of *legal method* lies at the core of legal education.

The above description of the sources of the law is based on the Civil law system, where codified (written) law is the prime source of law. The legal system in Norway is Civil law, with heavy emphasis being given to statute law. In other legal systems, such as the English Common law system, in general the system of England and all her former colonies, "the law" is a set of sources that give precedent for future decisions. Thus, case-law is treated somewhat differently based on the legal system. While Civil law places emphasis on prior decisions because of principles such as equality in front of the law, the prime source of the law is statute law. However, in Common law jurisdictions, a prior decision can be a prime source of law in itself. Regardless of whether the legal system is Civil or Common law, a junior court will attempt to follow the decisions of superior courts in order for the junior courts decisions not to be overturned.

When applied correctly, the sources of the law and various principles for applying the law to a specific case, allow for the skilled jurist to predict how the system of courts will interpret and apply the law. Armed with this knowledge, the lawyer will be able to advice the client on how to alter the details of the case so as to increase the likelihood of achieving the desired outcome. This is the fundamental part of lawyering.

There are countless legal textbooks explaining these matters in detail, one being §§ 1-5 in "Knophs oversikt over Norges rett"[15]. A somewhat more extensive approach to legal method is given by Professor Erik Boe in his "Grunnleggende juridisk metode"[16]. *Legal method* is a highly researched academic field.

A significant number of the Norwegian sources of the law are available in digital format via Lovdata, see subchapter 4.2.1. This includes all statute law, a large number of circulars, some legal theory, special publications and decisions by the various courts.

2.5 Segmenting the work of lawyers into individual tasks

We have thus far discussed how the lawyer applies his or her significant understanding of both the sources of and the principles for applying the law to specific cases, giving advice on how to proceed in a legal matter in order to maximise the chances of an opportune outcome. If machine learning is to have an impact on the legal industry, it must be capable of performing parts of these tasks in a way that reduces the time a lawyer spends on the same

tasks. This poses the obvious question of what tasks a lawyer typically performs during his or her day at work.

An obvious place to begin looking for a list of tasks would be The Norwegian Bar Association and The Royal Norwegian Ministry of Justice. However, this turns up a blank. A 2015 Norwegian Public Examination (NOU)[17] into the legal profession contains extensive elaborations on how the licensing of lawyers should be regulated, but states on page 100 that it has no intention of presenting an exhaustive summary of the tasks normally performed of lawyers. The Norwegian Bar Association [Den norske advokatforening] presents an annual report[18] on lawyer fees but does not delve further into the individual tasks of the lawyer; it only concerns itself with the types of case and average salaries. And a The Norwegian Bar Association annual exploration of the industry[19] equally does not present us with any list of tasks. A 2019 inquiry[20] into the reimbursement of lawyers' fees where the State provides free legal aid, performed by Manon Economics and Oxford Research on behalf of The Ministry of Justice, fails to explore the various tasks of the lawyer.

The 2019 study by Juristforbundet mentioned in chapter 2.2 sheds some light on some of the tasks lawyers believe are most likely to be affected by AI:

Casework: 56%
Administrative tasks: 75.8%
Court rulings: 3.2%
Simple legal advice: 47.8%
Due diligence: 18.3%

The list of tasks is not particularly detailed, and only lawyers that work with mergers and acquisitions need to perform a due diligence analysis.

The US Department of Labour *Occupational Information Network (O*NET)*[21] defines 22 different tasks typically performed by a lawyer. The list seems generally applicable to Norway, but there are several overlapping tasks, and a condensed list of typical lawyer tasks is shown below.

- Interpret the law.
- Predict the probable outcomes of cases.
- Prepare evidence.
- Argue a case in court and question witnesses and opposite parties.

- Advise clients on legal matters relating to business transactions, liability, advisability of prosecuting or defending lawsuits and other legal rights and obligations.
- Represent clients before government agencies and opposite parties.
- Study the sources of the law in order to determine ramifications for the case.
- Prepare or draft legal documents, such as wills, deeds, patent applications, mortgages, leases, and contracts, and documents relating to court proceedings such as subpoenas and appeals.
- Review documents, such as wills, deeds, patent applications, mortgages, leases, and contracts, and documents relating to court proceedings such as subpoenas and appeals.
- Negotiate settlements of civil disputes.
- Evaluate findings and develop strategies to uphold the client's interests.
- Tasks relating to notarising various legal positions.
- Perform administrative and management functions.
- Act as agent, trustee, guardian, or executor for businesses and individuals.

In the study by McKinsey introduced in chapter 2.2 the following lawyer-tasks were identified. It may seem as McKinsey missed out on the obvious tasks related to actually analysing the law and predicting outcomes of cases, and the complexity of those tasks, but these may be implicitly included in the listed tasks:

- Arbitrate disputes
- Draft legislation or regulations
- Evaluate information related to legal matters in public or personal records
- Identify implications for cases from legal precedents or other legal information
- Interview claimants to get information related to legal proceedings
- Meet with individuals involved in legal processes to provide information and clarity
- Prepare documentation of proceedings
- Prepare legal documents
- Provide legal advice to clients
- Represent the interests of clients in legal proceedings
- Research relevant legal materials to aid decision making

- Supervise activities of other legal personnel

2.6 Summary

Dramatic predictions have been made over the last 4-5 years about impending AI disruption of the legal industry due to the renaissance of a novel technology known as machine learning. In this chapter was explored some of those predictions, and in order to analyse the actual current impact of that technology on the Norwegian legal industry, the skills that are required in that industry and the tasks the industry performs, was also explored. The specific tasks will be compared to the technology itself in chapter 3.0.

3.0 Machine learning and why it's a game changer

3.1 Introduction: Why we need to explore the technologies

For the purpose of this thesis we could simply establish that there is a technology called machine learning, and that computers are learning how to solve some intellectual problems better than humans because of it. The interviews will likely not explore the finer points of machine learning or computer science terminology. However, a short and non-exhaustive overview of the terminology and basic concept of machine learning may be required in order to analyse the findings correctly. In addition, we should also explore the most important factors possibly limiting the development or application of machine learning. These assumptions can be tested against the findings.

> *"We find everywhere men of mechanical genius, of great general acuteness, and discriminative understanding, who make no scruple in pronouncing the Automaton a pure machine, unconnected with human agency in its movements, and consequently, beyond all comparison, the most astonishing of the inventions of mankind."*
>
> - "*Maelzel's chess-player*" by Edgar Allen Poe

3.2 Definition of Artificial Intelligence

Artificial Intelligence or just AI for short, is a term we will encounter when examining the technologies used in the legal industry.

When discussing Artificial Intelligence outside the field of computer science, we soon run into the problem of mythology and science fiction. From the ancient Greek poem about the beautiful *Galatea*, carved from ivory by Pygmalion of Cyprus and brought to life at the behest of the goddess Aphrodite, to HAL-9000 in the 1967 film by Arthur C Klarke and Stanley Kubrik. The myth may give the impression that AI somehow must be artificial thinking in the human sense. However, realised AI is so-far either rule-based arithmetic and logic operations, or advanced statistical models that infer the rules for performing arithmetic and logic operations. The feeling of what AI is, based on sci-fi and mythology, disconnects us from a real understanding. There are several interesting studies of this disconnection, e.g "Portrayals and perceptions of AI and why they matter"[22] by researchers at The Royal Society. And there is a possibility this problem will be encountered during the course of this thesis.

Human intelligence is defined as "*the ability to acquire knowledge, to think and reason effectively, and to deal adaptively with the environment*" according to Passer and Smith (2007)[23]. When John McCarthy and fellow applicants coined the phrase "Artificial intelligence" in a 1955 proposal for a summer conference on machine simulation of intelligence[24], they deliberately avoided defining intelligence, focusing instead on machine

simulation of intelligence. Wooldridge and Jennings (1995) [25] also avoid defining intelligence as a capacity of the mind, stating instead that "*an Agent is intelligent"* when it is capable of solving a problem in a way that maximises its potential for success. However, this creates all manner of problems, because even single cell amoeba are capable of maximising their potential for success [26],[27],[28]. Poole, MacWorth and Goebel (1998)[29] run into the same problem when defining AI as *"…any device that perceives its environment and takes actions that maximize its chance of successfully achieving its goals."*.

Kaplan, Andreas and Haenlein (2019) [30] are much more specific when they define AI as *"...a system's ability to correctly interpret external data, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation."*. This approach to AI seems quite similar to the approach taken by the highly influential United States Department of Defence agency DARPA [Defense Advanced Research Projects Agency]. DARPA defines AI in terms of complexity by segmenting AI into four aspects or categories[31];

- The ability to perceive,
- the ability to learn,
- the ability to form abstractions, and
- the ability to reason.

Through these four aspects of intelligence DARPA has segmented AI development into three waves. And we are currently in the early stages of the second wave.

3.3 First wave; transforming the World by following rules

DARPA defines the first wave AI as: *"engineers create sets of rules to represent knowledge in well-defined domains. The structure of the knowledge is defined by humans. The specifics are explored by humans"*. And this is where your computers are at, both at home, at work and in your hand provided you own a smartphone.

Digital computers have transformed the World by following specific rules on how to perform the arithmetic and logic operations required to generate output data from input data. Computers sent man to the moon, and computers allows us to write letters in Microsoft Word ®, send emails via Microsoft Outlook ® and play a vast number of highly realistic 3D computer-games. We can surf the Internet and access almost every imaginable piece of information there is. All because of digital computers following the specific rules laid down by humans. And over time humans have built on

prior programming, creating ever more advanced programming languages that simplify the complexity of the rules.

Digital computers following specific rules were able to defeat the World Chess Champion in 1997. The IBM computer *Deep Blue* defeated Gary Kasparov by combining the rules of chess (i.e. the rules on how each individual piece moves and a scoring system for each piece) with an alpha-beta algorithm that examined every possible outcome 6-20 moves ahead, and then deciding on the move that gave the highest likelihood of returning a superior score, according to Hsu and Campbell (1995)[32]. By exploring hundreds of millions of moves per second and choosing the move that gave the highest chance of success, the computer program defeated Kasparov.

However fantastic these things are, the main constraint of first wave AI is that computer programmes are bound by the limits of the rules envisioned by human programmers. If there are better rules than those envisioned by human programmers, the computer would neither know nor care; it would happily carry on performing the arithmetic and logic operations humans have told it to do.

3.4 Second wave: moving beyond the constraints of human rules

Machine learning is a technology that allows computer software to build predictive models based on experience from data rather than specific instructions given by human programmers. "*Clever new learning algorithms, combined with the availability of large training data sets and the relentless advance of computing horsepower"* are, according to Pillow and Sahani (2019) [33], driving computers into what DARPA defines as the second wave of AI. Computers are moving beyond the limitations of specific instructions. The technology is called "machine learning"; an old term likely coined by IBM engineer Arthur Samuel (1959)[34].

Machine learning is a technology that allows computer software to build predictive models based on experience from data rather than specific instructions given by human programmers.

There are two general approaches to machine learning.

- With *Classical learning* the computer extracts [or *mines*] latent rules that explain the relationship between sets of *labelled* or *unlabelled* data, and
- With *Reinforcement learning* the computer learns a problem-solving strategy through interaction with an environment in which it is placed.

3.4.1 Classical learning

The first approach to machine learning is classical learning. Classical learning is a statistical approach to modelling methods for organising information in some way or another. "*Machine learning algorithms use computational methods to "learn" information directly from data without relying on a predetermined equation as a model. The algorithms adaptively improve their performance as the number of samples available for learning increases"*[35]. This allows the machine learning model to predict or infer the output of new data based on a mathematical model built from a prior analysis of training data input, and is arguably the driving technology behind the Natural Language Processing technology being used by lawyers today.

An important aspect of classical learning is the Manifold Hypothesis. The Manifold hypothesis is a hypothesis that higher dimensional data, such as a random symbol, a picture, a matrix, a sentence or set of words etc, can be learned by examining the data that lie near the low dimensional variations [manifold] of the higher dimensional data. By separating manifolds, i.e. variants of the higher dimensional data, a prediction of the higher dimensional data can be made, as explained by Fefferman, Mitter and Narayanan (2016)[36]. This allows learning from training data and applying that learning to previously unseen data, and correctly predicting what it is.

Classical learning is either classification, regression, hard clustering, soft clustering, dimensionality reduction or association rule learning. Classical learning is either labelled or unlabelled, i.e. that the various datapoints that are used for training either have the correct categories pre-defined or not. Classification and regression require labelled data, and the rest do not.

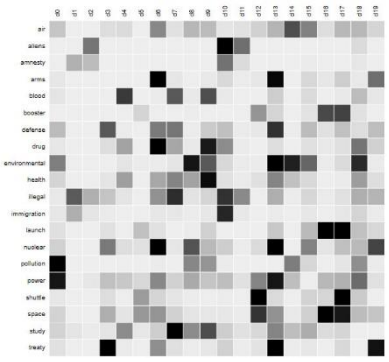
- *"Classification refers to the problem of identifying a category to which an input belongs to among a possible set of categories.",* according to Rebala, Ravi and Churiwala (2019) [37]. This could be used for identifying pieces of text based on the occurrence of words and phrases in that text, classifying the document based on algorithms such as the naïve Bayes, ibid page 62. This is also the case with "Bail Algorithms" deciding whether to incarcerate indictees ahead of trial [i.e. *varetektsfengsilng*], see Buskey and Woods (2018)[38]. Classification algorithms can even be used to transform complex decisions into decision trees that resemble human decision making, simplifying a decision-making process, according to Rebala, Ravi and Churiwala (2019) page 89.

- Regression analysis can be used to organise datapoints into categories based on the association between dependent and independent variables, ibid page 25. This could be used for risk analysis, predicting the effects of penalties or regulatory proceedings, see Fisher 1980[39].
- Clustering is a vector quantization method where the probability of data being one thing or another is based on clustering of similar data, based on its centroid point, se Burton, Shore and Buck (1983)[40]. Hard clustering forces a datapoint into one cluster or another, while soft clustering allows for a datapoint do exist in several clusters. This is obviously highly relevant in the natural language context of the law; semantics are a high dimensional conceptual space constructed from lexical co-occurance, see Cao, Song and Bruza (2004) [41]. Soft or "fuzzy" clustering "*is able to effectively retrieve semantic concepts, even from highly imbalanced datasheets"* according to Chen, Shyu and Chen (2016)[42].
- Dimensionality reduction is a method of reducing the number of random variables by exploring the features of the principal variables, see Roweis and Saul (2000)[43], and is particularly relevant in natural language processing (NLP), se Yoav (2016)[44]. The latent meaning of the words can be derived from the occurrence of words in documents, according to Deerwester, Dumais, Furnas, Landauer and Harshman (1990)[45] and how different words are used in similar context, see Blei, Ng and Jordan (2003)[46]. A covariance matrix can be created from the principle components of the occurrence matrix, i.e the matrix of how words occur in different documents, see Rebala, Ravi and Churiwala (2019) page 145. A captivating illustration of this process is illustrated in a facsimile below from the Univeristy of Koblenz-Landau, see figure 1.
- Association rule learning is a method of discovering significant associations between datapoints, see Agrawal, Imielinski and Swami (1993)[47] page 207. This approach was kicked off back in 1993 when Agrawal, Imielinski and Swami found a way to predict which items that you were likely to buy based on the occurrence of similar items in supermarket receipts, and this technology is one of the reasons why various webpages are able to suggest that you buy products or view specific information that seems to eerily occur with your interests, see Scime (2005)[48] page 282. An obvious application for legal tech would be to review documents and discover what was missing.

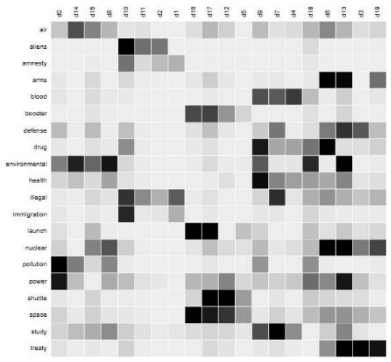
While it is true that these approaches are *simply* statistical models applied to numerical and lexical data, this is what is outperforming expert lawyers, see subchapter 2.2.

Figure 1

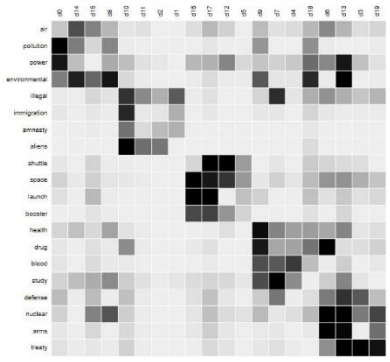
Extracting contextual meaning from the occurace of words based on the documents that contain the words. Univeristy of Koblenz-Landau[49].



The first matrix' shows the words on the vertical axis and the various documents on the horisontal axis.



In this matrix the words are rearranged based on document similarity.



And in the final matrix term similarity is grouped.

3.4.2 Reinforcement learning

The second approach to machine learning is reinforcement learning. This is programming a computer to do what comes naturally to us; to learn from experience. It is based on a sequential decision-making process model called the Markov decision process (MDP), see Rebala, Ravi and Churiwala (2019) page 223. The observer or learner in an MDP is the concept of a decision maker called the *intelligent agent*. The agent exists in an *environment* and makes some observation the available steps it can take at every timestep. With some understanding of the alternative steps the agent decides which states of the environment to change. The change of the state may result in a *reward*, either immediate or delayed, see Watkins (1989)[50]. This process is iterated until the agent learns the entire environment, or some or a sufficient approximation of it, every state the environment can be in, all the actions the agent can take in the environment and all the rewards the agent can receive from taking those actions. The agent's goal is to maximise its rewards, and learns how to do this through iteration, i.e. running simulations over and over again. The sequence of states, actions and rewards forms a trajectory that allows the agent to maximise the cumulative rewards over time, see Rebala, Ravi and Churiwala (2019) page 223 and Szita and Szepesvari (2010)[51].

3.4.3 Neural nets, deep learning etc

Classical and reinforcement learning can be performed in various ways. One way is to use a computer program to build approximations of algorithms and mathematical models. Another approach is to use a computer program that simulates the algorithm in a so-called neural network. The latter approach is commonly called *deep learning in artificial neural networks*.

Artificial neural networks are feed-forward networks in which information is fed forward through layers of nodes. The nodes are triggered in various degrees, not only 1 and 0, and can be arranged to simulate any logic gate. In other words, a artificial neural network consisting of dozens of layers of thousands of nodes and perform immensely complicated tasks. And the learning is organised through a system of back-propagation, in which the nodes are augmented by a system of weights and biases. See Rebala, Ravi and Churiwala (2019) pages 123-125.

3.4.4 Ensembling

A machine learning model may combine several or many dozen different algorithms depending on the complexity of the task. This is called *ensembling* and improve fitting; A model is fit when it is sufficiently trained to fit the data correctly. According to Rebala, Ravi, and Churiwala (2019) page 151 overfitting usually occurs when the data is insufficient to correctly define the features of that data, i.e. that the model may predict that a datapoint belongs to several categories. Correspondingly underfitting occurs when the model is incapable of correctly defining a category because the model has not been sufficiently trained on previous data, see Everitt and Skrondal (2010)[52] and Rebala, Ravi and Churiwala (2019) pages 95-96.

Ensembling methods are stacking (training algorithms to learn from different types of algorithms), bagging (having several algorithms vote on the outcome) and boosting (having algorithms learn from misidentification), see Rokach (2010)[53].

3.5 Terms commonly encountered

3.5.1 Big data

A common form of large datasets is Big Data, a term coined in 2000 by economist Francis Xavier Diebold[54] at the University of Pennsylvania to describe dealing with *the "explosion in the quantity (and sometimes, quality) of available and potentially relevant data, largely the result of recent and unprecedented advancements in data recording and storage technology"*. The term "Big Data" has evolved as a theoretical concept, it's essence being that

Big Data is *"Information that is characterised by the vast volume of the information, the variety of the information (numerical, textual and images) and the velocity by which the information is updated, requiring specialised technology and analytical methods to transform it to value"* according to De Mauro, Grexo and Grimaldi (2016) [55], see also Irahim et al. (2015) [56].

Machine learning does not require big data in order to learn, but Rebala, Ravi and Churiwala (2019) page 3 point to big data being one of the most important drivers of the rapid development of machine learning.

3.5.2 Data mining

Another common term is data mining, i.e. the horsepower hungry part of processing data. As Han and Kamber (2001) [57] explain, data mining is not extraction of data itself but the extraction of patterns from that data.

3.5.3 Digitisation, automation, machine learning and knowing the difference

Whether a computer application is a first or second wave AI according to the definitions above, it requires information in digital form. According to the Cambridge Dictionary, digital[58] means that the information is stored in *bits* of 0s and 1s, so that it is accessible by a computer. A bit is a portmanteau of the word's *binary* and *digits*, see MacKenzie (1980) [59]. Information stored in digital format can be accessed by many people at once and re-used with ease, in the same way as this thesis is written on a template that is reused by hundreds or even thousands of students.

Using a traditional typewriter instead of writing by hand increased the speed in which information could be transferred. Using a digital computer instead of analogue technology such as a typewriter allows for complex editing, re-use of information etc, and is an obvious improvement over typewriters and pen and paper but tells us nothing of whether there is machine learning involved.

When information is digital it is easily transferrable. When registering a business in the National business registrar, *Foretaksregisteret i Brønnøysund*, one would previously have to fill inn several extensive documents by hand, writing in capital letters so that the clerk reading the form and registering it didn't misinterpret the documents. The documents were sent by regular mail, and when the registration was official, a reply was sent in return by regular mail. The process could take weeks. Today, however, one simply logs in to a website www.altinn.no, uses some form of identification, typically the passcode generator Bank-ID,

and accesses an electronic form. Using the computer, you type in the various details, and shortly after you receive an email conforming the registration. The clerks in the other end likely only control the information received, and do not have to enter the information into their IT-system; it is done automatically, saving both the individuals wishing to register information and the registrars a lot of time. This kind of automation is an obvious improvement over physical documents, regular mail, and physical/analogue handling of the information, but tells us nothing of whether there is machine learning involved.

It is worth nothing that while many tasks can be automated, the major limitation of document automation is controlling that the end result is correct. Machine learning is changing that, as seen in the short reviews of software in chapter 4. It can be particularly hard to establish if machine learning has been integrated into various forms of automation, but there are likely two ways of finding out. First, machine learning and Artificial Intelligence being buzzwords, the software producers are likely to advertise the fact. Secondly, machine learning is likely integrated into the automation when the process involves having to make decisions.

It is in itself interesting to examine whether or not the legal industry has been able to exploit rudimentary digitisation and automation, or whether they still swear to pen, paper and the dictaphone.

3.6 Possible limitations of machine learning

3.6.1 Data requirement

The most obvious limitation of machine learning, it it's need for data, see e.g. Rebala, Ravi and Churiwala (2019) page 3. The internal documents of a law-firm are bound by various confidentiality constraints, and law-firms are unlikely to have informed clients that their cases may be used for training an AI (though there may be exceptions). In addition, there are various ethical constraints regarding the use of such documentation, such as the EU General Data Protection Regulation and the Code of Conduct introduced in subchapter 2.3 place strong limitations on the use of data. In addition, even the largest law-firms in Norway may not have enough samples for a classical machine learning system to be trained. And as Donaldson (2015) points out[60], data is in itself useless; you still need humans to make sure the machine learning model is correctly trained.

In subchapter 2.4 we saw how lawyers derive valid arguments. The legal environment is highly complex, the rules are not clearly defined, and the scoring system is complex. A reinforcement machine learning system can't learn to navigate the legal environment in the

same was as it can learn to navigate a 3D environment or a game with set rules such as Chess or Go. As we shall see in chapter 4, the practical application of machine learning legal tech is, however amazing, quite narrow. Having above looked closer at the technology itself, it seems unlikely that a machine learning model will be capable of running simulations of cases any time soon, severely limiting the AI to what they can be taught through direct interaction over some time, or by analysing large sets of data. See also subchapter 3.6.4 below.

3.6.2 Cost

An obvious limitation, at least for the time being, is cost of developing machine learning tools. A 2018 paper in *Natural Language Engineering* by Robert Dale[61] is listed the various funding gathered in order to set op the machine learning tools mentioned in the paper:

https://casetext.com	– funding US $ 30.8 million
https://rossintelligence.com	– funding US $ 13.1 million
https://vlex.com	– funding EURO € 4.0 million
https://www.exterro.com	– funding US $ 100.0 million
https://www.csdisco.com	– funding US $ 50.9 million
https://www.everlaw.com	– funding US $ 34.6 million
https://www.relativity.com	– funding US $ 125.0 million
https://kirasystems.com	– funding CA $ 65.0 million
https://www.seal-software.com	– funding US $ 43.0 million
https://www.lawgeex.com	– funding US $ 21.5 million
https://www.leverton.ai	– funding EURO € 15.0 million
https://ebrevia.com	– funding US $ 4.3 million
https://www.eigentech.com	– funding UK £ 13.0 million
https://www.legalsifter.com	– funding US $ 6.2 million
https://www.luminance.com	– funding US $ 13.0 million
https://www.rocketlawyer.com	– funding US $ 46.2 million

This seems to indicate that the cost of developing machine learning tools that are capable of more than rudimentary document automation and RPA, is quite expensive.

But it is worth mentioning that Joshua Browder, the self-taught teenager who programmed the DoNotPay chatbot over the course of one night (see subchapter 4.4) did so at a net cost of a UK £ 100 domain name server with an SQL-database and a few cans of Diet Coke. Some amazing solutions can promptly be established, and basic machine learning abilities do not require complex programming.

With the potentially significant costs of developing complex machine learning tools one must also take into consideration the size of the market. In the US there are around 1,350,000 lawyers[62]. There are around 140,000 practicing solicitors in England and Wales[63] and around 16,000 barristers[64]. In Norway there are around 9,500 practicing lawyers[65]. This is a significant limitation in the potential customer base compared to the market in the UK and US, which certainly would impact potential funding. Law is jurisdiction dependent, so software developed in one jurisdiction is not directly applicable in another jurisdiction. The significant market fuelling the 1,330 legal tech companies in the US[66] is much larger than the market in a small country such as Norway.

There are, of course, several other industries that potentially would be willing to invest in machine learning tools. There are more than 5,000 Public Accountants in Norway[67], more than 8,500 Authorised Accountants in Norway[68] and more than 2,850 real estate brokers in Norway[69]. The numbers are somewhat higher as not all accountants and brokers are members of these organisations. And there is a significant number of people with a legal degree who do not work as lawyers. Machine learning tools could enable these professions to provide legal services traditionally reserved for lawyers; they already have high qualifications in many areas of the law such as corporate law, taxation, VAT, real estate, inheritance, family law etc. And selling legal services as an integral part of an occupation outside the legal industry, is legal according to current regulation, see subchapter 5.4.

With more than 25,850 potential customers in several broad areas of legal services the individual cost of a subscription need not be significant. But the current absence of machine learning legal software seems to indicate that the will currently is lacking.

3.6.3 Culture in the legal profession

There are possibly three cultural factors in the legal profession that may limit the introduction of machine learning tools. Firstly, a sense of trust being something given by

obviously human lawyers to clients, face to face. Machines can't replicate that. Secondly, the industry itself is highly conservative, and likely reluctant to engage in technology, see subchapter 8.3.4. Third, the business is built on billable hours.

3.6.3.1 Trust and creativity

As seen in the list of tasks in subchapter 2.5 an important part of a lawyer's tasks is negotiation, facilitation and ensuring trust. A machine is incapable of earning trust; it either works or it doesn't. It is the lawyer who is ultimately responsible for the advice that is given, and the machine learning tools seem only to point out what should be examined further - not provide a complete answer. Thus, a lawyer still needs to do the job of deriving the correct answer, and the client's trust rests with the lawyer.

In his extensive article into machine learning and the law, associate professor at the University of Colorado Harry Surden (2014) [70] writes that most of the tasks performed by lawyers requires high-order cognition that current AI seem incapable of replicating, presenting a significant limitation on the potential impact of AI. In Surden (2012)[71] is explained how language changes that are trivial to humans may confuse machine learning tools. But Surden (2014) writes that he "*suggests that there [is] a subset of legal tasks often performed manually today by attorneys, which are potentially partially automatable given techniques such as machine learning, provided the limitations are understood and accounted for*." Surden then goes on to explain in detail how the various predictive machine learning models could assist lawyers in specific tasks.

3.6.3.2 Conservative industry

Citing a presentation by Toronto lawyer and legal teacher Simon Chester[72], Mark McKamey[73] claims that *"There are reasons to believe that the culture in the legal profession will significantly delay the integration of legal technology. Arguably, most of the legal profession is largely ignoring legal technology or engaging it in a merely symbolic sense in order to reassure clients. Even those who earnestly engage legal technology seem to only want to digitize current workflows"*. This seems to coincide with the statements of Professor Eric Vermeulen in subchapter 2.2 and Merete Nygaard, lawyer, and founder of Lawbotics. In an interesting interview Ms Nygaard explains how the legal industry is inefficient and not applying automation of tasks that easily could be automated[74]. The conservative nature of the industry is also a frequent topic in *Advokatbladet*, the monthly periodical from the Norwegian Bar Association *Advokatforeningen*.

3.6.3.3 Lawyers bill by the hour

As the above-mentioned reports on the Norwegian legal industry show that it has traditionally operated on billable hours (see subchapters 2.5). Robert Dale (2018) claims that this is a major barrier to introducing machine learning software in the legal industry. Developing machine learning software that could reduce the billable hours by a factor of hundreds or thousands, seems unlikely to receive massive funding from lawyers. It would not make sense for the legal industry to spend vast amounts of money developing tools that make lawyers obsolete. If would at least force the business to re-think its approach to billing, and as a consequence force the various public bodies regulating and controlling the legal industry to adjust. See also subchapter 8.3 regarding explanations as to the conservative approach

3.6.4 Understanding the letter of the law

A final possible limitation to the introduction of machine learning in the legal industry, is that there is evidence that deciding legal matters with the use of algorithms can result in unethical and illegal outcomes. Because of the complexity of law and ethics, even highly advance machine learning systems fail to correctly predict future outcomes on having learned prior outcomes. See also 3.6.3.1above and how working in the legal industry generally requires creativity and higher order cognition.

In subchapter 3.4.1 was mentioned a technology for setting bail that is used in more than 60 US jurisdictions. The system received massive criticism for being racist, according to Feller, Pierson, Corbett-Davies and Goel (2016)[75]. In general, the system learned from examining prior decisions and features such as age, race, educational level etc. that Blacks and Hispanics were more likely to be remanded in prison awaiting trial, than Caucasians. The reason for this is economy; Blacks and Hispanics generally belong to poorer segments of society than Caucasians, with the result that they can't afford even very reasonable bail. But Caucasians generally can. This led the system to predict that Blacks and Hispanics should be remanded in prison more often than Caucasians, which, naturally, caused outrage. And there are no provisions in the law itself to indicate that some people should go to jail more often than others simply because of their race. In other words, the data fooled the AI into drawing wrong conclusions.

3.7 Factors pushing the implementation of machine learning

3.7.1 Code of conduct

The legal industry Code of Conduct (advokatforskriften) includes several provisions in chapter 12 that may force lawyers to consider using machine learning tools. In article 1.2 the

lawyer is compelled to uphold the client's best interest, including disregarding personal gain. And it may likely be in the client's best interest that machine learning tools are used if it would make a significant difference in the number of billable hours. It is likely easer for a lawyer to disregard the possibilities of machine learning in this context if the tools are not widely used in the industry, and the opposite is the case if several competing lawyers use these tools.

3.8 Lawyer tasks that could be assisted by machine learning

In the subchapter 2.5 we have seen which tasks lawyers commonly perform, and so far in chapter 3 we have explored machine learning technology in some detail. None of the machine learning approaches seem capable of completely replacing lawyers, but the technology is certainly capable of assisting lawyers in some tasks.

In Robert Dale (2018) the various tasks are generalised into five points where machine learning is playing an increasing role.

3.8.1 Legal research

Legal research is the process finding the relevant laws, regulations, prior cases etc that are needed to support legal decision-making. This could be aided by searches aided by machine learning. The main challenge in using large archives of statute and case law such as *Lovdata* (see subchapter 4.2.1) is that the results that are returned are based on the keywords used. If, instead, questions could be put to the system in a natural language, or by entering entire parts of the case into the search query, the results could be returned in a more specific way.

3.8.2 Electronic discovery

Electronic discovery is the process of examining document evidence before making a decision. An example is the process of due diligence, i.e. the steps a responsible business is expected to take before entering into an agreement. Another example is patent dispute. Dale uses a Samsung vs Apple patent dispute in which 11,108,653 documents were examined at a cost of US$ 13 million over 20 months; it is impossible to know if a document contains relevant information before it has been examined, even though most documents probably don't contain relevant information. Machine learning is making a difference, providing two approaches; a fully automatic labelling approach, and an approach requiring initial manual labelling of some documents to learn from.

3.8.3 Document review

Document review [including contract review] could be aided by feeding the draft to a machine learning system, and having it highlight anomalies in the text based on having learned previous documents, provided the machine learning model has been trained on that specific type of document. Not only contracts could be reviewed, but also applications, claims, wills, and other documents that occur in abundance.

Also, document review is not limited to comparing a specific document to the patterns learned from training data, but also examining internal consistency in a document; words and phrases that stick out could be discovered (see subchapter 3.4.1).

3.8.4 Document automation

Dale explains that rule based document automation has been around for around a decade, and works by "*gathering relevant data from the user, either via form-filling or via a question-and-answer session […] The accumulated data is then used in a rule-based manner to craft a tailored document"*.

As Dale points out, this system is rule-based, i.e. *not* machine learning.

3.8.5 Legal advice

Legal advice could be enhanced in several ways by machine learning. One way could be for the client to give information regarding the case via a chatbot that extracts the specific information given, and another could be for the lawyer to use a machine learning system to ask the relevant questions after having gathered the information from the client. In this way, the process of lawyer-client communication may be sped up.

3.9 Summary

Computers and computer software are generally rule-based, i.e. they perform tasks by following specific instructions. Machine learning differs from this by following rules that it has learned either through examining large sets of data statistically or through interaction. The terminology and science fiction can cause some disconnection from the actual technology, but the technology is essentially statistical mapping that can generate stunning results when the data is too complex for humans easily to interpret it.

There are important limitations, mainly that any of the tasks performed in the legal industry require a high degree of creativity and cognition. However, some lawyer tasks are ripe for machine learning, such as legal research, electronic discovery, contract review and lawyer-client communication. In addition, rule-based document automation has been around for some time, and is being introduced in Norway.

4.0 Overview of common legal tech software

4.1 Introduction

There are a vast number of rule-based and machine learning based software (including web services) that are used by lawyers today. It is impossible within the scope of an MBA thesis to attempt to give an exhaustive list, but an overview of some of the common software is included below as they are likely to be mentioned in the interviews and because it lends weight to the final analysis.

4.2 Current Norwegian language legal tools

4.2.1 Lovdata

The most important Norwegian language legal tool is arguably *Lovdata* (www.lovdata.no). Lovdata is a foundation established by the Ministry of Justice and The Faculty of Law at the University of Oslo in 1981. Lovdata gives access to all laws and regulations, many circulars, court rulings, and other sources of the law. Machine learning does not seem to be integrated in any way, making searches a laborious task. The system is initially free to use, but professional uses get access to shortcuts and other functions that are not available for free. However, the search function is still laborious. Lovdata seems to be a purely rule-based reference work.

4.2.2 Gyldendal Rettsdata

Gyldendal Rettsdata (www.rettsdata.no) is a similar service to Lovdata, the main difference being that Rettsdata includes comments written by scholars and an extensive collection of document templates covering a wide area of legal services, and that there is no free access (save for a trial period). The templates are in Word format, with blank spaces for details such as names, dates etc, including some instructions. It seems a purely rule-based system.

4.2.3 Sticos, DIB, Infotjenester and Proff

Within the fields of corporate law, accounting, and human resource management there are several services that provide both step by step advice on how to perform various tasks such as generating documents and giving legal advice. The most important are probably *Sticos* (www.sticos.no), *DIB* (www.dib.no), *Infotjenester* (www.infotjenester.no) and *Proof* (www.proff.no). These services provide articles about specific legal topics, links to relevant statute law and prior judication, standardised documents, various specialised software such as accounting software and so-on. The services seem to be organised either by specific areas of law or information relevant to the law, and machine learning does not seem to be integrated in any way.

4.2.4 Lexolve Market

Lexolve Market (www.lexolve.no) is a Norwegian language document template system developed by Lawbotics AS (see the link to the interview in subchapter 3.6.3.2). A large set of pre-determined documents are combined with an editing tool that allows the user to answer various questions, filling in the blanks, resulting in a complete document, ready for print or other distribution.

4.2.5 Justify

Justify (www.justify.no) is a recent addition to the consumer market. The system allows anyone to log in via the Norwegian ID-system *BankID*, and by answering various questions the software combines parts of documents. There are three kinds of documents the site can create: Wills, Living Together Agreements (for cohabitating couples), and future Power of Attorney authorisations. The website claims to use machine learning to enhance its services, but it is unclear if the current solution is strictly rule-based or based in part on machine learning. To be sure nothing goes wrong; a lawyer will review the document as a part of the fee.

4.2.6 Summary

Though the various Norwegian language legal information software give both lawyers and other professionals easy access to information, organised in various ways, they do not seem to be based on machine learning. Some of the software, such as Sticos, DIB etc organise information for easy access, other software such as Lovdata give access to significant amounts of information, alas without any natural language processing system for easy searches, and Lexolve and Justify use questionnaires to fill in the details of documents created by lawyers.

All of the Norwegian language tools are rule-based first wave AI, with no learning ability and poor handling of uncertainty.

4.3 Current English language legal tools

The information about the software below is based on the information provided by the product website.

4.3.1 LawGeex Contract Review

Subchapter 2.2 covers the LawGeex contract review AI. By learning from vast amounts of previous legal documents the AI had developed a model that returned warnings of anomalous formulations, thus performing at a super-human level simply by applying clever statistical models to very powerful computing. Experienced lawyers spent an average of 92

minutes and discovered 85% of possible anomalies in the text, while the AI found 94% of the loopholes in 26 seconds [76]; that is 211 times faster than the average of the experienced lawyers and with a 9.6% improvement of quality.

However, it is important to note that the task of the AI is to point out the possible errors, but leaves the decision making on what to do with the errors to the lawyers.

4.3.2 Luminance

Luminance (www.luminance.com) is software developed at the University of Cambridge that allows lawyers to categorise, review and analyse documents at high speeds, taking the burden of low-level cognitive tasks such as due diligence, compliance, insurance and in-house contract management. By aiding the lawyers in discovering anomalous text or find specific pieces of information, it allows lawyers to work faster and more efficiently. According to the product website the software is both based on machine learning algorithms that allow an understanding of the contents of documents, but that also learns from interaction with the lawyers using it.

4.3.3 iBase

iBase (www.ibase.com) is a document sharing platform that allows for secure sharing with both colleagues internally in the firm and with external clients. According to the product website the system incorporates at least one machine learning element, i.e. fuzzy searching, see subchapter 3.4.1. The main emphasis of the system seems to be that it allows for secure sharing in virtual rooms that allows several people access at the same time – which could speed up lawyer-client interaction.

4.3.4 Kroll eDiscovery

In subchapter 3.8.2 Electronic Discovery was defined as determining the relevance of documents to an information request. Kroll (www.kroll.com) is a corporate investigation and risk consulting firm established in 1972, and they have developed various software including Kroll eDiscovery. Kroll eDiscovery is presented as a "*cyber risk, investigations, compliance, disputes and risk management*" system that amongst other things claims to use machine learning for searching large amounts of documents and for character recognition (e.g. for understanding low resolution text in photocopies, handwritten information etc). According to a 2015 study by a team from Kroll[77], some of the machine learning aspects are Boolean (regression) keyword searches, clustering based on similarity (likely fuzzy clustering), and latent semantic analysis (dimensionality reduction), see subchapter 3.4.1.

4.3.5 iManage

iManage (www.imanage.com) is another document sharing platform that allows the creation, management and collaboration on documents and emails from any device. According to the product website, it is used by a million professionals globally, and incorporates features from AI (i.e. machine learning) to "unlock more value from your information", and incorporates a secure cloud sharing system.

4.3.6 ROSS Intelligence

Founded in 2014 and developed in cooperation with the Vector Institute for Artificial Intelligence, with an initial funding of US$ 13 million, the contextual search engine ROSS Intelligence (www.rossintelligence.com) was launched in 2018. By utilising several algorithms at the same time ROSS examines the sources of the law, such as previous rulings by courts, legislative proposals and legal theory, finding relevant sources based on information regarding the specific case a lawyer is working on.

4.3.7 Neota Logic System "Perfect NDA"

Neota Logic "Perfect NDA" (www.neotalogic.com/product/perfectnda) claims to save the average customer 1,000 hours a year by allowing the use of "*own existing templates*", where "*Neota's award-winning no-code* […] *AI-logic ensures correct template selection*", and claims to be "*the only tool that combines document automation*" with AI. This extrapolates on simple rule-based document automation as explained in subchapter 3.8.4.

4.3.8 HighQ

Thompson Reuters "HighQ" is another document sharing platform that allows cloud-based secure file sharing, team collaboration and social networking software. Founded in the UK in 2001. In 2016 HighQ received US$ 50 million in investment from Goldman Sachs, Morgan Stanley and One Peak in order to expand into the US market[78].

4.4 Current machine learning tools directly competing with lawyers

When 18-year-old self-taught computer programmer Joshua Browder from London got his driver's licence in 2015, he quickly became frustrated over the number of parking-tickets he received. In a BBC interview [79] he claimed that parking-tickets unfairly target the disabled and elderly. To "The Daily Show" with Trevor Noah he said that he felt that competent legal assistance from lawyers were unreasonably expensive[80], so he set about taking on the parking companies himself.

He registered the domain www.donotpay.com, and over the course of an evening made a simple self-improving chatbot that would generate a complaint that the plaintiff could

print on his or her home computer and send to the parking company. During the first four months 30 000 parking tickets had been overturned by people using DoNotPay at a net sum of £2 million[81]. The service was launched in the United States the next year, and in short order caused parking-tickets to be overturned in the order of US$ 4 million[82]. The success helped Browder collect US$4,6 million in funding, and more than 1000 bots were integrated in a smartphone APP called DoNotPay[83]. So far, the completely free service has seen the overturn of more than US$ 25 million worth of parking-tickets[84].

Browder argues that the large amount of publicly available information makes it difficult for consumers to ascertain how and where to file a complaint, what to do if a decision is made in the consumers favour, and so-on. By compiling readily available information and combining it with details given by the consumer, Browder was able to give thousands of people access to legal solutions without the use of a single lawyer[85].

5.0 Opportunities and threats – a business strategy approach

5.1 Introduction

A new technology is being introduced into a monopolistic and conservative industry, potentially causing disruption. The purpose of this thesis is not to explore the individual strategies of law-firms and the approach to machine learning, but to explore how machine learning may affect the industry. Business strategy theory provides us with a framework for exploring this. Theory also provide us with approaches to predicting how the law-firms in general will react to possibly disruptive innovation.

In chapters 3 and 4 we have seen how machine learning tools can aid lawyers, but we have also seen how the technology may reduce the time a lawyer spends on his or her tasks and how it may allow others to deliver legal services. The history of legal industry's statutory monopoly on providing legal advice, is thoroughly covered in the Norwegian Justice Department "Norwegian Public Examination 2002:18" 3rd part chapter 5 [86]. The conservative nature of the industry is introduced in subchapter 3.6.3.2. The potentially disruptive nature of the technology being introduced is introduced in chapter 2.0.

5.2 Positioning strategy and industry structure

Business mission statements commonly include political claims such as being pro-environment, giving people opportunities etc. And while those statements may be true, there are specific criteria for defining a business, the most important being the search for profit. This applies to both solo-ventures and billion-pound corporations, Kaufman claims[87]. The claim is certainly true in finance, where investors share the same financial objective; "*They want the financial manager to increase the value of the corporation and its current stock price"*, according to Brealey, Myers and Allen[88]. Business strategy deals with the plan to achieve superior profitability, according to Michael Eugene Porter.

Michael E. Porter is one of the most significant thinkers within the field of business strategy, and in his influential 1996 Harward Business Review article "*What is strategy?*"[89] he explains how the strategy of superior profitability consists of analysing the two environments of a business; the business itself and how it positions itself in the market relative to its competition (positioning), and the composition of the industry that the business is in (industry structure).

Positioning is about finding out how to deliver a superior value proposition to the customers that the company decides to serve, in a manner that is uniquely well, so that the company receives value and a competitive advantage. The "*arithmetic of superior*

profitability" goes beyond simply improving on performance, Porter writes, it is about uniqueness. In order to move from the current lesser position to a superior future position, one needs to think about the whole company, how it fits together and the environment it is in. The strategy results in action steps, i.e. the things that must happen in order to realise the strategy, Porter explains. In "*What is strategy?"* Porter explains that the core of a positioning strategy is uniqueness, i.e. how to do something in a unique fashion that is not easily replicated by the competition. Simply competing to be more efficient than the competition results a mutually destructive *"a war of attrition"* and a mutual race to the bottom. Steps increasing the efficiency may soon become the standard of the industry, negating the possible advantages that operational effectiveness gives. Joan Magretta explores this in further detail in the book "*Understanding Michael Porter"*[90].

The other component of strategy, according to Porter, is **industry structure**. If positioning can be explained with the question "Is my business any good relative to my competition, and what can I do to be unique", industry structure can be explained with the question "Am I in a good industry, is it easy to get a good return in this industry?". In his highly influential 1979 HBR article[91] "*How Competitive Forces Shape Strategy*" Porter defines five forces defining if the industry's potential for profit. Porter writes that "*Moreover, in the fight for market share, competition is not manifested only in the other players. Rather, competition in an industry is rooted in its underlying economics*". The five forces Porter defines "*determines the ultimate profit potential of an industry*". The comments on the five forces are further elaborated on in this chapter.

- **Rivalry within the industry** deals with how hard the competition is. Porter explains that rivalry is about "*existing competitors […] jockeying for position – using tactics like price competition, product imitation*" etc.
- **Bargaining power of suppliers**, pushing up price of what they sell us, pulling profit out of the industry. In relation to this thesis, Porter explains that suppliers pose a credible threat when they are capable of integrating forward into the industry's business, providing "*a check against the industry's ability to improve on the terms which it purchases*".
- **Bargaining power of customers**, force down the price and ask for more, that we can't recover the cost of. Buyers are powerful when they are price sensitive, the products are (or seem) standard and undifferentiated etc.

- **Barriers of entry**, i.e. how easy it is for someone who is not already in the industry to begin competing with existing players. In relation to this thesis the most important barrier of entry is government policy, regulating who may provide legal services and how they may do it.
- **Substitutes**, i.e. a product or service that meets the same general demand, but that is provided by a rival industry.

The five forces do provide some aspects of examining the potential impact of machine learning on the legal industry.

5.3 Resource based view

An important criticism of the positioning theories is that they do not provide a lasting competitive advantage, according to Jay Barney (1991)[92]. This represents an alternative approach to strategy compared to Porter. If a superior positioning strategy can be formed, it may be upset by imitation the very next day.

Barney writes that "*A firm is said to have a sustained competitive advantage when it is implementing a value creating strategy not simultaneously being implemented by any current or potential competitors and when these other firms are unable to duplicate the benefits of this strategy"*, ibid page 102. According to Barney page 105- 111, when the "firm resources" such as assets, attributes, information, knowledge etc are providing a competitive advantage, that advantage is sustainable when these resources are:

- valuable, i.e. that the resource either exploits opportunities or neutralises threats
- rare, i.e. that that the resource is something it is something not everyone possesses or has access to, and
- hard to imitate, i.e. that they are not easily obtained by the competitors.

A highly unique machine learning tool could be such a resource.

5.4 The legal industry may be heading for limited deregulation

It is worth noting that the legal industry is heading towards a potentially significant change in the statutes regulating the industry. Norwegian Public Examination NOU 2015:3 suggested in chapter 28 a significant change to the legal industry, allowing anyone to provide legal services outside of trial. The suggestions have still not been put before Parliament, but the current legislation in The Courts Act (Domstolsloven) § 218 section 5 allows anyone to provide legal services as an integrated part of other services. In other words; if rival industries, such as accountants, developed a tool for creating contracts etc for their clients, they could compete with lawyers even under current regulation; if the suggestions in NOU

2015:3 are approved by Parliament, *anyone* could be given admission to provide legal services in this way.

Thus, individuals armed with document automation and machine learning software could begin providing legal services with no or little legal education. Depending on the complexity of the specific areas of the law and the software made available, this could flood the legal industry with semi-professional service providers offering the same general services for a lower price.

This is of course not only limited to low-end tasks such as small claims, writing complaints in welfare-cases and setting up wills. Other professionals such as accountants [revisorer & regnskapsførere] do have significant insight into the economies and inner workings of their business customers and could possibly begin providing legal services as an elongation of accounting. As is explained in NOU 2015:3 the current extent of The Courts Act § 218 section 5 is not well understood, as there is limited control of the degree of legal services included in the services of other industries. Armed with powerful machine learning and document automation software, accountants could begin providing complex legal services including mergers and acquisitions.

It is worth examining if lawyers believe this to be a threat.

5.5 Summary

The legal industry is heading towards some uncertainty due to machine learning tools, a challenge which to some degree could be compounded by simultaneous deregulation/demonopolisation of the industry.

6.0 Research method

6.1 Research problem and research questions

As introduced in chapter 1 the research problem is exploring the impact machine learning is having on the legal industry. This overall problem is divided into three specific research question:

1. How are law-firms responding to the machine learning phenomenon?
2. What do law-firms believe the future role of machine learning in the legal industry to be?
3. What are law-firms doing to be a part of that future?

6.2 Research paradigm

There are two paradigms or methods for conducting research: quantitative and qualitative. The research problem and research questions are exploratory into an area of limited research, dealing with a technology that was largely unknown ten years ago, and where the specific questions to ask were hard to determine. Asking close-ended questions that could be answered with a "yes", "no" or a number on a scale from 1 to 10 wouldn't provide a deeper understanding of the phenomenon and the industry's understanding of the phenomenon. Thus, a qualitative approach was chosen, in order "*to provide fuller descriptions"* and because this approach is *"particularly useful when we are exploring phenomena that one doesn't know that well, and which there is little research into*", according to Johannessen, Kristoffersen and Tufte page 37[93].

There are several challenges with a qualitative approach. First, generalisation, i.e. the act of drawing broad inferences from particular observations, is widely acknowledged in quantitative research, but more problematic in qualitative research. "*The goal of most qualitative studies is not to generalize but rather to provide a rich, contextualized understanding of some aspect of human experience through the intensive study of particular cases*"[94]. Thus, a qualitative approach is a good approach to preliminary studies for developing more pointed issues for use in future research, according to Kvale & Brinkmann.

Another challenge with qualitative research, is that both the researcher and informants at the onset may be biased towards confirming or dismissing existing theories. In addition, the biases of the researcher may influence the findings. Kvale & Brinkmann page 44 explain that an approach to mitigating this challenge, is to focus of the research on what the informants are actually describing themselves and their surroundings, and how they give meaning to the things they explain. This presents yet another challenge: masking sure it is the views of the

informants that is expressed in the findings, while at the same time objectively evaluating those views.

6.3 Confidence in the research

6.3.1 Reliability

The purpose of research is to describe, explore or solve problems. An important initial question is how, by what manner, am I supposed to describe, explore, or solve the problem at hand. The answer to that question decides the level of trust a practitioner, student, or practitioner may have to the results of the research, and as a consequence if the research may have an impact on future research.

This is the challenge of reliability, according to Johannessen, Kristoffersen and Tufte (2004) p. 46. Reliability is a challenge in the qualitative research paradigm because the data collection technique is not structured but determined by the conversation. Second, the research is contextual, and near impossible to replicate, ibid p. 228. Third, the researcher is an instrument in the research. Thus, there are no objective standards to measure the research against.

One approach to mitigating these concerns is to provide a through explanation into how the research was performed, ibid. p. 228. The purpose of this chapter is to provide that explanation.

6.3.2 Internal validity

Internal validity deals with the challenge of whether the research "*correctly reflects the purpose of the study and represents reality"*, ibid. Validity is improved via two methods: continuous observation and triangulation.

This study was conducted by exploring the technology and the industry through continuous observations in the form of reading articles, reviews, textbooks, studies, public examinations etc during the entire course of the study, and performing the interviews, and then triangulating the interviews against the continuous observations. This resulted in a gradual increase in understanding of the phenomenon.

6.3.2.1 Learning about the technology and industry

It is not hard to find opinions or articles about machine learning and Artificial Intelligence. Searching Google for the specific term "machine learning" returns 110 million hits, while the term "Artificial Intelligence" returns 109 million hits. Exploring Wikipedia and Youtube similarly returns vast amounts of information. The first step in learning about the

technology was to examine how information could be classified through the application of an algorithm. The next step was to examine what reinforcement learning was, and how a computer could learn how to play a game by running simulations against itself. With a general overview, I could approach a few textbooks, the most important being "*An Introduction to Machine Learning"* by Rebala, G., Ravi, A. and Churiwala, S. (2019). From that point I could organise the algorithms commonly used into the various modelling approaches with an emphasis on classical learning and reinforcement learning, how those models were built, and finally the architecture. With a general understanding I would then browse specific scientific papers to confirm or update my understanding, and explore the various problems related to disconnection between our understanding of the technology and the technology itself, definitions of AI, and the general history and technicalities of digital computers. This has since been condensed into a short overview, as the full extent isn't relevant for neither the findings nor the analysis. But it was necessary to form an understanding of what we in fact are dealing with, which aided my understanding of the answers given during the interviews.

Learning about the industry is both easy and complex. There is very limited research on the Norwegian legal industry, most of the research coming from either the Government or the industry itself. But the research deals mainly with the composition of the industry, competitive matters etc. Machine learning isn't mentioned in the Government investigations into the industry. There is, however, much research into the American legal industry and even the possible impact of machine learning on the American and British legal industries. Though Common law jurisdictions, the fundamentals of applying the law are the same in both the US, Britain and Norway, and the industries are highly comparable. The tasks are comparable, and the English language articles dealing with legal method and syllogisms explain the same process as Norwegian textbooks; A major premises is formed from an analysis of the sources of the law according to legal method, a minor premise is formed from the facts of the case at hand, and a logical conclusion is derived. The main difference is the emphasis on case law, se e.g. "*An introduction to Legal Reasoning*"[95] by Edward Levi, but this is likely of little or no relevance in relation to machine learning because case law is a significant source of the law in all jurisdictions. Thus, American and British research into the impact of machine learning on the legal industry is assumed to be applicable also to the impact of machine learning on the Norwegian legal industry.

It was assumed that textbooks and articles in various university or industry reviews found via Google Scholar are both authentic and reliable sources, and that Government and

industry research are equally both authentic and reliable. Where differences in approach were discovered, the sources were compared to other sources.

6.3.3 External validity

External validity, generality or transferability deals with applying the conclusions outside the specific context of the study.

The purpose of this research is to explore what is happening in the industry by talking to well-placed highly qualified persons in the part of the industry that likely has the highest focus on the phenomenon. At the onset of this thesis, I was unable to reveal any relevant research into the impact of machine learning on the Norwegian legal industry, and the research I did find was limited to structural matters such as industry composition, pricing strategies etc. There was no obvious research into which lawyer-tasks that could be affected by machine learning. The purpose of this thesis is *not* to evaluate or explore the strategies of individual firms. The main purpose of this thesis is to fill the gap in the understanding of the phenomenon, explaining how lawyers do their work, exploring the technology, and then exploring the possible impact.

In order to ensure generality within the scope of this thesis (see subchapter 6.5) several considerations had to be made to ensure that the findings would be representative of the industry. This resulted in five assumptions based on where the impact was believed to be experienced first.

- First, large law-firms are more likely than small to have the financial muscle of developing their own machine learning tools.
- Second, large law-firms are more likely than small to be able to use English language machine learning tools, as they to a larger degree than small firms have a global reach, and likely conduct more of their electronic discovery and document review in English.
- Third, large firms are more likely to have their own IT-departments, giving an in-house understanding of the technology that smaller firms are less likely to acquire.
- Fourth, large firms are more likely than small firms to have employed lawyers with multidisciplinary competence, including technological competence.
- Finally, most of the articles relating to machine learning and AI in the industry monthly *Advokatbladet* mainly deal with representatives from the very large firms.

The assumptions were not explored in further detail.

6.4 Research design and analytical units

According to Johannessen, Kristoffersen and Tufte (2004) pages 80-86 explain that a feature of a qualitative research design is the absence of a unified analytical direction. There are many ways of performing a qualitative study. This doesn't mean that "anything goes", they write. Decisions on how to gather data impact the alternatives for dealing with the data. Because of the different approaches to a qualitative design, transparency becomes even more important in order to ensure confidence in the findings, see subchapter 6.3 above. It is impractical to present an exhaustive list of potential qualitative designs, so Johannessen, Kristoffersen and Tufte list four; phenomenology, ethnography, grounded theory and case-study. These features of these four designs are:

- Phenomenology is "*studying and describing people and their experienced with and understanding of a phenomenon*" based on existing theory,
- Grounded theory is developing new theories based on the findings,
- Ethnographic design is describing a culture, a social group or a social system, and
- Case study is an in-depth study of one or a few specific cases

Machine learning is a specific phenomenon, and there is existing theory into business strategy when faced with technological and potentially disruptive innovation. This points towards a phenomenological design, and I wish to take a descriptive and exploratory approach to gathering data in order to understand how important actors in the legal industry view the phenomenon and its impact. Creswell & Poth (2018)[96] highlight two types of phenomenology, hermeneutic and psychological. Hermeneutic is: "*They write a description of the phenomenon, maintaining a strong relation to the topic of inquiry and balancing the parts of the writing to the whole.*". Creswell & Poth describe three major steps: Preparation, data collection and analysis and reporting.

During the preparation phase the researcher builds on existing knowledge and prior experience. "*Phenomenology is not only a description, but it is also seen as an interpretive process in which the researcher makes an interpretation"*, according to Creswell & Poth. The research questions are formulated in such a way as to understand the purpose or context of the phenomenon that is to be explored, and the informants are asked to describe their experiences.

During the data collection phase, data is collected from individuals who are experiencing the phenomenon Polkinghorne (1989)[97] recommends interviewing 5-25 individuals who are all experiencing the phenomenon. The questions are broad and open-

ended, with the purpose of providing an understanding of the common experience of the participants. With reference to Creswell & Poth, phenomenological research is not an objective observation, but an interpretive process of opinions. And the researcher brings his or her opinions to the process.

Bruce L. Berg (2001)[98] explains in chapter 11 an approach to analysing qualitative material. First, the interviews are completely transcribed. Then the text is coded via deduction from theory or induction from the material itself. The codes are categorised in nodes, and the nodes are organised to uncover similar statements, patterns etc. The organised material is re-examined in order to reveal meaningful patterns and processes. Then the identified patterns are considered in light of existing research and theories.

6.5 Recruiting informants and lack of prior research

Based on the considerations dealt with in the subchapter regarding external reliability/generality, I performed a Google search for the twenty largest law-firms in Norway. Armed with an overview from 2017[99], I began contacting the firms in order. I believed the firms would be more than willing to take part in the study, as it would entail both an opportunity to advertise their focus on technology and because they were likely to see the benefits the industry might have as a result of the research.

An email was sent to the 20 largest law-firms in Norway asking if they wished to take part in a qualitative study regarding machine learning and the legal industry. The emails included the information sheet and consent form. The initial response was limited. I then performed Google-searches where the names of the firms were combined with the words "maskinlæring", revealing individual lawyers in those firms that had expressed an interest in machine learning. I then proceeded to send personal emails. The initial round and follow-up round resulted in responses from eight firms, but after further contact only four were willing to participate. No-one rejected flat out, but they simply didn't reply.

This coincided with the SARS CoronaVirus 2 outbreak that saw significant attention in Norway towards the end of February 2020 and the following months, where most offices in Norway were closed and people worked from home. This may have affected the limited response. Also, because of the attention given to machine learning and artificial intelligence some law-firms may have been reluctant to reply out of wariness that answering questions wouldn't benefit their competitive situation. Regardless of the reason, only four firms responded, but these four firms represented the entire spectre of the twenty largest firms. In

addition, two of the firms met with two representatives, resulting in a total of six individuals taking part in the interviews.

According to Kvale & Brinkmann (2009)[100] page 133, the challenge with qualitative studies is that there is a tendency to interview either too many, which challenges the scope of the research, or too few, making it hard to generalise the findings. The challenge of generalising, and what to generalise, is dealt with in subchapter 6.3.3 (external validity).Interviewing six informants should provide enough data for that preliminary research, and also shed light on complex aspects relating to business strategy that may be generalised.

6.6 Developing an interview guide, and conducting the interviews

The informant recruitment method is covered in subchapter 6.5. This subchapter deals with the interview guide and the way the interviews were conducted.

Adhering to the broad and exploratory approach to phenomenology, the interview guide centred around the three research questions:

- In light of the advances in machine learning, what do you believe the future will look like for law-firms?
- What are you currently doing with regards to machine learning, in order to be competitive in the future?
- What are your thoughts on the significant attention given to Artificial Intelligence and the legal industry by your surroundings, and do you feel that this focus has been justifiable?

Save for the three themes, the interviews were unstructured/unstandardised, allowing the informants to speak freely in their own words, and exploring the various topics. Various additional questions were asked depending on what the informants said, in order to narrow or specify their explanations within the topics. Examples of follow-up questions were questions related to what machine learning tools in fact were doing, the degree of lawyer interaction, decision making and problem solving, thoughts on cost, both the cost of acquiring and developing machine learning, but also the reduction in billable hours and the various approaches to that challenge, if language matters, if machine learning is perceived as a threat or an opportunity, etc.

As the interviews progressed, I formed a broader understanding of the answers I received, leading to the interviews taking gradually longer because I had several follow-up questions. This resulted in the two group-interviews being more extensive and complex than

the two initial interviews. I took this into consideration when analysing the initial interviews; the two initial informants had answered the same essential questions, but I didn't fully understand the answers until after the final interviews.

The interviews were conducted in Norwegian.

The two first interviews, with Kielland (Ræder) and Grimsø Moe (Hjort), were conducted face to face at their offices. The interviews were recorded with the app *Nettskjema* on a mobile phone.

The interviews with Helboe and van Dam (Simonsen Vogt Wiig), and Olaussen and Weitzenboeck (Wikborg Rein) were conducted over the video conferencing applications Skype ® and Teams ® due to the Corona Virus outbreak.

All of the informants had received information about the study in advance, including the consent form, as stated above in subchapter 6.5.

6.7 Cases

Kyrre W Kielland is a partner with Advokatfirmaet Ræder AS. Mr Kielland's area of focus is corporate law, contract law, regulatory law, market law and insurance law. He was mentioned as a product liability expert in the international Legal 500 ranking and was named in the up-and-coming category in Finansavisen in 2019. His background includes a position as scientific assistant at the Institute of Private Law at the University of Oslo, participating in the International Visitor Leadership Program arranged by the US State Department, and he is a member of a voluntary board relating to blockchain technologies with The Norwegian Computer Society.

Eivind Grimsø Moe is a partner with Advokatfirmaet Hjort. Mr Grimsø Moe works in lawfirms office of Communication, Media and Technology, and works with corporate law with special focus on contracts, companies, financial markets, energy markets and transactions. His background includes being Associate General Counsel of NASDAQ OMX.

Peter van Dam is Chief Digital Officer at Advokatfirmaet Simonsen Vogt Wiig AS. He studied business administration for the financial sector at the Free University of Amsterdam. After 12 years in the consultancy industry he started in Norway in the IT sector and serviced many legal firms. The last 6 years he has been CDO for Simonsen Vogt Wiig.

Nicolai Halboe is partner with Advokatfirmaet Simonsen Vogt Wiig AS. His background includes 15 years within the IT Consultancy industry, including as legal director

with Capgemini Norge AS. His area of expertise is IT-contracts, contract negotiation and risk elimination.

Hanna Beyer Olaussen is Specialist Counsel with Wikborg Rein Advokatfirma AS, and part of the firm's Technology and Digitalisation practice. Her focus is advising the technology sector, with specialisation within IT, intellectual property law and marketing law. She works with contracts, M&A and litigation.

Emily M. Weitzenboeck, PhD (Oslo) is senior lawyer at Wikborg Rein Advokatfirma AS and part of the firm's Technology and Digitalisation practice. She is qualified to practice law in both Norway and Malta. Weitzenboeck works primarily with privacy and data protection law, information security, contract and e-commerce law. She has several publications in the field of IT law and her PhD-dissertation was on legal frameworks for emerging business models such as virtual businesses. Since September 2017, Emily has a part-time position at Wikborg Rein and is a full-time associate professor at Oslo Metropolitan University, where she teaches data protection law and contract law.

Firm	Informant	Size
Wikborg Rein	Hanna Olaussen	Wikborg Rein passed 1 bn NOK turnover in 2018, becoming the largest law-firm in Norway. With offices in Oslo, Bergen, London, Signapore and Shanghai.
	Emily M. Weitzenboeck	
Simonsen Vogt Wiig	Peter van Dam	Simonsen Vogt Wiig was 6th largest law-for in Norway in 2017, with offices in Oslo, Bergen, Trondheim, Tromsø, Stavanger, Kristiansand and Signapore.
	Nicolai Halboe	
Hjort	Eivind Grimsø Moe	Hjort was the 15th largest law-firm in Norway in 2017, with an office in Oslo.
Ræder	Kyrre W. Kielland	Ræder was the 20th largest law-firm in Norway in 2017, with an office in Oslo.

6.8 Organising the data

The interviews were fully transcribed and uploaded into Nvivo12, a qualitative analysis tool by alfasoft[101]. The interviews were initially coded through induction from the material itself, by consecutively creating nodes to deal with the various topics. However, as I progressed and began comparing the various nodes, it became obvious that this approach was too unstructured. The transcripts spanned 10,563 words, and the informants were generally very specific, leading to a large number of notes that were hard to combine. As I was coding

the transcripts, I read and re-read the interviews, and a deeper understanding was formed. I changed the approach to the matter, creating three categories of information based on the research questions, and began creating nodes from the transcripts based on the research questions. This resulted in a large number of organised nodes, making it easier to compare the contents of the nodes, and combining similar content.

Nvivo allows for several transcripts to be coded in the same tree, while still retaining information about which content came from which informant. This resulted in a single structure containing all the information from the six informants, organised by topic, while allowing for the identification of the statements from each individual informant.

6.9 The researcher

In subchapters **Feil! Fant ikke referansekilden.** and 6.3 is introduced the element of researcher bias within phenomenological studies, bringing his or her own opinions and biases to the process. Thus, the researcher becomes an integral part of the study being performed, and the researcher becomes a potential source of both qualification and disqualification of the findings.

I am a 2006 graduate from the University of Oslo with a professional degree in law (candidatus juris) specialising in tax law and business law. I have been employed in various legal functions in both the private and public sector, but not as an attorney (advokat). I see no obvious biases other than an initial disbelief regarding machine learning, which was soon replaced by confusion and curiosity.

6.10 Ethical reflection regarding the interviews

I base the data collection on the ethical guidlines addopted by the *National committee on research ethics in the social sciences, humanities, law and theology* (NESH)[102], and the notes by Per Nerdrum[103].

I have not discussed matters of a personal nature with the informants, nor enquired about or collected such information beyond their names, academic titles, positions of employment, employers, and ages.

Ahead of the interviews the informants signed or verbally acknowledged that they wished to participate in the study, and under what conditions. Due to the Corona Virus pandemic of 2020 some of the interviews were conducted via Microsoft Teams ® and Skype ®, and in those cases the acknowledgement was given verbally. No incentives were given or promised in return for participating in the interviews.

The summaries in the findings chapter were done anonymously in order to protect the integrity of the research, while specific quotes were laid before the informants prior to completion of the thesis. None of the informants opted out nor placed limitations on the use of quotes.

6.11 Notification

The Norwegian Center for Research Data (NSD) must be notified if research data may identify individual persons. NSD was notified of this project on February 3rd 2020, and approval was given on February 24th 2020, see attachment 1.

Important premises for the notification was that no part of the interviews will gather or relate to information regarding race, ethnicity, political, philosophical or religious affiliations or opinions, information regarding suspicion or conviction of criminal offenses, sexual matters or trade union membership.

The interviews are recorded via the Nettskjema-app installed on a private mobile phone. Nettskjema is a data collecting service provided by The University of Oslo, with access given to cooperation educational institutions. The data is fully encrypted, an is considered a safe means of storage by NSD. Temporary copies for the transcription process, written manually in Word ®, the complete transcriptions and the NVivo-files were stored in an encrypted memory stick with 128-bit AES encryption.

7.0 Results

7.1 Introduction

The chapter gives a condensed rendering of the results discovered in the transcripts. The text is an attempt to render the *meaning* of what the informants said. The initial interviews and the consequent rendering involve a triangulation against a continuous observation, see subchapter 6.3.2 which deals with validity and triangulation. An analysis of the rendering is given in chapter 8.0.

7.2 What is the current impact of machine learning on the legal industry?

7.2.1 Digitisation and automation

The informants explain that they are fully digital, i.e. that they use computers for writing all their documents and that their archives are fully digital. The most basic digital application is the Microsoft Office ®, and they use various platforms for sharing documents, both internally and with clients via *rooms* or *portals*. Several informants mention a range of data-roaming tools, especially High-Q. These applications are seeing increased usage, according to the informants.

One informant claimed that "many lawyers" in the large law-firms until recently didn't use computers themselves; they had secretaries transcribe documents and emails for them. Still several lawyers require assistance with Word, Excel and Powerpoint. According to that informant the most important step for increasing efficiency in a law-firm is for all lawyers to learn how to use those applications.

Several simple processes have been partially automated as a result of increased focus on digitisation in the public sector. The *Folkeregisteret* [a Norwegian registry of citizens], *Brønnøysundregistrene* [the Norwegian business registry], *Altinn* [a portal for communicating with various public offices, including the Tax Administration] are fully digital. This allows for both various degrees of integration with other systems and for digital communication, saving the lawyer time when exploring or registering information.

One informant claimed that the industry previously had looked at several ways of digitally sharing knowledge, e.g. creating standardised documents and libraries. According to the informant the law-firms were generally unable or unwilling to create such systems. According to that informant standardisation of information is a necessary prerequisite for developing systems for automation internally in a law-firm.

7.2.2 Machine learning

The informants explained that machine learning applications are generally not used in the legal industry.

When asked about the current use of machine learning applications, the informants initially gave mixed responses. Some were quick to down-play the importance of machine learning, while others quickly pointed out areas in which machine learning tools are being used to some extent. One informant explained that they were in the process of piloting some systems, and that they believed the technology may bear fruits in the future, but that the current level isn't improving effectiveness. Training the systems takes time, and if this work bears fruit it will take at least two or three years. Another informant said that they are experimenting, but that it seems to be experimenting for the sake of experimenting: they spend more time checking that the AI does what it is supposed to, than if they had just done the job themselves. A third informant said that they were playing with some of these applications, but that most of them had limited or no machine learning,

Though the initial focus varied from dismissive to quite optimistic, all the informants replied with varying degrees of detail, that machine learning tools were being used to explore the contents of large sets of text. Of the different potential applications mentioned in subchapter 3.8, the two that were explored in some detail were document review and electronic discovery. Document review is the process of reviewing documents ahead of entering into commitments and/or in relation to court proceedings. Electronic discovery is the process of exploring large sets of information in an attempt to find relevant information.

Practical application was explained in relation to due diligence, mergers and acquisitions, preventing overstepping antitrust-regulation, public investigations [e.g. into corruption], intellectual property violations etc. One of the informants explained how Luminance (se subchapter4.3.2) is used to examine tens of thousands of emails, searching for red flags and smoking guns. This process is immensely tedious and costly for the client when performed manually, the informant explained, but machine learning allows for training the model on a few hundred documents, and then the model will do the rest. Another informant explained how machine learning tools easily can discover red flags within a large body of contractual text, such as change of control-clauses, liability limitation clauses, intellectual property ownership-clauses etc, allowing the lawyer quickly to pinpoint challenging areas in the text. A third informant also confirmed the same application of machine learning.

It is worth noting that a fourth informant stated that such tools were quickly dismissed in relation to due diligence.

When challenged about what machine learning can do, all the informants confirmed that the decisions on what to do about the problems discovered via machine learning had to be solved manually by the lawyer or team of lawyers.

Other applications mentioned during the interviews were Kroll eDiscovery, iBase, and iManage, see chapter 4.

When asked about language and if English language machine learning software has problems understanding Norwegian, one informant explained that they had considered this a potential problem initially, but that much of their work is international and that there is a general tendency of using English more and more. Another informant confirmed that they do a low of the work in English, and that they hadn't really considered it a problem. A third informant explained, in relation to potential future application of machine learning, that improved language libraries would be needed in order for more English language software to have an application in Norway. A fourth informant supported that opinion and added that no one apparently had covered the expense of developing a purely Norwegian legaltech AI. It is important to note that these questions mainly relate to the potential future impact of machine learning, and that the current use of machine learning tools in document search and e-discovery doesn't seem to be limited by language barriers, as they are trained on sample documents.

When asked about the factors limiting a broader application of machine learning in the legal industry, the respondents generally responded in the same way. One respondent stated that an important obstacle with machine learning, is that the machine is incapable of understanding why something is wrong or insufficient; you may get a specific hit, document template etc, but the system incapable of capturing the situation and sorting out the irrelevant factors. Another informant explained that they believed it was unlikely to imagine an AI drafting contracts, writing subpoenas etc, and when prompted they replied that working as a lawyer required a high level of creativity. They seldom wrote the same sentence twice because, as a third informant stated, they must adjust the advice to the specific case. A client wishes to achieve a specific goal and finding what steps to take in order to reach that goal requires creativity. A fourth informant explained that they hoped that machine learning would be improved so that more boring and repetitive tasks could be performed by the machine, but

that the machine is generally incapable of understanding a contract and then adjusting it to fit the case. You still need a lawyer for that, a fifth informant claimed. Beyond advanced document review and e-discovery, they see limited near future application of machine learning in the legal industry.

One informant also pointed out, that to their knowledge, the *LawGeex AI* was designed to work with Non-disclosure agreements. These contracts are highly specialised, and relatively similar in structure and wording.

7.3 How are law-firms preparing for a machine learning future?

When asked, one of the informants stated that they didn't believe that the legal industry would develop machine learning legal tech itself. Some have tried, they claimed, establishing daughter companies to explore the technology, but have ended up becoming a subcontractor for the parent company, buying ready-made machine learning tools from third party developers. According to this informant third party developers are frequently offering new products to the legal industry, albeit with varying degrees of success, as many of the offers are quickly dismissed.

Another informant stated clearly that machine learning legal tech will not be developed by the legal industry itself. The informant then thought about it and said that some have tried, but the two examples they mentioned were not machine learning tools. All of the informants indicated that they were paying attention to the technology and what was going on.

Both informants pointed out the internal differences between the Norwegian law-firms, and one of the informants pointed out that they believed Norwegian law-firms lagged behind the most technologically progressive firms in the United States.

When asked about which areas of law that might be next in line for machine learning, two of the informants pointed out various types of repetitive tasks that don't require intellect, such as various types of conflict resolution, summary proceedings where there is a guilty plea [i.e. where the task before the court is simply administering the correct punishment] etc.

The informants were asked if they were preparing for machine learning by preparing training data, e.g. by digitising non-digitised documents, structuring, or standardising documents, etc. One of the informants staid that one challenge was volume; having ten agreements of a specific type is insufficient for training, you'd need ten thousand documents, and where would you get those, they asked rhetorically. Two other informants shared that

sentiment; a single law-firm doesn't have access to enough training data on its own. One of the informants pointed out that they use several different systems, all storing data in different formats. They perceived the lack of interchangeability between formats a potential problem for preparing for machine learning. However, a fourth informant pointed out that one wouldn't need large amounts of structured data; as long as the data is there in one form or another, a machine learning system could use it for learning, using review of large amounts of unstructured emails as an example of machine learning finding structure where it was hard for humans to do so.

7.4 Industry response to machine learning

7.4.1 Opportunity or threat?

The general opinion of the informants was that machine learning may pose an opportunity, but there were also some concerns. One informant said that they hoped AI might relieve them from tedious tasks and speed up certain processes. They were concerned about keeping up with the technology in order to stay competitive and because the market expects law-firms to be up to date. Another informant said they view machine learning as an opportunity, and that they wish to stay ahead, but that they are resigned to being passive recipients of machine learning tools, passively evaluating third party offers. A third respondent said they weren't concerned about industry fragmentation. If machine learning tools enable other industries in close proximity to the legal industry, such as the accounting industry, it would improve how the industries cooperate in order to provide a complete service to their clients. A fourth informant claimed that there is a significant unmet need for legal services; many individuals and businesses lack access to legal services, and they hoped that machine learning may allow for lawyers to tap into that "extremely large" market. And because this is a market that is currently unsupplied by the legal industry, no-one will be taking that market away from lawyers by tapping into it [e.g. via free and/or very reasonable web-based services, chat-services etc]. In addition, one informant explained that the large firms work business to business, representing a specific business entity in a business relationship with another business entity.

A fifth informant repeated the sentiment that machine learning tools will save them from time and workload so that they can focus on the interesting and most pressing tasks at hand. Several informants specifically mentioned that they believed machine learning would make the legal industry more interesting, by letting lawyers focus on the interesting parts of the law rather than perform work they considered menial.

The final factor motivating lawyers to examine machine learning tools, is allowing them to spend more time on the interesting and legally complex tasks, rather than menial sifting through large amounts of irrelevant information. This point was made by all the informants, in one way or another.

7.4.2 Limiting factors

Several informants mentioned that the industry is based on billing by the hour. In that respect, investing in machine learning tools that reduce the number of billable hours doesn't make sense.

Several informants stated that they don't believe that the industry itself will develop machine learning tools, and one obvious reason is the high cost associated with development. One informant said that there would have to be a significant volume of the same repetitive tasks in order for them to take the time and effort of training a system to perform those tasks. Another informant said that for this to be interesting, the increase in turnover would have to offset the potential drop in billable hours. They also mentioned the economic muscles of the industry; Norwegian law-firms are generally small by international standards, and the informant didn't believe that Norwegian firms were willing to foot the significant costs themselves. There seemed to be a unanimous belief that future machine learning tools would be either foreign white label software, i.e. software developed and licensed by a very large foreign firm, or by third party software developers. The informants didn't express opinions on whether third party software developers would be Norwegian or foreign, but several mentioned the importance of strong language libraries, i.e. software sub-components that allows the software system to run in different languages.

A final limiting factor are the regulatory requirements. Statute law require confidentiality of lawyers, and this extends to keeping confidential information secure. The informant didn't specifically mention GDPR.

7.4.3 Working around the challenges

As covered above, an important challenge posed by machine learning tools is that they reduce the number of billable hours. One informant believes that machine learning may necessitate productizing and package pricing. Another informant reaffirmed this, pointing out that if wouldn't make much sense investing in a costly system that cut billable hours down from months to hours, if there wasn't a way to increase revenue at the same time. Like the former informant, the latter believes in productizing and package pricing to offset the reduction in billable hours and investment cost.

7.4.4 Will AI replace lawyers?

The informants mentioned three types of AI that seem to be making an impact: document automation, electronic discovery tools and document review tools. The informants believed you still need a lawyer to control the results of automation or review. As one informant put it, the machine doesn't understand why the document it created is insufficient for the task. And as another informant put it, there is still no machine learning system that is capable of legal judgement. The lawyers will still be doing all the evaluations of what to do when machine learning discovers a red flag, "there is no doubt about it" said one informant. And a third informant reaffirmed that we still haven't reached a point where machines really are capable of doing our job.

The informants generally believe that machine learning may take over the tedious tasks of sifting through large amounts of information, and leading them straight to the point, enabling them rather than replacing them. This would free up their time to focus on solving the legal issues rather than looking for the legal issues. As one of the informants puts it; this may change the typical legal career. Or as another informant puts it; machine learning won't make being a lawyer more boring, it will only become more fun.

8.0 Analysis

8.1 The purpose of the analysis

The purpose of the analysis is to relate the results from chapter 7 to the theory covered in prior chapters. This will result in conclusions being drawn in chapter 9.

8.2 The machine learning impact

8.2.1 The current impact

Based on the responses from the informants, a broad representation of all the tasks discovered by McKinsey is given on the vertical axis (see subchapter 2.5), while the areas Robert Dale identified as being affected by machine learning tools is represented on the horizontal axis (see subchapter 3.8). The reason for using the areas identified by Dale on the horizontal axis, is that they span broadly based on what the technology can do.

The responses from the informants are ranged from:

- Green – no impact
- Yellow – limited impact
- Red – significant impact

tasks/tools	Legal research	Electronic discovery	Document reveiw	Document automation	Legal advice
Arbitration					
Draft regulations					
Evaluate records					
Researching sources of law					
Interviews and meetings					
Prepare documents					
Research documents					
Provide legal advice/services					
Proceedings					
Supervision of staff					

The table seems to indicate that the impact of machine learning is very limited, to a few areas, and only one are of significance. This view was also generally expressed by the informants, stating that the impact is very limited at the moment, except electronic discovery.

As demonstrated by both Justify and DoNotPay, machine learning aided document automation systems are practical, but the informants didn't believe this was having a real impact on the industry. These clients are clients currently not served by the industry, with the consequence of having no impact, several claimed.

The informants consistently explained that machine learning rarely gives the answer, but may aid and augment the way the lawyer performs a task. These tools do, in general, not replace lawyers, but some of the tools represent a significant reduction in the amount of time spent at certain repetitive and tedious tasks.

8.2.2 The potential near-future impact

While the table indicates that the impact of machine learning is very limited at the moment, there are large areas of obvious improvement potential. Obvious examples, based on the machine learning chapter, are:

- Augmenting reference work searches and document templates with machine learning clustering algorithms, allowing for improved searchability
- Using chatbots to gather information through interaction with the lawyer, reducing the amount of time spent

Areas that are likely not to be affected by machine learning, are the tasks that rely on trust and confidence in the lawyer, and that require a creative approach – which lies at the core of how lawyers work, see e.g. subchapter 3.6.3.1. Faced with a client wishing for a specific outcome, one of the tasks of a lawyer is to explore the various options that may lead to that outcome. This is a highly creative task, unlikely to be mimicked by machines any time soon. Machines require either prior data in order to learn or an environment to interact with. When the client approaches the lawyer the specifics of the case may still be susceptible to alteration based on the advice given by the lawyer and according to the desired outcome of the client.

It is also worth noting that there is an increased use of data-roaming tools, i.e. tools that allow for information gathering and sharing over a secure channel. Some of these tools to integrate machine learning elements, e.g. iBase that uses fuzzy-clustering in order to improve searching, see subchapter 4.3.3. The informants seem to consider data-roaming tools as an

obvious way of conducting their work, as it improves the speed and interaction between lawyer and client, and this experience is likely to gain ground even in the lower tiers of the industry.

8.2.3 Time-frame

The informants had somewhat different approaches to how long it would take until machine learning makes a significant impact. As mentioned above, most of the informants knew about or used themselves systems for document review and electronic discovery. One of the informants stated that in this context, the future is already here. Another informant explained that they were training systems with the hope of reaping benefits in a few years; at the moment they were putting more time and effort into the system than they were getting out of it, but they believed in a future reward. A third informant stated that "*They have been talking about this* [AI in the legal industry] *for years, but nothing has happened thus far, so I believe it will still be a few years*".

This apparent discrepancy between the informants may be the results of how they perceive AI as a phenomenon. As explained in chapter 2, AI is understood in a variety of contexts, resulting in some degree of disconnection. If one perceives machine learning in the context of an Artificial General Intelligence, i.e. an AI which is capable of generally solving all tasks as well or better than humans, a development which is commonly referred to as *the singularity*, see e.g. Ben Goertzel (2007)[104], one is likely to be disappointed.

However, the informants interviewed are highly specialised within the field of law and technology, and this is unlikely to be the case. It is more likely that the current machine learning tools aren't having the massive impact that was predicted just a few years ago, following the LawGeex demonstration. The impact is limited, and because of the way the technology works, there is limited potential for its application to the legal industry. One simply can't program a reinforcement learning model to simulate arguing a case, because the rewards and the contextual understanding of the law is simply too complex. See especially subchapter 3.6.3.1.

However, if one views machine learning in relation to what the technology is capable of performing, especially within the area of electronic discovery, the future is certainly here. And because of the expert level of the informants, they are all likely to have a deep understanding of this.

The informants consistently focused on the importance of understanding their surroundings. In order to be relevant in the business/corporate tier of the industry, one is required to have a deep understanding of the respective industries being served. It is simply not sufficient to have a general understanding of the law; one also needs to have an intrinsic knowledge of the industry sector, e.g. the oil industry, retail industry etc. Several informants said that they didn't believe a machine would be able to gain a sufficient level of knowledge to become relevant in many of the tasks that they perform.

I suppose, having to draw a conclusion, the impact of machine learning on the legal industry is already demonstrated, and the real question is the extent to which these tools will be developed. Thus, the time-frame deals not mainly with technological advances, but with the implementation of the technology already available.

8.3 The innovators dilemma – industry forces and disruption

8.3.1 Invoicing by the hour and the consequences regarding machine learning

The informants consistently explained that they were invoicing by the hour, that machine learning is reducing the number of billable hours within specific tasks, and that it doesn't make sense for the industry to develop its own tools because of this. The only way it would make sense to develop or use these tools, is if they provide a competitive advantage or if they provide a net increase in profits. Several informants explained that they might have to change their pricing strategies to deal with the impact of machine learning. They also consistently explained that the legal industry would not be developing machine learning tools; if would be acquired from third party or as white label tools from abroad.

In subchapter 2.5 was introduced two studies that deal with the invoicing practices of lawyers. Lawyers in general charge by the hour, and the hourly rate doesn't necessarily indicate an objective quality standard (see especially the 2019 inquiry into the reimbursement of lawyer's fees). In the book "*Trender og utfordringer i regnskap og revisjon*" chapter 9[105] which deals with the hourly rate as an indicator of quality in the services provided by chartered accountants, a high price is perceived as a sign of high quality in urban areas, but the opposite is the case in rural areas. The natures of the two industries are highly comparable, though the legal industry to a much higher degree works on behalf of private individuals, and it is reasonable to assume that the perception of price and quality is of similar character in both industries; the hourly rate is not an objective measurement of quality, but is nevertheless perceived as such. Further research is required to confirm or reject this assumption, and this thesis does not explore this further, but assumes that there is a comparability.

However, the practice of invoicing by the hour does present the legal industry with several major challenges in respect to the introduction of both machine learning software and document automation; the latter requires a machine learning component in order to be really effective, see subchapter 3.8, especially 3.8.4.

In chapter 8 of "*Trender og utfordringer i regnskap og revisjon*"[106] Åmo and Gårseth-Nesbakk explore the pricing strategies of the accounting industry.As Åmo and Gårseth-Nesbakk explain in subchapter 8.2.1, the cost-plus pricing strategy involves pricing the products and services based on the cost of producing these, and then adding a premium that covers both indirect costs and the profit. This method is called "cost-plus". And this represents a potential major double-edged sword; on the one hand it may force lawyers to reduce prices and reduce profits, and on the other hand it may lead to industry fragmentation from the bottom up; both increasing internal rivalry and letting other industries provide complex legal services.

8.3.2 Innovating may reduce profits

One of the possible outcomes of this may be that law-firms continue billing by the hour, but with the added cost of machine learning software licenses. Due to the Code of Conduct introduced in subchapter 3.7.1 lawyers are required to provide their services in a manner which is not unnecessarily costly, and this may force the industry into using tools that dramatically cuts the amount of billable hours.

In the best-case scenario, the software increases the complexity of the service provided in such a way as to allow for retaining or even increasing the number of billable hours. This is not a very likely scenario due to the nature of machine learning, which is to explore data created by us, and find patterns that allow for speeding up the process. But if the number of billable hours somehow remains the same or even slightly increases, so does the cost of providing the service, as software becomes an increased expenditure. And if the number of billable hours fall drastically, at least in some areas, the cost of software still remains, further reducing the profits when the profits are based on billable hours.

8.3.3 Not innovating may causes loss of turf

The second possible outcome of this double-edged sword is that law-firms don't adopt, and gradually loose turf to new-comers such as rival industries and new entrants with little or no legal qualification. By underestimating there rivals the legal industry may not only lose a little turf, but in the same process allow a market for rivals using machine learning tools such as document automation combined with a learning system. And as the rivals establish

themselves, the new market is demonstrated, fuelling further development. This may lead to a spiral of losing turf to newcomers, and by the time the law-firms discover that they have lost a significant part of their market, it may be too late to attempt to catch up; parts of the unique product the legal industry provides may have been democratised due to machine learning, for ever removing the traditional high profits of that industry.

Two obvious examples of such rivals were mentioned above.

In chapter 4 was explored a man called Joshua Browder, who out of pure annoyance created the chatbot DoNotPay that freely lets people contest claims. This is a great example of new technology that gives consumers at the bottom of a market access to legal services that otherwise would be provided by lawyers.

The same can be said for *Justify*, also explored in chapter 4, that gives customers at the bottom of the market access to legal services. Though *Justify* is in fact a law-firm, the system demonstrates how easily a complex process can be automated, and the approach is likely to catch on in other contexts; there is no reason why someone couldn't create a completely free version with basically the same content, as the considerations are strictly steered by statutory law. The same goes for lease agreements, prenuptial agreements etc; it's just a matter of putting in the effort, and a completely free version could be provided for those willing to take the risk of not having a lawyer draw up the documents.

8.3.4 What does theory tell us?

Law-firms could view machine learning as an opportunity to work differently, expanding into the possibly significant unmet demand for legal services claimed by the previously mentioned Norwegian Public Examinations, e.g. NOU 2015:3 pages 77 and 82. But according to the influential theories of Clayton Christensen, law-firms will be reluctant to change their pricing strategies. In his book "*The Innovator's Dilemma: When new technologies cause great firms to fail*"[107] law-firms will consider machine learning only as efficiency tools, not as an opportunity for working differently.

Christensen explains that businesses tend to innovate at the highest tiers of their market because that's where the highest profits traditionally have been achieved *"by charging the highest prices to the most demanding and sophisticated customers at the top of market."*. According to Christensen, businesses have learned over time that this approach works, developing a cultural bias against change; if it works, don't fix it. Christensen calls this being

"*Held captive by their customers"*, forcing businesses into "*sustaining innovation*" because that is where the profits are. And this strategy is vulnerable to "*disruptive innovation*".

Christensen explains that disruptive innovation gives "*a whole new population of consumers at the bottom of a market access to a product or service that was historically only accessible to consumers with a lot of money or a lot of skill*".

With direct relevance to this thesis, Christensen's theories explain that disruptive innovation strategies aren't attractive to successful businesses. Disruptive innovation returns lower margins initially, target smaller markets, and the products are simpler than sustaining innovation strategies. Applied to the legal industry, traditional and well-established law-firms are likely to incrementally improve services for their most profitable customers rather than tap into a potentially large market that is unmet, because the returns aren't as great. Even though the business itself might still be quite profitable by serving more clients by charging each individual client less.

As rivals enter the market armed with machine learning software, the low-end lawyers delivering the easy legal services, risk being outcompeted on possibly both price and quality. Used to charging 1,500-2,000 NOK an hour they might be outperformed by accountants, financial advisers, IT-developers and students charging 500-1,500 NOK an hour. But the low-end lawyers won't go away just because their bread-and-butter work disappears. They will start competing with the higher end lawyers on more complex tasks. They do, after all, have the same education, possibly further fuelling industry disruption.

8.3.5 Possible approaches to curbing these challenges

There are two factors that should be taken into consideration when looking at possible approaches to curbing the challenges above. First, in subchapter 8.3.4 we looked at studies dealing with the unmet demand for legal services, i.e. that there is a large group of potential legal service customers that are currently not having their needs met. Second, the value of legal services to the customer does not lessen just because the lawyer has achieved the level of advice in shorter time than previously. If machine learning and document automation tools aiding the lawyer speed up the process, leading to fewer billable hours, the value to the customer is still the same.

The price of a unit of legal service, is hourly rate multiplied by number of hours. When there is an increase in available units because of technological innovation, the law of

supply and demand predicts that the price will fall. However, holding all else equal, the net profit for a lawyer may remain the same.

Example: The hourly rate is 2,000 NOK and the average probation of a will takes 25 hours, giving a unit price of 50,000 NOK. The lawyer sells 50 of these units per year, at a total gross profit of 2,5 million NOK. Machine learning document automation and machine learning information gathering is developed that allows the lawyer to reduce the time spent on an average probation of a will to 10 hours. Billing by the hour means that the lawyer only charges 20,000 NOK for the same job. The immediate impression is that the lawyer has "lost" 30,000 NOK on the job, which is bad business. However, because there is an unmet demand for legal services, the lawyer is capable of probating 125 wills a year. With the same hourly rate, he still makes 2,5 million NOK.

Example 2: The hourly rate is 2,000 NOK and the average probation of a will takes 25 hours, same as above. The lawyer acquires machine learning tools, but stops charging by the hour. The service that previously cost 50,000 NOK is now sold at a fixed price of 29,999 NOK. The lawyer probates 125 wills a year but increases his profits from 2,5 million to 3,75 million NOK. Changing the pricing strategy from an hourly rate to a fixed price both increases profits while saving time.

This approach, charging the customer for the value created rather than the time consumed, is one of the business concepts of Lawbotics. Merete Nygaard, founder of Lawbotics, states that "*I believe one will see a very different World if one automates half or more of the routine work done by lawyers. […] Law-firms might have to look if they should charge for the value created, rather than hour by hour, as today"*[108].

Though machine learning and various processes such as filling in documents and combining predetermined text in order to create complete documents may be automated, which require machine learning to be really effective, are impacting the industry, it is too early to say if the technology is advanced enough to really make a difference and if there really is an unmet demand for legal services. In the higher end tier of legal service customers, the demand for services is likely met to a higher degree than in the lower end tier because the customers aren't as price-sensitive – which is the premise for Christiansens theory. This may imply that only low-end legal service providers could benefit from shifting pricing strategy, but it is too early to say.

The main objective of this thesis is to explore what the impact of machine learning is. Future research is required to explore the pricing strategies relative to that impact, and especially relative to the possible near future impact of yet to be introduced machine learning tools.

8.4 Machine learning as a resource

A fifth informant repeated the sentiment that machine learning tools will save them from time and workload so that they can focus on the interesting and most pressing tasks at hand. Several informants specifically mentioned that they believed machine learning would make the legal industry more interesting, by letting lawyers focus on the interesting parts of the law rather than perform work they considered menial.

The informants saw machine learning as an opportunity to keep up with the competition. Two of the informants explained that their clients commonly ask for various contract templates for free. Machine learning may allow for increased standardisation, satisfying client requirements, allowing the firms to stay competitive. Another informant stated that a lawyer might become liable if they didn't use tools that were common in the industry. A fourth informant explained that even though using machine learning tools might reduce the number of billable hours, a concern raised by other informants, using machine learning tools might give increased access to clients allowing for both increased turnover and more interesting work. This includes not only machine learning, but all digital tools.

When discussing machine learning with the informants, several mentioned that they are passive recipients of technology, either from third parties or as white label from foreign law-firms. The Norwegian legal industry is generally not developing machine learning tools itself, at least not in the context of moving beyond automating specific documents. The industry considers technology based on the perceived benefit of acquiring it.

According to the resource-based view on strategy, see subchapter 5.3, a law-firm gaining a competitive advantage by acquiring a valuable new machine learning tool, the tool thus also being rare, is only able to sustain that competitive advantage if the resource is hard or impossible to imitate. But when the tools are developed externally, there is likely no rarity, as the developers are themselves pursuing a profit. And there is no need to imitate a tool which isn't rare; you just simply also get a copy.

When *Lovdata* (see subchapter 4.2.1) was introduced, it represented a significant benefit over manually exploring the extensive physical libraries commonly seen in law-

offices. With a few clicks, the relevant sources of the law could be gathered, and explored in peace. Competitors without access to *Lovdata* would still manually turn pages for hours, looking for the same information. Of course, *Lovdata* was in part free, and in part available for purchase to anyone. The competitive advantage of using *Lovdata* was non-existant.

The competitive advantage of using third-party machine learning legal tech will be short lived, if existent at all; competing law-firms are likely to soon catch on to the new industry standard of using a specific tool, negating any benefit.

9.0 Conclusions

9.1 Research questions

9.1.1 How are law-firms responding to the machine learning phenomenon?

Based on the findings, the law-firms are responding to machine learning by keeping the technology under close observation, and they are acquiring tools with machine-learning elements, but so far it seems that there has been little or no innovation.

One reason for the lack of innovation, is the limited impact machine learning is perceived to have.

Though machine learning tools are capable of making a significant impact on certain processes, this has so far not lead to innovation, e.g. changing pricing strategies so that the price reflects the perceived value of the service rather than (the reduced) time it takes to perform the service.

It is my clear impression that the informants are highly aware of the current level of technology, that they are sharing that knowledge, and that the tools in use are being used in an attempt to command the industry standard. Thus, the lack of innovation is likely not caused by a lack of knowledge of the technology itself or tools utilising it.

9.1.2 What do law-firms believe the future role of machine learning in the legal industry to be?

Based on the findings, the informants believe that machine learning will have an impact on the future of the industry, but that the impact may be limited by the complexity of the tasks that they perform, the deeper understanding of their business environment that is required in order to understand what the correct solution is, and the level of creativity that is required by the industry in order to provide custom advice.

9.1.3 What are law-firms doing to be a part of a machine learning future?

There are fragmented examples of actors attempting to do things in a new way, such as Lawbotics, Justify, and DoNotPay. The general tendency, however, is that the legal industry isn't innovating beyond gradual and industry-wide adaptation of tools that improve on existing methods of working.

The industry seems to be behaving in accordance with the theories of Clayton Christensen covered in subchapter 8.3, not innovating because it isn't in their interest. This

allows for what Christensen calls “The innovators dilemma”, potentially allowing disruption. However, it is too early to make any reasonable predictions on whether there will be any disruption, as the current impact – and the seeming potential of machine learning – is more limited than the initial hype of just a few years back seemed to imply.

In addition, law-firms are not pursuing a competitive strategy based on machine learning tools being a valuable resource. Though the largest firms are attempting to command the industry standard, the industry standard only gives short-lived competitive advantages, as they are soon acquired by the competition. The theories of Barney indicate that this strategy will not result in a competitive advantage.

9.2 What did they say that didn’t fit the assumptions / theories

Expected findings that were confirmed:

- That machine learning is incapable of performing legal judgement, creatively planning a process in order to maximise a specific coal.
- That lawyers are somewhat reluctant in relation to investing in technology that reduces billable hours, in accordance with business strategy theory.
- That lack of technical aptitude is a factor in the lower tier of the legal industry, possibly limiting adaptation of even rudimentary technology. In spite of the machine learning hype, the majority of computer software applied in the legal industry are first wave AI, i.e. software that follows specific instructions rather than learns from data.

The biggest surprise is that machine learning is making such an impact in document review / e-discovery, and that it is possibly pushing law-firms towards changing its pricing strategies.

9.3 Conclusions and potential for future research

A strong case has been made that machine learning tools are impacting the legal industry and will continue to do so, over time reducing the current requirements of manually performing the various routine tasks in a law-firm. Machine learning tools are already assisting lawyers in discovering errors and potential risks within narrow areas with a higher degree of precision/quality than human lawyers are capable of on their own, and certainly speeding up the processes. However, the complexity of the legal industry, its environment and the creativity required to perform the tasks of lawyers is an important limiting factor.

This thesis provides a preliminary exploration into various areas that could be researched in the future.

The degree of impact rests with the ability to develop software that becomes more broadly applicable, beyond the narrow tasks currently served. Future research is recommended to deal with this problem.

If the software being developed becomes broadly applicable, spanning many or all of the tasks currently performed by lawyers and augmenting the way lawyers perform those tasks, the industry isn't adapting its strategies to cope. According to the theories of Christensen, within the framework of Porter, and the theories of Barney, there certainly is a potential for industry disruption. Advantages in machine learning software will not primarily be gained by the legal industry, but by software developers, rival industries and new entrants. Future research is recommended to deal with this problem, because the data isn't sufficient to draw definite conclusions.

9.4 What is the impact of this?

One important impact is that we are still not seeing signs of the impending disruption that has been predicted. The technology may cause disruption if its full potential is unleashed, but so far machine learning has seen limited application in the industry. The industry is still considering various rule-based document automation systems, seeming to support the assumption that the cost of developing sufficiently effective machine learning tools is simply too high, and the complexity of the tasks to great, to cause rapid development of tools that could unleash the potential of machine learning.

9.5 What have we learned?

This research has provided insight into how the Norwegian legal industry views the current technological advances and the impact on the industry. This problem has not been tackled qualitatively before.

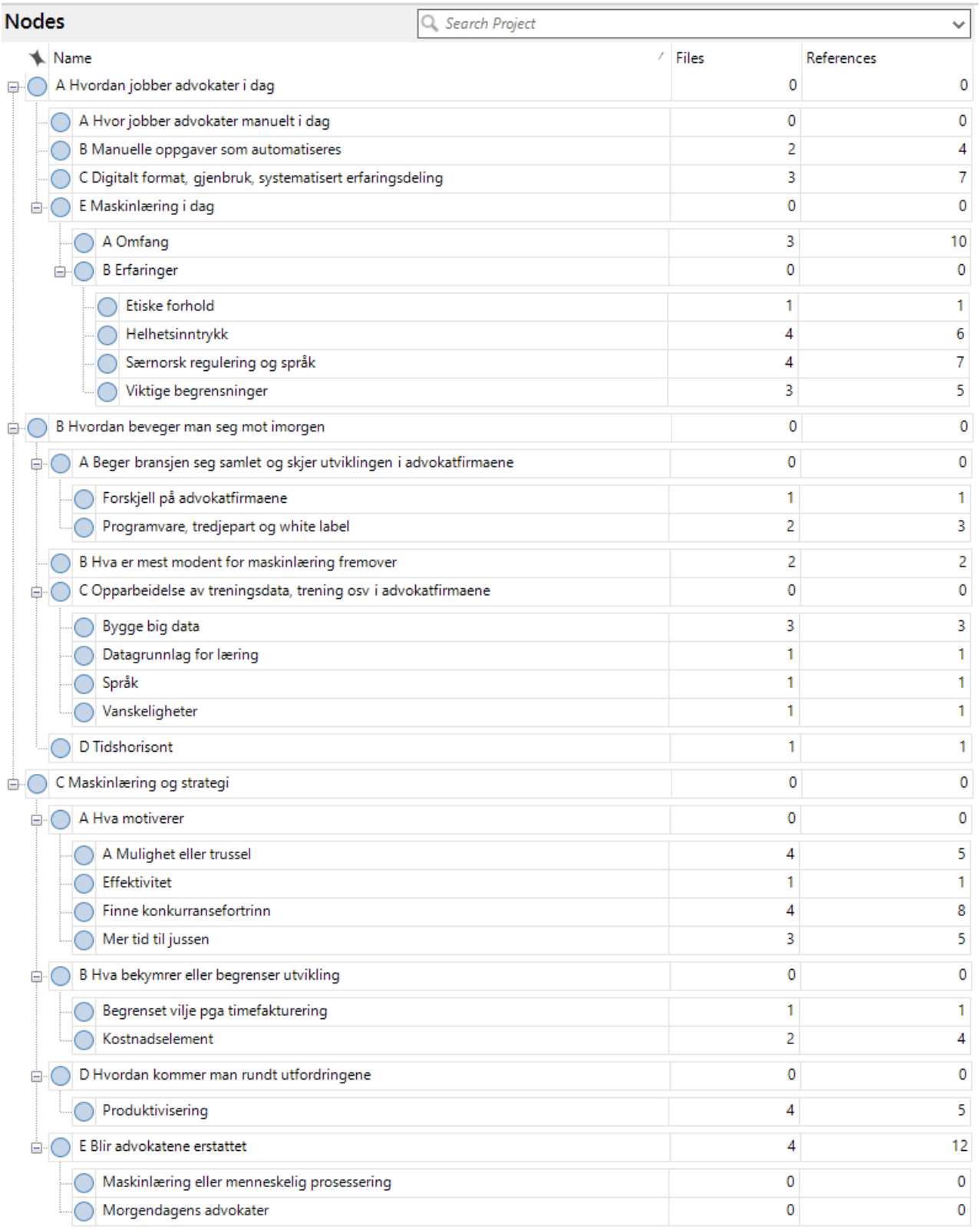
9.6 Weak points / the weakest points in the thesis

A weak point in the thesis is transferability. Though there is much research dealing with the pricing strategies of law-firms, further research is required to give the findings generality.

Another weak point is the broadness of the research; the scope of an MBA thesis isn't sufficient to sufficiently explore the research questions.

Figures and illustration

Figure 1



Name	Files	References
A Hvordan jobber advokater i dag	0	0
A Hvor jobber advokater manuelt i dag	0	0
B Manuelle oppgaver som automatiseres	2	4
C Digitalt format, gjenbruk, systematisert erfaringsdeling	3	7
E Maskinlæring i dag	0	0
A Omfang	3	10
B Erfaringer	0	0
Etiske forhold	1	1
Helhetsinntrykk	4	6
Særnorsk regulering og språk	4	7
Viktige begrensninger	3	5
B Hvordan beveger man seg mot imorgen	0	0
A Beger bransjen seg samlet og skjer utviklingen i advokatfirmaene	0	0
Forskjell på advokatfirmaene	1	1
Programvare, tredjepart og white label	2	3
B Hva er mest modent for maskinlæring fremover	2	2
C Opparbeidelse av treningsdata, trening osv i advokatfirmaene	0	0
Bygge big data	3	3
Datagrunnlag for læring	1	1
Språk	1	1
Vanskeligheter	1	1
D Tidshorisont	1	1
C Maskinlæring og strategi	0	0
A Hva motiverer	0	0
A Mulighet eller trussel	4	5
Effektivitet	1	1
Finne konkurransefortrinn	4	8
Mer tid til jussen	3	5
B Hva bekymrer eller begrenser utvikling	0	0
Begrenset vilje pga timefakturering	1	1
Kostnadselement	2	4
D Hvordan kommer man rundt utfordringene	0	0
Produktivisering	4	5
E Blir advokatene erstattet	4	12
Maskinlæring eller menneskelig prosessering	0	0
Morgendagens advokater	0	0

Attachment 1 – Reporting the study to *Norsk Senter for forskningsData (NSD*)

Meldeskjema 128159

Sist oppdatert 23.02.2020

Hvilke personopplysninger skal du behandle?

Navn (også ved signatur/samtykke)

Lydopptak av personer

Bakgrunnsopplysninger som vil kunne identifisere en person

Andre opplysninger som vil kunne identifisere en fysisk person

Type opplysninger

Du har svart ja til at du skal behandle bakgrunnsopplysninger, beskriv hvilke

Navn og hvilket advokatfirma vedkommende er ansatt i.

Du har svart ja til at du behandler andre opplysninger som vil kunne identifisere en person, beskriv hvilke

Om vedkommende er ansatt advokat eller partner i det aktuelle advokatfirmaet.

Skal du behandle særlige kategorier personopplysninger eller personopplysninger om straffedommer eller lovovertredelser?

Nei

Prosjektinformasjon

Prosjekttittel

A qualitative exploration into the possible impact of Machine Learning (Artificial Intelligence) on large Norwegian law-firms

Dersom opplysningene skal behandles til andre formål enn behandlingen for dette prosjektet, beskriv hvilke

Ikke aktuelt.

Begrunn behovet for å behandle personopplysningene

For å kunne undersøke hvordan advokatfirmaene forholder seg til ny teknologi er det nødvendig å snakke med representanter for disse firmaene. Jeg begrenser personopplysningene som samles inn til kun å gjelde navn, hvilken posisjon vedkommende har i firmaet (ansatt advokat, partner eller liknende), og den autoriserte tittelen "advokat" eller "advokatfullmektig".

Ekstern finansiering

Type prosjekt

Studentprosjekt, masterstudium

Kontaktinformasjon, student

Anthony Christopher Caffrey, tony@caffrey.no, tlf: 47443396

Behandlingsansvar

Behandlingsansvarlig institusjon

Nord Universitet / Handelshøgskolen / Marked, organisasjon og ledelse

Prosjektansvarlig (vitenskapelig ansatt/veileder eller stipendiat)

Bjørn Willy Åmo, bjorn.w.amo@nord.no, tlf: 75517245

Skal behandlingsansvaret deles med andre institusjoner (felles behandlingsansvarlige)?

Nei

Utvalg 1

Beskriv utvalget

Advokater med kompetanse på strategi/utvikling

Rekruttering eller trekking av utvalget

Representanter for de 20 største advokatfirmaene i Norge, rangert etter omsetning i 2017.

Alder

25 - 67

Inngår det voksne (18 år +) i utvalget som ikke kan samtykke selv?

Nei

Personopplysninger for utvalg 1

Navn (også ved signatur/samtykke)

Lydopptak av personer

Bakgrunnsopplysninger som vil kunne identifisere en person

Andre opplysninger som vil kunne identifisere en fysisk person

Hvordan samler du inn data fra utvalg 1?

Personlig intervju

Vedlegg

intervjuguide.docx

Grunnlag for å behandle alminnelige kategorier av personopplysninger

Samtykke (art. 6 nr. 1 bokstav a)

Informasjon for utvalg 1

Informerer du utvalget om behandlingen av opplysningene?

Ja

Hvordan?

Skriftlig informasjon (papir eller elektronisk)

Informasjonsskriv

informasjonsskriv studien.doc

Tredjepersoner

Skal du behandle personopplysninger om tredjepersoner?

Nei

Dokumentasjon

Hvordan dokumenteres samtykkene?

Manuelt (papir)

Hvordan kan samtykket trekkes tilbake?

Per epost tony@caffrey.no eller telefon 47443396.

Hvordan kan de registrerte få innsyn, rettet eller slettet opplysninger om seg selv?

På forespørsel blir transkriberte intervjuer umiddelbart oversendt per epost. Transkripsjonene inneholder selve intervjuene samt navn og firma. Personopplysningene kan rettes eller anonymiseres frem til innleveringsprosessen.

Totalt antall registrerte i prosjektet

1-99

Tillatelser

Skal du innhente følgende godkjenninger eller tillatelser for prosjektet?

Behandling

Hvor behandles opplysningene?

Maskinvare tilhørende behandlingsansvarlig institusjon

Private enheter

Retningslinjer/tillatelse til å behandle opplysninger på private enheter

retningslinjer for informasjonsbehandling.docx

Hvem behandler/har tilgang til opplysningene?

Prosjektansvarlig

Student (studentprosjekt)

Interne medarbeidere

Tilgjengeliggjøres opplysningene utenfor EU/EØS til en tredjestat eller internasjonal organisasjon?

Nei

Sikkerhet

Oppbevares personopplysningene atskilt fra øvrige data (kodenøkkel)?

Nei

Begrunn hvorfor personopplysningene oppbevares sammen med de øvrige opplysningene

De transkriberte intervjuene påføres et nummer som korresponderer med et nummer på samtykkeerklæringene i papir. Disse oppbevares ikke sammen. Selve masteroppgaven kan inneholde navn på informanten og navn på firmaet informanten representerer dersom informanten ønsker det, og informanten blir sitert.

Hvilke tekniske og fysiske tiltak sikrer personopplysningene?

Opplysningene krypteres under forsendelse

opplysningene krypteres under lagring

Andre sikkerhetstiltak

Opplysningene anonymiseres

Adgangsbegrensning

Hvilke

De transkriberte intervjuene oppbevares ikke sammen med personopplysningene, som kun foreligger i form av samtykkeerklæringene på papir.

Varighet

Prosjektperiode

14.02.2020 - 15.06.2020

Skal data med personopplysninger oppbevares utover prosjektperioden?

Nei, data vil bli oppbevart uten personopplysninger (anonymisering)

Hvilke anonymiseringstiltak vil bli foretatt?

Personidentifiserbare opplysninger fjernes, omskrives eller grovkategoriseres

Lyd- eller bildeopptak slettes

Vil de registrerte kunne identifiseres (direkte eller indirekte) i oppgave/avhandling/øvrige publikasjoner fra prosjektet?

Ja

Begrunn

Informantene tilbys anonymisering, men fordi utvalget er så spesifikt (de aller største advokatfirmaene i Norge) og fordi disse firmaene i noen tilfeller utvikler tekniske løsninger som er av stor betydning for studien, kan relevante sitater og angivelser av teknologi/firma bli gjengitt i oppgaven, under forutsetning av aksept fra de respektive informantene.

Tilleggsopplysninger

--- fine ---

Attachment 2 – Information to informants about the study

Vil du delta i forskningsprosjektet
"A qalitative inquiry into the possible impact of Machine Learning (Artificial Intelligence) of large Norwegian lawfirms"?

Dette er et spørsmål til deg om å delta i et forskningsprosjekt hvor formålet er å avdekke hvordan store advokatfirmaer forholder seg til kunstig intelligens, og hvordan disse firmaene innretter seg med tanke på fremtiden. I dette skrivet gir vi deg informasjon om målene for prosjektet og hva deltakelse vil innebære for deg.

Formål

Formålet med prosjektet er å undersøke hvordan store advokatfirmaer forholder seg til maskinlæringsfenomenet, hvilken rolle de tror at fenomenet vil ha for fremtidens advokatkontor, og hvordan de innretter seg for å være en del av denne fremtiden.

Prosjektet er en masteroppgave til tittelen Master of Business Administration ved Nord Universitet.

Opplysningene brukes ikke til andre formål enn nærværende prosjekt.

Hvem er ansvarlig for forskningsprosjektet?

Nord Universitet/HHN ved professor Bjørn W Åmo er ansvarlig for prosjektet.

Hvorfor får du spørsmål om å delta?

Du er en medarbeider/partner med spesiell kompetanse om temaet i et av de 20 største advokatfirmaene i Norge rangert etter omsetning i 2017. Du er enten kontaktet direkte fordi du tidligere har uttalt deg offentlig om temaet for forskningsprosjektet, eller fordi du er utpekt av det aktuelle advokatfirmaet som en medarbeider/partner med spesiell kompetanse om temaet.

Hva innebærer det for deg å delta?

Intervjuet gjennomføres ansikt til ansikt og tar ca 30 minutter. Det gjøres lydopptak av intervjuet, og opptaket transkriberes. Hvis du velger å delta kan du velge å svare anonymt. Hvis du ikke tar forbehold om anonymitet, kan svarene dine bli gjengitt sammen med navn/tittel/firma i masteroppgaven. Svarene dine vil inngå i en analyse av hvordan store advokatfirmaer generelt ser på fenomenet maskinlæring. Sitater fra intervjuet kan bli brukt for å belyse funnene.

Det er frivillig å delta
Det er frivillig å delta i prosjektet. Hvis du velger å delta kan du trekke deg helt fra studien innen 25. mars 2020. Du kan også avbryte intervjuet mens det pågår, hvis du ombestemmer deg. I så fall vil ingen del av intervjuet bli benyttet i studien, og alle opplysninger destrueres.

Etter 25. mars 2020 vil opplysningene være innarbeidet i analysen, men du kan fremdeles kreve anonymisering av alle opplysninger frem til 10. mai. Etter 10. mai er oppgaven i en innleveringsprosess.
Ditt personvern – hvordan vi oppbevarer og bruker dine opplysninger

Vi vil bare bruke opplysningene om deg til formålene vi har fortalt om i dette skrivet. Vi behandler opplysningene konfidensielt og i samsvar med personvernregelverket.

- Det er bare MBA-student Anthony Christopher Caffrey og veileder/forskningsansvarlig, sensor og eventuelle administrativt ansatte ved Nord Universitet som har tilgang til hele lydopptaket og transkripsjonen.
- Opptak og transkripsjonen oppbevares som datafiler hos Universitetet i Oslo (nettskjema.no) eller hos Nord Universitet (Canvas).

Selve masteroppgaven er offentlig tilgjengelig etter sensur.

Hva skjer med opplysningene dine når vi avslutter forskningsprosjektet?

Prosjektet skal etter planen avsluttes 15. mai 2020. Lydopptakene vil bli slettet etter at masteroppgaven er sensurert, ca 15. juni 2020. Transkripsjonene blir oppbevart i anonymisert form hos Nord Universitet på ubestemt tid. Andre kopier av transkripsjonene destrueres/slettes. Funnene fra intervjuet samt enkeltsitat vil leve evig i den ferdige masteroppgaven. Funnene fra masteroppgaven kan bli gjengitt andre steder, for eksempel i *Advokatbladet*.

Dine rettigheter

Så lenge du kan identifiseres i datamaterialet, har du rett til:

- innsyn i hvilke personopplysninger som er registrert om deg,
- å få rettet personopplysninger om deg,
- få slettet personopplysninger om deg,
- få utlevert en kopi av dine personopplysninger (dataportabilitet), og
- å sende klage til personvernombudet eller Datatilsynet om behandlingen av dine personopplysninger.

Hva gir oss rett til å behandle personopplysninger om deg?

Vi behandler opplysninger om deg basert på ditt samtykke.

På oppdrag fra Nord Universitet har NSD – Norsk senter for forskningsdata AS vurdert at behandlingen av personopplysninger i dette prosjektet er i samsvar med personvernregelverket.

Hvor kan jeg finne ut mer?

Hvis du har spørsmål til studien, eller ønsker å benytte deg av dine rettigheter, ta kontakt med:

- Nord Universitet ved professor Bjørn Willy Åmo, bjorn.w.amo@nord.no telefon 75517245 og student Anthony Christopher Caffrey, tony@caffrey.no telefon 47443396.
- Vårt personvernombud: Toril Irene Kringen, personvernombud@nord.no, telefon 75517000
- NSD – Norsk senter for forskningsdata AS, personverntjenester@nsd.no, telefon: 55 58 21 17.

Med vennlig hilsen

Bjørn Willy Åmo
professor/veileder

Anthony Christopher Caffrey
student

Samtykkeerklæring

Jeg har mottatt og forstått informasjon om masteroppgaveprosjektet *A qalitative exploration into the possible impact of Machine Learning (Artificial Intelligence) on large Norwegian lawfirms*, og har fått anledning til å stille spørsmål.

Jeg samtykker til:

- ☐ å delta i intervju og at intervjuet inngår i en kvalitativ studie
- ☐ at mitt navn, yrkestittel, arbeidsgiver og eventuelt partnerskap i det aktuelle advokatkontoret kan bli knyttet til sitater og liknende i studien,
  ELLER
- ☐ at intervjuet inngår i studien under forutsetning av at jeg og firmaet jeg er tilknyttet ikke blir identifisert,
- ☐ at disse opplysningene blir publisert i en offentlig tilgjengelig masteroppgave, samt at Nord Universitet oppbevarer transkripsjonene i anonymisert form på ubestemt tid

Eventuelle andre betingelser/forutsetninger knyttet til samtykket:

Jeg samtykker til at mine opplysninger behandles frem til prosjektet er avsluttet, ca. 15. juni 2020

(Signert av prosjektdeltaker, dato)

Attachment 3 – Approval from *Norsk Senter for forskningsData (NSD)*

Det innsendte meldeskjemaet med referansekode 128159 er nå vurdert av NSD. Følgende vurdering er gitt: Det er vår vurdering at behandlingen av personopplysninger i prosjektet vil være i samsvar med personvernlovgivningen så fremt den gjennomføres i tråd med det som er dokumentert i meldeskjemaet 24.02.2020 med vedlegg, samt i meldingsdialogen mellom innmelder og NSD. Behandlingen kan starte. MELD VESENTLIGE ENDRINGER Dersom det skjer vesentlige endringer i behandlingen av personopplysninger, kan det være nødvendig å melde dette til NSD ved å oppdatere meldeskjemaet. Før du melder inn en endring, oppfordrer vi deg til å lese om hvilke type endringer det er nødvendig å melde: nsd.no/personvernombud/meld_prosjekt/meld_endringer.html Du må vente på svar fra NSD før endringen gjennomføres. TYPE OPPLYSNINGER OG VARIGHET Prosjektet vil behandle alminnelige kategorier av personopplysninger frem til 15.06.2020. LOVLIG GRUNNLAG Prosjektet vil innhente samtykke fra de registrerte til behandlingen av personopplysninger. Vår vurdering er at prosjektet legger opp til et samtykke i samsvar med kravene i art. 4 og 7, ved at det er en frivillig, spesifikk, informert og utvetydig bekreftelse som kan dokumenteres, og som den registrerte kan trekke tilbake. Lovlig grunnlag for behandlingen vil dermed være den registrertes samtykke, jf. personvernforordningen art. 6 nr. 1 bokstav a. PERSONVERNPRINSIPPER NSD vurderer at den planlagte behandlingen av personopplysninger vil følge prinsippene i personvernforordningen om: - lovlighet, rettferdighet og åpenhet (art. 5.1 a), ved at de registrerte får tilfredsstillende informasjon om og samtykker til behandlingen - formålsbegrensning (art. 5.1 b), ved at personopplysninger samles inn for spesifikke, uttrykkelig angitte og berettigede formål, og ikke viderebehandles til nye uforenlige formål - dataminimering (art. 5.1 c), ved at det kun behandles opplysninger som er adekvate, relevante og nødvendige for formålet med prosjektet - lagringsbegrensning (art. 5.1 e), ved at personopplysningene ikke lagres lengre enn nødvendig for å oppfylle formålet DE REGISTRERTES RETTIGHETER Så lenge de registrerte kan identifiseres i datamaterialet vil de ha følgende rettigheter: åpenhet (art. 12), informasjon (art. 13), innsyn (art. 15), retting (art. 16), sletting (art. 17), begrensning (art. 18), underretning (art. 19), dataportabilitet (art. 20). NSD vurderer at informasjonen som de registrerte vil motta oppfyller lovens krav til form og innhold, jf. art. 12.1 og art. 13. Vi minner om at hvis en registrert tar kontakt om sine rettigheter, har behandlingsansvarlig institusjon plikt til å svare innen en måned. FØLG DIN INSTITUSJONS RETNINGSLINJER NSD legger til grunn at behandlingen oppfyller kravene i personvernforordningen om riktighet (art. 5.1 d), integritet og konfidensialitet (art. 5.1. f) og sikkerhet (art. 32). For å forsikre dere om at kravene oppfylles, må dere følge interne retningslinjer og eventuelt rådføre dere med behandlingsansvarlig institusjon. OPPFØLGING AV PROSJEKTET NSD vil følge opp ved planlagt avslutning for å avklare om behandlingen av personopplysningene er avsluttet. Lykke til med prosjektet! Kontaktperson hos NSD: Tore Andre Kjetland Fjeldsbø Tlf. Personverntjenester: 55 58 21 17 (tast 1)

Attachment 4 – Interview guide

The interview is centred around these three themes:

- In light of the advances in machine learning, what do you believe the future will look like for law-firms?
- What are you currently doing with regards to machine learning, in order to be competitive in the future?
- What are your thoughts on the significant attention given to Artificial Intelligence and the legal industry by your surroundings, and do you feel that this focus has been justifiable?

Allow the respondent to speak freely at first, and then point in the direction of the specifics from the preliminary research. Possible additional questions are:

- Which tasks are the machine learning tools *specifically* performing?
- What is the degree of lawyer-AI interaction, decision making and problem solving?
- What are the thoughts on costs?
- How does the potential reduction in billable hours affect your willingness to explore AI?
- Does language or jurisdiction matter in the application of AI?
- Do you perceive this as a threat or as an opportunity?
- Are you actively pursuing AI?

List of terms covered in this thesis

Agent – In Computer Science an agent someone who performs a task. It can be the computer system itself (hardware and software), the model or a conceptual entity such as the agent in a reinforcement learning model.

Algorithm – Sequential instructions on how to solve a particular class of problems. One example is Euclid's algorithm for the greatest common divisor of two numbers invented in 2,300 years ago.

Analogue – Data, which is represented in physical form, the opposite of digital data represented as numbers, usually binary.

Artificial Intelligence - The science and engineering of technology that solves problems that usually require thinking.

Big Data – Information that is characterised by the vast volume of the information, the variety of the information (numerical, textual and images) and the velocity by which the information is updated, requiring specialised technology and analytical methods to transform it to value.

Binary – Mutually exclusive states such as 0 and 1, yes and no, on and of, true or false. Binary states are commonly represented as 0 and 1 in computer science. Any digital data can be represented in binary form.

Business – A specific commercial enterprise, offering services and/or goods within a specific area of an industry.

Chatbot – Software that is designed to interact via a conversation with the user. Some use machine learning in the form of natural language processing while others return queries based on keywords or similar wording. Both can learn from interaction, e.g. by asking the user to rate the relevance and quality of the response.

Computer Science – The study of processes that interact with data (information) and that can be represented as data in the form of programs.

Data – Information, such as letters, numbers, pictures, sounds and any other information however complex.

Data Mining – The process of extracting usable data patterns from Big Data, usually through algorithms.

Data Point – A piece of information that a machine learning system is trained to predict. It can be anything from a simple 0 or 1, or a complex item such as an aeroplane, depending on what the machine learning system attempts to predict.

Data Science – A multi-disciplinary that deals with the theories and methods of how insights are extracted from Big Data.

Digital – Representing data in numerical form, usually binary.

Digitization – Information that used to be analogue (i.e written on paper) is transformed to digital format, enabling computer software to use it.

Deep learning – A machine learning method which applies artificial neural networks and various algorithms, where learning can be supervised, unsupervised or partially supervised.

GPU – Graphics Processing Unit, the ultra-fast processors of 3D graphics cards that can be run in parallel, allowing for powerful mining.

Industry – A category of businesses generally offering the same kinds services or goods

Jurist – In a broad sense a jurist is someone who knows the law. In a narrow sense a jurist is someone who holds a post-graduate degree in jurisprudence.

Lawyer – A licensed practitioner of law. In Norway a lawyer is a jurist who practices law on behalf of clients.

Machine Learning – Artificial intelligence that predicts answers to hitherto unasked questions without specific instructions, by having learned patterns for solving such questions by examining previous data and/or through interaction with its environment and improving upon its own learning over time.

Natural Language Processing – A subfield of linguistics, computer science and machine learning that allows for interacting with a computer using natural language.

Pixel – Commonly used to describe the individual dots/squares of colour hue and saturation that make up a larger picture, but in the terminology of neural nets a pixel is used in the same way as “Data point”.

Robotic process automation – Software that step by step performs repetitive tasks usually performed manually. Each step may follow a predetermined instruction, and/or may involve machine learning if a decision has to be made.

Citations (APA)

[1] LawGeex (February 2018.). Comparing the Performance of Artificial Intelligence to Human Lawyers in the Review of Standard Business Contracts. Retrieved from: https://images.law.com/contrib/content/uploads/documents/397/5408/lawgeex.pdf

[2] Dahl, T. (2018, 31. mai) Advokatbransjen vil endre seg totalt. Advokatbladet. Retreived from https://www.advokatbladet.no/advokatens-fagdager-2018-digitalisering-legal-tech/advokatbransjen-vil-endre-seg-totalt/113676 September 10th 2019.

[3] Brynjolfsson, E. and McAfee, A. (2017/7/21) The business of artificial intelligence. Retrieved 1. August 2019 from https://hbr.org/cover-story/2017/07/the-business-of-artificial-intelligence

[4] Robertson, D. (2017, August 1). This is how Facebook's shut-down AI robots developed their own language – and why it's more common than you think. The Independent. Retrieved from https://www.independent.co.uk/voices/facebook-shuts-down-robots-ai-artificial-intelligence-develop-own-language-common-a7871341.html

[5] Nussey, P. (2016, January 20). How much of what lawyers do can be automated? A look at new research. Retrieved August 1, 2019, from https://www.linkedin.com/pulse/how-much-what-lawyers-do-can-automated-look-new-research-peter-nussey

[6] Chui, M., Manyika, J. and Miremadi, M. (2015) "Four fundamentals of workspace automation" published in McKinsey Quarterly

[7] Remus, Dana and Levy, Frank S., Can Robots Be Lawyers? Computers, Lawyers, and the Practice of Law (November 27, 2016). Available at SSRN: https://ssrn.com/abstract=2701092 or http://dx.doi.org/10.2139/ssrn.2701092

[8] Hill, C. (2016, March 16). Deloitte Insight: Over 100,000 legal roles to be automated. Legal Insider. Retrieved from https://legaltechnology.com/deloitte-insight-100000-legal-roles-to-be-automated/

[9] Bønke Grønning, T., & Gangnes, O. M. (2019, November 15). Er juristene klare for den nye digitale revolusjonen? Juristen. Retrieved from https://juristen.no/nyheter/2019/11/er-juristene-klare-den-nye-digitale-revolusjonen?_cldee=dG9ueUBjYWZmcmV5Lm5v&recipientid=contact-b04cf8965be4e811a820005056a3763c-2c3ac933d64d455c89c833f4b15ad546&esid=23d37bc6-0c0d-ea11-a811-000d3a228e6c

[10] Ab Rahman, Airini et al. (2017). "Emerging Technologies with Disruptive Effects: A Review". Perintis eJournal. 7 (2). https://www.researchgate.net/publication/321906585

[11] Spaak, T (2003) "Legal Positivism, law's normativity, and the normative force of legal justification" Ratio Juris. Vol. 16 no 4 December 2003 (469-85)

[12] Black, H. C., Nolan, J. R., & Nolan-Haley, J. M. (1990). Blacks law dictionary: definitions of the terms and phrases of American and English jurisprudence, ancient and modern by Henry Campbell Black. 6th ed. St. Paul, MN: West.

[13] The University of Oslo Faculty of law curriculum, accessed January 20th 2020 https://www.uio.no/studier/program/jus/index.html

[14] Frank, J. (1933) "Why not a clinical lawyer-school?" University of Pennsylvania Law Review and American Law Register, Vol 81 No 8, June, pages 907-923. doi:10.2307/3308391

[15] Lilleholt Kåre, Eckhoff, T., Mestad, O., Arnesen, F., & Selmer, K. S. (2004). Knophs oversikt over Norges rett §§ 1-5. Oslo: Universitetsforlaget

[16] Boe, E. M. (2012). Grunnleggende Juridisk Metode: en introduksjon til rett og rettstenkning (3. utg.). Oslo, Universitetsforlaget.

[17] Webster, Bergljot et al. (2015) Norway's public investigation into the legal industry NOU 2015:3 Advokaten i samfunnet, DIFI, Oslo

[18] The Norwegian Bar Association *Advokatforeningen* 2017 report into the salaries of lawyers. Retrieved November 20th 2019 from https://www.advokatforeningen.no/globalassets/advokatforeningen_salaerrapporten2017_web.pdf

[19] Nielsen, M. & Sørgaard Djupvik, H / Advokatforeningen & Kantar TNS. (2019) Bransjeundersøkelsen. Retreived from https://www.advokatforeningen.no/medlemstilbud/tall-og-statistikk/bransjeundersokelsen/

[20] Midttømme, K, Wifstad, K, Forsberg Johnsen, P. F, Aalen, P, Grimsby, G og Viblemo, T. E/Menon Economics og Oxford Research (januar 2019) Undersøkelse av salæroppgaver i saker med fri rettshjelp. Menon Publikasjon 5/2019. Hentet fra https://www.menon.no/analyse-stykkpris-medgatt/

[21] O-net summary report for: 23-1011.00 - Lawyers. Retrieved January 21, 2020, from https://www.onetonline.org/link/summary/23-1011.00

[22] Cave, S., Craig, C., Dihal, K., Dillon, S., Montgomery, J., Singler, B., & Taylor, L. (2018). Portrayals and perceptions of AI and why they matter. (DES5612)https://royalsociety.org/~/media/policy/projects/ai-narratives/AI-narratives-workshop-findings.pdf

[23] Passer, M. W., & Smith, R. E. (2007). Psychology: the science of mind and behaviour. New York, USA: McGraw-Hill Companies, Inc. page 313

[24] McCarthy, J., Minsky, M. L., Rochester, N. & Shannon, C. E. (1955). A proposal for the Dartmouth summer research project on Artificial Intelligence http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html

[25] Wooldridge, M. and Jennings, N. R. (1995) Intelligent agents: theory and practice, The Knowledge Engineering Review, Vol. 10:2, 1995, 115-152

[26] Nakagaki, Toshiyuki; Yamada, Hiroyasu; Tóth, Ágota (2000). "Intelligence: Maze-solving by an amoeboid organism". Nature. 407

[27] Nakagaki, Toshiyuki; Kobayashi, Ryo; Nishiura, Yasumasa; Ueda, Tetsuo (November 2004). "Obtaining multiple separate food sources: behavioural intelligence in Physarum plasmodium". Proceedings of the Royal Society B. 271 (1554): 2305–2310

[28] Barone, Jennifer (2008-12-09). "Top 100 Stories of 2008 #71: Slime Molds Show Surprising Degree of Intelligence". Discover Magazine. Retrieved October 21st 2019

[29] Poole, Mackworth & Goebel 1998, p. 1.

[30] Kaplan, Andreas and Haenlein, Michael (1 January 2019). "Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence". Business Horizons. 62 (1): 15–25.

[31] Launchbury, John (2017) A DARPA perspective on Artificial Intelligence. A DARPA Information Innovation Office statement released on Youtube on February 15th 2017. https://www.youtube.com/watch?v=-O01G3tSYpU&feature=emb_title

[32] Hsu, Feng-hsiung; Campbell, Murray (1995). "Deep Blue System Overview". Proceedings of the 9th international conference on Supercomputing. ACM. pp. 240–244.

[33] Pillow, J. W. and Sahani, M. (2019) Editorial overview: Machine learning, big data, and neuroscience. Current Opinion in Neurobiology 55:iii-iv May 2019

[34] Samuel, Arthur (1959). "Some Studies in Machine Learning Using the Game of Checkers". IBM Journal of Research and Development. 3 (3): 210–229.

[35] MathWorks free ebook on Machine Learning retrieved October 25th 2019 https://se.mathworks.com/campaigns/offers/machine-learning-with-matlab.html

[36] Fefferman, C., Mitter, Sanjoy and Narayanan, H. Testing the manifold hypothesis. The journal of the American Mathematical Society Vol 29 no 4 October 2016 p 983-1049

[37] Rebala, G., Ravi, A. and Churiwala, S. (2019) An Introduction to Machine Learning. Springer Nature Switzerland AG, Cham, Switzerland

[38] Buskey, B., & Woods, A. (2018). Making Sense of Pretrial Risk Assessments. Retrieved May 29, 2020, from https://www.nacdl.org/Article/June2018-MakingSenseofPretrialRiskAsses

[39] Fisher, Franklin M. Multiple Regression in Legal Proceedings. Columbia Law Review Vol. 80, No. 4 (May, 1980), pp. 702-736

[40] Burton, D. K.; Shore, J. E.; Buck, J. T. (1983). "A generalization of isolated word recognition using vector quantization". IEEE International Conference on Acoustics Speech and Signal Processing ICASSP. 8: 1021–1024

[41] Cao G., Song D., Bruza P. (2004) Fuzzy K-Means Clustering on a High Dimensional Semantic Space. In: Yu J.X., Lin X., Lu H., Zhang Y. (eds) Advanced Web Technologies and Applications. APWeb 2004. Lecture Notes in Computer Science, vol 3007. Springer, Berlin, Heidelberg

[42] Chen, C., Shyu, M. & Chen, S. Weighted subspace modeling for semantic concept retrieval using gaussian mixture models. Inf Syst Front 18, 877–889 (2016).

[43] Roweis, S. T.; Saul, L. K. (2000). "Nonlinear Dimensionality Reduction by Locally Linear Embedding". Science. 290 (5500): 2323–2326.

[44] Goldberg, Yoav (2016). "A primer on neural network Models for Natural Language Processing", Journal of Artificial intelligence Research, volume 575

[45] Scott Deerwester; Susan T Dumais; George W Furnas; Thomas K Landauer; Richard Harshman. (Sep 1990) "Indexing by Latent Semantic Analysis". Journal of the American Society for Information Science. 41 (6): 391–407

[46] Blei, David M.; Ng, Andrew Y.; Jordan, Michael I (January 2003). Lafferty, John (ed.). "Latent Dirichlet Allocation". Journal of Machine Learning Research. 3 (4–5): pp. 993–1022.

[47] Agrawal, R.; Imieliński, T.; Swami, A. (1993). "Mining association rules between sets of items in large databases". Proceedings of the 1993 ACM SIGMOD international conference on Management of data - SIGMOD '93. p. 207.

[48] Scime, Anthony (2005). "Web Mining: Applications and Techniques". Hershey, PA: Idea Group Publishing. pp. 282

[49] Institute for Web Science and Technologies at the University of Koblenz-Landau and GESIS, Leibniz Institute for the Social Sciences in Cologne http://topicmodels.west.uni-koblenz.de/ckling/tmt/svd_ap.html extracted October 28th 2019

[50] Watkins, C.J.C.H. (1989), "Learning from Delayed Rewards" (Ph.D. thesis), Cambridge University

[51] Szita, Istvan; Szepesvari, Csaba (2010). "Model-based Reinforcement Learning with Nearly Tight Exploration Complexity Bounds" ICML 2010. Omnipress. pp. 1031–1038

[52] Everitt B.S., Skrondal A. (2010), Cambridge Dictionary of Statistics, Cambridge University Press

[53] Rokach, L. Ensemble-based classifiers. Artif Intell Rev 33, 1–39 (2010). https://doi.org/10.1007/s10462-009-9124-7

[54] Diebold, F. X. (2000) Big Data Dynamic Factor Models for Macroeconomic Measurement and Forcasting. University of Pennsylvania and NBER

[55] De Mauro, A.; Greco, M.; Grimaldi, M. (2016). "A Formal definition of Big Data based on its essential Features". Library Review. 65 (3): 122–135.

[56] Ibrahim Abaker Targio Hashem; Ibrar Yaqoob; Nor Badrul Anuar; Salimah Mokhtar; Abdullah Gani; Samee Ullah Khan (2015). The rise of "big data" on cloud computing: Review and open research issues". Information Systems vol 47 pages 98–115.

[57] Han, J.; Kamber, M. (2001). Data mining: concepts and techniques. Morgan Kaufmann Publishers, Massachusetts, USA

[58] DIGITIZE: meaning in the Cambridge English Dictionary. Retrieved from https://dictionary.cambridge.org/dictionary/english/digitize?q=digitisation

[59] Mackenzie, C. E. (1980). Coded Character Sets, History and Development. The Systems Programming Series (1 ed.). Addison-Wesley Publishing Company, Inc. p. x. ISBN 978-0-201-14460-4.

[60] David Donaldson "Big data useless without human element" The Mandarin (9 October 2015), online: archived at , which emphasizes the importance of human curiosity in making large datasets useful.

[61] Dale, R. (2018). "Law and Word Order: NLP in Legal Tech." Natural Language Engineering, 25(1), 211–217.

[62] Mazareanu, E. (2019, August 9). U.S.: number of lawyers 2019. Retrieved from https://www.statista.com/statistics/740222/number-of-lawyers-us/

[63] King, K. (2017, October 23). There are now 140,000 practising solicitors in England and Wales. Retrieved January 22nd 2020, from https://www.legalcheek.com/2017/10/there-are-now-140000-practising-solicitors-in-england-and-wales/

[64] About the Bar. Retrieved January 22nd 2020 from https://www.barcouncil.org.uk/about-the-bar/

[65] Advokatforeningen. Retrieved January 22nd 2020, from https://www.advokatforeningen.no/om/presserom/bli-advokat/

[66] Discover Legal Technology. (n.d.). Retrieved January 22, 2020, from https://techindex.law.stanford.edu/

[67] Revisorforeningen. Retrieved January 22, 2020, from https://revisorforeningen.no/om-oss/om-revisorforeningen/

[68] Regnskap Norge. Retrieved January 22, 2020, from https://www.regnskapnorge.no/om-oss/jubileumsaret/regnskap-norge/

[69] Norsk Eiendomsmeglerforbund. (2018, December 10). Retrieved January 22, 2020, from https://www.nef.no/om-nef/

[70] Harry Surden (2014) Machine learning and law, Washington Law Reviev Vol 89:87

[71] Surden, H. (2012). Computable Contracts. University of California, Davis Law Review, 46, 629–700. Retrieved from https://lawreview.law.ucdavis.edu/issues/46/2/Articles/46-2_Surden.pdf

[72] Simon Chester, "How Tech is Changing the Practice of Law: Watson, AI, Expert Systems, and More" (Debate presented at the Pacifc Legal Technology Conference, Vancouver, 2 October 2015).

[73] McKamey, Mark, Legal Technology: Artificial Intelligence and the Future of the Law Practice (December 6, 2016). APPEAL Review of Current Law and Law Reform - (2017) 22 Appeal 45. Available at SSRN: https://ssrn.com/abstract=3014408

[74] Bryhni, S. (interviewer) and Nygaard, M. (interviewee). (n/d) Videointervju med Lawbotics. Sprint Consulting AS, Oslo, Norway. https://sprint.no/artikler/videointervju-med-lawbotics

[75] Feller, Avi; Pierson, Emma; Corbett-Davies, Sam, and Goel, Sharad. A computer program used for bail and sentencing decisions was labelled biased against blacks. It's actually not that clear. The Washington Post. October 17, 2016. http://www.cs.yale.edu/homes/jf/Feller.pdf

[76] Comparing the Performance of Artificial Intelligence to Human Lawyers in the Review of Standard Business Contract. Hentet 5. mars 2019 fra https://www.lawgeex.com/resources/AIvsLawyer/

[77] Losey, R. C., Sullivan, J., & Reichenberger, T. (2016, February 21). Discovery Team's 2015 TREC Report. Retrieved from https://e-discoveryteam.com/ai/mr-edr/e-discovery-teams-preliminary-report/

[78] Hill, C. (2016, January 27). HighQ receives $50m in investment from Goldman Sachs, Morgan Stanley and One Peak. Legal Insider. Retrieved from https://legaltechnology.com/highq-receives-50m-in-private-equity-investment/

[79] Browder, J. (interviewee). (2017, March 17) BBC World Interveiw. BBC, London UK. Retrieved from https://www.youtube.com/watch?v=pXozW2BUg9s

[80] Noah, T. (intervieweer) and Browder, J. (interviewee). (2018, March 10) The Daily Show with Trevor Noah. Retrieved from https://www.youtube.com/watch?v=VkizYljxcD8

[81] https://www.dailymail.co.uk/news/article-3341645/The-18-year-old-student-saved-British-drivers-2MILLION-just-four-months-Teenager-sets-free-appeals-website-bedroom-overturned-30-000-parking-fines-already.html

[82] Cuthbertson, A. (2016, June 29). Robot Lawyer Overturns $4 Million in Parking Tickets. Newsweek. Retrieved from https://www.newsweek.com/robot-lawyer-chatbor-donotpay-parking-tickets-475751

[83] Mannes, J. (2017, July 12). DoNotPay launches 1,000 new bots to help you with your legal problems. TechCrunch. Retrieved from https://techcrunch.com/2017/07/12/donotpay-launches-1000-new-bots-to-help-you-with-your-legal-problems/?guccounter=1

[84] Titcomb, J. (2019, July 4). British 22-year-old's 'robot lawyer' app raises $4.6m from Facebook backers. The Telegraph. Retrieved from https://www.telegraph.co.uk/technology/2019/07/04/british-22-year-olds-robot-lawyer-app-raises-46m-facebook-backers/

[85] Monticello, J. (intervieweer) and Browder, J. (interviewee). (2018, September 20) Meet the Robot Lawyer Fighting Fines, Fees, and Red Tape. ReasonTV . Retrieved from https://www.youtube.com/watch?v=xbXM-aNRNlY

[86] https://www.regjeringen.no/no/dokumenter/nou-2002-18/id145716/sec4

[87] Kaufman, J. (2012). The personal Mba: a world-class business education in a single volume. London: Portfolio., p. 36

[88] Brealey, R. A., Myers, S. C., & Allen, F. (2017). Priciples of corporate finance. 12. Edition. New York, NY: McGraw-Hill Education, p 1.

[89] Porter, M. E. (1996) What is strategy? Harvard Business Review.

[90] Magretta, J. (2012). Understanding Michael Porter: the essential guide to competition and strategy. Boston, MA: Harvard Business Review Press.

[91] Michael E. Porter, "How Competitive Forces Shape Strategy", Harvard Business Review, May 1979 (Vol. 57, No. 2), pp. 137-145.

[92] Barney, Jay (March 1991). "Firm Resources and Sustained Competitive Advantage". Journal of Management. 17 (1): 99–120.

[93] Johannessen, A., Christoffersen, L., & Tufte, P. A. (2004) Forskningsmetode for økonomisk-administrative fag. Abstrakt forlag, Oslo, Norway

[94] Polit, D. F., & Beck, C. T. (2010). Generalization in quantitative and qualitative research: Myths and strategies. International Journal of Nursing Studies, 47(11), 1451–1458. doi: 10.1016/j.ijnurstu.2010.06.004 page 1.

[95] Levi, E. (1948). An Introduction to Legal Reasoning. The University of Chicago Law Review, 15(3), 501-574. doi:10.2307/1597535

[96] Creswell, J. and Poth, C.N (2018) Qualitative Inquiry and Research Design. Choosing among five traditions. Fourth edition. Sage Publications, Inc. California, USA

[97] Polkinghorne, D.E. (1998) Phenomenological research methods. In Valle, R.S & Halling, S. (Eds.), Existential-phenomenological perspectives in psychology (pp. 41-60) Plenum Press, New York, USA

[98] Berg, B. L. (2001) Qualitative Research Methods for the Social Sciences.

[99] https://rett24.no/articles/dette-var-de-20-storste-firmaene-i-2017

[100] Kvale, S. and Brinkmann, S. (2009). InterView – Introduksjon til et håndværk. Hans Reitzels Forlag. Copenhagen, Denmark

[101] https://www.alfasoft.com/no/produkter/statistikk-og-analyse/nvivo.html

[102] https://www.etikkom.no/forskningsetiske-retningslinjer/Samfunnsvitenskap-jus-og-humaniora/

[103] Nerdrum, Per «Mellom sannhet og velferd. Etiske dilemmaer i forskning belyst ved et eksempel». Notat, Høgskolen i Oslo, avdeling for økonomi-, kommunal- og sosialfag.

[104] Goertzel, B. (2007). Human-level artificial general intelligence and the possibility of a technological singularity. Artificial Intelligence, 171(18), 1161–1173. doi: 10.1016/j.artint.2007.10.011

[105] Kolvereid, L and Åmo, B. W (2019) «Timepris som kvalitetssignal for et regnskapskontor» in Gustavsen, T. (Ed.) Trender og utfordringer i regnskap og revisjon. pp. 135-149. Fagbokforlaget, Bergen Norway

[106] Åmo, B. W and Gårseth-Nesbakk, L. (2019) «Prissetting av regnskapstjenester – dagens praksis møter den digitale transformasjonen» in Gustavsen, T. (Ed.) Trender og utfordringer i regnskap og revisjon. pp. 135-149. Fagbokforlaget, Bergen Norway

[107] Christensen, C. M. (2016) The innovators dilemma: when new technologies cause great firms to fail. Boston, MA: Harvard Business Review Press.

[108] Tobiassen, M. (2017, August 19). Bygger automatisk advokat. Dagens Næringsliv. Retrieved from https://www.dn.no/teknologi/automatisering/advokater/bygger-automatisk-advokat/2-1-153594