UPPSALA
UNIVERSITET

Department of Law

Spring term 2020

Dissertation in Procedural Law

30 higher education credits

The Robot as a Judge

- *dangers, challenges and opportunities.*

Author: Felix Wideroth

Supervisor: Professor Eric Bylander

Abstract

The practical use of AI in law has increased significantly in recent years and we have probably only seen the beginning of a breathtaking development. The paper analyses Artificial Intelligence in the judiciary in general and AI as a judge in particular. The purpose is to find out what it takes to use AI in the judicial system, but above all and most interestingly, what challenges we face and what we must consider in that development. The paper addresses how machine learning, artificial neural networks and natural language process work in a legal context. Furthermore, it addresses challenges such as manipulation of AI, bias in data, evidence and evidence evaluation in the light of for example deepfakes as well as challenges related to the implementation of AI systems in society in general and the judiciary in particular. What I particularly want to convey with the paper is that AI in the judicial system is not as far away as we might think, but most importantly that we must take an active role in the development of AI. It is we who have the opportunity to create useful and beneficial AI and counteract its negative aspects. We can no longer turn a blind eye to the transformation that is taking place. We can no longer be a spectator, but instead we must be actors in this development. It is our opportunity, but also our responsibility to be involved so that we can develop ethical and fair AI created to promote human everyday life, not to harm it!

Keywords: AI as a judge, legal AI, automated decision making, artificial intelligence, machine learning, natural language processing, manipulation of AI, evidence and deepfakes

“We are not spectators of the future, we are actors of the future. So, we cannot just cross our arms and say AI are so bad. Why is AI bad? You are the one that can make AI good. So, you are actor of the future.”

- *Manuela M. Veloso*

Herbert A. Simon University Professor at the School of Computer Science at Carnegie Mellon & Head of J.P. Morgan AI Research

Foreword

This dissertation is written as the final work on the Law Programme at Uppsala University. My hope is that it may constitute a first proper analysis of artificial intelligence in the judicial system, along with the challenges, opportunities and dangers that it entails.

I became interested in the field of artificial intelligence about two years ago. That interest led me to begin more detailed study in computer science and in particular the use of artificial intelligence in the law. The lack of research and specific studies meant that I was forced to seek knowledge using methods that differed from those normally used for a Master's thesis but that were incredibly fruitful in such a new field, namely conversation and discussion. As in all major changes, these discussions begin as casual conversations between colleagues and friends, who exchange thoughts and ideas that in turn lead to new thoughts and ideas. This has given me the opportunity to spend several hundred hours in discussions with other students, researchers and professionals in both artificial intelligence and law from countries including Sweden, Singapore and the USA. Much of the analysis I describe comes from such discussions and from my analysis when I subsequently summarised the discussions. During those subsequent reflections, I have had a chance to put things into perspective and begin to frame the problem from a legal point of view.

I hope you enjoy reading it!

Felix Wideroth

Stockholm, January 2020

Table of Contents

ABSTRACT ..... **II**

FOREWORD ..... **I**

TABLE OF CONTENTS ..... **II**

ABBREVIATIONS ..... **IV**

1 INTRODUCTION **1**

1.1 General 1
1.2 Research aims and questions 3
1.3 Demarcations 4
1.4 Methodological approach 5
1.5 Research situation, materials and selection of materials 6
1.6 Outline 8
1.7 Terminology 8

2 ARTIFICIAL INTELLIGENCE **10**

2.1 General 10
2.2 What is Artificial Intelligence? 11
2.3 Machine learning 15
2.4 Data and its importance for machine learning 20
2.5 The importance of hardware 21
2.6 Summarising observations 22

3 CONDUCT OF PROCEEDINGS IN COURT **22**

3.1 General 22
3.2 The structure and functions of the courts 23
3.3 Parts of court proceedings 24
3.4 Evidence and assessment of evidence – what the court may decide on 25
3.5 Judgments and reasoning in support of judgments 26
3.6 A word on legal certainty 28

4 ARTIFICIAL INTELLIGENCE IN THE JUDICIAL SYSTEM **29**

4.1 General 29
4.2 Legal method in a calculation-based world 32
4.3 Manipulation of AI 33
4.4 Transparency in the case of automated decision-making 36
4.5 Evidence and evaluation of evidence 37
4.5.1 Evaluation of evidence by the AI judge 37
4.5.2 Manipulation and fabrication of evidence with the aid of Artificial Intelligence 38
4.6 Bias in the courts and in the case of machine learning 39

5 THE AI JUDGE **43**

5.1 General 43
5.2 The three instances of the courts 44
5.3 Judges and jurors 45
5.4 Implementation and acceptance 46
5.5 Should AI judges exist? 47

6 CONCLUSION **49**

SOURCE AND BIBLIOGRAPHY **52**

Official publications 52

LEGAL CASES 52
LITERATURE 52
INTERNET SOURCES 58

Abbreviations

AI	Artificial Intelligence such as a computer program/program code/software
ML	Machine Learning
CJP	Swedish Code of Judicial Procedure
NLP	Natural Language Processing
NLU	Natural Language Understanding
ANN	Artificial Neural Networks

1 Introduction

1.1 General

Inefficiency, lengthy procedures, high costs and miscarriages of justice are a few of the things that we sometimes associate with the judicial system in general and perhaps courts in particular. But what if that were not the case at all? What if court proceedings were always efficient, cheap and completely free from incorrect conclusions and miscarriages of justice. That is far from impossible with a robot as a judge. The law must be objective, and we must be able to carry out objective assessments based on pre-established parameters without being influenced by preconceived ideas and emotions. That makes it possible to apply the law in mathematical models that are able to make decisions and hand down judgments for us. With a robot as a judge, we can eliminate the consequences of tiredness and conscious and unconscious bias and ensure that the judge always has access to the most up-to-date knowledge of the current legal situation. We can deal with cases in parallel, instead of one by one, and we can thus also become significantly more efficient. A computer has almost unlimited capacity for work and, in theory, does not suffer from all the "weaknesses" we humans suffer from. The possibility of using a robot as a judge in the future is spelled “Artificial Intelligence”.

AI[1] is something we come into contact with every day. We find it in everything from self-driving cars to the spam filter in your inbox. In health care, it is becoming increasingly common for it to be used in diagnostics and patients are even beginning to demand it in certain types of assessments. This mainly relates to analysis of various types of X-ray data from Magnetic Resonance Imaging (MRI),

[1] The use of the term Artificial Intelligence as opposed to AI is described in more detail in section 1.7 Terminology.

for example. AI models trained to detect tumours have been shown to be much more accurate than a doctor when it comes to diagnosis.[2] However, the legal profession is still relatively unaffected by Artificial Intelligence, though that is in the process of changing. In the United States, for example, various AI systems are used to predict a prisoners' risk of reoffending and there is also another type of AI that has been shown to be able to predict the outcome of cases in the US Supreme Court with a surprisingly acceptable degree of certainty.[3] AI is beginning to play an increasing role in the field of business law in particular and a lot of time and money is being invested in developing programs to support lawyers in their work.[4] This mainly involves automation of contracts, analysis of contracts and other documents as well as legal research. If we look more closely in public authorities in Sweden, the use of automated decisions, which is a type of AI, albeit extremely simple, is fairly widespread. Projects are being implemented at the Swedish National Courts Administration to look at how Artificial Intelligence could change their work. At present, however, it is mainly a question of automation of more administrative tasks such as scheduling, archiving and similar administrative work. The prospect of having an AI system as a judge seems remote at the moment.[5]

As I described at the start, a robot as a judge opens up enormous possibilities, but there are many major challenges to be faced along the way. Those challenges are not just technical, but also moral and ethical and challenge our current legal methods.[6] The introduction of AI into the administration of justice is not aimed at

[2] McKinney et al p. 1 ff.
[3] Katz et al. 2017 p. 1 ff.
[4] Öster & Ekholm, AI standard på storbyråerna [AI Standard at Large Firms], 2019 and Öster & Ekholm, Algoritmernas nya era är här [The New Era of Algorithms is Here], 30/11/2019 and Öster & Ekholm, AI nu självklar på advokatbyråerna [AI Now Goes Without Saying at Law Firms], 30/11/2019
[5] Öster & Ekholm, Robotdomaren ännu långt bort [The Robot Judge Still a Long Way Off], 30/11/2019.
[6] Katz 2013 p. 1 ff.

changing the system itself, but aims to partially replace existing participants with new technologies in order to improve the system. We therefore need to re-ask ourselves key questions about the judicial system, its function and its purpose. However, we also need to look at how Artificial Intelligence works in order to deal with the challenges that arise from a legal point of view. In that context, we then need to ask ourselves whether we want to implement AI in the judicial system, how we will go about it and when we will do it.

1.2 Research aims and questions

The practical use of AI in general has increased exponentially in recent years and what we are seeing is probably just the beginning of a breathtaking development. In Sweden, most major law firms now use some form of AI solution to make their day-to-day work more efficient. Swedish public authorities are increasingly beginning to use AI to make certain decisions and we are now in a situation where a robot as a judge is no longer just a fantasy, but a real possibility. The fact that a computer can perform complex tasks that previously could only be done by humans is absolutely revolutionary. However, the development and subsequent implementation of AI models is an incredibly difficult and complex issue. To begin with, it is about technical challenges but, perhaps most importantly, it is also about ethical and moral positions that force us to go back to basics on many issues. As far as the law is concerned, this means that we need to question the new technologies from a legal point of view, but it also means that we need to start questioning and challenging previously established legal methods and approaches in order to examine how this new technology can best be used to promote the rule of law.

Procedural law is an excellent starting point for this critical analysis, since procedural law is the backbone of the legal method. Court proceedings are our ultimate tool for resolving legal problems and establishing and claiming liability.

Court proceedings therefore contain all the elements in which Artificial Intelligence is likely to be used in law and also highlight many of the ethical elements of law and the law's role in society that are becoming topical. Simply put, procedural law is almost a formalisation of ethical positions on how we solve legal problems that have arisen in our interaction with one another. In view of that, this presentation seeks to carry out a more detailed analysis of Artificial Intelligence in the judicial system in general and AI as a judge in particular. In the work, I wish to analyse what is required for AI to be used in the judicial system and, above all and most interestingly, what challenges we face and what we need to think about in its development. In order for us to develop AI that is both beneficial and fair, it is crucial for us to be involved in its development.

The following questions are discussed within the context of this study:

1. What opportunities, dangers and challenges are there with AI judges?
2. Should AI judges exist?

1.3 Demarcations

The use of Artificial Intelligence in the judicial system is not a simple matter. It is extremely complex, involving many smaller areas that could in themselves form their own area. Since it is also very new, there is almost nothing written about AI in the judicial system, particularly in a Swedish context. The focus of the work is therefore generally broad, but some boundaries must still be established. AI and the law can be divided into two aspects. On the one hand, *regulation* of the use of Artificial Intelligence and, on the other, *how we use* Artificial Intelligence as a tool in the law. I have chosen to focus on the latter aspect and, in particular, on the use of Artificial Intelligence in the judicial system. This means that we step into an interdisciplinary field in the presentation that requires an understanding of the law and how the law works and also of computer science and how Artificial Intelligence works. I assume that the basics of the legal method are well known to

the reader, but I will touch upon and challenge them later in the presentation. Knowledge and understanding of AI, on the other hand, is slightly alien to many people, which why I will go through the basics of Artificial Intelligence and the use of AI. This is important because the subsequent presentation requires the reader to understand the basics in order to be able to follow the analysis.

Despite the paper's conceptual nature, the work is very technical. Unfortunately, it is not possible to avoid that in an area such as this. One of the reasons I decided to keep it broad-based is because people, in general, have a limited understanding of how Artificial Intelligence works. A narrower focus in any of the areas I deal with below in my presentation would require analysis that is even more technical because it would require an even more detailed understanding of how AI works in that particular area.

Even though most of the examples given in the work fall within the area of responsibility of the ordinary courts, the work is not just limited to them, but may be applied to courts of all kinds. The ordinary courts generally deal with more complex and detailed cases than the administrative courts, which is why they are best placed to be used as examples.

1.4 Methodological approach

As I have already stated, the work focuses on an area where technology and law interact. For the first time, this interaction takes place not just on a practical level by, for example, offering the possibility of sound and image recording or the use of technological aids in the presentation of a case. The use of Artificial Intelligence in the law means that this interaction can now take place on an intellectual level, where AI can take over and replace the work of analysis previously carried out by a trained lawyer. Analysing this using generally-accepted legal methods such as legal dogmatics does not therefore appear to be the best approach. I have therefore chosen to apply an interdisciplinary research method in which I combine law and

technology.[7] The method aims to analyse and provide a critical review of the use of technical solutions in the practice of law. It is therefore *not* a question of analysing technical solutions in general, but of analysing specific technical solutions that are capable of making legal decisions or carrying out qualified legal work.

1.5 Research situation, materials and selection of materials

Significant research is being conducted in the field of Artificial Intelligence, both on the technical side and with regard to how ethical and sustainable AI is to be developed. However, at present there is limited research in the field of Artificial Intelligence in the judicial system. In the private sector, research and development is being carried out to develop AI solutions to make day-to-day work more efficient. What this research has in common is that it is primarily about tools that are able to take over the more repetitive parts of a lawyer's work. It is mainly about research in databases, analysis of contracts and other extremely time-consuming tasks. There is some research from the USA in which AI has been used to predict the outcome of court cases, as well as research into the legal method in a calculation-based world. But these have only scratched the surface in a field that has the potential for much more. If we look at Sweden, research on Artificial Intelligence in the Swedish judicial system is almost non-existent. Perhaps the most important efforts in beginning to analyse AI's role as a judge have been made by Eric Bylander in the article entitled "Den datoriserade domaren" [The Computerized Judge] in SvJT [Svensk Juristtidning – Swedish Law Review].[8] It also says something about the research situation when it is "only" an article that

[7] Gräns p. 428.
[8] Bylander.

makes the largest current contribution to the field. Above all it is a clear sign that more research is needed.

The material I have used consists of the classic legal sources when it comes to the law, such as laws, government bills, studies and doctrine. However, a significant part of my material consists of research on Artificial Intelligence and machine learning. I also use articles from industry magazines and research publications which deal to a varying extent with the field of Artificial Intelligence and Artificial Intelligence in the judicial system.[9] One important aspect of the research situation needs to be highlighted. Artificial intelligence is now probably one of the hottest research areas in the world. That also means that enormous resources are being spent on research and development. This, combined with the fact that Artificial Intelligence is a relatively new area of research, means that development takes place extremely quickly and new breakthroughs follow closely on the heels of others. For example, while this dissertation was being written, Google introduced a new type of algorithm in its search engine that proved to be much better at understanding language and linguistic context than previous models. What is presented in this paper is what was known in December 2019. In the same way, this presentation only includes what was known at the time of writing. That, of course, applies to all types of scientific presentations, but I think it is particularly important to point it out in this case. I wish to emphasise this because it is important for you, the reader, to be up-to-date in the field and to bear in mind that there are large-scale changes in this field taking place on a continuous basis.

[9] I have used arXiv for research in computer science and AI. This is a database for digital pre-publication, so-called pre-prints, of as yet unreviewed scientific articles. The archive is provided by Cornell University and, although it is open, the material is of high quality due to stringent moderation and a requirement that a new account needs to be confirmed by an already known and established author. All major research institutions regularly use this archive to publish their research.

My goal is to contribute to this important issue, and it is my hope that this thesis may serve as a starting point for future discussion about Artificial Intelligence in the Swedish judicial system. The focus is therefore on carrying out an overall analysis of Artificial Intelligence in the judicial system in general and in the courts in particular to identify not only the opportunities, but also the dangers and, above all, the challenges we face.

1.6 Outline

The first half of the study consists of two sections, each of which introduces the reader to the subject by going through the basics. I begin by providing an overall explanation of what Artificial Intelligence is and how it works through the use of specific examples. I then move on to a section on conduct of proceedings in court, how the Swedish courts are organised, what divisions there are and the bases for conduct of proceedings. The basic knowledge provided in these two sections is then crucial for the subsequent analysis in the section on Artificial Intelligence in the judicial system. In the section on Artificial Intelligence in the judicial system, I carry out a methodical review of the key elements of Artificial Intelligence, focusing on the challenges of implementing AI in view of the principles of conducting court proceedings.

1.7 Terminology

Computer science in general and Artificial Intelligence in particular is a scientific field that many people do not often come into contact with, even though we come into contact with AI-based tools on a daily basis via our phones/smartphones/computers. Therefore, in order to simplify matters for the reader, I wish to begin by clarifying and explaining some of the terminology that will be used in subsequent parts of the work. The overwhelming majority of all research in Artificial Intelligence today is based on the English language. I have

therefore chosen to continue to use the English terms as well as the associated abbreviations of these terms. This is because it is often the case that no sufficiently good translations to Swedish are available and I have no wish to complicate matters for a reader who is already familiar with the field or who will in future continue to research by him or herself.

Artificial intelligence, when used to describe a subject area, will be referred to as "Artificial Intelligence". Artificial Intelligence, when used to describe a computer program/program code/software, i.e. to describe a specific capability achieved by implementing techniques in the field of Artificial Intelligence, will be– abbreviated to "AI".[10] AI is the overall term commonly used to refer to this new technology in everyday language and in research. AI will also be used in conjunction with the terms “program”, “model” or “application” and describes the algorithms and/or computer code that represent specific AI with a specific task.

The term “data” is key in any discussion about Artificial Intelligence. In a computer context, data is used to describe the technical representation of information stored in files on a computer. In this context, we also distinguish between programs and data, even though they are stored in the same way in a computer's memory. Thus, program files are not referred to as data, though data is what the programs produce and, above all, use in their tasks.

The term “bias” is used in two different contexts in the work: in relation to persons and in relation to data. Bias can be described as a person's way of reacting and evaluating the outside world based on his or her own experience. Bias in a person manifests itself in the person's approach and therefore also their actions at the expense of other possible approaches which, in turn, could be as acceptable as their own. Bias is therefore an evaluation of a person, a group or a thing, for

[10] A similar approach for the division of the terms Artificial Intelligence and AI has been adopted by the Swedish Innovation Agency Vinnova.

example, that is based on their preconceptions, preferences or ideologies, which is referred to as cognitive distortion. Bias in data constitutes the digital representation and formalisation of the bias that exists in us humans.

2 Artificial Intelligence

2.1 General

In this section, I will provide the reader with an overall introduction to the field of Artificial Intelligence, what it is and what its various major components are. To provide a broader context, I will present the reader with a brief historical background. I will also review various machine learning models and explain why and how information, or data as it is often called, plays an important role in AI and why the various components of a computer and the technological development in this area affect AI's capabilities at present and will affect them in future. The purpose of the section is to familiarise the reader with the functions that can be created with the aid of Artificial Intelligence because that is crucial to understanding the subsequent discussion of Artificial Intelligence in the judicial system and the introduction of AI.

I am aware that the section is heavy going in technical terms, but it serves a higher, more long-term purpose. I could have further simplified how AI works and only discussed it extremely superficially and at an even more general level, but that would have been to do the reader a disservice and would not have served the quieter and more implicit purpose of the dissertation, which is to be the reader's first contact with Artificial Intelligence in the law. Although the parts discussed initially may appear to be peripheral to the main subject of the dissertation, they

are crucial to how Artificial Intelligence works and are extremely important for future understanding.

2.2 What is Artificial Intelligence?

At present, there is no clear-cut definition or generally-accepted demarcation of Artificial Intelligence. Artificial Intelligence basically consists of digital technologies and methods that make possible the performance of automated information processing and decision-making that in the past could only be carried out by humans because it required human mental activity.[11] In this presentation, I will use the Swedish innovation agency Vinnova's definition of Artificial Intelligence:

"The ability of a machine to mimic intelligent human behaviour. In other words, machines' ability to function in meaningful ways in relation to the specific tasks they are intended to perform and the situations they are intended to operate in. Artificial intelligence is also the area of science and technology that is about studying, understanding and developing machines with intelligent behaviour."[12]

The opposite of Artificial Intelligence is natural intelligence, in other words the intelligence that we humans and other carbon-based life forms possess. Artificial Intelligence can thus perhaps most easily be described as intelligence that is created in a machine to make it act appropriately and with an understanding of its surrounding environment. AI is often divided into specific AI and general AI. Specific AI consists of, for example, advertising tools, the program AlphaGo[13]

[11] Vinnova 2018 p. 28.
[12] Vinnova 2018 p. 28.
[13] Created by the company DeepMind.

which beat the world's top Go[14] player, Le Sedol, or IBM's Watson, which took part in Jeopardy. These specific AI systems also include legal tools such as ROSS[15], a legal research tool and MODRIA[16], a tool for dispute resolution. The long-term goal in the development of AI is to develop what is referred to as *Artificial General Intelligence* (AGI). This means a program that is not programmed by humans to perform a specific task, but that is able to learn and take on new tasks that were previously completely unknown to it without human intervention. The goal is to develop a machine that can imitate and achieve the cognitive ability that exists in humans. However, there is still a long way to go to achieve AGI. There is not even agreement within the scientific community on whether it will happen in 10, 50 or 100 years' time. They are not even in agreement as to whether it is possible, though neither can they say that it is impossible.

At this point, it is also appropriate to clarify one important difference between AI and robotics. AI is primarily about cognitive functions, whereas robotics is about motor functions. However, in practical applications, the difference is not as clear-cut because robotics requires an ability to analyse the environment, as well as sensory capabilities that can be achieved with the aid of AI. Developing motor skills is also more difficult, more expensive and more time-consuming than developing cognitive functions.[17]

Although we may think, based on what we are told by the media, that Artificial Intelligence is something new, the field itself is old.[18] The Artificial Intelligence research field was created as early as 62 years ago in the summer of 1956 under

[14] Go is one of the world's most complex board games with more than 2x10^170 potential positions on the board. For a long time, it was considered too difficult a problem for machine learning to solve. It was thought that human intuition was required in order to be successful.
[15] For more information, see: https://rossintelligence.com/
[16] For more information: https://www.tylertech.com/products/modria
[17] Vinnova p. 28.
[18] Press Forbes 30/12/2016.

the name of "The Dartmouth Workshop".[19] However, the idea, or rather the question, "Can computers think?", was one that scientists were asking themselves far earlier. “Can computers think?” is not about a computer being able to carry out certain calculations – which in itself could be interpreted as thinking – but is aimed at the more complex and philosophical aspects of the human intellect. One of the first people to seriously ask himself this question was Alan Turing, the father of modern programming. In 1950, he published his work, “Computing Machinery and Intelligence”, in which he presented a test that he called "The Imitation Game", which today is more widely known as “The Turing Test”.[20] Turing believed that if a human being could converse with a machine without being able to tell whether it is a machine or a human, the criterion of human intelligence had been fulfilled and that also answered the question of whether computers could think. The Turing Test is widely criticised today, but it has also been hugely influential as a concept in the philosophy concerning Artificial Intelligence.[21]

So how does AI work? In simple terms, it is pattern recognition.[22] Pattern recognition is also, in general, how we humans work when we learn something new. Signals are sent to your brain through your eyes and then through your optic nerve. These signals then become the image you see. Your senses of hearing and touch also work in the same way. But how can you understand and determine what is what? Let us take a bird as an example. Your friend Alex has just been on safari in Tanzania. Alex proudly shows you a photograph he took on the trip. You immediately see that it is a picture of a bird, and Alex tells you that it is an African hoopoe. When you look at the picture, you also understand that it probably eats insects which it catches in trees or other cavities in order to survive. You

[19] Sharkey BBC 21/6/2012 and Vinnova p. 28.
[20] Turing p. 433.
[21] Pinar Saygin et al. 463 f.
[22] Bishop p. 1 ff.

understand this because you see that it has a long beak, which you know is a characteristic of birds that feed in that way. You understand all this despite the fact that you had never seen an African hoopoe before and did not even know that it existed up to then. Why is that? Without any prior knowledge of Tanzania's biotope, you can obtain that much information from a single image just through your own previous experience. In very simple terms, it is a question of pattern recognition. You have probably seen a lot of birds in Sweden: seagulls, woodpeckers. etc. You also read about birds in school and remember parts of it. You therefore have a fairly good idea of what a bird looks like. So, when you see a picture of this animal, you can easily with precision say that it is a bird. That is because you recognise the shapes and characteristics of a bird. You can also say that it most likely feeds on insects found in tree trunks or other cavities since you have learned that long beaks are good for that. AI works in the same way.[23]

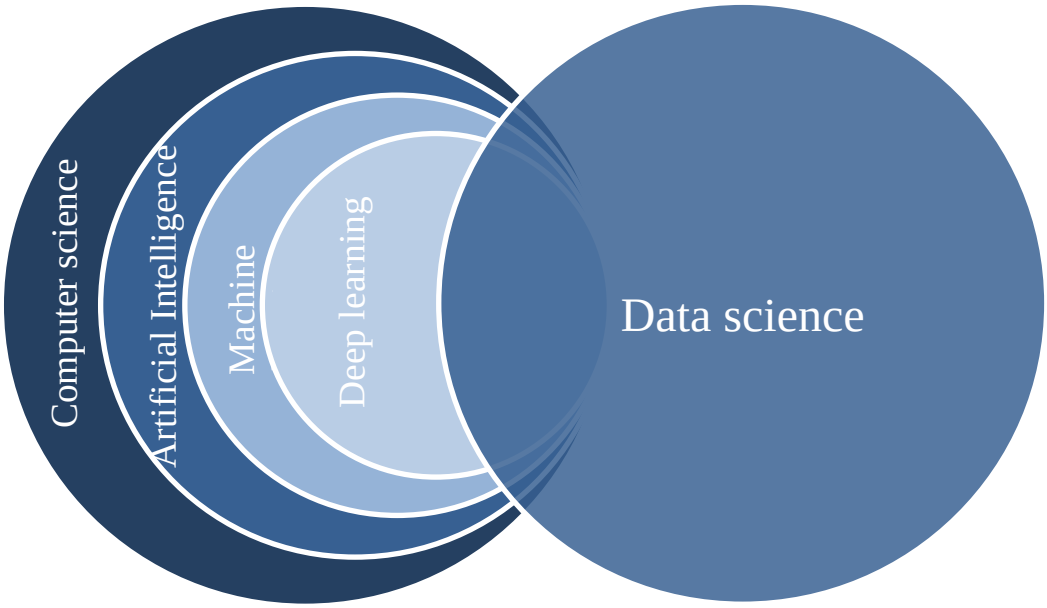

Figure 1Representation of the distribution of scientific fields in computer science

[23] Bishop p. 3 f.

Artificial Intelligence, as I state above, is a collective term for a field of science with extremely broad contents. It forms part of the much larger main field of computer science (see Figure 1). In the field of computer science, but also extending beyond that field, we find an area known as Data Science. Data science is a field in which Artificial Intelligence interacts with other fields such as law. While computer science makes up the model and the algorithms themselves, data science is the field where we connect these with other scientific fields. We use mathematical models from computer science, along with knowledge and data from another field in order to create a deeper understanding with the aid of mathematical models and Artificial Intelligence. One really interesting field is machine learning, which forms a subgroup of Artificial Intelligence. When we talk about Artificial Intelligence in everyday speech, it is machine learning that we are usually referring to. It is precisely in this field that the biggest breakthroughs have taken place in recent years. In the next section I will discuss machine learning and its basic principles and current models for application in the law.

2.3 Machine learning

Machine learning is the scientific study of algorithms and statistical models that computers use to perform a specific task without having been expressly programmed to do that task.[24] Instead, machine learning relies on patterns and other conclusions. Machine learning algorithms use different mathematical models to make predictions or make decisions without being expressly programmed to perform that particular task. Machine learning is used in a wide range of applications such as e-mail filtering and computer vision, where it is difficult or even impossible to develop a conventional algorithm to perform the task

[24] The definition "without being expressly programmed" is often attributed to Arthur Samuel, who coined the term "machine learning" in 1959.

efficiently.[25] E-mail filtering, for example, uses so-called Bayesian Networks to determine whether or not an e-mail should be classified as spam.

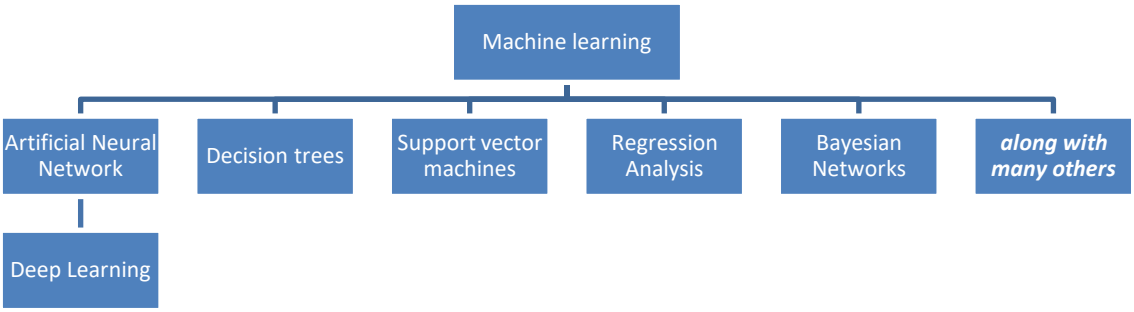

Figure 2 Figure to visually describe the many different machine learning methods

One machine learning method is the use of what are referred to as Artificial Neural Networks (hereinafter referred to as Neural Networks). They are called Neural Networks because the model attempts to imitate how our brain solves problems, i.e. how our neurons and synapses work. This approach to machine learning is inspired by how our brain processes the signals from our eyes. A Neural Network is distributed in layers, with the simplest model having an input layer and an output layer. The use of multiple layers is what is referred to as Deep Learning. Each layer or group of layers aims to solve a specific task, such as finding circles or right angles in an image. Layers in the network are based on the output from the previous layer and together these layers form a Neural Network. This approach is just one of many different methods for building models for machine learning. Other methods include regression analysis for recommendations for films on YouTube and Netflix and, as mentioned above, Bayesian Networks in e-mail filters.

The use of machine learning can also be divided into various fields in which these methods are then used. One of the most difficult and complex areas of machine learning and one that is particularly relevant to the law is the area known

[25] Bishop p. 3 f.

as Natural Language Processing (NLP) and its sub-area of Natural Language Understanding (NLU).[26] Interest in NLP and NLU has increased considerably in recent years due to their potential in cognitive AI models.[27] Known applications that use NLP include Alexa and Siri, our AI assistants that we find in everything from phones to smart speakers in the home. NLU aims to go beyond the structural understanding of language and begin the interpretation of intention, context and ambiguities in order, thereby, to go so far as to generate well-formulated human language by itself. NLU algorithms must address the extremely complex problem of semantic interpretation – i.e. understanding the intended meaning of spoken or written language – including all the subtleties, contexts and conclusions that we humans are able to understand. NLP and NLU are what we refer to as an Artificial Intelligence-complete problem or Artificial Intelligence-hard. The term is derived from the mathematical problem formulation P vs NP[28], where NP-complete problems belong to the problems we are unable to solve within polynomial-time (reasonable time) using a computer, but whose correctness is easy to check if we have the answer.

There are two good examples of tremendous progress within NLP and NLU that has been made in recent years. The first example is a search in Google. NLP is used when you carry out a search in Google. We often use Google when we want to learn something new. Sometimes it may be the case that we do not really know

[26] SAS Institute and Radford et al p. 1 ff.
[27] Metz The New York Times 18/11/2018.
[28] P vs NP is a mathematical problem formulation and also one of the millennium problems. P represents relatively simple problems that can be solved within polynomial-time, whereas NP represents difficult problems that cannot be solved within polynomial-time. Modern encryption is one example. Modern encryption is based on multiplication of large prime numbers. If you have the answer, i.e. the prime numbers that a particular encryption contains, it is easy to check whether it is correct by means of multiplication. However, to try to calculate which prime numbers an encryption consists of involves testing large numbers of combinations, which requires time and computing power. Modern encryption works on this principle.

how to formulate our question, what words to use or even how to spell them.[29] For this, Google has developed a model called BERT (Bidirectional Encoder Representations from Transformers).[30] In simple terms, BERT is a way of building and training deep neural networks[31]. The special thing about BERT is that they have managed to build a model that has a tremendous ability to understand linguistic context. They do this by looking at the text from above and not in a linear fashion, from start to finish. By looking at the text from above, they are able to analyse one word's relationship to other words in the same sentence and thereby understand the context. The model still faces challenges and there is much room for development, but it has taken a big step towards increasing our ability to express ourselves in natural writing when we search in Google. The second example is from a company called OpenAI. OpenAI is a non-profit company in the USA founded by Elon Musk, among others[32]. Their stated goal is to develop safe AI that will benefit people. OpenAI is one of the leading research companies when it comes to Artificial Intelligence. In July 2019, Microsoft, among others, invested a billion dollars in OpenAI and their business. One of their main advances in Artificial Intelligence and NLP is a model called GPT-2. It is a relatively small model which consists of only 1.5 billion nodes and which is trained on only 40GB of data.[33] The model proved to be so good at producing text from only small sections of text that they initially chose not to release the trained model to the public. They were criticised for that decision. Their concern was that the model could be used for negative purposes including creating large amounts of fake news and similar content of such high quality that it was indistinguishable from real

[29] Nayak Google 25/10/2019.
[30] Devlin et al p. 1 f.
[31] “Deep neural networks” is the term for neural networks with one or more "hidden" layers for deep learning, hence the term “deep”.
[32] Elon Musk has founded other well-known companies such as Tesla, SpaceX and PayPal
[33] Radford et al 2019 and Vincent The Verge 14/02/2019.

news.[34] Similar and actually even more frightening is the presence of so-called Deep Fakes, i.e. images or video that have been manipulated to imitate someone else. These models are already very good now and will only get better.[35]

Creativity is something we have long considered unique to us humans. With AI, that is no longer the case. In the German Tübingen, researchers have managed to create an AI system that is able to imitate Vincent van Gogh using a photograph for inspiration.[36] Can you tell which picture was painted by Van Gogh in the image above?[37] AI's ability to be creative is not limited to imitating well-known artists. A company in Luxembourg has created AIVA (Artificial Intelligence Virtual Artist), which does an excellent job of composing music, whether classical, rock or hip hop.[38] In subsequent tests, research subjects were unable to tell whether a piece was composed by a human or a machine.

Figure 3 One picture was painted by Vincent Van Gogh and the other was painted by an AI system.

[34] Radford et al 2019 and Vincent The Verge 14/02/2019.
[35] More on this in section 4.5.
[36] Leon et al p. 5.
[37] The painting on the left is by Vincent van Gogh and the painting on the right is by an AI system. In the same research, AI was able to imitate artists such as Edvard Munch, Wassily Kandinsky and Pablo Picasso.
[38] Go to https://youtu.be/Emidxpkyk6o to listen to a piece composed by AIVA (Artificial Intelligence Virtual Artist).

2.4 Data and its importance for machine learning

Data is a common concept in computer science. As I stated above, data is simply information. It can be anything and it can be made up of text, images or anything that a computer can convert to machine language with or without tools. For example, image recognition uses the value in the pixels in the image to determine what the image shows. Data is the foundation of all AI. A model cannot produce any results itself; it is only when we have trained the model that it becomes usable. Training an AI model means using what is referred to as "training data" to teach the model what we want to teach it. Training data thus contains both the question and the correct answer, and we enter them both into the model in order to teach it. In the next step, we use test data to verify that the model produces acceptable results.

As we now understand, data is incredibly important, particularly the data we use when we train a model. A model is only as good as the data used to train it.[39] Because there is one thing we need to be clear about: computers only do exactly what we tell them to do. A computer obeys the instructions in its code – we could also refer to them as rules. If X, then Y, for example. However, it is usually not as simple as that. Things get considerably more complex when the computer needs to keep track of and sort several thousand different variables. If we use good data, it leads to a good model. However, if we use poor data, it leads to a poor model. The problem lies in the fact that we still do not fully understand how a model works or how certain data will affect the model. As I have described above, it is all about patterns in one way or another. The model looks for patterns in the input data and compares them to what it already knows. It is therefore very important for us to keep track of our data and for us to understand how we use data to train an AI model.

[39] Vinnova 2018 p. 77 f.

2.5 The importance of hardware

Artificial intelligence is not possible without the hardware to carry into effect the mathematical models developed. The core of a computer is its processor, also known as the CPU (Central Processing Unit). The processor is primarily responsible for all the calculations carried out. The processor in turn works together with what is referred to as RAM (Random Access Memory). The RAM memory is where the computer holds all the essential data for ongoing processes in the computer to which the processor requires fast access. A computer naturally contains a number of other components, but these two, in a way, form the basis of a computer's ability to function. Enormous resources are currently being poured into research to develop smaller, better and faster computers. One interesting parallel is the discovery that the processor in graphics cards (GPU – Graphical Processing Unit) used in computers was particularly good at running machine learning models. It was therefore common about five years ago, when the Playstation 4 was launched, to link large numbers of PS4s together to run AI applications. It is less common today because the large processor manufacturers now manufacture special processors specifically suited to AI applications.

To place the development of hardware in a simpler context, I will share one example. The computer aboard Apollo 11 (the first manned journey to the moon) had a processing capacity of 0.043 MHz. It is stated that an iPhone X has a processing capacity of 2490 MHz, which is 100,000 times more powerful than the computer that put man on the moon over 50 years ago. Artificial Neural Networks are one example of the importance of hardware to Artificial Intelligence. The idea behind Artificial Neural Networks existed as early as the 1950s, but was abandoned fairly quickly because there was no way of training the network. In 1969, two researchers from Princeton also proved mathematically that it was impossible to train such networks for anything other than the most menial tasks. However, there was a breakthrough in 1986 when Hinton showed how a multi-

layered network could be trained to perform complex tasks using what is known as backpropagation[40]. However, it took 26 years for that breakthrough to take effect because training of that kind required large computational resources that had only just become available at that time.[41]

2.6 Summarising observations

The purpose of this section is to introduce you, the reader, to Artificial Intelligence in order to provide you with an understanding of the basics. I therefore started out by going through what Artificial Intelligence is, before then going into the important aspect of machine learning. The successes in Artificial Neural Networks in recent years are of specific interest and the field of Natural Language Processing is of particular interest to us lawyers. It is important to bear in mind that Artificial Intelligence and the use of AI is an extremely complex area and my account only touches on the basics. However, it is important to understand the basics for the discussion of AI in the judicial system below. I understand that this part is highly technical and it may be difficult to gain a full understanding of everything. What I want you to specifically take away from this section is what was said about data and the role of data in AI. That is crucial for the analysis of AI in the judicial system below.

3 Conduct of proceedings in court

3.1 General

The purpose of the work – to analyse Artificial Intelligence and AI in the role of judge – requires a brief overview of how the Swedish courts operate, how they are

[40] Backpropagation is a method of automatically configuring networks in order to optimise their tasks.
[41] Somers 2017.

structured and their various divisions. The types of cases in a court vary greatly. At one extreme we have, for example, large, complex criminal cases which may mean that main proceedings go on for several months whereas, at the other extreme, we have smaller cases that are settled without the need for main proceedings carried out orally.

3.2 The structure and functions of the courts

One of our most fundamental rights in Sweden is the right to have your case heard by an impartial court. That means that the courts fulfil one of our most important social functions and form the backbone of the judicial system. The judicial system consists of three different types of courts: the ordinary courts (the district courts, the courts of appeal and the Supreme Court), the general administrative courts (the administrative courts, the administrative courts of appeal and the Supreme Administrative Court) and the specialist courts (e.g. the Labour Court and the Foreign Intelligence Court). In the ordinary courts and general administrative courts, there are three instances, with the district courts and the administrative courts as the lowest instances and the Supreme Court and the Supreme Administrative Court as the highest instances. The ordinary courts deal primarily with criminal and civil cases, though they also deal with other cases such as bankruptcies, trustees and liquidators. The general administrative courts, for their part, deal with cases between the public and the individual. The courts' job description contains a requirement that they should achieve high levels of quality and efficiency when settling cases.[42]

[42] Ministry of Justice, Det svenska rättsväsendet [The Swedish Legal System].

3.3 Parts of court proceedings

Court proceedings usually begin by one party filing a document instituting the proceedings. The court then decides whether the specific requirement for a document instituting the proceedings has been met. If so, a document instituting the proceedings is then issued. There will then be a period of preparatory work, etc. up to the main proceedings. During the main proceedings, the parties put forward their views on the matter, after which the court makes a decision on the basis of applicable law. Conduct of proceedings in court – with some differences between criminal and civil proceedings – is based on certain fundamental principles which were established as early as 1948 in the current Code of Judicial Procedure. These principles, which consist of the duty to act objectivity, the duty to prosecute and the duty to carry out a preliminary investigation, along with the principle of orality, the principle of immediacy and the principle of concentration, form the basis for how proceedings are conducted in court.[43] The latter three principles aim to guarantee free evaluation of evidence and to provide the best basis for right, fair decisions.[44] Pursuant to the principles of immediacy and orality, as a main rule the court may only take into consideration what is brought up in the main proceedings. However, due to the reform, "En modernare rättegång" [A More Modern Judicial Process], the strict application of the principles has been relaxed and softened. This has meant, in particular, that it has become possible to a greater extent, in specific cases and circumstances, to refer to documents and sound and image recordings made during the preparations for the lawsuit. The investigators justify this by stating that it goes hand in hand with the view of a modernised judicial process which is suited to the lawsuit.

[43] Government Bill 2004/05:131 p. 80 f and SOU 2017:7 p. 23.
[44] Government Bill 2004/05:131 p. 81.

3.4 Evidence and assessment of evidence – what the court may decide on

In criminal cases in particular, the most difficult thing to decide on is usually not the law, but how the evidence in the case should be evaluated. In accordance with 35:1 RB [Rättegångsbalken – the Swedish Code of Civil Procedure], the court must decide what has been proved after a careful examination of everything that has been submitted. We understand from this that what we refer to as "free production of evidence" and "free examination of evidence" exists in the Swedish courts. Nevertheless, that does not mean that the court is able to evaluate the evidence according to its own discretion. In NJA [Nytt Juridiskt Arkiv – New Juridical Archive] 2015 p. 702, the Supreme Court stated that only rational reasons may be allowed to affect the judgment and that analysis must be structured and based on objective grounds. The court's conclusions and considerations must then be set out in the reasoning in support of the judgment in such a way that it is easy for a reader to follow how the matter has been examined.[45] That procedure must be guaranteed by a method for how evidence is evaluated and how it can be checked afterwards. The check takes place, inter alia, in the reasoning in support of the judgment, in which the court must justify how it evaluated the evidence, and also at higher courts acting as review courts.[46] There is no restriction on what may or may not constitute evidence in the physical sense. It may consist of an image, a document or a program (computer code). However, there are of course requirements for what may constitute evidence in the legal sense and what it is therefore possible to include in the main proceedings. These rules are mainly set out in Chapter 35 RB. The standard of proof required also differs according to the type of case and whereabouts in a legal process a ruling is issued. The lowest

[45] NJA 2015 p. 726 f.
[46] Andersson, "Bevisvärdering i brottmål, några reflektioner" [Evaluation of evidence in criminal cases, some reflections] 06/12/2016.

standard of proof that can be applied against an individual is “may be suspected” and the highest is “established beyond reasonable doubt”. In the case of rulings on, for example, coercive measures during a preliminary investigation, the lower standard of proof may be sufficient, whereas the highest standard of proof, “established beyond reasonable doubt” is required for a conviction in a murder trial.

Witness testimony is a common type of evidence in court, but it also constitutes a complex form of evidence. The court must adopt a position on a person's story, but factors such as tone, body language and choice of words play an important part. Therefore, a witness often needs so-called supporting evidence that supports the contents of the testimony. Nowadays, in principle, testimony is always recorded at a trial. That has the positive effect that testimony does not necessarily have to be retaken in a higher court unless there is something new to add. It is considered sufficient to replay the testimony from the lower court.

3.5 Judgments and reasoning in support of judgments

The requirements imposed with regard to transparency in a particular judgment are set out in the rules and regulations for how judgments must be written. It is in the judgment that the court gives reasons for its decision by explaining what has been proved and why. 17:7 RB sets out the minimum requirements for the contents of a judgment in civil cases and 30:5 RB sets out the requirements for criminal cases, which are largely the same. The European Court of Human Rights has also ruled in some cases that the courts must specify the grounds for their rulings to a greater extent.[47] According to the ordinary procedure, a judgment must be handed down in writing. In addition to practical information on the time and place for notification of the judgment, parties and representatives, etc., the judgment must

[47] SOU [Statens offentliga utredningar – Swedish Government Official Reports] 2001:103 p. 275.

contain the reasoning in support of the judgment, the parties' claims and objections and the facts on which they were based and the reasoning in support of the judgment must state what has been proved in the case.[48] The Processlagsberedningen [procedural law preparatory materials] stated that it was of the utmost importance for the grounds on which the court relies with regard to what can be regarded as proved to be clearly set out in the judgment. The survey and presentation also serve as a kind of review of the court's examination of the evidence and the degree of clarity depends on the circumstances of the case in general. The Processlagsberedningen also stated that the courts' reasons for their judgments were often excessively brief and even unclear, an objection which was also put forward by the previous procedural law commission.[49] Well-reasoned, clear judgments increase confidence in the judicial system because they enable the court's reasoning to be reviewed. Well-reasoned judgments also make it easier for the losing party to decide whether there are grounds for appeal to a higher court.[50] In addition to what has been proved, the judgment should also explain what is considered to have been proved. It is of the utmost interest to obtain information on what evidence or testimony, individually or collectively, forms the basis for a particular conclusion or, on the contrary, why a particular item of evidence or testimony cannot form the basis for a particular conclusion.[51]

The grounds and reasons for a particular decision are set out in the written judgment. The reasoning and discussions that take place during the deliberations in court are protected by what is referred to as secrecy of deliberations in accordance with secrecy legislation. Parties occasionally contact the court to obtain more information, which in practice is a waste of time because any

[48] SOU 2001:103 p. 277.
[49] SOU 2001:103 p. 278.
[50] Ekelöf 2011 p. 204 f.
[51] SOU 2001:103 p. 279.

reasoning and conclusions not included in the judgment are protected. We can therefore verify that the written judgment constitutes the primary source and reasoning for a judgment. The requirements imposed on the court as far as transparency is concerned are fully met with the written judgment. There is also a certain logic to this because the judicial system must be easy to understand and must be easily accessible to all citizens.

3.6 A word on legal certainty

In advance of the analysis of Artificial Intelligence in the court system and the use of AI set out below, something should be said about the term "legal certainty". Among lawyers, the term is often used to mean "predictability in legal affairs".[52] Legal certainty is key in a State governed by the rule of law such as Sweden and our courts and procedural law are designed to guarantee us it. In the doctrine, two meanings have generally been ascribed to the use of the term "legal certainty". Firstly, it is used to denote the value of a State governed by the rule of law, in other words the legal protection afforded to the individual against public power, and secondly, it is used to denote the ability of the legal system to maintain security and the right to life and property. In an essay, Ekelöf argues in favour of the second meaning, whereas in a later analysis, Frändberg states that it should instead be used in the sense of "legal peace" and that legal certainty should be used exclusively as a term for the value of a State governed by the rule of law.[53] [54] There is no clear definition of the concept of legal certainty and I believe that both meanings can be included.[55] What we can establish is that it generally means – and thus is highly relevant in the discussion on Artificial Intelligence – that a country has legislation

[52] Frändberg p. 284.
[53] Ekelöf 1942, p. 7 ff.
[54]Frändberg p. 284.
[55] Frändberg p. 285.

and a system whereby the individual citizen has protection against arbitrary interference by society itself. This means, for example, that it is not possible to be prosecuted or convicted without sufficient evidence, that there must be support in law, and that everyone, regardless of their origin or social status, is judged on equal terms.[56]

4 Artificial Intelligence in the judicial system

4.1 General

A comprehensive reform of the Swedish Code of Judicial Procedure entered into force on 1 November 2008. In the course of the study that was carried out, the legislator stated as follows: *"It is necessary for technological developments in the courts to keep pace with developments in society in general to enable the court system to meet the requirements imposed by modern society."*[57] In the current era of digitalisation and technological development, is it relevant to question what that really means? In practice, it has meant, above all, that video and audio recording have come to be used to an increasing extent along with the ability to participate using video in specific cases. The latest step in the Swedish National Courts Administration's digitalisation efforts is an e-service for filing for divorce that was launched on 15 October 2019.[58] I consider that we can begin by asking why, in practice, it does not mean more. In my opinion, it is not enough for video documentation to have been implemented. A digital e-service for filing for divorce is a step in the right direction, but it is also not enough in my view. In SOU 2016:89

[56] Ministry of Justice, Det svenska rättsväsendet [The Swedish Legal System].
[57] Government Bill 2004/05:131 p. 79 f.
[58] Swedish National Courts Administration press release 15/10/2019

– "För digitalisering i tiden" [Digitisation in time] – the researchers predict that Artificial Intelligence is the fourth industrial revolution and the second machine age. The latter concept means that machines are now being developed to take over humans' intellectual and mental abilities. That leads to incredible opportunities for developing the courts' activities and functions and thus increasing legal certainty in the long run.[59]

In a very interesting study from the USA, three American researchers demonstrate a general method for using AI to predict outcomes at the U.S. Supreme Court. With the aid of machine learning, the model was able to predict outcomes with 70% accuracy.[60] A similar study, carried out by another team of researchers, was able to predict outcomes at the European Court of Human Rights with 79% accuracy with the aid of machine learning.[61] Although the models themselves are not applicable in a real court, it is the first evidence that it is far from impossible for an AI system to be a judge.[62] The ability to predict requires the ability to interpret and understand. So, the model itself is a strong indicator that development is on the way. Estonia is one of the countries that has come furthest on the road towards introducing an AI system as a judge. Estonia itself is one of the world's most digitally-adapted countries, where the State is investing heavily in being at the forefront of the wave of digitalisation that is sweeping the world. At the end of 2019, they began beta testing an AI judge whose purpose is to deal with minor civil-law disputes involving sums not exceeding around SEK 70,000. The parties will be able to upload documents and other relevant information to the system and the AI judge will then issue a ruling on that basis. Just as in a normal court, there is a chance to appeal.[63] China is investing heavily in Artificial

[59] SOU 2016:89 p. 133.
[60] Katz et al, 2017 p. 1 ff.
[61] Alteras et al p. 1 ff.
[62] Alteras et al p. 1 ff and Katz et al, 2017.
[63] Niiler Wired 25/03/2019.

Intelligence in its judicial system. *Beijing No 1 Intermediate People's Court* (Approximately: Beijing Court of Appeal No. 1) uses Xiaofa. Xiaofa is an AI robot that has the answers to more than 40,000 legal questions and can help with over 30,000 different types of cases.[64]

The end of carbon-based life forms in the judicial system may not be as far away as we think.[65] We are therefore well on our way to being able to apply Artificial Intelligence to the court system but, as I have already pointed out, there are many challenges to be faced along the way. Below, I will focus on the existing dangers and challenges associated with AI and its use in the court system. The analysis begins with how the legal method can work in a calculation-based world and then provides a relatively comprehensive analysis of important aspects such as requirements for transparency, the risk of manipulation of an AI function, evidence and evaluation of evidence and bias. This section focuses on examining the technical aspects of AI and discussing them specifically in the light of legal principles and procedures. The ethical, moral and legal philosophical aspects will be dealt with in the next section, but some aspects will still be touched upon in this section because they are so closely linked. The analysis takes place on a continuous basis in the text and has been interwoven with the fact-based presentation. Of key importance in this section is the question of how stringent the requirements we impose should be. Should the requirements for automated decision-making be more stringent than those we impose on human decision-making and, if so, in what way? The reasonable answer is yes, otherwise we must question why we would be interested in using AI for decision-making if we cannot require a higher level of

[64] Harris, World Government Summit 11/07/2018.

[65] "The end of carbon-based life forms in the judicial system" refers to the fact that we humans are made of the element carbon. This is as opposed to computers and, in particular, components such as transistors and microchips, which are mostly made of silicon. Hence also the name "Silicon Valley" for the Mecca of the technology and computer industry in California, USA.

reliability. The most interesting and important question is therefore how stringent and how far-reaching can and should the requirements we impose be?

4.2 Legal method in a calculation-based world

Having gained a basic understanding of how Artificial Intelligence and machine learning work, it is now time to focus on how they work in combination with the law. It is already an established fact that Artificial Intelligence is changing the way we work in the law. A serious analysis of Artificial Intelligence in the judicial system should therefore include a discussion of what the law is and why the law exists. Let us begin with the second question: Why does the law exist? It is a complex issue and I will try to be brief. Legal decisions deal with countless issues and a list of the functions of the law quickly grows very long. The question is then whether we can provide a comprehensive yet detailed explanation and a reasonable answer for why the law exists. Let us therefore look at it from a practical point of view. The reason why most people resort to a lawyer is because they have a problem or because they want to avoid problems in future. This also applies at the level of society. The law has no purpose in itself, but exists to fulfil practical functions, and the reason for the existence of a regulation is because a problem exists. It becomes extremely obvious when it is written out like that and a trained lawyer may consider it to be self-evident. The fact that the law is ultimately about dealing with problems is fundamental to any understanding of why the law exists. That fact also has far-reaching implications for how the legal method should be developed and improved along with AI. Above all, however, it forms the basic premise on which we must base any analysis of AI in the judicial system in general and in the court system in particular.

In this work, I initially stated that the legal method is objective, which may almost be described as a qualified truth. It is true that the law is objective in a way. The law consists of an enormous regulatory framework whose purpose is to deal

with different types of problems. However, that regulatory framework is only a starting point for subsequent investigations, legal cases and doctrine that we use when laying the foundations for a decision. However, there is rarely a simple answer to a legal question and when you ask a lawyer a question, you often get the answer "it depends", followed by a request for more information about the circumstances, etc. If we take this one step further and look at legal philosophy and how morality and ethics should form part of judicial decisions, it becomes even more complex and difficult to understand. This makes the law an extremely complex system – perhaps one of the most complex systems in existence – precisely because it has this moral and ethical component. There is more about the more legal philosophical aspects in section 5.

We have now established that the law is a complex system designed to solve problems. Furthermore, the legal method is of course adapted to how we humans work and are able to solve problems; not to how computers solve problems. In future, when research and development of AI judges really start to take off, it will be important for us to take this into account. It is only when we have identified the underlying functions both in the law and in the medium in which the law must be interpreted – human or computer – that progress will be possible. It is only then that we will be able to see, in a meaningful way, how the application of the law can be automated using AI.[66]

4.3 Manipulation of AI

Just like other IT systems, an AI system can be manipulated. Because AI is so complex, there are also more opportunities and methods for that manipulation than in traditional IT systems.[67] Conscious manipulation of AI algorithms and

[66] Wahlgren p. 401.
[67] Schwarz p. 18.

interaction with AI are two areas in which Artificial Intelligence is vulnerable. Deliberate manipulation can take place in ways as simple as placing a sticker on a stop sign so a self-driving car fails to stop, for example, which we humans would do if we were behind the wheel.[68] Other examples include image recognition, where the value of each pixel in an image is changed in such a way that the algorithm sees something different, though the image remains unchanged for us humans (see Figure 4). For us, the two pictures below are the same. We see a panda in both pictures and have no difficulty identifying it as a panda. For an Artificial Neural Network, however, the two images are completely different. An Artificial Neural Network classifies the first image as a panda with 57.7% certainty whereas, after a filter is overlaid, it classifies the second image with 99.3% certainty as a gibbon (a kind of ape).[69]

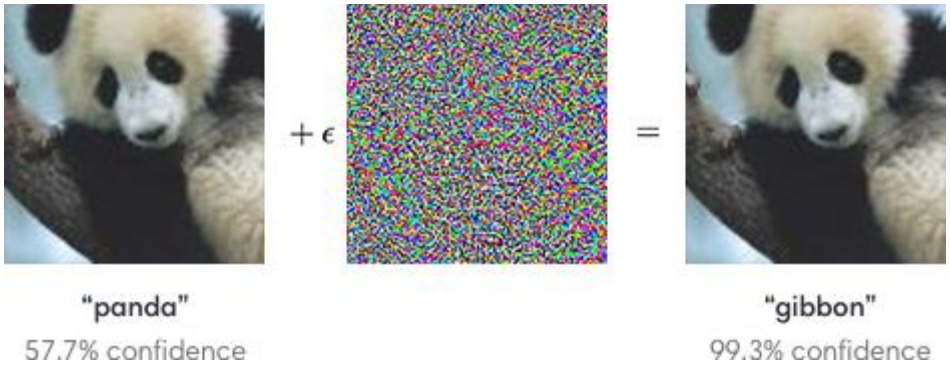

Figure 4 An opposite input, overlaid on a typical image, can cause a classifier to incorrectly categorise a panda as a gibbon (ape).

Manipulation through interaction is about adversely affecting the behaviour of an algorithm. The type of advanced AI that is able perform extremely complex tasks is taught by us with the aid of data and then keeps on learning more after that in order to develop its knowledge. In manipulation through interaction, irrelevant or incorrect data is entered, which leads to the AI making certain decisions.[70] One

[68] Eykholt et al, Section 4.3 and Schwarz p. 18.
[69] Goodfellow et al, p. 3f.
[70] Schwarz p. 18.

concrete example of this is when Microsoft released its intelligent Twitter bot, Tay. Within half a day, Tay was exhibiting behaviour typical of a Nazi and an anti-feminist and Microsoft was forced to close it down. This took place due to interaction, i.e. the people (or other robots) that Tay was communicating with on Twitter.[71] In this sense, AI can be likened to a child during its most important years of development. We "bring up" the algorithm through the interaction we have with it.[72]

There are vulnerabilities in AI in the same way as there are vulnerabilities in IT systems in general and also in humans. One particular feature of vulnerabilities in AI is that they are dependent on detailed knowledge of how the model is structured and even what training data was used when teaching it.[73] In the example of image recognition of a panda, this is the knowledge that was used to manipulate the image so it was identified as a gibbon instead.[74] We humans are vulnerable to manipulation in different ways and manipulation of the judicial system exists in the form of corruption or other influence on the various parties at the court. There will always be attempts to manipulate the judicial system, no matter how it is structured. The introduction of AI really only changes the form that the attempts at manipulation take and the important thing about the change is that we must create an understanding of how it can happen. We can then, on the basis of that understanding, develop methods to counteract the attempts. We will never be able to fully control people in the judicial system because we cannot fully control a person. We have, at least in theory, the ability to fully control and affect the methods of manipulating AI or to make it so difficult that the risk of it happening becomes almost non-existent.

[71] Hunt The Guardian 24/04/2016.
[72] Szegedy et al p. 1 ff.
[73] Papernot et al p. 1 ff.
[74] Goodfellow et al p. 3f.

4.4 Transparency in the case of automated decision-making

The use of Artificial Intelligence must be transparent, and it must be used in such a way as to be understandable. In Artificial Intelligence, this is known as "The Black Box Problem". The term comes from the fact that AI is usually likened to a "Black Box" where, without knowing how it works or how to do it, we apply the contents of the box and get a result. However, we will not always want to know or even care how a particular AI model has arrived at a particular result. One good comparison is with an ordinary washing machine.[75] We wash clothes almost daily and we usually have a washing machine to help us. However, we do not care how our clothes become clean. The only important thing is that they do become clean. Even though you can certainly work out how a washing machine works, there is no overriding reason to do so. How our clothes get clean is not really important to us; the important thing is that they do get clean. It is the same for some AI programs. We will not always care how they have arrived at a particular result. However, things are different when it comes to the law. The law is basically about values – what is right and what is wrong. How and why a legal decision is made is therefore extremely important to us. In a democratic society, it is extremely important for citizens to have confidence in the judicial system. The public's ability to gain insight into the administration of justice by the courts and their activities for that purpose is therefore vital in order for a society based on the rule of law to function. In the above section on judgments and reasoning in support of judgments, I describe the requirements for transparency of court proceedings and, in particular, of a judgment. What can be verified is that the judgment and the

[75] The comparison was made by Manuela M. Veloso, *Herbert A. Simon University Professor in the School of Computer Science at Carnegie Mellon & Head of J.P. Morgan AI Research,* during a discussion with her on 18 October 2019.

reasoning in support of the judgment constitute the entire possibility of understanding a particular outcome from the court. This therefore imposes incredibly stringent requirements on how a judgment is written in order for it to contain everything that is essential to the case and also for it to be easy to understand even if the reader is not a trained lawyer.

The challenge for Artificial Intelligence and perhaps for Neural Networks in particular is that it is almost impossible to look at the code in the model in a given situation and understand why it made the decision it made. In practice, each single node in a Neural Network works as a weight in a wave that often depends on thousands of inputs. Furthermore, a Neural Network can consist of several million or even billion nodes which, in practical terms, makes it impossible to survey in a concrete way. We must therefore take the lead in the demands we make on people. We humans work in a similar way to Neural Networks. We cannot "go in" to another person, look at the neurons and synapses, and understand what the person has been thinking and why. We therefore impose a requirement on the person to give reasons for his or her conclusion, which is done by text or speech. We should therefore reasonably impose at least the same requirements on the AI judge in terms of transparency.

4.5 Evidence and evaluation of evidence

4.5.1 Evaluation of evidence by the AI judge

Evidence forms a very important part of all cases before a court and, as I stated above, free production of evidence and free examination of evidence exists in Swedish courts. I have already discussed the use of Artificial Intelligence for analysing and understanding text, but evidence can also take many other forms, particularly images, video and testimonies. For the AI judge, this means that methods need to be developed that can evaluate evidence by mechanical means. This presents a challenge, particularly when it comes to interpreting testimony.

The task is far from impossible, but will require extensive research and development of digital tools that, for example, interact with sensors in combination with analysis of body language and facial expressions in relation to speech. For example, machines that can tell whether or not we are lying already exist. We know them by the collective name of lie detectors. These machines produce data that an AI system can access and evaluate just as we humans do when we use them.

4.5.2 Manipulation and fabrication of evidence with the aid of Artificial Intelligence

Let us take a step away from the use of Artificial Intelligence in court and instead explore how something generated using Artificial Intelligence can affect court proceedings. Court proceedings often aim to establish who is responsible for something or to settle a disagreement between different parties. Evidence forms a central part of all such proceedings. The evaluation of evidence and examination of evidence is a difficult, complex field and I have discussed the basis of it above. However, so-called Deepfakes[76] are now putting these principles to the test and challenging the free production of evidence and evaluation of evidence. Deepfakes are images, videos and sounds that have been altered or even generated with the aid of technology deriving from Artificial Intelligence. "Deep" comes from the type of advanced Artificial Intelligence used, often referred to as “deep learning”. Deepfakes are of very high quality – such high quality that it is very difficult for us to tell whether or not what we see is genuine.[77] The process of creating deepfakes involves two things: an engine to build the content and an engine to detect whether or not it is genuine. This type of deep learning is referred to as a “Generative Adversarial Network” (GAN).[78] GAN networks thus consist of two neural networks: one programmed to generate content in order to trick the other

[76] For examples of deepfakes https://www.youtube.com/watch?v=gsv1OsCEad0 (28.12.2019).
[77] Kight, M. Meet the fake celebrities dreamed up by AI and Karras et al p. 4 ff.
[78] Karras et al p. 4 ff.

neural network into thinking that it is a real image. The two networks learn from each other to enable them to create the perfect copy. Unfortunately, it is also the case that the ability to detect deepfakes using technological tools is still limited. The ability to detect deepfakes is thus closely linked to the ability to create them. New innovative models for analysing images and video are being developed, but there is a constant risk that they will be used to create even better models.

The fact that images, video and audio can be manipulated and created in such a way that it is almost impossible for us humans to tell whether they are real throws up completely new challenges for the court. In combination with the ability to use AI to understand and create text documents, fabricated "evidence" can be created to an alarming degree. Fortunately, the technological developments and the digital revolution of the past decade have already begun to challenge how evidence should be interpreted and evaluated. However, I consider that the new era that we are now entering, with deepfakes and other electronically-generated content, makes a thorough review of the Swedish principles of production of evidence and evaluation of evidence necessary so we can continue to guarantee a legally-secure process in court in future.

4.6 Bias in the courts and in the case of machine learning

Bias, both conscious and unconscious, is something that occurs in the courts, even though it goes against much of what the court system stands for.[79] There are many practical examples of bias having a clear impact in a trial in recent years. One of the most recent examples is the trial of Elin Ersson who, on 23 July 2018, stopped a flight at Landvetter Airport from taking off in order to prevent a deportation. There had to be a retrial after the first trial since it was revealed that one of the jurors had posted a comment on the matter of her guilt on social media before the

[79] Harris, World Government Summit 2018.

start of the trial. The Court of Appeal later disqualified the juror in question and ordered a retrial.[80] Another example is an extremely high-profile judgment by Solna District Court in March 2018. A man was acquitted of assault on the grounds that "he came from a good family". The judge and one of the jurors wanted to convict, but the other two jurors wanted to acquit him, and the outcome therefore went in favour of the accused. After the judgment, the prosecutor stated that she considered it to have been based on "assumptions and prejudices".[81] It is certainly true that both these examples involved jurors, which perhaps also says something about the juror system, but in my view they clearly highlight the fact that we humans are not without prejudice and find it difficult to be completely impartial.

Even a trained and skilled judge may be influenced in one way or another in his or her professional practice.[82] When it happens, it is probably mainly unconscious and completely without any intent. Bias and prejudice are unfortunately extremely widespread and pervasive in society and can be found in everything from the salary differences between men and women to how we treat our fellow human beings based on how they look or how they act. In a 2011 study on judges' parole decisions, a clear correlation could be shown between food intake and breaks and the chance of parole being granted. [83] In the study, in which the judges had three work sessions preceded by rest and food, it was shown that the chance of parole fell steadily from 65% at the beginning of the session to almost 0% at the end of the session, to then go back up to 65% after a break and intake of calories. This clearly shows the impact of something as trivial as food and rest on our ability to make objective judgments. In her doctoral thesis entitled "Confirmation Bias in Criminal Cases", Moa Liden establishes that Swedish judges who sat in both the

[80] Wedin & Håkansson Sverige Television 06/11/2019.
[81] Sverige Television 05/03/2018.
[82] Harris, World Government Summit 2018.
[83] Danzigera et al p. 6889 ff.

detention proceedings and the main proceedings had a greater tendency to convict than if they only sat in the main proceedings.[84] This is another clear sign that a judge, in his or her professional practice, cannot be completely freed from the bias that always exists in human beings.

Based on the above, we can verify that the impact of bias and prejudice in the judicial system is fairly widespread. Therefore, not everyone is equal before the law. Naturally, methods have been developed to counteract the influence of bias and guarantee everyone's equality before the law and legal certainty in treatment. One of the aims of the three instances in which the courts are divided is to counteract bias and prejudice and guarantee legal certainty. Many errors are corrected in this way. In the two examples I described above, the Court of Appeal either issued a different ruling from the district court or referred the case back to the district court for a retrial. However, it leads to considerable costs for the public and the individual – both the plaintiff and the defendant – when this occurs.

The follow-up question we therefore ask ourselves is whether this can be solved using an AI-based robot as a judge. In theory, the answer is yes. Computers only do exactly what we programme and teach them to do. However, the problem here is in the latter part of that sentence: what we *teach them* to do. This throws up major challenges that we need to solve.[85] Above, I described how AI works and particularly emphasised the importance of the data we use. Another aspect of teaching an AI judge is that the larger the dataset (a structured quantity of information) of training data we use, the more secure the trained model becomes in reality. The data that is considered for teaching purposes mainly consists of ordinances and regulations, precedents, Government bills, public investigations, doctrine and previous judgments. Other types of data may also be relevant in some

[84] Lidén p. 3.
[85] Metz The New York Times 11/11/2019.

cases, but these make up the most important sources. Previous judgments, investigations, doctrine and precedents are particularly full of various types of bias. There is therefore a considerable risk that when we train an AI model, it will access and learn from the prejudices and bias that exist in the data.

Bias can be divided into two different types: specific and systematic. Examples of specific bias include the above two examples of the trial of Elin Ersson and the trial at Solna District Court. Specific bias is often clear and easy to identify, which also makes it easy to correct any errors that occur. For example, if we use all existing legal cases from all instances as training data, it is likely that specific bias will disappear. That is because cases of specific bias are extremes which will be cancelled out by what may almost be referred to as the mean value. Systematic bias, on the other hand, presents a completely different type of challenge and can be very dangerous because it is difficult to detect and is represented consistently throughout the dataset. That also means that it is all-pervading throughout the law and its application. The question we should then ask is whether in that case it is bias or not. Google recruitment adverts are one clear example of how data rules and what effect social structures have. After one of Google's AI systems had studied quantities of data, it concluded that it was more efficient to show adverts for highly-paid jobs to men. That was neither the purpose nor the intention when the system was first developed. However, the data used to train the model, to target adverts at a particular group of recipients, identified those patterns in the data and therefore obeyed the instructions it had been given.[86] The computer does not understand the fact that we consider it morally wrong, it only goes by what it learned in the data provided.

Google's advert robot is also a good example of how bias can be expressed in data. We often relate prejudice and bias to simple labels such as ethnicity, gender,

[86] Gibbs The Guardian 08/07/2015 and Schwarz p. 19.

religion or other broad generalisations. One easy way to remedy this could be to cleanse the data of such information. But that would not happen. Systemic bias in society runs so much deeper than just labels. Structures and other patterns that can reflect bias can be found in information such as home addresses, occupations, education and other apparently trivial things. Nor is it a good idea, depending on the purpose for which the data is used. For example, take an AI model that diagnoses patients. Information on aspects such as gender and ethnicity or other sensitive information is very likely to be of decisive importance for making an accurate diagnosis and therefore for deciding on the correct treatment.

In conclusion, we can now verify that the challenge lies in how we process the data used. That requires understanding in order to understand and develop methods not to cleanse data, but to ensure that the models understand and are given an ability to determine what is important and when. That is a difficult task that requires a great deal of research.

5 The AI Judge

5.1 General

What I have discussed so far has dealt almost exclusively with the technical challenges that Artificial Intelligence involves and the implementation of AI in the light of key issues of procedural law. However, in the above analysis I have also often discussed morality, ethics and the more philosophical parts of the law. In this section, I will discuss those aspects of the AI judge in more detail. AI judges also raise issues of a more practical nature, such as the function of the courts' three instances, the occasions when there is more than one judge sitting in the court as well as implementation. The section begins with a discussion of this.

5.2 The three instances of the courts

As we know, the courts in the Swedish legal system are structured in three instances. Does that also mean that we must implement the AI judge at all three instances? If all the potential advantages of the AI judge can be put into practice, the spontaneous answer could be "yes". The hypothetical AI judge depicted: super-efficient, knowledgeable and free from bias and other shortcomings that can be found in us humans, undeniably sounds like the ultimate judge. However, I wish to argue in favour of a different approach. I consider that it would be appropriate to implement it mainly at the lowest instance. District courts and administrative courts in particular must strictly apply and comply with applicable law and should not carry out far-reaching new interpretations.[87] However, things are more difficult when it comes to courts of appeal and, in particular, the Supreme Court for example. At this point we need to ask ourselves what the role of the courts of appeal and the Supreme Court is. What we want from the Supreme Court is for it to interpret the law when it is unclear. The Supreme Courts precedential function is of tremendous value. For my part, I would not want to hand over that function to an AI judge. Our laws and rules are based on our values – they are a formalisation of morality in society. Our moral perception that it is wrong to steal from another person has led to a mountain of rules to regulate this in different situations. Another alternative solution would be to introduce an additional instance or even review the entire function of instances. The starting point with the AI judge must be to help the court system we already have, not replace it. The responsible solution that is sustainable in the long term would therefore be to see how and where it would be appropriate to introduce AI judges. A first step would

[87] Of course, all courts, regardless of the instance, have an obligation to apply the applicable law. The higher courts, nevertheless, have a greater ability to issue more far-reaching new interpretations in unclear cases.

be to follow in Estonia's footsteps and implement the AI judge in smaller civil cases for a limited value.

5.3 Judges and jurors

Most cases in the district courts are decided by a single judge through a written procedure, though it is often the case that a court consists of more than just a single judge.[88] In criminal cases and family law cases under civil law, for example, the judge is accompanied by jurors in both the district courts and the courts of appeal. Nor is it uncommon that multiple judges participate in the main proceedings. This occurs mainly in the courts of appeal and the Supreme Court.[89] The function of the jurors in the Swedish legal system is intended to ensure that the people have their voice heard and to provide the general public with insight into the courts and thereby generate confidence.[90] The juror is a kind of judge and has the same right to vote as a qualified judge in the sense that his or her vote is of equal worth. The juror participates in the question of guilt but not in the question of the penalty. That is decided by the qualified judge. The jury system poses a challenge to the use of an AI judge because it requires a human component – the very component that AI judges are intended to replace. Simply replacing a qualified judge with an AI system would not serve any great purpose because many of the potential benefits consist of the absence of bias and unrivalled efficiency. In that situation, the AI judge's contribution would mainly consist of an unprecedented knowledge base. Implementation of AI judges in divisions where jurors form part of the court thus requires a basic analysis of the function of the jury along with that of an AI judge.

[88] The Code of Criminal Procedure, Chapter 1, sections 1–3(d) and The Swedish Courts, www.domstol.se, 30/12/2019.
[89] The Code of Criminal Procedure, Chapter 1, sections 1–3(d) and The Swedish Courts, www.domstol.se, 30/12/2019.
[90] Bli nämndeman, Om uppdraget [Be a Juror, About the Post], www.blinamndeman.se, 30/12/2019.

Another important and interesting practical question arises when we have more than one judge sitting in main proceedings. Does that mean that we should also have more than one AI system which all carry out their own assessment of the case, just as judges do at present? Let us begin by examining why there is a system in which more than one judge sits. Firstly, we must assume that there is more than one judge sitting in main proceedings in order to reassure us that the process is being carried out properly and because a judge cannot be required to know everything. One judge may have one opinion, whereas the other judge has a different opinion. This difference of opinion and views is extremely healthy because it leads to constructive discussions and more legal certainty in the outcome.

5.4 Implementation and acceptance

Yet another important aspect of the AI judge is the question of implementation and acceptance by users, i.e. the citizens who will be subjected to the robot's activities, as well as how secure an AI judge must be in order to be implemented or rather, perhaps, how much more legal certainty it must provide. Is there any justification for using an AI judge that makes 99.9% correct decisions free from bias and other considerations, whereas the corresponding figure for a human being is 99%.[91] Allow me to give an example: researchers in self-driving cars believe that self-driving cars will be able to reduce the number of road traffic deaths by 90%. 253 people died on the Swedish roads in 2017. With self-driving cars, that figure would have been as low as 25.[92] It will never be possible to achieve 100%, because there will always be unforeseen events and situations that mean that road traffic deaths can never be completely avoided. Although we know that self-driving cars will run

[91] These figures are purely hypothetical in order to serve as examples.
[92] The Swedish Transport Administration.

people over and injure other people, we can nevertheless all agree that we should introduce them because we will reduce the number of injuries by 90%. If we now look at an equivalent situation for a court, the use of a robot as a judge will mean that legal proceedings are fairer and are 90% more accurate. However, it will not be 100%. There will still be cases in which an innocent person is convicted, a guilty person is acquitted or someone is treated unjustly. Will we still use an AI judge? How much better does it have to be to be justified? There are no answers to these questions as yet. But they form an important part of the process of being able to implement AI judges and they also provide answers to the questions concerning the extent to which and where AI judges should be implemented.

The next step is when we as citizens accept that an AI judge will judge us. Basically, this is about acceptance of the use of Artificial Intelligence in different parts of society. At present, the rate of acceptance is still low, but it is increasing every day. Self-driving cars are a first step and as more and more aspects of society are taken over by Artificial Intelligence, we will probably gain more and more trust in functions based on AI. For example, in health care we are already seeing that patients increasingly ask for assessments of x-ray images, etc. to also be carried out by an AI system. The question we should really ask ourselves is therefore when will we start demanding to be judged by an AI judge instead of a human being or at least demanding a supplementary opinion from an AI judge?

5.5 Should AI judges exist?

In his article “Den datoriserade domaren” [The Computerized Judge], Bylander concludes by saying that there is something uniquely human in the task of a judge and that, in the light of that approach, the view of the AI judge takes on a new dimension. Should AI judges exist? It is not a simple question. Above, I have described the possibilities opened up by AI judges and also the challenges and dangers associated with them. The first thing we need to ask ourselves is why we

want to apply Artificial Intelligence to the law. What is the purpose of that endeavour? Is it about making the process more efficient because we can see that the current system is inefficient? Is the aim to make the process fairer in order to make the principle that everyone is equal before the law a reality?

This presentation does not aim to answer those questions, but they form a crucial part of the general discussion of whether the AI judge should be or not be. We are still at the stage where there are more questions than answers and there are many uncertainties surrounding the AI judge. Let us therefore, from this point on, ignore the challenges and dangers that exist and assume that they can be solved and that we can achieve that super-judge that is free from bias and that always has the latest knowledge and is extremely efficient. In the above section on the three instances of the courts, I discussed how the law functions by giving examples of the difference between what cases are dealt with in district courts and in the Supreme Court. There is something very human in the most fundamental parts of the law. It is a question of ethics, morality and the values on which we base our lives, of which a great deal of law has become a formalisation. On the one hand, the AI judge encroaches on precisely that human aspect of the judge's function and, on the other, we see how it is precisely that human aspect that leads to injustices and situations in which everyone is not equal before the law. Where should we draw the line between the computer and the human? When does the human do more harm than good in court?

One answer to the question of whether there should be AI judges is, unsurprisingly, the classic answer often given by lawyers: it depends. The Swedish court system is extensive and complex and deals with a large number of different cases of a changeable nature. To generalise by giving a simple answer of yes or no is therefore almost reckless. There are areas of the court system that AI judges will form a natural part of, where their enormous capacity will benefit all parties involved and will be of great value. However, there are also parts of the court

system where AI judges have no place, where that uniquely human aspect needs to be expressed and given space. However, where to draw the line between human and machine is another much more important question.

6 Conclusion

Much of what has been discussed in this work is extremely hypothetical. At present, there is still no working or even close to working AI robot that could act as a judge to the fullest extent. However, we cannot say that there never will be one. Research into Artificial Intelligence is making increasing strides forward and we cannot say when or if we will make the necessary breakthroughs. However, one thing we can be sure of is that we must be proactive in its development. There are too many nightmare scenarios in which AI is used for harmful purposes for us to just sit back and wait for technological development to happen. If we ignore the technical part of Artificial Intelligence, it becomes primarily a question of morality and ethics.

Where will this development take us? AI will not become part of the court system overnight. It will happen in the form of a gradual transformation. Even though a robot as a judge may not be suitable for replacing the entire court itself, that does not mean that it should not happen to a limited extent. The robot as a judge is not a question of all or nothing. There are plenty of intermediary options that could significantly improve the court system. The first roles that AI will take on are likely to be in scheduling and other administrative aspects.[93] There are many parts of the legal process that can be automated using AI to increase efficiency. For example, an AI function could be used for approving or rejecting the institution

[93] Öster & Ekholm, Robotdomaren ännu långt bort [The Robot Judge Still a Long Way Off], 30/11/2019.

of actions and/or requesting additional information so actions can be brought. The institution of an action is, in technical terms, a rather simple process in which a number of requirements must be met. At present, this is done manually by a person who goes through each action and either approves, rejects or requests additional information. That could definitely be taken over by an AI system to free up those skills for other things. The next step after that would probably be the implementation of solutions suited to simpler types of cases. One example of this would be cases that are settled on the basis of documents, without going to main proceedings.

In addition to the potential role of AI in the legal system, this work also contains a great deal of discussion on the role of Artificial intelligence in society as a whole and the importance of developing AI in a sustainable manner. There are incredible possibilities with Artificial Intelligence and AI. It can be of great help to humanity, but there are also enormous dangers if we use it the wrong way. In the end, I ask more questions than I answer in this work. However, one of its major underlying purposes is to spark the reader's interest in looking towards the horizon and seeing what is coming. We are facing a technological shift in society and in the law in particular. Future research is crucial in this process and there are three areas that I consider to be particularly important:

- Bias – how we can prevent bias from being developed in the AI models used,
- Transparency – what requirements we can and should impose regarding the transparency of AI decisions with legal effect and
- Manipulation of AI models – to identify possible methods and approaches and how we can protect ourselves from them.

What I would particularly like to say and convey with this presentation is that AI in the court system is not as far away as we might think, but above all that we

must take on an active role in the development of AI right now. We are the ones who have the opportunity to create useful, beneficial AI and counteract its negative aspects. We can no longer turn a blind eye to the transformation that is taking place. We can no longer be spectators, but must instead be actors in this development. It is our opportunity, but also our duty to be involved so we can develop ethical, fair AI, created to improve people's everyday lives, not harm them!

Source and bibliography

Official publications

Prop. 2004/05:131 *En modernare rättegång*

SOU 2001:103 *En modernare rättegång*

SOU 2016:89 *För digitalisering i tiden*

SOU 2017:7 Straffprocessens ramar och domstolens beslutsunderlag i brottmål – en bättre hantering av stora mål

Bylander, E, *Den datoriserade domaren*, SvJT 2019 p. 1-7

Ekelöf, P O, *Rättssamhälle och rättssäkerhet*, SvJt 1942 7–19 (cit: Ekelöf 1942)

Artificiell intelligens i svenskt näringsliv och samhälle — Analys av utveckling och potential, Final report, Vinnova 2018 Reg. no. 2017-05616

Legal cases

NJA [Nytt Juridiskt Arkiv – New Juridical Archive] 2015 p. 702

Literature

Aletras, N, Tsarapatsanis, D, Preoμiuc-Pietro, D, Lampos, V, *Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective*, PeerJ Computer Science, 2016.10.24

Bishop, C M, Pattern Recognition and Machine Learning, Springer, 2006

Danzigera, S, Levavb, J, Avnaim-Pessoa, L, *Extraneous factors in judicial decisions*, Proc Natl Acad Sci U S A. 2011, Volume 108, no. 17, p 6889–6892.

Devlin, J, Chang, M-W, Lee, K, Toutanova, K, *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*, arXiv, 2019, arXiv:1810.04805v2

Ekelöf, P O, Edelstam, H, Pauli, M, *Rättegång. Femte häftet*, edition 8, 2011, Norstedts Juridik (cit: Ekelöf et al 2011)

Eykholt, K, Evtimov, I, Fernandes, E, Li, B, Rahmati, A, Xiao, C, Prakash, A, Kohno, T, Song, D, *Robust Physical-World Attacks on Deep Learning Models*, arXiv 2018, arXiv:1707.08945

Frändberg, Å, Rättsordningens Ide – En antologi i allmän rättslära, Iustus förlag, 2005.

Gibbs, S, Women less likely to be shown ads for high-paid jobs on Google, study shows, The Guardian, 08/07/2015 (11/11/2019)

Goodfellow, I, Shlens, J, Szegedy, C *Explaining and Harnessing Adversarial Examples*, arXiv, 2015, arXiv:1412.6572v3

Gräns, M, *Allmänt om användningen av andra vetenskaper inom juridiken*, in Nääv, M, Zamboni, M (ed.), *Juridisk Metodlära*, 2015, Studentlitteratur, Edition. 1

Harris, B, *Could an AI ever replace a judge in court?*, World Government Summit, 11.07.2018 (30.11.2019), https://www.worldgovernmentsummit.org/observer/articles/could-an-ai-ever-replace-a-judge-in-court

Hunt, E, Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter, The Guardian, 24/04/2016 (30/12/2019) https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter?CMP=twt_a-technology_b-gdntech

Katz DM, Bommarito MJ, II, Blackman J (2017) *A general approach for predicting the behavior of the Supreme Court of the United States*, PLoS ONE 12(4): e0174698. https://doi.org/10.1371/journal.pone.0174698 (cit: Katz et al, 2017)

Katz Dm, *Quantitative Legal Prediction—Or—How I Learned To Stop Worrying And Start Preparing For The Data-Driven Future Of The Legal Services Industry*, Emory Law Journal Vol 62:909 2013 p 909-966 (cit: Katz 2013)

Knight, Will, *Meet the Fake Celebrities Dreamed Up by AI*, MIT Technology Review, 31.10.2017 (30.11.2019) https://www.technologyreview.com/f/609290/meet-the-fake-celebrities-dreamed-up-by-ai/

Leon A. Gatys, Alexander S. Ecker, Matthias Bethge, *A Neural Algorithm of Artistic Style*, arXiv, 2015, arXiv:1508.06576v2

Lidén, M. 2018. *Confirmation Bias in Criminal Cases*. 284 pp. Uppsala: Department of Law, Uppsala University.

McKinney, S.M., Sieniek, M., Godbole, V. et al.International evaluation of an AI system for breast cancer screening. Nature 577, 89–94 (2020) doi:10.1038/s41586-019-1799-6

Metz, C, *We Teach A.I. Systems Everything, Including Our Biases*, The New York Times, 11/11/2019 (20/11/2019), https://www.nytimes.com/2019/11/11/technology/artificial-intelligence-bias.html#click=https://t.co/SbJWcnrehL (cit: Metz, The New York Times, 11/11/2019)

Metz, C, *Finaly, a Machine That Can Finish Your Sentence - Completing someone else's thought is not an easy trick for A.I. But new systems are starting to crack the code of natural language*, The New York Times, 18/11/2018 (23/11/2019), https://www.nytimes.com/2018/11/18/technology/artificial-intelligence-language.html?module=inline (cit: Metz, The New York Times, 18/11/2018)

Nayak, P, Understanding searches better than ever before, Google, 2019-10-25 (30.12.2019)

Niiler, Eric, *Can AI be a Fair Judge in Court? Estonia Thinks So*, Wired, 2019.03.25 (20.11.2019) https://www.wired.com/story/can-ai-be-fair-judge-court-estonia-thinks-so/

Papernot, N, McDaniel, P, Goodfellow, I, Jha, S, Berkay Celik, Z, Swami, A, *Practical Black-Box Attacks against Machine Learning*, arXiv, 2017, arXiv:1602.02697v4

Pinar Saygin, A, Cicekli, I, Akman, V, *Turing Test: 50 Years Later*, *Minds and Machines*, Kluwer Academic Publishers, 2001

Press, G, *A very short history of Artificial Intelligence (Ai)*, Forbes, 2016.12.30 (30.11.2019) https://www.forbes.com/sites/gilpress/2016/12/30/a-very-short-history-of-artificial-intelligence-ai/#1d153fe66fba

Radford, A, Wu, J, Amodei, Dar, Amodei, Dan, Clark, J, Brundage, M, Sutskever, I, Askell, A, Lansky D, Hernandez, D, Luan, D, *Better Language Models and Their Implications*, OpenAI, 14/02/2019 (15/11/2019), https://openai.com/blog/better-language-models/

Schwarz, Å, *Hur vi förbereder oss för riskerna med AI som lurar runt hörnet*, i Akenine, D, Stier, J (ed.), *Människor och AI – En bok om artificiell intelligens och oss själva*. BoD- Books on Demand, 2019

Sharkey, N, *Alan Turing: The experiment that shaped artificial intelligence*, BBC, 21/06/2012 (11/11/2019) https://www.bbc.com/news/technology-18475646

Somers, J, *Is AI Riding a One-Trick Pony?*, MIT Technology Review, 29.09.2017 (30.11.2019). https://www.technologyreview.com/s/608911/is-ai-riding-a-one-trick-pony/

Szegedy, C, Zaremba, W, Sutskever, I, Bruna, J, Erhan, D, Goodfellow, I, Fergus, R *Intriguing properties of neural networks*, arXiv, 2014, arXiv:1312.6199v4

Karras, T, Laine, S, Aila, T, *A Style-Based Generator Architecture for Generative Adversarial Networks*, arXiv, 2019, arXiv:1812.04948v3

Turing, A, *Computing Machinery and Intelligence*, Mind, Volume LIX, no. 236, October 1950, p 433–460 (cit: Turing)

Vincent, J, *OpenAI's new multitalented AI writes, translates, and slanders*, The Verge, 14/02/2019 (15/11/2019), https://www.theverge.com/2019/2/14/18224704/ai-machine-learning-language-models-read-write-openai-gpt2

Wahlgren, P, *Automatiserade juridiska beslut*, i Nääv, M, Zamboni, M (ed.), *Juridisk Metodlära*, 2015, Studentlitteratur, Edition 1

Wedin, H, Håkansson, A, *Aktivisten Elin Ersson döms till dagsböter – igen*, Sveriges Television 06/11/2019 (23/11/2019), https://www.svt.se/nyheter/lokalt/vast/dom-faller-mot-aktivisten-elin-ersson-igen

Öster, U, Ekholm, E, *Algoritmernas nya era är här*, Advokaten no. 4 2019 annual volume 85 (30/11/2019), https://www.advokaten.se/Tidningsnummer/2019/nr-4-2019-argang-85/algoritmernas-nya-era-ar-har/ (cit: Öster & Ekholm, Algoritmernas nya era är här, 30/11/2019)

Öster, U, Ekholm, E, *Robotdomaren ännu långt bort,* Advokaten no. 4 2019 annual volume 85 (30/11/2019), https://www.advokaten.se/Tidningsnummer/2019/nr-4-2019-argang-85/new-page22/ (cit: Öster & Ekholm, Robotdomaren ännu långt bort, 30/11/2019)

Öster, U, Ekholm, E, *AI standard på storbyråerna*, Advokaten no. 4 2019 annual volume 85 (30/11/2019), https://www.advokaten.se/Tidningsnummer/2019/nr-4-2019-argang-85/ai-nu-standard-pa-storbyraerna/ (cit: Öster & Ekholm, AI standard på storbyråerna, 30/11/2019)

Öster, U, Ekholm, E, AI nu självklar på advokatbyråerna, Advokaten no. 4 2019 annual volume 85 (30/11/2019), *https://www.advokaten.se/Tidningsnummer/2019/nr-4-2019-argang-85/new-page20/*
(cit: Öster & Ekholm, AI nu självklar på advokatbyråerna, 30/11/2019)

Internet sources

Andersson, P-E, *Bevisvärdering i brottmål, några reflektioner*, 06/12/2016 (28/11/2019) http://www.domarbloggen.se/sodertorns-tingsratt/bevisvarderingen-i-brottmal-nagra-reflexioner/

Bli nämndeman, om uppdraget, 27/06/2019 (30/12/2019) http://www.blinamndeman.se/Bli-namndeman/Om-uppdraget/ (part of the Swedish National Courts Administration)

The Swedish National Courts Administration, *Ny e-tjänst – gemensam ansökan om skilsmässa*, Press release 15/10/2019, (13/11/2019), https://www.domstol.se/nyheter/2019/10/ny-e-tjanst--gemensam-ansokan-om-skilsmassa/

Ministry of Justice, Det svenska rättsväsendet, (09/11/2019) https://www.regeringen.se/contentassets/f7691b53fffc4f7c9edeeddf4cbb2e2f/det-svenska-rattsvasendet.pdf

Man ”av bra familj” friad för misshandel, Sveriges Television, 05/03/2018 (20/11/2019), https://www.svt.se/nyheter/lokalt/stockholm/man-av-bra-familj-friad-for-misshandel (cit: Sverige Television 5/3/2018)

SAS Institute, *Natural Language Processing (NLP) - What it is and why it matters*
https://www.sas.com/en_us/insights/analytics/what-is-natural-language-processing-nlp.html (30/11/2019)

The Swedish Courts, Allmänna domstolar 19/06/2019 (30/12/2019) https://www.domstol.se/om-sveriges-domstolar/sa-fungerar-domstolarna/allmanna-domstolar/
(cit: The Swedish Courts, www.domstol.se)

The Swedish Transport Administration, *253 döda i trafiken 2017 lägsta antalet sedan 40-talet,* Press release 25/04/2018 (10/01/2020) https://www.trafikverket.se/om-oss/pressrum/pressmeddelanden/Nationellt/2018/2018-04/253-doda-i-trafiken-2017-lagsta-antalet-sedan-40-talet/