Artificial Intelligence as a Commons – Opportunities and Challenges for Society

Amelie Salameh

Suggested citation:

Salameh, Amelie. 2017. "Artificial Intelligence as a Commons – Opportunities and Challenges for Society." Hannover: Hochschule Hannover. https://doi.org/10.25968/opus-1153.

Abstract

The development of Artificial Intelligence (AI) has profound implications for improving human and computational productivity in the future. However, it also is an existential risk to human life because it could exceed human capabilities. As such, information about the technology, the direction of the development and its purpose is important. This can be achieved through openness and transparency of processes. Indeed, companies hold property rights over AI and monopolies of software, data and experts. As a countermovement to leading AI companies, the "Open AI Movement" has evolved to push open-source AI research and products, to empower users, and to bridge the digital divide through participation and access. In this thesis, the implications of the declaration of AI as a commons have been analyzed through interviews with AI experts in the United States. The legal placement of AI is controversial but it could be seen as a basic human right. Other findings are that this field is very competitive and that the best approach is to collaboratively develop software that adds additional value on the edge of the commons.

Hochschule Hannover

Fakultät III – Medien, Information und Design

Abteilung Information und Kommunikation

Artificial Intelligence as a Commons – Opportunities and Challenges for Society

Bachelorarbeit

im Studiengang Informationsmanagement

vorgelegt von

Amelie Salameh
Matrikelnummer: 1265885

Berlin, der 02.08.2017

Erstprüfer:	Prof. Dr. Christian Wartena
Zweitprüfer:	Dirk Jehmlich

Abstract

The development of Artificial Intelligence (AI) has profound implications for improving human and computational productivity in the future. However, it also is an existential risk to human life because it could exceed human capabilities. As such, information about the technology, the direction of the development and its purpose is important. This can be achieved through openness and transparency of processes. Indeed, companies hold property rights over AI and monopolies of software, data and experts. As a countermovement to leading AI companies, the "Open AI Movement" has evolved to push open-source AI research and products, to empower users, and to bridge the digital divide through participation and access. In this thesis, the implications of the declaration of AI as a commons have been analyzed through interviews with AI experts in the United States. The legal placement of AI is controversial but it could be seen as a basic human right. Other findings are that this field is very competitive and that the best approach is to collaboratively develop software that adds additional value on the edge of the commons.

Keywords: open-source software, OSS, Artificial Intelligence, AI, OpenAI, commons, digital divide, digital commons, regulation, transparency, openness

Table of contents

Index of acronyms and abbreviations

OSS	Open-Source Software
AI	Artificial Intelligence
AGI	Artificial General Intelligence
ANN	Artificial Neural Network
ML	Machine Learning
API	Application Programming Interface
IPR	Intellectual Property Rights
U.S.	United States (of America)
MIT	Massachusetts Institute of Technology
ILSVRC	ImageNet Large Scale Visual Recognition Challenge
BA	Bachelor of Arts
BS	Bachelor of Science
MS	Master of Science
UDHR	Universal Declaration of Human Rights
UTSA	Uniform Trade Secrets Act
FOIA	Freedom of Information Act
GDPR	General Data Protection Regulation
EP	European Parliament

Index of images and tables

1 Introduction

The goal of this thesis is to conduct a preliminary analysis of transparency and openness in the development of Artificial Intelligence (AI) as well as explore the possibility of the declaration of AI as a commons. A commons has several characteristics and one of the first question would be if AI fulfills these traits. Other questions include: What would happen if everyone gained access to AI? How can we distribute software equally? If there are no monopolies and no owners, who would regulate it and who would allocate the wealth generated by AI? How can we ensure transparency about means and purposes in the making of AI? What other challenges might society face? What opportunities emerge for society?

Definitions of several technical terms have been examined and established at the beginning. Furthermore, AI experts in the United States (U.S.) were interviewed and the results have been analyzed and summarized in this paper.

Artificial Intelligence is a unique technology in that sophisticated AI can pose an existential risk to human life (Bostrom 2016, p. 9). A so-called "superintelligence" is one that would exceed human intelligence and therefore pose a potential threat to human existence by surpassing human capabilities. A human-level (or above) Artificial Intelligence could develop its own will and it could be in conflict with the will of humans. Autonomous robots – AI paired with robotics – could be a great danger, too. Autonomous weapons, disruption of financial markets and oppression are some of the dangers humanity might face as soon as a sophisticated AI has been developed, according to Hawking (Fuller 2016). As Russell and Norvig (1995, p. 3) state:

> "Although no one can predict the future in detail, it is clear that computers with human-level intelligence (or better) would have a huge impact on our everyday lives and on the future course of civilization."

We are not there yet – a lot of research and development still needs to be done. However, experts give warnings, hold talks, get inquired by governments, and form committees to explore these risks.

Presently, AI can already program machine-learning software which can executes tasks. As such, the barrier of the prerequisite of expertise to program AI software has greatly

reduced while the pace of AI development has increased (Simonite 2017). Soon there could be AI which can program its own code. Recursive self-improvement is risky as human developers may have problems understanding the code. Experts even predict that these sophisticated AIs will be able to communicate in languages not understandable by ordinary citizens with the consequence of exclusion. It is unforeseeable but predictable that superhuman intelligence, or superintelligence, will first be created by computers within the AI space and not by enhancing human intelligence somehow (Yudkowsky 2012, p. 32). An exponential technological growth that is beyond human control is called "technological singularity" by futurologists (Bostrom 2016, p. 9).

Having an understanding of background processes as well as capabilities, and knowing the means and purposes of AI could keep humans in control. Knowledge especially plays a role in transparency and openness. Only informed people can make informed choices, and in order to be informed, it is important to have access to information (Bostrom 2016, p. 13). But leading AI companies like Google and Facebook have built monopolies around the software and data including property rights and expert accumulation. And expertise seems like the trading good these days because big AI companies are fighting over talented software engineers (Metz 2017). With a lot of power often comes a lot of responsibility. Bill Hibbard (2016, p. 3) even declares the responsibility of AI developers and announces them as "representatives for the future of humanity".

There are initiatives trying to bring more transparency and balance into the market surrounding this field of science. For example the OpenAI organization has "openness explicitly built into its brand identity" (Bostrom 2016, p. 2). It pushes openly accessible and modifiable AI. Indeed, the "Open AI Movement" can be observed worldwide with more and more developers as well as companies' open-sourcing their AI's.

This thesis tries to combine the phenomena of AI as an existential risk, "Open AI" as a multiplicator of transparency and openness, and the idea of the commons as a stepping stone for participation and inclusion.

It begins by rendering a definition, history, methods, and other technical terms of AI. In the third chapter, commons will be explained and the theories around human participation and involvement in software development and common distribution. The fourth chapter will combine both phenomena, AI and commons, explore open-source AI and the "Open AI Movement," represent and summarize the AI expert interviews, outlining opportunities and challenges facing society as well as the legal placement of AI. The fifth chapter

will outline an assessment on competition within the AI market and finally the last chapter will conclude the findings of this thesis.

2 Overview of Artificial Intelligence

This chapter will give a short overview on how Artificial Intelligence can be defined, the condensed history of Artificial Intelligence, and different techniques to develop AI.

2.1 Definition of Artificial Intelligence

Artificial Intelligence as a field of science can trace its roots to 1956. Artificial intelligence, or AI, is the attempt to understand human intelligence and to build similar intelligent entities with machines. AI is inextricably linked to other fields of science including Neuroscience, Mathematics, Psychology, Computer Engineering and Linguistics. Neuroscience, the scientific study of the nervous system, is a fundamental element of AI. Artificial Neural Networks (ANNs) are computing systems inspired by the biological neural networks that constitute brains. Neural networks are the base of connection in the human brain and makes us able to understand entities, put them into context, able to feel pain and other emotions. These are tried to be recreated artificially nowadays (Russell & Norvig 1995, pp. 4-8).

Language also plays a big role in the building process of AI. Humans communicate with each other through language making it necessary for an intelligent machine to understand widely-spoken languages, too. Theories about how children can learn a language and create their own sentences from listening in their environment are the foundations of Natural Language Processing and Computational Linguistics. AI is mainly utilized in understanding problems of language use (Russell, Norvig 1995, pp. 11–16). The code is also a language which the developer must learn and use comprehensively. This means that AI is an interdisciplinary science which combines the discipline of Cognitive Science and the discipline of Computer Science. The goal of AI is to construct a computer which can learn to do tasks by considering examples, generally without task-specific programming, and thus replace the need for human intelligence to do those same tasks. A meta-goal is to understand human intelligence. As such, AI is differentiated by “Strong AI” and “Weak AI”. Strong AI, or Artificial General Intelligence (AGI), is AI that is capable of under-

standing human intelligence and imitating it. Weak AI is not able to explain human cognitive functions and simulate precisely human-level intelligence or beyond (Kurzweil 2012, p. 161).

There is no unique definition of AI, but there are different approaches to explain artificial intelligence. An illustration of four main categories made by Stuart Russell and Perter Norvig (1995, p. 4) can be seen in figure 1.

Thinking Humanly	**Thinking Rationally**
"The exciting new effort to make computers think . . . *machines with minds*, in the full and literal sense." (Haugeland, 1985) "[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning . . ." (Bellman, 1978)	"The study of mental faculties through the use of computational models." (Charniak and McDermott, 1985) "The study of the computations that make it possible to perceive, reason, and act." (Winston, 1992)
Acting Humanly	**Acting Rationally**
"The art of creating machines that perform functions that require intelligence when performed by people." (Kurzweil, 1990) "The study of how to make computers do things at which, at the moment, people are better." (Rich and Knight, 1991)	"Computational Intelligence is the study of the design of intelligent agents." (Poole *et al.*, 1998) "AI . . . is concerned with intelligent behavior in artifacts." (Nilsson, 1998)

Figure 1: Definitions of Artificial Intelligence[1]

Presently, the differences in definitions lie in the methods or techniques used to construct the machines. There are several established methods which reach and exceed human intelligence in particular tasks. The goal of creating a machine which behaves like a human in the full range of cognitive function was abandoned early; instead practitioners concentrated on specific tasks for real-world problems (NSTC 2016, p. 5).

Machine programs consist of algorithms which function as instructions for the machine. A computer is made out of different pieces such as micro-processors or RAM storages, which themselves are made of billions of tiny switches called transistors. These transistor are able to be "on" or "off" – one (1) for on and zero (0) for off. That is how programmers

[1] Source of image: Russell & Norvig 1995, p. 5

can instruct the computer to execute tasks – through algorithms. The algorithms tell the computer either to switch off or on the transistor A, or to switch on the transistor A if the transistor B is switched off, or to switch off transistor A and B if C is switched on (see figure 2).

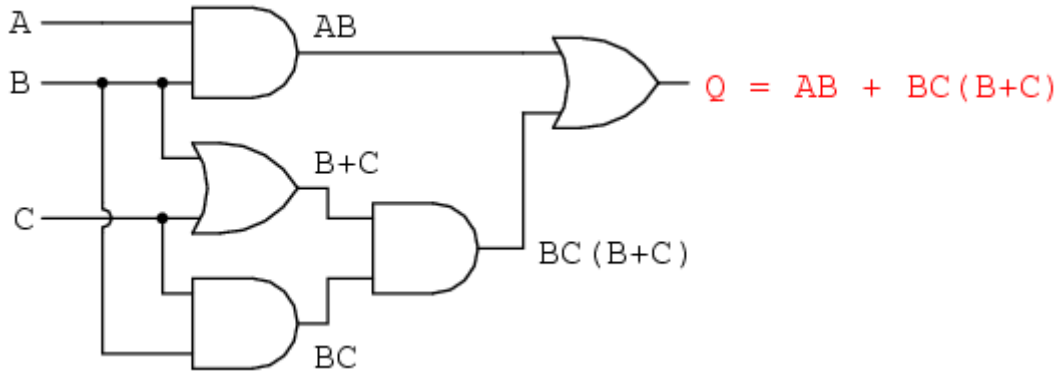

Figure 2: Logical reasoning / Boolean connectors[2]

This logical reasoning was first discovered by Claude Shannon in 1940 at the Massachusetts Institute of Technology (MIT). He discovered that the algorithm, no matter how complicated, can be simply expressed by three logical connectors: AND, OR and NOT (Domingos 2015, pp. 1-4).

2.2 Brief history of AI

The literature concerning the history of AI is vast, exploring its birth with the first computers, funding history, computational theorems, and timelines of development. This chapter will give an overview over the most significant events in the history of AI regarding this thesis.[3]

Alan Turing is among the most renowned AI pioneers. In the 1940's, when Claude Shannon published his Master of Science (MS) thesis "Relay and Switching Circuits," Turing published a journal article titled "Computing Machinery and Intelligence" in 1950 where

[2] Source of image: https://www.ibiblio.org/kuphaldt/electricCircuits/Digital/DIGI_7.html, by Tony R. Kuphaldt, date: 2000-2017

[3] A more detailed history of AI is available in the appendix of the AI 100 Report. Peter Stone, Rodney Brooks, Erik Brynjolfsson, Ryan Calo, Oren Etzioni, Greg Hager, Julia Hirschberg, Shivaram Kalyanakrishnan, Ece Kamar, Sarit Kraus, Kevin Leyton-Brown, David Parkes, William Press, AnnaLee Saxenian, Julie Shah, Milind Tambe, and Astro Teller, "Artificial Intelligence and Life in 2030," One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel, Stanford University, Stanford, CA, September 2016, http://ai100.stanford.edu/2016-report.

he asked the question “Can machines think?”. He proposed that the answer to this question could be determined by a test, which he called “the imitation game,” later known as the Turing Test (Turing 1950, pp. 433-460).

The challenge of the Turing Test was for an interrogator to listen to a dialogue between a machine and a human and being able to distinguish the machine from the human (Stahl 2004, p. 68). The test was designed to probe whether the AI could execute human-level (or higher) tasks in all cognitive functions. For being able to pass the test, the AI would need to possess the “following capabilities: natural language processing, knowledge representation, automated reasoning and machine learning” (Russell, Norvig 1995, p. 5).

Turing also raised the thought of the possibility that a machine could be programmed to learn from experience as children do, which was later adopted into the theorems of Machine Learning (Turing 1950, p. 444).

As previously mentioned, the field of science called “Artificial Intelligence” became formally initiated in a proposal for a summer workshop at Dartmouth College in 1956 by John McCarthy, Claude Shannon, Marvin Minsky and Nathaniel Rochester. This two-month summer workshop brought together ten researchers to explore Artificial Intelligence on the basis of “every aspect of learning or any other feature of intelligence (…) in principle be so precisely described that a machine can be made to simulate it” (McCarthy 1955).

These researchers knew that, in 1955, their current computing capacities were not sufficient enough for machines to mimic human intelligence. Despite the fact that they agreed on the need to take full advantage of the available technology, they were not able to write programs which could perform as they wanted them to (McCarthy 1955). Some of the most fundamental research questions proposed during this conference remain unsolved and actively researched today. How machines can think creatively or develop a human-like intuition or how they self-improve is one such question. Other ideas, e.g. how machines can use language or how neural nets can be simulated, have been solved by various individuals and companies.

AI’s development went through highs and lows. There were funding stops around the middle of the 1970’s called the “AI winter” because investors and other developers concluded that the technological infrastructure was not mature enough to develop any highly sophisticated software. Hope was renewed in 1997 when IBM’s chess-playing computer Deep Blue won a chess game against world champion Garry Kasparov (Bostrom 2014,

p. 21; NSTC 2016, p. 6). Since then, the Research and Development (R&D) of AI has flourished and attracts investors as well as individuals to take part in this field of science. The field especially flourished because of the explosion of the amount and the availability of data from various sources, e.g. Social Media, electronic commerce, electronic publications and the digitalization of the government. One notable phenomena called "Big Data," for example, improved the error rate of computer programs for image recognition from 26 percent in 2011 to only 3.5 percent in 2015 (NSTC 2016, p. 6). The next significant achievement happened in March 2016 when Google DeepMind's AI called "AlphaGo" defeated the best player Lee Sidol in a game called "Go" – an ancient Chinese board game which has near-infinite number of board positions and usually incorporates human intuition. This was achieved with the method of Reinforcement Learning, a technique of Deep Learning (see chapter 2.3.) (Burger 2016).

Nowadays Artificial Intelligence, also sometimes referred to as Machine Intelligence, is implemented in various industry sectors and businesses. Figure 3 shows in which areas Machine Intelligence is implemented and applied as of 2016, as well as different stakeholders and technology providers involved in the field. Artificial intelligence can be best explained through the mostly used state-of-the-art method Machine Learning, and the sub category Deep Learning as well as Reinforcement Learning (Metz 2016).

2.3 Machine Learning, Deep Learning and Reinforcement Learning

Machine Learning (ML) is a technical approach to AI, which can be found in most applied AI systems as of 2017, as the figure 3 shows.

ML starts with a dataset as a foundational input, e.g. images of animals or a register of convicted persons. This is followed by an attempt by the system to create a generalization model for these examples, in order to predict future outcomes of yet-unseen input. These models are statistical calculations of entities with predictions as a result. For instance, in order to be able to find a cat on a picture, the algorithm is trained by being fed 200,000 images of cats. Eventually, the machine is able to figure out what a cat looks like. Algorithms are instructed on how to process these images by breaking them down into individual pixels, structuring and detecting these pixel into objects, and categorizing these objects with determined tags.

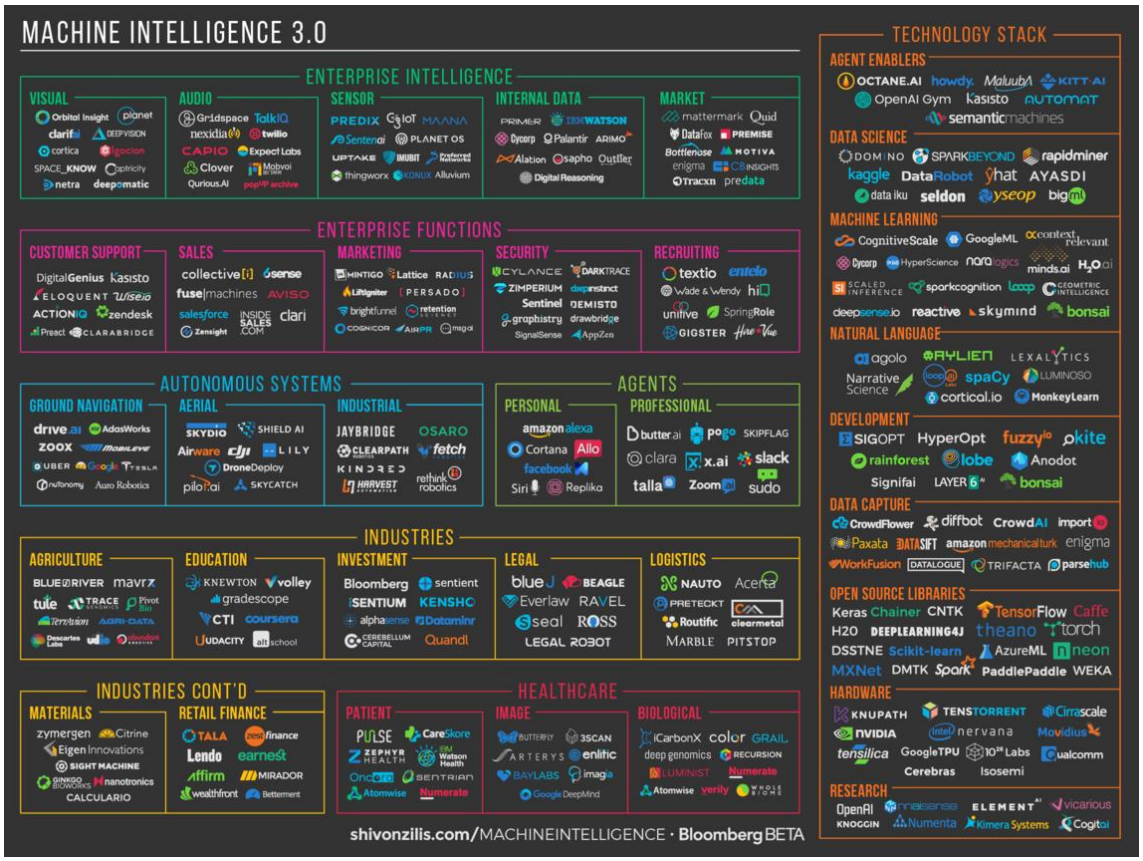

Figure 3: **Landscape of Machine Intelligence** [4]

If such a machine receives a picture as new input, it will examine the structures of the objects pixel-precisely and then calculate the probability of that object in the picture being a cat, expressed in percentages. This input is called “training data” because it is used to train a ML model, e.g. the ImageNet dataset.

ImageNet is the outcome of the annual ImageNet Large Scale Visual Recognition Challenge (ILSVRC), started in 2010. A team of researchers from different American universities published a training dataset of manually annotated images. The challenge was to submit algorithms for object detection and localization from images and videos which use the manually annotated images as training data. Another source of an uncategorized image dataset is given to test if the algorithm can detect and localize objects within images and videos automatically. The results are submitted to the evaluation server which evaluates the precision and recall (Russakovsky et al. 2015, pp. 1-6).

Deep Learning is based on the neural network scheme of the human brain. An ANN passes on the data from neuron to neuron to find patterns and patterns-of-patterns within

[4] Source of image: https://www.oreilly.com/ideas/the-current-state-of-machine-intelligence-3-0, posted by Shivon Zilis and James Cham, designed by Heidi Skinner, date: 11/07/2016

the data (NSTC 2016, p. 9). In 2014, The GoogLeNet team from *Google* won the ILSVRC and doubled the precision of the results in both tasks – classification and detection. The team was named after Yann LeCun's convolutional network "LeNet," which is a subtype of neural networks. Convolutional networks are most often used within the area of object recognition. They use images as input and pass them on to layers of the network where each layer puts a filter on the image to distinguish objects and then forwards the processed image to the next layer. Each artificial neuron receives as input the outputs of neurons in a previous layer. How the input conveys from each featured map (layers) and goes through subsampling stations until all convolutions are connected to generate output can be seen in figure 4.

Reinforcement Learning is the iteration of a task which a machine performs by trial and error. If, for example, the computer should find out how a three-dimensional figure can walk, this task will be repeated as long as the figure loses balance and falls down. As soon as the computer finds out how to move the figure into a direction without errors, it creates an algorithm that can be used in robotics to move robots from one location to another.

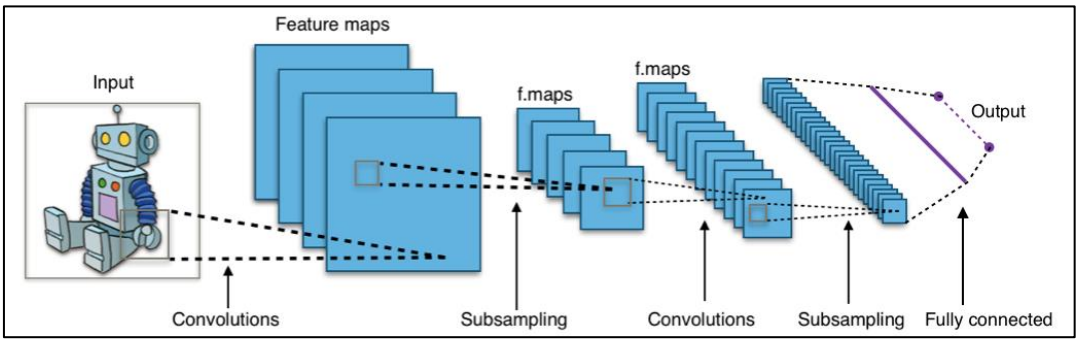

Figure 4: **Convolutional network of the GoogLeNet team**[5]

The specialty is that Reinforcement Learning does not need data because the figures can be created. The OpenAI project uses this technique to play simple open-source Atari computer games (see chapter 4.1.3.). The Stanford Artificial Intelligence Laboratory (SAIL) study panel's report, the "100 Year Study on AI," describe it as follows: "Reinforcement learning is a framework that shifts the focus of machine learning from pattern recognition to experience-driven sequential decision-making. It promises to carry AI applications forward toward taking actions in the real world" (Stone et al. 2016, p. 9).

[5] Source of image: https://commons.wikimedia.org/wiki/File:Typical_cnn.png?uselang=de, posted under the Wikimedia Commons, date: 12/16/2015

The different stages of AI that machines can develop can be seen in figure 5. The figure tries to explain the difference between Machine Learning, Machine Intelligence and Machine Consciousness conceptually. The illustrations on the left "try to show the evolution conceptually from single to multi-level hierarchy and finally self-sustained" (Paka, 2016). A machine that is able to do ML is one that handles data in a task-oriented way. A machine that is able to perform "Machine Intelligence" is one that learns from data and builds models on its own. The highest level is called "Machine Consciousness," which is the ability to teach oneself and through a sort of a conscious mind. As marked in the figure 5, as of 2016 technology has reached the state of Machine Intelligence but has not yet reached Machine Consciousness.

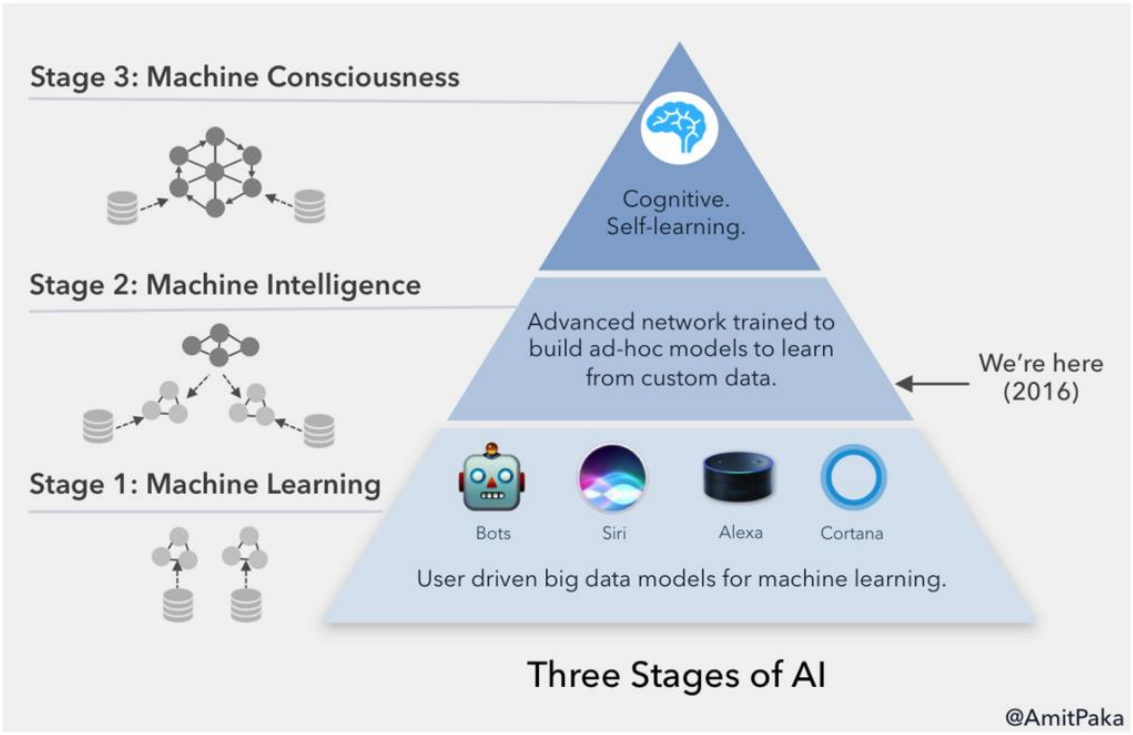

Figure 5: **Three stages of AI in 2016**[6]

3 Definition and distinction of commons

The following chapter will give an overview of the theory of the commons – including several definitions and distinctions between the common and public goods as well as pure

[6] Source of image: https://amitpaka.com/three-stages-of-ai-9d2df56dbd08, posted by Amit Paka, date: 09/30/2016

and impure public goods. Theories, like the "tragedy of the commons", and phenomena like "free-riding" will be explained. Finally, digital commons will be discussed.

3.1 Difference between commons and public goods

Before entering the sphere of the commons, or common goods, it is important to discuss the difference between a common good and a public good. James Quilligan (2012), an advocate for Global Common Goods and policy advisor, wrote that the distinction between these two forms of collective property are not well-defined. He explains that society understands the difference between public and private goods as berries sold at the market versus berries found in the wild. Expounding on that, private goods are commercial products or services created by businesses and public goods are things such as parks, roads, public safety or education. Public goods are ruled by the government and thus the rules of the government apply and the public must align themselves with these rules. Commons, on the other hand, are resources that are best maintained by the public. Quilligan (2012) explains:

> "One of the great challenges before us is to create powerful and broadly recognized distinctions between public goods and commons – the *shared resources which people manage by negotiating their own rules through social or customary traditions, norms and practices*. (…) For the commons to be embraced in economic, ecological and social policy, *their immediacy should be apparent to everyone.* The cognitive apprehension of common goods must quicken our capacity to experience and understand the *things we share beyond the enclosed spaces of private and public property.* Formal categories may help clarify distinctions among private, public and common goods, but they do not convey the sense of human meaning, being and intersubjectivity that lie at the heart of any commons."

Furthermore, he states that most Western governments adopted the theories of John Maynard Keynes who believed that the intervention of the government into individuals' consumption of goods was a way to control needs and satisfy people through more jobs and higher wages. In this theory, the human desire for common goods is non-existent. Instead, it is about personal preference satisfaction and correlates with the belief that goods – like food, water, knowledge, and social technologies – are market goods that are governed by the state but influenced and distributed by the market. James Quilligan (2012) summarizes:

> "In short, state provision of public goods *fails to account for the higher total net benefit that consumers would receive through self-organized and socially negotiated production, use and protection of their own resources*. Hence, the commons have no definitional reality in Keynesian thought."

Since the 1980's many states have increased the rights of private property, free markets, and free trade. Public goods, which are governed by the state, also imply hierarchies and top-down systems. However, it can be argued that decentralized self-governing systems of co-production seem to offer fairer and more direct access to resources compared to commercial delivery chains or the bureaucratic provision of goods. Decision-making processes and the production as well as distribution expands to a wider community. Direct access also enhances the efficiency (Quilligan 2012).

3.2 Pure and impure public goods

A public good, as well as a commons, should be predestined to function and be available for everyone without consequences. If a good is limited to only a certain group of people and a certain amount of usage, it would fall under the definition of a private good. A pure public good is characterized as non-rivalrous in consumption and non-excludable. Non-excludable means that no person can be excluded from the good – goods that exist outside of the marketplace. For example, "air" – no person is excluded from air or misses out on air if another person uses it. This concept extends to more abstract ideas as well, such as peace – if one person enjoys or uses it, another person will not be excluded. Additionally, the usage will not reduce it or let it disappear, which is why there is no need for rivalry. An example of a pure public good is education. Education is provided by the state and is free for everyone. Not only is it available to all, but it is compulsory - in most countries children must go to school until a certain age. Educating citizens brings many benefits, some of which, like the correlation between higher education level and higher income level, are more visible than others. Another benefit of education, which is often taken for granted, is that citizens can read and understand basic laws and other procedures.

Public goods are known for having ambiguous externalities, or side-products of actions. There are positive and negative externalities. For example, educating women can lead to positive effects on child survival, which would be considered a positive externality. An example for a negative externality would be the result of utilizing a river for releasing pollutants, which will deteriorate the environment. Both are public goods, the river as

well as education, but the examples of actions concerning each have different side-effects as outcomes (Kaul et al. 1999, pp. 5-6).

Impure public goods are missing at least one of the characteristics required for an entity to be considered a pure public good. Either the good is non-excludable but introduces rivalry, which is called a "common pool resource," or it is non-rivalrous but introduces excludability, which is called a "club good". Examples of impure public goods would be public transportation or national defense. They do not satisfy the criteria of a pure public good to all extents (Kaul et al. 1999, p. 6).

3.3 The tragedy of the commons

Despite that by definition, commons are the ideal state of communality and inclusiveness, it is hard to enforce them within society. Phenomena like the so-called "tragedy of the commons" is a well-established occurrence.

Aristotle (Politics. Book II, ch. 3) observed long ago that "what is common to the greatest number has the least care bestowed upon it. Everyone thinks chiefly of his own, hardly at all of the common interest." This behavior, where people use commons but care more about their individual well-being and their good instead of the common good, is known as the tragedy of the commons.

Garett Hardin (1968, pp. 1243-1248) wrote a journal article in 1968 about the tragedy of the commons, where he explained how common goods reach the point of exploitation because individuals take care of their individual good instead of the common good. He cites an example of a common pasture which is open to all. In this case, social stability is at a point where no wars or other conditions minimize the people or animals using the pasture. The example begins with shepherds that use the pasture for their sheep and can at least add one sheep each to their herd to maximize their gains. The advantage of an additional sheep is higher than the disadvantages, e.g. less forage and smaller grazing area, because the disadvantages are shared between all, but the advantages only apply to the individual. Eventually, the impact on the environment is quite high as the herds increase progressively. The pasture cannot renew itself as fast as the sheep are grazing the weed. Therefore, the tragedy lies in the pollution of sheep through the shepherds who do not take care of the pasture. Hardin calls this the "tragedy of freedom in a commons" (Garret Hardin 1968, p. 1244).

Common examples are about how the population is overgrown and is overusing resources – most of the time non-renewable ones – e.g. water depots. In these cases over-population is the main crisis which leads to the tragedy of the commons (Greco & Floridi 2004, p. 5).

3.4 Free riding

Free riding is the phenomena of people using the collective good without necessarily contributing to it, so that each user is able to free-ride on the efforts of others. And if everyone would decide to free-ride instead of supplying, then there would be no collective good (Ostrom 2015, pp. 6-7).

Hardin (1968, p. 1244) expressed the free-riding problem with the example of the shepherds sharing a common pasture and are "locked into a system that compels (each one) to increase his herd without limit," thus leading to overgrazing and land degradation. Olson (2015, p. 6) argues that groups need some benefit as a reward for their collective work, otherwise they would not be motivated. This means that they do not work out of internal gratification, but because of incentive.

In some cases, free-riding can be useful. This is true within the space of digital commons, e.g. operating systems like Linux, where the more people use it, the better the product (see chapter 3.5). Especially in the case of Linux, the more users make it to their default system, the easier it becomes for all users to compile and use programs or to transfer data from one system to another. The positive side effects of the product being freely used were bug reports and other suggested improvements, which has helped the developers to enhance the product and fix issues (Weber 2005, pp. 154-156). Therefore, free-riding does not have to be condemned but thought of while implementing commons.

3.5 Digital commons

When it comes to digital products, especially operating systems, software is always non-rival. There is usually no rivalry in consumption within digital resources because one person using it does not restrict another person from using it. This is especially relevant concerning open-source software because it can be copied without a maximum on the quantity (Weber 2005, p. 154). Even though digital products, e.g. the Internet or software, are free to everyone, one cannot make use out of it without the appropriate hardware. Digital resources always exclude individuals who cannot afford digital devices and the

necessary infrastructure if there are no public and free access points. As Bulatov (2016, p. 5) clarifies:

> "*Software is very cheap to replicate* – so if you can give it to one person, you can give it to the next person for almost free. There is the second issue of computational resources. Those ones are limited. (…) So, there is already a bit of an *inequality because of the access to computational resources*."

Consequently, hardware is afflicted with rivalry and exclusion. Similarly, language can also be a factor of exclusion. According to the Human Development Report of 2016 (UNDP 2016, p. 147), "individuals unable to read or write English are thus excluded from most of the opportunities for participation on the Internet." In the report of 2001 (UNDP 2001, p. 40), which was focused on new technologies and making them work for human development, it was found out that most internet users at that time were predominantly "urban and located in certain areas, better educated and wealthier, young, male." New statistics show that the number of women has increased, and in America has even exceed the number of male users (see figure 6). Nevertheless, a digital divide exists. Greco and Floridi (2004, p. 9) describe the dilemma in their paper on the "Tragedy of the Digital Commons" as "a discrimination between those who can be denizens of the infosphere and those who cannot, between insiders and outsiders, between information rich and information poor [people]."

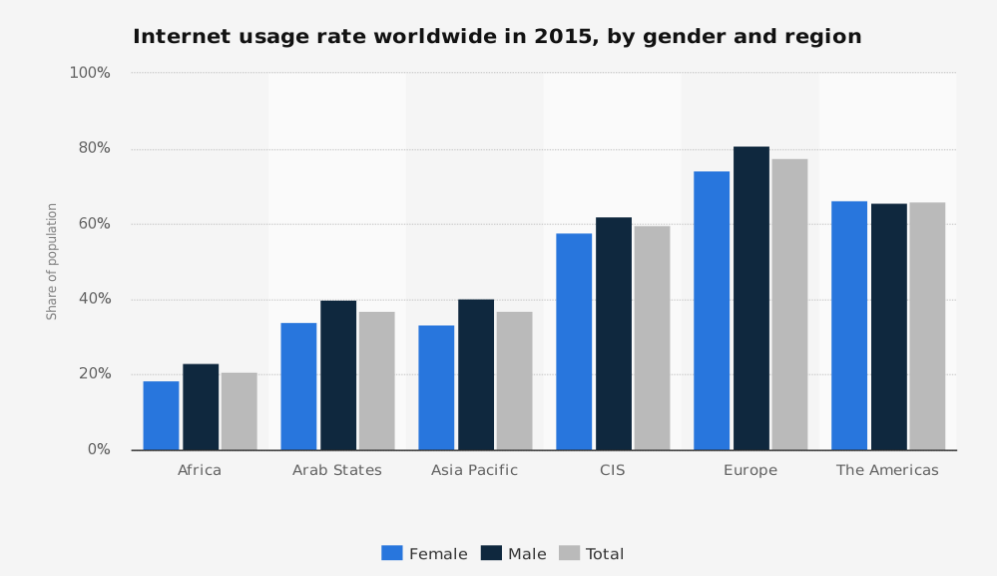

Figure 6: Internet usage worldwide in 2015, by gender and region[7]

[7] Source of image: https://www.statista.com/statistics/491387/gender-distribution-of-internet-users-region/, original source: Union internationale des télécommunications op. 2015, date: 2015

4 AI as a common property

This chapter tries to combine AI technology with the theory of the commons and to identify analogies with other commons. First, open-source software will be examined, then "Open AI". The survey which has been conducted through interview series will be presented. The legal placement of this theory and the current legal status of AI will be addressed. Finally, a short analysis of societal implications will be addressed as well as an assessment of the "free" market of AI.

4.1 Open-source and AI

Developers are often mistaken as people who use their expertise to solve a problem. Software can be a tool of expression and a product of human creativity (Weber 2005, p. 47). It is comparable to writing a novel or building a car – it is a messy process with multiple iterations. Code is always evolving. More than 12 million lines of code have been written for the sophisticated operating system Linux[8]. It is hard work and it would have been hard to achieve alone (Weber 2005, pp. 54-57). Developers are used to co-operating on tasks like quality assurance or coding complex systems. For larger projects, tasks are broken down and distributed between them. Typically, an IT project manager leads the process to make sure all parts get done and that the system functions well as a whole in the end. It seems unlikely to achieve the same outcome without any project managers, no clearly defined organization, and a team spread out around the globe with no monetary rewards as a benefit.

Steven Weber (2005, p. 9) wrote:

> "Intuition tells us that thousands of volunteers are unlikely to come together to collaborate on a complex economic project, sustain that collaboration over time, and build something that they give away freely, particularly something that can *beat some of the largest and richest business enterprises in the world at their own game*."

But projects like these have happened and have been successful. The next chapter will give an overview on open-source software and open-source software commons, and what kind of role AI plays within this domain.

[8] A comparison of different computer software programs and the number of lines of code visualized in an infographic can be found here: http://www.visualcapitalist.com/millions-lines-of-code/, posted by Jeff Desjardins, date: 09/24/2015

4.1.1 Open-source software

As the name indicates, open-source software (OSS) means the source of the software is open. It is the opposite of proprietary software, belongs to no one and empowers users to design their own system by using what is available and by altering it.

To use a metaphor: the soft drink Coca-Cola lists all ingredients on the back of their bottles, but does not hand out their recipe for free. No one is allowed to copy the recipe because if that were possible, competitors could produce, modify or improve upon it causing the company to lose its Universal Selling Point (USP). Because this recipe is as essential as nothing else, why should they hand it out for free? Coca-Cola would probably get bankrupt if they do so. Open-source software is different. OSS is a unique form of non-commercialization. It could cost something, but it remains free (Weber 2005, p. 3).

Open-source software is defined as "free" which does not necessarily mean that it does not cost. It means that the software code is free in regards to freedom. "Freedom includes the right to run the program for *any purpose*, to study how it works and adapt it to your own needs, to redistribute copies to others, and to improve the program and share your improvements with the community *so that all benefit*" (Weber 2005, p. 5). The benefit for the wider online community is a very important factor of the open-source thinking. Everyone who downloads open-source software should be able to use it for any purpose and it should help others by redistributing the source code and the documentation. The redistribution should also be made available under the same terms – therefore licensing fees are eliminated. This leads to the open-source software commons.

4.1.2 The open-source software commons

Owning something inherently means excluding non-owners. The open-source software commons are a set of rules on how to avoid ownership and retain the freedom of software.

Even though, OSS developers are not hostile to Intellectual Property Rights (IPR). They actually defend IPR because they do not just upload their work to the public domain. The Intellectual Property regime is codified in a series of licenses which is different than the mainstream IPR approach. The goal is to maximize the growth, use and further development of free software, and enable future generations. OSS licensing schemes generally try to create a social structure to empower users by ensuring access to source code, not reserving rights for the author only. It also constrains users from putting restrictions on other users. Creativity is hindered if developers cannot access the resources they need –

distribution of raw materials and tools is fundamental for the success of any software (Weber 2005, pp. 84-87). Within the OSS community, exclusive property seems like a "moral bad" as it encourages competition. Cooperation is an essential gift of OSS development. In the heart of every OSS project is the creative essence which must be fully supported. No procedure, better division of labor or new technology would be able to replace the creative essence to the project (Weber 2005, pp. 47-62).

Developers commonly refer to code as ugly or clean when they evaluate the structure. The style of the solution is important. Because their motivation is often to enhance their skills and get better, to have fun while coding, to have a community which cares about technology, and to develop better software, instead of money (Weber 2005, pp. 133-143). Communication plays a key part within the OSS culture because the developers depend on their network for complex code and questions which may arise. Developers in large corporations are not able to make use of these networks because they are often not allowed to talk about their codes (Weber 2005, p. 177).

These are basic principles of the OSS culture:

- access to computers should be unlimited
- information should be free
- mistrust authority and promote decentralization
- judge people only on the value of what they create, not who they are or what credentials they present
- people create art and beauty on computers
- computers can change human life for the better (Weber 2005, pp. 144-145)

Despite, the OSS community is organized in its own way – more of a conceptual integrity. In most cases, one "architect" of the system designs a "master plan" and shares tasks that need to be done through channels of the OSS community. There are teams and sub-teams with leaders who are in charge of the jobs who distribute smaller tasks and redistribute the finished jobs to the bigger project. Therefore, it is better to have small, reliable teams who constantly work on code. But there is no forced division of labor – it is all voluntary work. Developers are able to join as well as leave at any time and choose the tasks they want to work on or not (Weber 2005, pp. 47-62).

> "[It is] a community that *specifies its own norms and values* in the digital world. [And also] a *political movement* aimed at replacing obsolete nineteenth-century

> capitalist structures with new 'relations of production' more suited to the Information Age." (Weber 2005, p. 7)

This new work model is being used in the domain of AI as well. Developers started building their own artificially intelligent machines from home in cooperation with teams spread around the globe. As Bostrom (2016, p. 4) explains:

> "In the open source software sector, significant contributions are made by individuals who are volunteering their own free time. One motive for such contributions is that they *enable a programmer to demonstrate skill, which may raise his or her market value.* Such a skill-signaling motive appears to be a strong influence among many AI researchers."

From this model, the "Open AI Movement" was born.

4.1.3 The "Open AI Movement"

In December 2015, the OpenAI Non-Profit Artificial General Intelligence Research Company opened in San Francisco (U.S.). The company has been founded by well-known AI start-up investors. Elon Musk, Sam Altman, Peter Thiel, Reid Hoffmann, Jessica Livingston and Greg Brockman are the lead investors. Furthermore, Microsoft, Amazon Web Services, Infosys, Yresearch and the Open Philanthropy Project are referred on the website as sponsors. Besides that, 24 companies donated access to their software and/or their user data. The OpenAI office hired around 60 researchers and engineers (May 2017) for exploring AI capabilities without depending on financial return. The company publishes their research, represented in papers and blog posts, on their website as well as systems, tools, and AI platforms on GitHub. GitHub is a non-profit software development platform where developers can host and review code and which makes it easy to share code with others (GitHub 2017). The progression that the OpenAI organization has made since 2015 is striking. Some of this progression is supported by the active engagement with the OSS community – on GitHub and other platforms. They are practicing long-term research and their mission is to influence the future of AGI by developing and researching it themselves (OpenAI 2017).

In November 2015, shortly before the time when OpenAI introduced itself, the Google Brain team published their open-source software library for Machine Intelligence called "TensorFlow". TensorFlow is a toolkit developed by researchers and engineers of the

Google Brain team within Google's Machine Intelligence research organization for sharing ML source code, Deep Neural Networks and other research advancements internally. According to their website, they released TensorFlow because it is applicable to various other domains, too. The XLA, which stands for Accelerated Linear Algebra, of the Google Brain team has been released in March 2017 together with the new version TensorFlow 1.0. The XLA should improve the performance by decreasing memory usage and execution speed, improving portability to other hardware platforms and reducing the mobile footprint by changing magnitudes. But for the development of the XLA, the Google Brain team released the alpha-version, which is the first and not completed version, of this tool. They encourage the OSS community to take part in the development process (Google Brain 2017). Both toolkits can be found on GitHub and can be contributed to, modified, and used for any purpose.

Besides the fact that the Google Brain team gains knowledge and an improved software library by these contributions, e.g. bug adjustments or source code amendments, Moore (2016, p. 6) states:

> "(…) specifically, for TensorFlow that is because they think it is to their advantage to build good tools for themselves, but it is also to their advantage to get everyone else using those tools, so then it is *easy to hire people* that already know how to use Google tools."

Some of the other AI giants opened up their research as well as software, such as Facebook, Microsoft, and IBM (Bostrom 2016, p. 2). The company Apple started publishing their internal AI research as of December 2016 after being criticized for not being open, for example by Facebook's AI Director Yann LeCun (Shead 2016). Before Apple decided to publish their work, interviews with experts presented an idea of why they found the motivation to do so. Moore (2016, p. 6) explained that "one reason why companies like Apple have a hard time recruiting talented people [is] because they do not let them publish papers."

The question is which motivations and incentives dominated these decisions. The answer to this question can partially be commented by the results of the following qualitative interviews.

4.2 Interviews about artificial intelligence as a common property

I first became aware of the debate around ethics and AI in my Information Management studies. During my internship at a creative strategy agency, I learned about a problem the team faced with AI Application Programming Interfaces (APIs). APIs are toolsets for building application software and provide components which make programming new software easier. The agency used an AI API to analyze pictures from social media for a project. The project team was concerned that the API publisher might not be stable enough to support the company, to maintain the service and to provide free access long-term. Therefore, open-source AI became an interesting starting point for this thesis.

From October 16th to December 22nd 2016, I travelled to San Francisco to do research on the "Open AI Movement". I interviewed six experts on the possibility of AI as a commons and attended talks, trade fairs and other events linked to AI.

The following chapter will explain the initial situation, the interview technique, how the results of the interviews were analyzed, and a short reflection on the procedure.

4.2.1 Statement of the problem

Discussions about the access of AI showed that APIs are not counted as real openness and that only free software which is modifiable and stays open can be used as a basis for any business. If a business decides to use AI for their own needs APIs can provide the necessary requirements but not long-term because of eventual lock-ins. Open-source software on the other hand can be used for long-term business architectures (Weber 2005, pp. 74-75 & 227-231).

API.AI is an online platform for Natural Language Processing API's which can be integrated in applications, devices, services and robots. The following extract is written in the Terms of Use and Privacy Policy under the User License Agreement of the website API.AI to demonstrate the difference between APIs and open-source licensed software:

> "The licenses granted herein are only valid if:
>
> (i) the Services are NOT modified in any manner;
> (ii) all copyright and proprietary notices or labels in the Services are maintained in their original format;
> (iii) the Services are only installed and used in accordance with your network security policies,

(iv) you possess the necessary authority and power to install the Services,

(v) you promptly pay all license fees when due,

(vi) this Agreement is accepted without any modification, and

(vii) you have not breached the terms of this Agreement in any manner" (API.AI 2017)

Discussions I had also showed that the motives behind a developer's decision on how they want to publish their software was interesting enough to warrant further research.

Given the background of AI as an existential risk, "Open AI" could seem like an amplifier of this risk by opening the sources to the public as many people still do not know or understand fully how AI is developed or what it is capable of achieving. Even developers have hard times understanding the full range of capabilities of AI, simply because they have not been explored yet (Bostrom 2016, pp. 9-12). Also, that AI-powered companies own these systems and the rights for it is not being treated as an important topic for public discussions. At the end, new milestones are achieved by the expertise of developers. No matter how many talks or discussions are held about AI as an existential risk, the question is if developers can build these very powerful systems (Bostrom 2016, p. 11; Hibbard 2008, p. 3). It is paramount that these developers need to be asked about the future of AI.

Although there are arguments not to publish under OSS licenses, the thought of opening as a requirement was interesting to me. What would happen if everyone gained access to AI? If there are no monopolies and no owners, who would own it and who would regulate it? Maybe some inequality issues could be solved. Maybe AI would not just be a tool of the privileged. Maybe there is a democratic way to regulate AI. A more detailed examination of these questions can be found in chapter 4.4.

A lot of AI developers and companies publish their code and documentation under open-source software commons (Bostrom 2016, p. 2). They share it and allow others to contribute, to modify and to distribute it into the indefinite future. As such, I had three main questions I wanted to understand from experts: What is the motivation of the developers to open-source their codes? Should some institution/government/person regulate, own and rule over AI? Should AI be declared as a commons?

4.2.2 Methodology and target group

In order to test the underlying research question and to capture subjective views of experts[9] and their real-life context, I collected data through semi-structured and guideline-based interviews with industry experts in San Francisco, USA. A conscious decision was made not to conduct a highly structured survey, instead to conduct semi-structured interviews as qualitative interviewing has greater flexibility and allows the opportunity to capture the industry expert's perspective on the subject of study in greater detail while at the same time allowing for quantitative analysis of the interview responses (Leech 2002, p. 665). The overarching aim of the expert interviews was to collect data on (i) developers' motivation to develop open-source AI and (ii) their perspective on AI as a commons. In order to guarantee a certain degree of comparability between the different interviews, I conducted them on the basis of an interview guide with a number of open-ended questions depending on interesting aspects during the conversation. The interview design was the result of comprehensive desk research: prior to the interviews, I analyzed the websites of AI companies, conducted a market overview, assessed reports of experts and conducted personal conversations. The interview guide can be found in the Appendix of this thesis.

The interview partners were selected based on the key informant technique in order to attain expert knowledge, since key informants are characterized as individuals possessing meaningful knowledge and are willing to communicate with the researcher (Marshall 1996, pp. 92-96). The target group had mixture of diverse fields and business backgrounds whose current work is related to open-source AI as well as experts who were not new to the AI space and had insights about recent business developments. Six AI experts were interviewed. Most studied Computer Science or related fields.

That was particularly helpful in assessing the first part of the research question about the incentives and motives of involved actors. The high response rate underlines the importance and relevance of the topic. The initial contact mode was an interview request via e-mail with attached information about the purpose of the study. As part of the sampling technique, I made use of snowball sampling asking at the end of each interview: "Is there anyone else in the field of AI you know or can think of who I should definitely speak with?" Though controlling for this, I was not referred to any new interviewee which was

[9] An expert in this study is defined as a person who has specialized knowledge, experience in the field of open-source AI and/or has privileged access to information about decision processes regarding open-source AI.

not already part of the initial population, thus indicating that a representative sample had been chosen. The interviews were conducted in English with predetermined open-ended questions. Each interview consisted of about 12 questions and had a duration of about 30 minutes. Two interviews have been conducted via telephone. The interviews were recorded, transcribed and analyzed. The interview questions can be found in the Appendix of the thesis. In general, the interviews were structured along the following four question blocks: (1) personal motivation, (2) open-source AI movement, (3) AI as a commons and (4) the future of AI. A tabular representation of the interview partners can be found in Table 1.

Given the possibility that some of the experts came from multiple disciplines before entering the field of open-source AI and/or general AI research, questions about their motivation and their journey were asked first. Questions about the reasons developers chose to make their work open-source in regard to the OpenAI company, were of interest. Crowdsourcing, the accessibility of data for AI researchers and developers, the ownership of powerful AI systems, and regulation of AI were also discussed. Furthermore, the accessibility of AI systems, the contingency of AI as a commons and the future of AI were talked about at the end of the interview.

Table 1: List of expert interview profiles

Name of interviewee (last, first) and current position	Date and location of interview	Biography
Gourley, Sean *CEO of augmented intelligence company, Physicist*	12/13/2016 *San Francisco, California, USA*	Sean Gourley is a physicist, entrepreneur, and TED fellow. He is originally from New Zealand where he ran for national elected office and helped start New Zealand's first nanotech company. Gourley studied at Oxford as a Rhodes Scholar where he received a PhD for his research on the mathematical patterns that underlie modern war. This research has taken him all over the world from the Pentagon, to the United Nations and Iraq. Previously Gourley worked at NASA on self-repairing Nanocircuits. In 2009 Sean founded Quid where he served as the CTO until 2015. He is based in San Francisco where he is the Founder and CEO of Primer, a machine intelligence company.
Hibbard, Bill *Emeritus Senior Scientist, AI Research Associate*	10/31/2016 *Phone Interview (San Francisco, California and Madison, Wisconsin, USA)*	Bill Hibbard is an Emeritus Senior Scientist at the University of Wisconsin-Madison Space Science and Engineering Center (SSEC) currently working on transparency for Artificial Intelligence and as a Research Associate for the Machine Intelligence Research Institute. He published a lot of his work, for instance his books "Ethical Artificial Intelligence" and "Super-Intelligent Machines", among other AI related papers and talks. Prior, he was involved in the Machine Intelligence Project of the SSEC which focuses on how to create machine

		intelligence and on the social consequences. Before that he spent most of his career working on the SSEC Visualization Project and his open-source visualization systems Vis5D, Cave5D and VisAD. He has a BA in Mathematics and MS and PhD in Computer Science.
Song, William *Co-founder of AI start-up, Software Engineer*	10/01/2017 *Phone Interview* *(Berlin, Germany* *and* *San Francisco,* *California, USA)*	William Song is a Co-founder of the start-up drive.ai where he is working as a Software Engineer and applies deep learning to self-driving cars. Before that he worked as a Software Engineering Intern at Google where he helped to improve, amongst other things, the perception for self-driving cars, the request simulation for Maps Engine and the Street View object detection using deep learning. He also worked as an Intern at Facebook, Tagged and Blue Coast Systems. He has a BA (Honors) from the University of Waterloo in Software Engineering and a MS from the Stanford University in Computer Science.
Lai, Morgan *Founder of stealth mode AI-powered start-up, AI Scientist*	11/23/2016 *San Francisco,* *California, USA*	Morgan Lai is currently in the stealth mode of an AI-powered start-up. Before, she worked as an AI Scientist/Software Engineer at Loop AI for 8 months. She also worked at Flux Factory, ShopKeep and XL Hybrids. She was the CTO and Co-Founder of Zelda&Scotch. After her studies, she worked as a Software Engineer for the MIT. She has a BS in Computer Science and Electrical Engineering and a MA in Computer Science and Artificial Intelligence from the MIT.

Moore, David *PhD student in Computer Science*	12/13/2016 *Berkeley, California, USA*	David Moore is a PhD student in Computer Science at the University of California, Berkeley, and advised by Stuart Russell. His primary current project is the application of Bayesian inference to nuclear weapons test monitoring – given seismic waveforms from a global network of stations, they want to infer a set of seismic events that plausibly explains the observed signals. Portions of this work have been funded by the CTBTO and DTRA. Before coming to Berkeley, he was an undergraduate student at Williams College, where he majored in Computer Science and Mathematics.
Bulatov, Yaroslav *Research Software Engineer at OpenAI*	10/27/2016 *San Francisco, California, USA*	Yaroslav Bulatov is currently working as a Research Software Engineer at OpenAI since September 2016. Previously, he worked as a Software Engineer at Google Brain for 2.4 years, building the open-source software library for machine intelligence called TensorFlow. He was also a part of the Google Street View team for 3 years improving Optical Character Recognition (OCR) from Street View imagery using deep convolutional neural networks. Previously, he has worked at Strands for almost 2 years to improve OCR from videos and at ViewPlus Technologies for 2 years which develop software for the blind. He graduated in Computer Science at the Oregon State University in 2006.

4.2.3 Data analysis

Data analysis consists of two steps. The first involves the transcription of the recorded interviews in order to map and order key themes and to find patterns of similarities and differences (Breuer 2010, pp. 40-43). I did not use a coding strategy for the interviews since partial transcripts already assess what is relevant and stands out. However, in the analysis, categories were formed in Microsoft Excel and the existing data classified in a structured way to interpret the findings – i.e. find patterns of similarities and disparities, thereby generating new knowledge. To ensure sufficient rigor of the analysis, the answers were summarized for each question and displayed beneath with underlining statements of the experts.

In the second step of the analysis, the findings of the interviews were compared with the findings of existing studies in the field (triangulation). This cross-checking enhances reliability and validity of the data gathered in the thesis. These results can be found in chapter 5 – the "free" market of AI.

Of course, analysis relies to some extent on the interpretation of the author which is an unavoidable limitation. As Gummesson (2012, p. 33) framed it:

> "When studying complex and ambiguous phenomena, like the ones that are studied in marketing often are, *intuition plays a vital part* and is in fact required as it often is impossible to know exactly how to process data and arrive at conclusions."

4.2.4 Results and interpretation

The very first question of the interview was about the expert's current position and current projects they are working on. The answers to this question have been summarized in the Table 1 in chapter 4.2.2. The results of each question will be summarized and underlined by direct quotations of the expert's answers. The interpretation follows at the end. A detailed cross-checking of the results can be found in chapter 5 – the "free" market of AI.

Question 1: Why did you engage in open-source AI?

The motivation of the developers was quite different but they all shared a fascination for the idea of AI and the science behind it. For some of the experts it is important to understand human intelligence and the world better – by amplifying human intelligence or by building machines which will be sophisticated. Some were basically just fans of open-

source development and the unpredictability of AI development. Most were against monopolization from big companies like Google or Facebook.

Question 2: Why do you think some developers open-source their AI systems, for instance the OpenAI office? Could you see benefits in it?

It was generally agreed that open-sourcing codes helped people advance their careers because it gave them an advantage if they were able to present their work and show successful research and software which they created. That is why leading AI companies let their employees publish papers and even let them publish the codes, e.g. the program TensorFlow. Another major benefit is the massive online community for developers where they commonly contribute and share code snippets to improve systems. But the participation grows the bigger the project is and the more attention it gets. Experts mentioned that developers also love developing social status within these online communities. As Moore (2016, p. 6) concludes:

> "That is one reason why companies like Apple have a hard time recruiting talented people because they do not let them publish papers. Open-source is also a big part of that where people feel much more fulfilled by their work if you are able to put it out there and see it contributing to a larger body of sort of accumulated tools and accumulated code that other people can use. *It is definitely a hiring advantage."*

Lai (2016, p. 5) states:

> "(…) with TensorFlow – yes, it is great – but it *has to be trained* to become a useful model and if you do not have the training data to train the TensorFlow, then you will not be able to have a useful model. I mean it would be really awesome for more technology companies like Apple, Google or Facebook to kind of like *start to really outsource all this AI technology*. But there is still a very fundamental barrier that other competitors need to overcome in order to make use of all this technology."

Question 3: What do you think about the accessibility of data for your research?

There are many open-source datasets available online. The real problem for the experts is that the data is often not as valuable as it might seem. Good training data which can be used for models, is not as widely available. However, it depends on the project. Another issue is money for early stage start-ups. These start-ups have big data clusters available but are not able to compete in any way with the data handling processes at Google. Google owns various fleets of self-driving cars which are collecting and interpreting data every

day. The company also has teams in India which are categorizing and tagging raw data like images and videos all day long to produce useful training data. These processes need a lot of money and therefore only wealthy AI enabler are capable of providing these sets. Reinforcement learning simulators are available on a larger scale and as a result researching in this area easier. The problem with these companies having a very special infrastructure and massive resources is that thus they have a recruitment advantage too. Developers are able to work with huge amounts of data compared to the open source datasets on the internet. A lot of AI developers are working inside big global AI players such as Google or Facebook. They have a lot of opportunities and advantages working for these companies. Even so, some developers prefer to stay in academia, while others join or found start-ups or simply enjoy developing AI as a hobby.

Question 4: To what extent does crowdsourcing help to solve specific problems in the field of AI?

Crowdsourcing seems to be a strategy which only bigger companies or academic laboratories can use. People need incentives to take part in the labeling process and thus the publisher makes use of gamification or prizes as a reward for the work. As Bulatov (2016, p. 2) clarifies:

> "There is this dataset called ImageNet and because it came out of this public competition where a lot of people worked on it – it now *exceeds human accuracy*. So, I think this is really the way to get results. If you have a benchmark that people get excited about, you put it out there and eventually you get the human accuracy just by people slowly improving it. And if you make a benchmark that is sufficiently complicated, hopefully, when people solve that benchmark – they will *also solve AI*. So that is another thing we are interested in."

Question 5: Can you picture a national copyright online library for AI systems which anyone can access?

The answers to this question were divided. On one hand, experts think that it would not be practical because code is always evolving. Each day there are 40 changes made to the code base of TensorFlow. Additionally, code is not useful after some time – it is pointless to freeze code to one specific moment in time because of its evolving character. On the other hand, transparency could be enforced through this idea. There are already platforms which function as libraries for open-source code and for scientific papers. For software,

developers use GitHub which is a private company. For scientific papers, researchers publish and search the website arXiv.org which is a non-profit. But definitely, old code is not in the interest of any developer and they would appreciate a platform where everything comes together – scientific papers, the codes and documentations – and is not privately owned and can handle evolving content.

Question 6: Do you think someone can or should own an AI system which influences humans at a large scale?

The experts agreed that a form of intellectual property applies to AI in this case. Additionally, companies or individuals who developed an AI could own it as well. However, they raised concerns about wealth distribution and regulation, specifically, that after the automation of human tasks, jobs will be replaced and the wealth originated by the algorithm will most likely not be distributed fairly. Inequality increases as employees of big AI enterprises, as well as individuals, train an algorithm which will replace them or treats them as products to increase the gains of the enterprise's owners. As Gourley (2016, pp. 13-14) points out:

> "You know the Uber drivers. They are driving around, picking up data from their cars, they get paid a little less than a taxi driver, right? And they are doing all that work to train an algorithm that will be the heart of self-driving cars that will replace them that Uber will own. They are *techno[logy] servants*. (…) And let us step down a thing – *at least on Facebook we have been treated as products*. But if you are training algorithms – you are working for them."

Some think that there will be no ownership in the future because of the open-source software movement and the fact that more and more scientists openly publish their work. Others think that the concept of ownership must be redefined as taking over responsibility for the things someone produces, distributes and/or is making available to others.

Question 7: Could you also imagine that in the near future it would be beneath human dignity if someone does not own a virtual assistant or another kind of AI?

Experts pointed out that software is easy to replicate and therefore this scenario is unlikely. Most think that AI will increase inequality. Hibbard (2016, p. 4) summarizes:

> "Technology gives an advantage to those who are already wealthy and powerful. And that will accelerate. One of the things I say in my talks quite often is that *human brains are distributed according to biology*. We do not have any control over the

distribution of the human brain. So, every human is born with a brain which is about as good as the brain of other humans. I mean there is obviously variation in IQ and math ability and verbal ability. But we all have roughly the same intelligence because intelligence is meted out by a natural process. But when intelligence can be created by an artificial process – AI – then we can expect the distribution of intelligence to follow a curve much more *like the distribution of other artificial things* like buildings, ships or computers. Which all have, what you might call a power law distribution, where there are a few very large ones. If you look at buildings, there are a very few large buildings in the world and then there is a very large number of much smaller buildings. *And we can expect the same distribution of intelligence when it becomes an artefact.* There will be few very large brains and a lot more ordinary brains like the brains that you and I have. So, inequality will be hard to escape. I think that any effort to enforce equality of intelligence is going to be tough. And *along with intelligence goes income*. There is this positive feedback cycle of the more intelligent you are, the more money you can get. And the more money you have, the larger brain you can afford. Once brains become an artefact then the size of the brain you have depends on how much money you have. And then how much money you have depends on how smart you are."

Question 8: Do you think an institution, no matter which one, is able or should be enabled to regulate AI or do you think not?

The opinions on regulation were divided, especially in regards to trust in the government. Some experts believe that AI is not mature enough to start thinking about how to regulate the software. Some believed that it definitely had to fall on the government to regulate, but depended on the kind of algorithm which was being used and inspected. Licensing schemes similar to existing ones could be developed. Again, the distribution of money and power was an important regulation issue. According to Gourley (2016, p. 12):

"(…) as a society if you think very, very carefully about the *redistribution of wealth and gains* that come from these algorithms because these algorithms ultimately come from science which is a production of everything we as a society have created. So, you have to acknowledge that you got there because you benefitted from society which means you *need to find ways to give back*."

Question 9: Could you think of AI as a commons to which everyone gains a right?

On one hand, experts believe that it is hard to push equal rights and access by law enforcement. According to the experts, people are the key of politics. For example, transparency and openness at Google got pushed by the scientific researchers working there. The employees demanded the right to publish their work and received support because Google would have had difficulties finding equally skilled new hires. Furthermore, a lot of responsibility lies in the hands of the developers of AI. Hibbard (2016, p. 3) considers:

> "(…) the developers of AI are really acting as representatives for the future of humanity. That all future generations of human beings are going to live in a world that is largely created and determined by AI. The way that we create *AI now will create the circumstances of human life into the indefinite future.* (…) Where you cannot escape it. (…) So, AI, whether some people want to admit or not, is a sort of a commons and *we should have it ruled as a commons*."

Experts hope that companies set aside competitive behavior and co-operate. Research results are only withheld if the companies stand in direct competition with another company, e.g. object recognition software is important for self-driving car technology or image recognition on social media platforms. The ultimate goal then should be to work towards the public good and to share the wealth with everyone. Technology alone is not mature enough to consider the distribution of wealth. Gourley (2016, pp. 12-13) concludes:

> "It is not whether it is given, it is what use you can make of it. And one of the things we mistake is that if something is freely available we should all have equal opportunities to build something with it. But the reality is things which are freely available like high frequency trading data on financial markets – what would you do with that? (…) If you got a lot of money to trade, it might be very valuable. (…) But what good is it to me if I have my job at Home Depot for $7.25 bucks an hour selling wallpaper? What am I going to do with that? *We are in this kind of fallacy like the freedom of things means that we will freely benefit from it.*"

Question 10: How do you envision the future of AI?

There are extremely pessimistic and optimistic views on the future of AI. All experts are sure about the fact that AI and robotics will replace human labor. Humans that train algorithms will probably be replaced afterwards without getting any benefit from it. Bulatov (2016, p. 7) reasons that "it will definitely enrich the society monetarily because a lot

of the work that people are doing can be done by robots. So, humans will have more time and more resources than they want." Other experts, like Hibbard, are more concerned about an upcoming social unrest. AI will be integrated into people's everyday life and children will be used to talk to robots and use other speech interfaces at an early age, e.g. Apple's Siri or Amazon's Alexa. The technological knowledge will accelerate within society – the development, the usage and the application will be integrated into education (Hibbard 2016, pp. 6-7).

According to speculations of Gourley, an expert in calculating war and served the U.S. Pentagon as a consultant, in the next five to ten years there will be the first AI battles. Countries will fight each other by using AI technology, e.g. drones, and eventually without any human soldiers. Additionally, Gourley thinks that the definition of the word "expert" will change because it will be extremely complicated to distinguish human-machine collaboration with solely human intelligence. The pairing with artificial intelligence will be the norm. Furthermore, people will be pacified by algorithms which adapt to their needs. As an example, Gourley (2016, pp. 15-16) names the Facebook News Feed which adjusts to personal preferences and makes people content but keeps their attention in a filter bubble.

4.2.5 Reflection on the procedure

Overall, I am very satisfied with the interviews and their outcome. The experts were drawn to share insights and interested in discussing the topic. Questions about speculative future scenarios cannot be answered in a rational manner and therefore often left confusion with the experts. The questions were important to understand the future of open-source AI developments but more recent developments could have been discussed in detail instead. The outcome of the interviews reflects most papers on open-source AI and recent articles about the "Open AI Movement".

4.3 Opportunities and challenges for society

4.3.1 Equality through access

Water is a commons provided by nature and a necessity to human life. But water is about to process from a freely available resource to a highly priced commodity. The discussion about water as a non- or excludable good is therefore interesting because water has always

been a commons. Often the management of access to water by the government is still not equitable nor efficient but does not lack availability of water (Watkins 2006, p. 2). But since the commercialization of water, especially from high developed countries where a lot of times the demand is higher than the availability, this fact changes. Pumping and selling water is only allowed if one owns the land where the water was found. That is why big companies like Nestlé buy land to make the water resources on that land their property to sell it again in plastic bottles. The result of bottled water merchandising is less freely available water supply and access - therefore a water crisis because water is a life-giving resource (Schnell, Gehriger 2013). What mostly though changed was the way humans saw water. In a lot of countries, e.g. India, water is sacred. Shiva (2016, p. 138) points out the route of this water crisis this way:

> "The water crisis results from an erroneous equations of value with monetary price. However, *resources can often have very high value while having no price*. Diverse cultures have different value systems through which the ethical, ecological, and economic behavior of society is guided and shaped. Similarly, the idea that life is sacred puts a high value on living systems and prevents their commodification. Protection of vital resources cannot be ensured through market logic alone. It demands a recovery of the sacred and a recovery of the commons."

Considering digital products, products created by humans instead of nature, the access is also still a problem. As Weber (2005, p. 248) explains the digital divide:

> "The digital divide between developed and developing countries is now a central feature of international politics and the global economy. The slogan captures *a fundamental disparity in access to and the ability to use new technologies*, a reflection of long-standing divides of poverty, education, and freedom to make choices."

What seems inadequate is how technology companies treat developing countries. On one hand, companies depend on the cheap labor from countries like India or Malaysia because they could not find comparable low-paid workers for data entry, tagging or simple programming where their headquarters are based. On the other hand, they use these resources, build up on them, apply patents and prices, and then re-distribute highly priced software products to these developing markets. That is also why people from developing countries demand OSS in government infrastructures (Weber 2005, pp. 247-251).

> "From an efficiency perspective, the possible *exclusion of four billion people from the next era of wealth creation makes no sense*. From an ethical standpoint, it is

> even more problematic than was the exclusion of previous eras, because there is no intrinsic environmental or resource-based reason" (Weber 2005, p. 249).

OSS though shifts the decision-making prerogative into the hands of people in the developing countries and enables as well as empowers them to build their own products. Facilitating the accessibility of computational resources and the needed digital infrastructure, could allow a wealth transfer to poor countries but also a surplus of inexpensive low-value IT labor. But users will find out with increasing familiarity what new technologies are good for and what added value they get out of them. And these novel possibilities will open up a new market in many places in the world. "The empowerment that comes with free access to source code is not then simply a gratifying emotional experience; it is a necessary economic prerequisite of evolving demand" emphasizes Weber (2005, p. 252).

4.3.2 Innovation through openness

Innovation often does not derive from new inventions but from perfectly tailored products which fit the users' needs (Norman 2013, pp. 295-297). Especially if it comes to open-source software, there happens to be user-driven innovation. Because of self-made software solutions which are not only made for users but also made by future users. These self-made solutions from the user are often meeting the needs of other users. It usually fits better than when so-called outsiders try to understand the users' needs and design products for them because they cannot look inside their heads and understand their thoughts and actual needs as easy (Weber 2005, pp. 138-139). Also, the needs of minorities and people from developing countries can be incorporated into OSS easily. Because software solutions can come from everywhere, not only from developed countries, and particularly from emerging markets where the demand for innovation is high (Weber 2005, p. 252).

Most of the times, progress can only be reached if people commit to work on making progress and there has to be a motivation. The ILSVRC is one example of how innovation got fostered until even exceeding human accuracy. The winner of the competition of 2014, the GoogLeNet team (Szegedy 2015), wrote on its website:

> "The team participated with an open submission, meaning that the exact details of its approach are shared with the wider computer vision community to foster collaboration and accelerate progress in the field."

That means that the submitted codes are accessible to anyone on the website of the ILSVRC and can be used in R&D. Innovation can therefore be amplified by approaches being created by their own users and being shared with others – signifying through openness.

Also, Norman (2013, p. 279) explains in his book "The Design of Everyday Things" that the most successful and powerful form of innovation is incremental and small instead of radical bringing along major changes. As follows, open AI development, which builds upon the work of others, each person contributing small changes, could have a big chance of bringing forth innovation. And openness is an important factor like in the development of YouTube. Freely accessible online videos with educational or just entertaining content – a platform which innovated the online video landscape incrementally. As well as e-books and online papers innovated the sphere of publishing, especially self-publishing (Norman 2013, p. 296). Open freely accessible information developed voluntarily by individuals. "Open AI" could have a similar educational and empowering effect.

4.3.3 Sincerity through transparency

Bill Hibbard (2008, p. 2) makes an important argument in his short paper called "Open Source AI" which summarizes the workshop on the "Sociocultural, Ethical and Futurological Implications of Artificial General Intelligence" at the first conference on Artificial General Intelligence in 2008 (AGI-08). He stated out that with the introduction of transparency to the development of AI, corruption could be prevented. He argues that we only were able to reach such a high expertise in AI because of specialization. And that the people who specialized in a particular area are working as the representatives of this area. These people are usually called experts within western societies and unknowingly people depend on their advice and decisions from time to time. Hibbard calls them "agents" and gives an example of political experts who got elected by people which they should represent. If these agents work only in their own interest, then corruption is a common conclusion. To prevent this kind of corruption, it is indispensable to make decisions transparent. Hibbard (2008, p. 2) concludes:

> "Protecting the interests of humanity will require that law and public opinion change to *recognize that AI designers and managers are humanity's agents*. Once this agent relation is recognized, preventing corruption will require transparency, which includes an open source design for AI."

Also, transparency means providing citizens with information to equip them with knowledge to participate in debates and to make informed choices (see chapter 4.4.3.). Halperin and Hoffman (1976, p. 132) state out that "[a] healthy democracy requires public participation in the formulation and administration of government policy. Full information about policy proposals and debate over alternatives ensure popular consent."

4.4 Legal placement of AI

Elinor Ostrom (2015, p. 1) states that neither the state nor the market were successful in governing resource systems or in enabling individuals to sustain resource systems long-term. In AI, private people already started self-governed AI projects with little regulation. While leading AI companies still have the monopoly of rights over software and data, especially proprietary rights, in some countries governments have started to regulate the monopoly and create laws which strongly affect companies' freedoms.

In April 2016, the European Parliament (EP) adopted the General Data Protection Regulation (GDPR) and added new laws. These laws will take effect in 2018. Some will put pressure on companies and their handling of data while others will empower users to request explanations or object to automated processes.

In September 2016, the United Kingdom's House of Commons Science and Technology Committee issued a report on Robotics and AI that demanded a general digital strategy on behalf of the government and legal regulations within the Robotics and AI space. The report concludes that it is too soon to manifest any law but that a standing interdisciplinary committee should form to consider any present or future development of AI - and its ethical, social and legal implications. The committee should also work together with other parties, such as the Council of Data Ethics. The objective of the House of Commons was start discussions to prepare the way for manifestations, even if these might hinder progression in the field (Science and Technology Committee 2016).

Shortly after that report in January 2017, a press release from the EP got published, where it calls for European Union-wide rules for Robotics and AI. In which is written that:

> "MEP's [(Members of the European Parliament)] urge the Commission to consider creating a European agency for robotics and artificial intelligence to supply public authorities with technical, ethical and regulatory expertise. They also propose a voluntary ethical conduct code to regulate who would be accountable for the social,

environmental and human health impacts of robotics and ensure that they operate in accordance with legal, safety and ethical standards" (European Parliament 2017).

In June 2017, the Ethics Commission on Automated and Connected Driving released a report. Initiated on the 30th of September 2016 by the Federal Ministry of Transport and Digital Infrastructure of Germany, the committee examined and suggested ethical regulations for automated vehicles and traffic (Di Fabio, Udo et. al 2017).

It is noteworthy that different committees arise around AI regulation. While the right approach of legislating AI is controversial and diverse, it will be tried to use existing laws[10] around the topics of property rights, informational rights, and personality rights to challenge the aforementioned monopoles of rights concomitant with non-transparency.

4.4.1 Basic human rights and AI

Some of the basic human rights, which got manifested worldwide within the Universal Declaration of Human Rights (UDHR), can play a role in the considerations of an AI bill. The following articles are listed according to relevance from the viewpoint of the author.

For example, the described right to education in article 26. In paragraph 1, it manifests that elementary education, which is seen as fundamental education, shall be free and compulsory for any human. And "higher education shall be equally accessible for all on the basis of merit" (The United Nations 1948, art. 26, 1). Furthermore, in paragraph 2, the UDHR states that education in their sense is contributing to peace in form of reinforcement of "respect for human rights and fundamental freedoms" as well as "understanding, tolerance and friendship among all nations, racial or religious groups, and shall further the activities of the United Nations for the maintenance of peace" (The United Nations 1948, art. 26, 2). The question would be if educational AI could ease the access to that basic human rights or if education about AI is a matter of fundamental or higher education.

The subsequent article, article 27 paragraph 1, seems interesting as well. Because it adjudges the "shar[ing] in scientific advancement and its benefits" to all humans (The United Nations 1948, art. 27, 1). This could mean that benefits from the advancements of AI should also be shared with the general public.

[10] Any cited law is referred to the laws of the United States (U.S.). Because most research for this thesis has been conducted in the U.S. and most AI R&D has and is being realized in the U.S. If any law is not from the USA, it will be indicated.

Also, article 25 paragraph 1, expresses the "right to security in the event of unemployment, (…) or other lack of livelihood in circumstances beyond his control" (The United Nations 1948, art. 25, 1). This could be carried over to the lack of control of humans to future unemployment through AI. It is of course possible to prepare for any future unemployment through AI by further educating, providing alternative jobs or by offering Universal Basic Income by either the government or the company which replaces the employee with an AI.

The chosen rights from the UDHR mostly refer to rights to participate in or access goods. The description of how to handle these goods, meaning that they should be equally accessible to and shared with everyone, is similar to the description of a commons (see chapter 3). This might strengthen the argument to think about providing equal access to AI for all humans.

4.4.2 The "Right to Object" and "Automated Individual Decision-Making"

As mentioned above, one of the added paragraph by the EP to the GDPR is the paragraph 71 which is being called the "right to explanation." The right enables users of software that makes algorithmic decisions without human intervention to issue a request for an explanation from companies about these decisions. The right especially got implemented to protect minorities from racial profiling. Because algorithms automatically process data from natural persons to analyze or predict conditions like health status, economic status, work performance, personal preferences, behavior or even reliability. Especially when minorities just make such a small percentage that the algorithm puts them into dangerous categories and significantly affects them. For example if they get automatically refused on online recruitment platforms or during online credit applications (Parliament and Council of the European Union 2016).

The European Parliament (Parliament and Council of the European Union 2016) expressed the right to an explanation as follows:

> "In any case, such processing should be subject to suitable safeguards, which should include specific information to the data subject and the right to obtain human intervention, to express his or her point of view, to obtain an explanation of the decision reached after such assessment and to challenge the decision."

The EP adds the special cause that "automated decision-making and profiling based on special categories of personal data should be only allowed under special conditions", for

example criminal offences like fraud, in accordance to the law. That means that any other solely automated individual decision-making on user-level predictors are restricted if they affect users significantly (Parliament and Council of the European Union 2016). Despite the assumption of some articles and experts, there is no "right to explanation" traceable in that document. The law is actually divided into two articles – article 21 called "Right to Object" and article 22 called "Automated Individual Decision-Making, including Profiling" (see figure 7 and 8).

In chapter III "Rights of the Data Subject," article 12, paragraph 1, the EP makes sure that any legal entity controlling personal data is providing the data to affected members of the public in a transparent, understandable and easily accessible way (Parliament and Council of the European Union 2016).

Bulatov commented the "right to explanation" during the interview and gave an assessment of it as "unenforceable" and "useless" because of uncertain or ambiguous terms, e.g. the word "explanation". He (Bulatov 2016) argues that "nobody knows what exactly an explanation is – you can make up some sentences which sound plausible, you can not verify if this explanation is correct or not. It is very subjective." He even speaks for other AI researchers and data scientists in that matter.

But as objection to Bulatov' comment, the judicial discretion law empowers a judge or court guided by the principles of law to judge based on their personal assessment in the case of any lack of certainty. The discretion is the possibility to choose between various holdings (Bingham 2011, pp. 28-29).

Article 21

Right to object

1. The data subject shall have the right to object, on grounds relating to his or her particular situation, at any time to processing of personal data concerning him or her which is based on point (e) or (f) of Article 6(1), including profiling based on those provisions. The controller shall no longer process the personal data unless the controller demonstrates compelling legitimate grounds for the processing which override the interests, rights and freedoms of the data subject or for the establishment, exercise or defence of legal claims.

2. Where personal data are processed for direct marketing purposes, the data subject shall have the right to object at any time to processing of personal data concerning him or her for such marketing, which includes profiling to the extent that it is related to such direct marketing.

3. Where the data subject objects to processing for direct marketing purposes, the personal data shall no longer be processed for such purposes.

4. At the latest at the time of the first communication with the data subject, the right referred to in paragraphs 1 and 2 shall be explicitly brought to the attention of the data subject and shall be presented clearly and separately from any other information.

5. In the context of the use of information society services, and notwithstanding Directive 2002/58/EC, the data subject may exercise his or her right to object by automated means using technical specifications.

6. Where personal data are processed for scientific or historical research purposes or statistical purposes pursuant to Article 89(1), the data subject, on grounds relating to his or her particular situation, shall have the right to object to processing of personal data concerning him or her, unless the processing is necessary for the performance of a task carried out for reasons of public interest.

Figure 7: Excerpt from the General Data Protection Regulation, article 21[11]

Article 22

Automated individual decision-making, including profiling

1. The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her.

2. Paragraph 1 shall not apply if the decision:

(a) is necessary for entering into, or performance of, a contract between the data subject and a data controller;

(b) is authorised by Union or Member State law to which the controller is subject and which also lays down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests; or

(c) is based on the data subject's explicit consent.

3. In the cases referred to in points (a) and (c) of paragraph 2, the data controller shall implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision.

4. Decisions referred to in paragraph 2 shall not be based on special categories of personal data referred to in Article 9(1), unless point (a) or (g) of Article 9(2) applies and suitable measures to safeguard the data subject's rights and freedoms and legitimate interests are in place.

Figure 8: Excerpt from the General Data Protection Regulation, article 22[12]

[11] Source of image: Parliament and Council of the European Union 5/24/2016, article 21, p. 45

[12] Source of image: Parliament and Council of the European Union 5/24/2016, article 22, p. 46

4.4.3 Freedom of Information

The freedom of information, as the freedom of speech, are basic human rights. In article 19 of the UDHR it is written that "this right includes freedom (…) to seek, receive and impart information and ideas through any media and regardless of frontiers" (The United Nations 1948, art. 19).

The "Freedom of Information Act" (FOIA), which is a law that enables citizens to request information from their government and got implemented to "keep citizens in the know". But the FOIA is not applicable to companies or other institutions than the government in the USA at the moment – it applies to government bodies only (U.S. Department of Justice 2017). In other countries similar laws apply to private bodies as well.

The Human Development Report (2000, p. 78) of 2000 found out in a survey:

> "Often people have little information about decisions by the government or large businesses that have profound effects on their lives – about building schools, roads, water supplies and irrigation systems or about setting up businesses that would create employment or pollute the environment."

The freedom of information is therefore important to all citizens for keeping them knowledgeable, to safeguard them against mismanagement and corruption, and to enable them to make informed choices (Siraj 2010, p. 213). It could also be conveyed to other informational goods like decisions about and developments of AI.

The FOIA though also stands in competition with the right to secrecy. The government can withhold information if it endangers the national security. National security is a rather wide-spread term and can be misused by members of the executive branch to justify secrecy. To not let that happen, Halprin and Hoffman suggested to weigh out secrecy and disclosure reasons and to "establish a rule that information should be made public unless a convincing case can be made for withholding it from foreign governments" in 1976 already (Halperin & Hoffman 1976). And then there are also IPR which challenge the free flow of information.

4.4.4 Intellectual Property Rights

There are various forms of IPR, but in general it should give the creator of the intellectual work the power to control what happens with its work or how the asset gets used (Information Commissioner's Office, p. 2). In the Constitution of the United States, article 1,

section 8, is written that IPR, especially copyrights, are "promoting the progress of science and useful arts, by securing for limited times to authors and inventors the exclusive right to their respective writings and discoveries" (USTPO 2015). Two IPR are most important to the discussion about OSS and AI and will be explained in the following – patents and copyrights. In the core, a patent is granting right over property to the creator, approved by the government, to guarantee that others are not permitted to make, use, offer for sale or sell the invention for a specific time. Depending on the patent, this applies to the state, a continent or worldwide. But in exchange, the creator is still able to disclose its invention publicly when the patent got approved and conceded by the government (USTPO 2015). The application of IPR to a scientific knowledge has been analyzed by Murray and Stern. So-called patent-paper pairs are scientific articles which come with a patent that ensures the property right over that disclosed knowledge. Means the scientific paper becomes intellectual property which cannot be reused by other scientists while it stays open and accessible to them. A copyright ensures that the exact expression of the work is protected. So, a copyright would rather protect the written lines of code while a patent would protect the final product itself. Researchers (Murray, Stern 2007, p. 649) claim that:

> "[A] more recent 'anti-commons' perspective argues that the expansion of IPR (in form of patents and/or copyrights) is *'privatizing' the scientific commons and limiting scientific progress*. Specifically, the anti-commons hypothesis states that IPR may inhibit the free flow and diffusion of scientific knowledge and the ability of researchers to *build cumulatively on each other's discoveries*."

It is inevitable that there are arguments for and against IPR. To make it more distinct, Weber compares the work of software developers to the work of lawyers or veterinarians. As soon as lawyers or veterinarians develop a new method, it makes more sense to spread the information instead of patenting it. The sources, like the law or books about diseases, are freely accessible to anyone. And as soon as these practitioners are confronted by exceptional cases, they inquire input from their colleagues in an inter-exchange between experts. They can ask about their experience with similar cases and how they would approach the problem. Weber is reasoning that "a lawyer or veterinarian who could not access their communities easily, quickly, and at low cost would soon fall behind the curve and end up considerably less competent than the norm" (Weber 2005, p. 177).

Also, patenting is a costly and time-consuming process which would be eliminated by openness according to Bostrom (2016, p. 4).

Besides patents and copyright, which protect a companies’ intellectual property, there are trade secrets. The Uniform Trade Secrets Protection Act (UTSA) from 1985 ensures that information which has market relevance and lets a company stay competitive will not be unintendedly disclosed (see figure 9). It shall prevent misappropriations (American Bar Association 1985).

> (4) "Trade secret" means information, including a formula, pattern, compilation, program, device, method, technique, or process, that:
>
> (i) derives independent economic value, actual or potential, from not being generally known to, and not being readily ascertainable by proper means by, other persons who can obtain economic value from its disclosure or use, and
>
> (ii) is the subject of efforts that are reasonable under the circumstances to maintain its secrecy.

Figure 9: Excerpt from the Uniform Trade Secrets Protection Act, article 1, 4[13]

This law as well as the other IPR can be very important for companies and other parties from the private sector because it awards rights to them. But on the downside, IPR can worsen the state of openness and transparency, especially if misused or exploited.

5 The “free” market of AI

One of the main reasons companies do not practice open-sourcing AI components is competition. They compete with code, data and experts. In the example of the ILSVRC, one might assume participants were motivated by competition. The results of the competition were astonishing. Nevertheless, it is the option of this author that AI competitions should probably include more ethically urgent assignments. Some AI research laboratories already endeavor to design moral software which can distinguish from right or wrong, such as the project “Moral Machine” by MIT, which lets the public participate in autonomous

[13] Source of image: American Bar Association 1985, article 1, 4

vehicle decision-making. Interested folks can virtually decide how an autonomous vehicles might behave in extreme dilemmas of various social collocations – e.g. killing two passengers including a child or five adult pedestrians. The project tests public opinion for their designed dilemmas as well as makes it possible for participants to create their own dilemmas. The results are also open to the public and can be compared by anyone. The start-up called "drive.ai" tries to build a software for self-driving cars that can communicate with and handle pedestrians and other road users as humans who might behave unpredictable (Knight 2016). When William Song (2017), Co-Founder and Lead Software Engineer at drive.ai, was asked if he would open-source the start-up's code, he responded:

> "Investors will be very concerned if a start-up is open-sourcing their technology because basically you have no defensibility. Because right now *the* space is still very competitive. Everyone is trying to keep secrets."

However, Song said he would appreciate it if companies agreed on open-sourcing their technologies because self-driving is complicated to solve alone. He also mentioned that he would not be surprised if a company would open-source their self-driving technology, that there are some open-source engines available to train self-driving systems, and that the software is not even the problem. Data collection and processing depends a lot on money. He stated that as a start-up, it is very hard to build clusters of the size that companies like Google own already. Song (2017) affirms:

> "I do not think anyone can match Google's capability of both, collecting and handling data. They basically redefined how data should be processed in the last decade. They definitely have the best infrastructure and data collection capability in self-driving cars because they have fleets and fleets of vehicles. (…) It is not very comparable at the moment but hopefully as we get more funding, as we grow, it will become smarter."

The start-up scene is even more competitive than leading AI companies because they depend a lot more on funding from investors and existential security. As a result, it is difficult to open-source their technology as well as exchange ideas or programs with the public. Start-ups would therefore greatly benefit if leading AI companies open-source their technology and data clusters.

Another aspect is the competition for talented experts. The fight over AI experts got compared to the competition for players from the National Football League (NFL). AI experts

also make news if they sign a contract with a new company. Their income usually increases with each new contract as well. These experts fluctuate between different corporations. Yaroslav Bulatov, a software research engineer at OpenAI, changed his position from Google to OpenAI because he wanted to engage in more openly accessible research. He enjoys the fact that he is able to re-use pre-existing code or let other interested developers use his modified software. He sees the development of AI behind closed doors as a major problem. If anyone else needed AI capabilities, due to the time and money required because nothing would be easily accessible for free.

OpenAI's research is not connected to the commitment of generating financial return and can therefore focus on the research itself. For this reason, the project is quite unique compared to any other start-up or corporation where the main focus is to supply market demands.

The owners of code in industrial settings have considerable rights over future developments – employment relationship and status, product development tendencies or the allocation of profits. In the world of OSS, some of these privileged rights are abolished and a very strong power is being transferred from the "leader to the follower," especially through the "right to fork code". Because leaders depend on the continuous contribution of the community but are not directly able to acquire gains (Weber 2005, p. 180).

As mentioned earlier, exclusive property rights strengthen competition and seem like a "moral bad" according to Weber (see chapter 4.1.2.). They make "pirates out of neighbors" and seal any chance for collaboration or collective exchange. He concludes that "in this guise law effectively forbids the realization of a co-operating community" (Weber 2005, p. 47).

Competition in the field of AI is just another complication to existing research problems and brought along social dilemmas.

6 Conclusion

The most prevalent finding is that openness in AI stems from a variety of reasons. As mentioned, researchers value the ability to publish their work and employees appreciate companies that allow this. Especially for young talents within the AI field, openness is a

huge gain for their career (Bostrom 2016, p. 4). Learning from and teaching others is another benefit.

However, the field of AI is very influenced by businesses and business leaders. Money is the primary factor, which will determine what happens in the future. The world will change depending on the future developments of AI, whether it comes to industries, governments or social life.

Matters of interest at the moment are recruiting outstanding talent, driving autonomy forward, advancing techniques and models, solving important societal problems, and reforming the field by introducing more transparency.

Declaring AI as a commons would not have immediate implications but would transfer the focus to its immediacy in human life. Not only looking at the existential risk it poses, but also looking at the improvements it could offer humanity if handled in a transparent way. Findings show that the risk of corruption within the AI space is quite high and that wealth emerges exponentially, but is concentrated instead of distributed equitably. Therefore, many experts recommend a new model of ownership to block gains concentration within the hands of a few and to promote a fair sharing system (see chapter 4.2.4.).

Considering law and rules surrounding AI technology, the findings show that it will be most important to guide wealth distribution and technological infrastructure to alleviate the digital divide. Additionally, it will be important to create laws which guide developers as well as practitioners to work on and produce AI properly and ethically. The easiest mechanism is through transparency. Transparency maximizes sincerity towards the general public and reduces corruption within the AI space. It also leads to higher interaction between experts and between other individuals, by strengthening the communal co-operation and therefore the community (Hibbard 2016, p. 2; 2008, p. 2). Software by itself is barrier and border free, but it is unclear if this could happen worldwide. Law makers and experts should sit together and discuss future strategies. There are already committees, and other forms of expert consortiums, which discuss ethical issues surrounding AI. Most often they discuss practical topics like the application of self-driving cars or how to make a safe code which acts morally right (Stone et al. 2016, pp. 20-21). New committees could form with the aim to close the digital divide and also transform AI to a commons which would presumably lead to a higher common good.

The distinction that something which is freely accessible is not useful to all, is very important. As Gourley mentioned, there are still a lot of humans which cannot make any use

of AI software. Others are concerned that their jobs might get automated by AI (see chapter 4.2.4). As a result, “Open AI” cannot fight inequalities of the global digital divide. It will only accelerate it if nothing is being done to change that fact (Gourley; Hibbard 2016).

Declaring AI as a commons would not change that fact. It is more important to include developers in decisions, recognize their significance, and even take away some of their power in the establishment within the creation of AI.

As Weber remarks, “[a] good business model is simply one which succeeds adding additional value on the edge of the commons” (Weber 2005, p. 218).

7 References

American Bar Association (1985): Uniform Trade Secrets Act. UTSA, revised 2/11/1986. Available online at http://www.uniformlaws.org/shared/docs/trade%20secrets/utsa_final_85.pdf, checked on 7/24/2017.

API.AI (2017): Natural language understanding platform. Available online at https://api.ai/, checked on 6/5/2017.

Bingham, Thomas Henry (2011): The business of judging. Selected essays and speeches. Oxford: Oxford University Press. Available online at http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=694200.

Bostrom, Nick (2014): Superintelligenz. Szenarien einer kommenden Revolution. 1. ed. Berlin: Suhrkamp. Available online at http://ebooks.ciando.com/book/index.cfm/bok_id/1680604.

Bostrom, Nick (2016): Strategic implications of openness in AI development. Available online at www.nickbostrom.com, checked on 6/7/2017.

Breuer, Franz (2010): Reflexive Grounded Theory. Eine Einführung für die Forschungspraxis. 2. ed. Wiesbaden: VS Verlag für Sozialwissenschaften / Springer Fachmedien Wiesbaden GmbH Wiesbaden. Available online at http://dx.doi.org/10.1007/978-3-531-92580-6.

Bulatov, Yaroslav (2016): Interview of the author with Yaroslav Bulatov on the 10/27/2016. Topic: the future of AI from an ethical perspective. AI as a commons. OpenAI office, San Francisco, California, USA.

Burger, Harald Christoph (2016): Google DeepMind's AlphaGo: How it works. TasteHit. Available online at https://www.tastehit.com/blog/google-deepmind-alphago-how-it-works/, updated on 3/16/2016, checked on 6/21/2017.

Di Fabio, Udo et. al (2017): First guidelines in the world for self-driving computers. Ethics Commission on Automated Driving presents report. With assistance of Alexander Dobrindt. Federal Ministry of Transport and Digital Infrastructure. Available online at https://www.bmvi.de/SharedDocs/EN/PressRelease/2017/084-ethic-commission-report-automated-driving.html, updated on 6/20/2017, checked on 7/24/2017.

Domingos, Pedro (2015): The master algorithm. How the quest for the ultimate learning machine will remake our world. New York: Basic Books.

European Parliament (2017): Robots. Legal Affairs Committee calls for EU-wide rules. Lietzen, Lina, lega-press@europarl.europa.eu. Available online at http://www.europarl.europa.eu/news/en/news-room/20170110IPR57613/robots-legal-affairs-committee-calls-for-eu-wide-rules, published on 1/12/2017, checked on 4/17/2017.

Fuller, Steve (2016): Stephen Hawking is wrong. Humans will not compete with AI - we will merge with it. Available online at http://www.telegraph.co.uk/technology/2016/10/20/stephen-hawking-is-wrong-humans-wont-compete-with-ai--we-will-me/, updated on 10/20/2016, checked on 7/16/2017.

Garret Hardin (1968): The tragedy of the commons. The population problem has no technical solution; it requires a fundamental extension in morality. In *Science* 162 (3859), pp. 1243–1248. Available online at https://www.sciencemag.org/site/feature/misc/webfeat/sotp/pdfs/162-3859-1243.pdf, checked on 6/25/2017.

GitHub (2017). GitHub. Available online at https://github.com/, checked on 7/15/2017.

Google Brain (2017): TensorFlow. Available online at https://www.tensorflow.org/, updated on 4/26/2017, checked on 5/15/2017.

Gourley, Sean (2016): Interview of the author with Sean Gourley on the 12/13/2016. Topic: the future of AI from an ethical perspective. AI as a commons. San Francisco, California, USA.

Greco, Gian Maria; Floridi, Luciano (2004): The tragedy of the digital commons. In Ethics and Information Technology 6 (2), pp. 1–26. DOI: 10.1007/s10676-004-2895-2.

Halperin, Morton H.; Hoffman, Daniel N. (1976): Secrecy and the right to know. In Law and contemporary problems 40 (3), pp. 132–165. Available online at http://scholarship.law.duke.edu/lcp/vol40/iss3/5, checked on 7/24/2017.

Hibbard, Bill (2008): Open Source AI. Edited by University of Wisconsin - Madison. Wisconsin, Mad. Available online at http://www.ssec.wisc.edu/~billh/g/hibbard_agi_workshop.pdf, checked on 9/26/2016.

Hibbard, Bill (2016): Interview of the author with Bill Hibbard on the 10/31/2016. Topic: the future of AI from an ethical perspective. AI as a commons. San Francisco, California, USA, Madison, Wisconsin, USA.

Information Commissioner's Office: Intellectual property rights and disclosures under the Freedom of Information Act. Freedom of Information Act. 2.2nd ed. Available online at https://ico.org.uk/media/for-organisations/documents/1150/intellectual_property_rights_and_disclosures_under_the_foia.pdf, checked on 7/3/2017.

Kaul, Inge; Grunberg, Isabelle; Stern, Marc A. (Eds.) (1999): Global public goods. International cooperation in the 21st century. New York, NY: Oxford Univ. Press. Available online at http://www.loc.gov/catdir/enhancements/fy0604/99010940-d.html.

Knight, Will (2016): New self-driving car tells pedestrians when it is safe to cross the street. How will self-driving cars respond to all those unpredictable humans? MIT

Technology Review. Available online at https://www.technologyreview.com/s/602267/new-self-driving-car-tells-pedestrians-when-its-safe-to-cross-the-street/, updated on 8/30/2016, checked on 7/22/2017.

Kurzweil, Ray (2012): How to create a mind. The secret of human thought revealed. New York: Viking.

Lai, Morgan (2016): Interview of the author with Morgan Lai on the 11/23/2016. Topic: the future of AI from an ethical perspective. AI as a commons. San Francisco, California, USA.

Marshall, M. N. (1996): The key informant technique. In Fam Pract 13 (1), pp. 92–97. DOI: 10.1093/fampra/13.1.92.

McCarthy, John (1955): A proposal for the Dartmouth Summer Research Project on Artificial Intelligence. With assistance of Claude Shannon, Marvin Minsky, Nathaniel Rochester. Stanford University. Available online at http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html, updated on 4/3/1966, checked on 6/21/2017.

Metz, Cade (2016): The Rise of Artificial Intelligence and the End of Code. What the AI Behind AlphaGo Can Teach Us About Being Human. Wired. Online. Available online at https://www.wired.com/2016/05/google-alpha-go-ai/, updated on 9/5/2016, checked on 6/16/2017.

Metz, Cade (2017): The battle for top AI talent only gets tougher from here. Available online at https://www.wired.com/2017/03/intel-just-jumped-fierce-competition-ai-talent/, updated on 3/23/2017, checked on 7/16/2017.

Moore, David (2016): Interview of the author with David Moore on the 11/11/2016. Topic: the future of AI from an ethical perspective. AI as a commons. Berkeley, California, USA.

Murray, Fiona; Stern, Scott (2007): Do formal intellectual property rights hinder the free flow of scientific knowledge? In *Journal of Economic Behavior & Organization* 63 (4), pp. 648–687. DOI: 10.1016/j.jebo.2006.05.017.

Norman, Donald A. (2013): The design of everyday things. Revised and expanded edition. New York: Basic Books. Available online at http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=658202.

NSTC (2016): Preparing for the future of Artificial Intelligence. With assistance of Executive Office of the President. Edited by National Science and Technology Council Committee on Technology. Executive Office of the President. Washington D.C. Available online at https://www.whitehouse.gov/blog/2016/10/12/administrations-report-future-artificial-intelligence, checked on 11/16/2016.

OpenAI (2017): OpenAI. Edited by OpenAI. OpenAI. Available online at https://openai.com/, updated on 6/4/2017, checked on 5/15/2017.

Ostrom, Elinor (2015): Governing the commons. The evolution of institutions for collective action. Cambridge: Cambridge University Press (Canto Classics). Available online at http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&AN=1077401, checked on 12/1/2016.

Parliament and Council of the European Union (5/24/2016): General Data Protection Regulation. GDPR. Source: European Union. Available online at http://data.consilium.europa.eu/doc/document/ST-5419-2016-INIT/en/pdf, checked on 7/24/2017.

Quilligan, James (2012): Why distinguish common goods from public goods? In David Bollier, Silke Helfrich (Eds.): The wealth of the commons. A world beyond market and state /ed. by David Bollier and Silke Helfrich. Amherst, Mass: Levellers Press. Available online at http://wealthofthecommons.org/essay/why-distinguish-common-goods-public-goods, checked on 6/25/2017.

Russakovsky, Olga; Deng, Jia; Su, Hao; Krause, Jonathan; Satheesh, Sanjeev; Ma, Sean et al. (2015): ImageNet Large Scale Visual Recognition Challenge. Available online at https://arxiv.org/pdf/1409.0575v3.pdf, checked on 5/22/2017.

Russell, Stuart J.; Norvig, Peter (1995): Artificial intelligence. A modern approach. Upper Saddle River: Prentice Hall (Prentice Hall series in artificial intelligence).

Russell, Stuart J.; Norvig, Peter (2010): Artificial intelligence. A modern approach. 3. ed. Upper Saddle River, NJ: Prentice-Hall (Prentice-Hall series in artificial intelligence).

Schnell, Urs; Gehriger, Res (2013): Bottled Life. Original title: The truth about Nestlés business with water. Starring Silvia Seuboth-Radtke. Urs Schnell, Laurent Stoop, Dodo Hunziger.

Science and Technology Committee (2016): Robotics and artificial intelligence. Fifth Report of Session 2016–17. Edited by House of Commons. House of Commons. Available online at http://www.publications.parliament.uk/pa/cm201617/cmselect/cmsctech/145/14502.htm, checked on 10/25/2016.

Shead, Sam (2016): Apple is finally going to start publishing its AI research. Available online at http://www.businessinsider.de/apple-is-finally-going-to-start-publishing-its-artificial-intelligence-research-2016-12?utm_source=feedly&utm_medium=webfeeds&r=US&IR=T, updated on 12/6/2016, checked on 7/15/2017.

Shiva, Vandana (2016): Water wars. Privatization, pollution, and profit. Berkeley, California: North Atlantic Books.

Simonite, Tom (2017): AI Software Learns to Make AI Software. Google and others think software that learns to learn could take over some work done by AI experts. MIT Technology Review. Available online at https://www.technologyreview.com/s/603381/ai-software-learns-to-make-ai-software/, updated on 1/18/2017, checked on 7/16/2017.

Siraj, Mazhar (2010): Exclusion of Private Sector from Freedom of Information Laws. Implications from a Human Rights Perspective. In Journal of Alternative Perspectives in the Social Sciences 2 (1), pp. 211–226. Available online at http://www.japss.org/upload/interiorspecialissue.pdf, checked on 7/24/2017.

Song, William (2017): Interview of the author with William Song on the 1/15/2017. Topic: the future of AI from an ethical perspective. AI as a commons. San Francisco, California, USA.

Stahl, Bernd C. (2004): Information, Ethics, and Computers. The Problem of Autonomous Moral Agents. In Minds and Machines 14 (1), pp. 67–83. DOI: 10.1023/B:MIND.0000005136.61217.93.

Stone, Peter; Brooks, Rodney; Brynjolfsson, Erik; Calo, Ryan; Etzioni, Oren; Hager, Greg et al. (2016): Artificial Intelligence and Life in 2030. One Hundred Year Study on Artificial Intelligence (AI100). Report of the 2015 Study Panel. Stanford University. Stanford, California (1). Available online at http://ai100.stanford.edu/2016-report, checked on 9/20/2016.

Szegedy, Christian (2015): Building a deeper understanding of images. With assistance of Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Drago Anguelov, Dumitru Erhan, Andrew Rabinovich. Google. Google Research Blog. Available online at https://research.googleblog.com/2014/09/building-deeper-understanding-of-images.html, updated on 9/5/2015, checked on 5/22/2017.

The United Nations (12/10/1948): The Universal Declaration of Human Rights. UDHR, revised 12/10/1948. Source: The United Nations. Available online at http://www.ohchr.org/EN/UDHR/Documents/UDHR_Translations/eng.pdf, checked on 7/24/2017.

Turing, Alan (1950): Computing Machinery and Intelligence. In Mind: A Quarterly Review of Psychology and Philosophy 49, pp. 433–460. Available online at https://www.csee.umbc.edu/courses/471/papers/turing.pdf, checked on 5/22/2017.

U.S. Department of Justice (2017): What is FOIA? Available online at https://www.foia.gov/about.html, checked on 7/15/2017.

UNDP (2000): Human rights and human development. Human development report 2000. New York: Oxford Univ. Press.

UNDP (2001): Making new technologies work for human development. Human development report 2001. New York: Oxford Univ. Press.

UNDP (2016): Human development for everyone. Human development report 2016. New York: United Nations Development Programme.

Union internationale des télécommunications (op. 2015): Measuring the information society report 2015. Geneva: International telecommunications union.

USTPO (2015): General information concerning patents. United States Trademark and Patent Office. Available online at https://www.uspto.gov/patents-getting-started/general-information-concerning-patents#heading-3, updated on October 2015, checked on 7/23/2017.

Watkins, Kevin (2006): Beyond scarcity. Power, poverty and the global water crisis. New York: UNDP (Human development report, 2006). Available online at http://hdr.undp.org/en/media/HDR06-complete.pdf, checked on 7/2/2017.

Weber, Steven (2005): The success of open source. 1. ed. Cambridge, Mass.: Harvard Univ. Press.

Wold Economic Forum (2016): Global Risks Report 2016. Edited by Wold Economic Forum. Wold Economic Forum. Geneva. Available online at https://www.weforum.org/reports/the-global-risks-report-2016/, checked on 11/16/2016.

Yudkowsky, Eliezer (2012): Artificial Intelligence as a positive and negative factor in Global Risk. In Nick Bostrom (Ed.): Global catastrophic risks. Repr. Oxford: Oxford University Press, pp. 308–343. Available online at https://intelligence.org/files/AIPosNegFactor.pdf, checked on 4/16/2017.

Appendix

Interview guide

1. What are you working on at the moment?
2. Why did you engage in the research of AI?

Or: Why did you co-found/found (…)?

Trigger/Motivation

3. Did you set any principles for your work at some point?

Perhaps what you want to achieve someday or something you would never want to work on?

4. Why do you think some developer open-source their AI systems, for instance the OpenAI office?

Or: Did you ever think about open-sourcing your code?

Could you see benefits in it?

5. To what extent does crowdsourcing help to solve specific problems in the field of AI?

Have you had maybe any help from outside for your research?

When to use/when not? Are there other ways to involve others?

6. What do you think about the accessibility of data for your research?

Are there a lot of open-source databases?

Is it more difficult for most developers compared to companies like Google/Facebook/Uber…?

7. Do you eventually know what deposit copies are?

It means that any copyrighted work that is published must be submitted in two copies to the United States Copyright Office at the Library of Congress.

Can you picture a national copyright online library for AI systems which everyone can access?

8. Do you think someone can own an AI system which influences the society at a large scale?

9. Could you imagine that in the near future it would be beneath human dignity if someone does not own a virtual assistant (or another kind of AI)?

Maybe even on the same level as the right to education for children

10. *Democratizing AI:*

- *Freedom of information enables citizens to make informed choices and safeguards them against mismanagement and corruption*

- *Some countries extended the legislation to private bodies*

- *Access information about internal processes leads to higher transparency*

Could you think of AI as a commons to which everyone gains a right?

11. *The House of Commons Science and Technology Committee released a report on Robotics & AI:*

They demanded a digital strategy on behalf of the government and law regulations for AI

Do you think an institution, no matter which one, is able or should be enabled to regulate AI?

(Biological and toxin weapon convention)

12. How do you envision the future of AI?

Is there a way to enrich the society? Is there any problem you could think of someone who works in the field of AI should address at the moment?

Eidesstattliche Erklärung / Affidavit

Name / Last name:	**Vorname / Surname:**
Salameh	Amelie
Matrikelnummer / Student number:	**Studiengang / Studies:**
1265885	Informationsmanagement / Information Management

Deutsch: Hiermit versichere ich, Amelie Salameh, an Eides statt, dass ich die vorliegende Bachelorarbeit mit dem Titel „Artificial Intelligence as a Commons – Opportunities and Challenges for Society“ selbständig und ohne fremde Hilfe verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe. Die Stellen der Arbeit, die dem Wortlaut oder dem Sinne nach anderen Werken entnommen wurden, sind in jedem Fall unter Angabe der Quelle kenntlich gemacht. Die Arbeit ist noch nicht veröffentlicht oder in anderer Form als Prüfungsleistung vorgelegt worden.

English: I, Amelie Salameh, hereby assure under oath that I wrote the present Bachelor thesis with the title “Artificial Intelligence as a Commons - Opportunities and Challenges for Society” independently and without any help, and used no other tools or resources than those indicated. The parts of the work, which have been taken from the wording or the meaning of other works, are indicated in each case with reference to the source. The work has not yet been published or presented in any other form as an assignment.

Berlin, den _____ _____

Ort / Place, Datum / Date Unterschrift / Signature