

TALLINN UNIVERSITY OF TECHNOLOGY

School of Business and Governance

Department of Law

Henri Mart Pihl

**THE RISKS AND CHALLENGES OF LEGAL PERSON- OR
AGENTHOOD OF ARTIFICIAL INTELLIGENCE
TECHNOLOGIES**

Bachelor's thesis

Programme Law, specialisation European Union and International Law

Supervisor: Tanel Kerikmäe, PhD

Tallinn 2020

I declare that I have compiled the paper independently and all works, important standpoints and data by other authors have been properly referenced and the same paper has not been previously been presented for grading. The document length is 9248 words from the introduction to the end of summary.

Henri Mart Pihl..... (signature, date)

Student code: HAJB15995

Student e-mail address: pihl.henri@gmail.com

Supervisor: Tanel Kerikmäe, PhD:

The paper conforms to requirements in force

..... (signature, date)

Chairman of the Defence Committee:

Permitted to the defence (name, signature, date)

TABLE OF CONTENTS

ABSTRACT 4

LIST OF ABBREVIATIONS 5

INTRODUCTION 6

1. AI AS LEGAL PERSONS OR AGENTS – NATURE AND SUBJECT MATTER..... 7

 1.1 Examples of currently used AI technologies which may warrant the granting of legal person-
 or agenthood 7

 1.2 Understanding legal personhood and legal agenthood 9

 1.3 Philosophical and historical context of personhood. 11

 1.4 AI as a natural person 12

 1.5 AI as a legal person..... 13

 1.6 AI as a legal agent..... 15

2. THE EUROPEAN UNION PERSPECTIVE ON LEGISLATION CHANGES TO AI AND
SMART ROBOT TECHNOLOGIES 16

3. THE PRACTICAL AND ETHICAL BENEFITS AND DETRIMENTS OF GRANTING AI
TECHNOLOGIES LEGAL PERSON- OR AGENTHOOD 20

 3.1 The practical benefits of granting AI technologies legal person- or agenthood. 20

 3.2 The ethical and moral benefits of granting AI technologies legal person- or agenthood. 21

 3.3 The practical detriments of granting AI technologies legal person- or agenthood. 22

 3.4 The ethical and moral detriments of granting AI technologies legal person- or agenthood. 22

CONCLUSION 24

LIST OF REFERENCES 26

APPENDICES 29

ABSTRACT

Artificial intelligence technologies are quickly becoming a major part of modern societies in many fields, such as finance, health care, production, services, just to name a few. With the rapid development of artificial intelligence technologies, the legislation related to it has to be able to keep up with the practical and ethical demands of the society. One of the hardest things to do for legislators and legal scholars is to predict what is going to happen in the future and to come up with appropriate policy and legislation changes ahead of time. To have any chance of success in that, one has to look at the current trends in the field and also consider what is likely to happen in the near future. The granting of legal person- or agenthood rights to artificial intelligence technologies is one of the potential methods of dealing with the gaps or uncertainties in the legal systems related to said technologies and the ramifications of these methods are discussed in this research. The author also provides an overview of the currently discussed policy and regulation changes of the European Union on this matter as well as the EU's approach in general. Finally the author will give his opinion on the practical and moral benefits and detriments which may rise as a result of granting artificial intelligence technologies legal person- or agenthood status and why the author thinks that there is currently no need to grant either of those statuses to artificial intelligence technologies but a separate categorization should be considered.

Keywords: Artificial Intelligence, Legal Personhood, Legal Agenthood, Liability, European Union.

LIST OF ABBREVIATIONS

AI – Artificial Intelligence

EU – European Union

HLEG – High Level Expert Group

ICCPR – International Covenant on Civil and Political Rights

UDHR – Universal Declaration of Human Rights

VITAL - Validating Investment Tool for Advancing Life Sciences

INTRODUCTION

The goal of this research is to find out what are currently discussed and proposed policy and legislation changes related to the legal person- and agenthood of artificial intelligence (AI) technologies, with the focus being on the European Union. As different AI technologies become more common in all facets of life, it makes sense to analyze the potential need for granting different rights and responsibilities to those technologies. Whether there is a need for it at all, only in specific scenarios or widespread need, has to be analyzed and compared. Differing opinions all over the world can create a lot of friction and inefficiency when it comes to the development of these technologies and their implementation, thus a better overview of the current and future developments in this field could lead to better understanding and cooperation, both nationally and internationally.

The aim of this research is to analyze the risks and challenges which could stem from the granting of legal personhood or agenthood to AI technologies. The author intends to create an overview of the different policy approaches and practices of legal personhood and agenthood of AI technologies as well as give an overview of some of the opinions that the experts in the field have expressed. The author intends to focus on the legal perspective but other perspectives, such as moral or philosophical, will also be discussed as necessary. The main purpose for this thesis is to be a guidepost to understanding the current and future situation in regards to the granting of legal person- or agenthood to AI technologies.

The research question for this thesis is: what are the practical and moral reasons for potentially granting AI technologies legal person- or agenthood and what are the risks and challenges that rise from that?

The research method for this thesis is primarily the analysis of legal literature and opinions of various high-level expert groups.

1. AI AS LEGAL PERSONS OR AGENTS – NATURE AND SUBJECT MATTER

In order to analyze the risks and challenges which would arise from granting AI technologies legal person- or agenthood, one must first understand what is meant when we talk about AI legal personhood and legal agenthood in the context of this research.

1.1 Examples of currently used AI technologies which may warrant the granting of legal person- or agenthood

In order to better understand the AI technologies currently used in different aspects of life, which may warrant the granting of legal person- or agenthood, the author will give a non-exhaustive list of some of the more well-known solutions used around the world and a short description of what they are capable of doing. But before that, the author would like to define AI. The definition that the author will be basing this paper on is the definition provided by the EU Commission's communication on AI, which is the following: "Artificial intelligence (AI) refers to systems that display intelligent behavior by analyzing their environment and taking actions –with some degree of autonomy –to achieve specific goals."¹ The author wants to note here that while the definition of AI in this research is rather broad, the AI what we are talking about is only the most advanced technologies available to us today and when we are talking about AI in the context of this research we do not mean simple smart devices, vacuum robots etc. But now onto the examples:

Sophia

Sophia is an advanced human-like robot, developed by a Hong Kong based company Hanson Robotics, who was created in 2016. She is the first robot to be granted citizenship of any country, which took place in 2017, when Sophia was granted the Saudi Arabian citizenship. She is also the first robot Innovation Ambassador for the United Nations Development Programme and has been on TV shows such as the Tonight Show and Good Morning Britain. Sophia's design and purpose is to be able to communicate with humans. She is able to read and use facial expressions, hold

¹ Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions on Artificial Intelligence for Europe, Brussels, 25.4.2018 COM(2018) 237 final.

conversations on predetermined topics, walk and even draw.² For additional reading on what Sophia is and what she could mean for intellectual property rights see:³

Nadine

Nadine is an advanced human-like robot developed by the Institute for Media Innovation in Singapore's Nanyang Technological University. Nadine is able to answer questions in different languages, remember the people she has interacted with, show emotions through expressions and gestures. She also has a “personality”, meaning her mood and answers can differ depending on what you say to her and how you say it. Nadine is designed to be a companion robot, she can assist people with special needs, read stories, show images, put on Skype sessions, send emails, and communicate with the family. She can play the role of a personal, private coach always available when nobody is there.⁴ To read more about Nadine and AI in the medical field, see:⁵

VITAL

VITAL, which stands for Validating Investment Tool for Advancing Life Sciences, does not have a human-like physical appearance unlike Nadine and Sophia. VITAL is essentially a risk assessment AI, which was developed by a Hong Kong venture capital firm, Deep Knowledge Ventures, to help them make investment decisions in the biotech field and in 2014 became the first AI to be appointed as a director to a corporate board.⁶ VITAL is treated as a member of the board with observer status, meaning that while it does not take part in the discussions and overall leadership of the company, it does have a vote when it comes down to deciding on whether to invest in a company or not.⁷

Alicia T

Alicia T is an AI which is being developed by Tieto, a Finnish software and communications company, for its data-driven businesses unit, of which she was also made a part of the leadership

² Hansen Robotics: Sophia Accessible: <https://www.hansonrobotics.com/sophia/> 14 March 2020.

³ Rocha, E., Sophia: Exploring the Ways AI May Change Intellectual Property Protections, DePaul Journal of Art, Technology and Intellectual Property Law, Vol. 28, Issue 2 (Spring 2018), pp. 126-146.

⁴ Nadine Social Robot Accessible: <https://imi.ntu.edu.sg/IMIResearch/ResearchAreas/Pages/NadineSocialRobot.aspx> 14 March 2020

⁵ Zapusek, T., Artificial Intelligence in Medicine and Confidentiality of Data, Asia Pacific Journal of Health Law & Ethics, Vol. 11, Issue 1 (November 2017), pp. 105-126.

⁶ Artificial intelligence gets a seat in the boardroom Accessible: <https://asia.nikkei.com/Business/Artificial-intelligence-gets-a-seat-in-the-boardroom> 15 March 2020.

⁷ Robots in the Boardroom: Artificial Intelligence and Corporate Law Accessible: <https://www.law.ox.ac.uk/business-law-blog/blog/2017/11/robots-boardroom-artificial-intelligence-and-corporate-law> 15 March 2020.

team. Alicia T's role in the leadership team is to support data-driven decision-making and innovate new data-driven ideas with the help of machine intelligence and advanced data analytics. While she is fitted with a conversational interface, meaning you could communicate with her, and she can cast votes in meetings, her main role is to be used as a data analyst, but her role will expand as the technology gets more sophisticated.⁸

The examples provided here can be categorized roughly into two different types. The first kind are robots or systems which have a tangible, physical form, like Nadine and Sophia. Self-driving cars, for example, could also be considered in this category if we were to expand it. Their tasks generally revolve around human interaction and doing physical things. In a sense, they are poster-boys, –girls and –cars for the future of robotics. They are a tangible, physical proof of the progress and innovation being made in the robotics and AI fields. The second kind of AI technologies are the VITAL's and Alicia T's of the world, technologies which do not have a physical body, nor are their tasks similar to those of the physical robots. One would not be wrong to call these technologies algorithms or programs. The main purpose for these types of AI are generally data compilation, analysis and using those abilities to project and find correlations between that data in ways which is either impossible or impractical for humans to do themselves. While these kinds of AI are not as visible and talked about as the androids of the world, the author would say that they are, at least currently, much more capable in their intended roles and much more valuable in a commercial sense, being able to analyze data in ways which is too difficult or time consuming for humans to do.

1.2 Understanding legal personhood and legal agenthood

In order to analyze the risks and challenges which would arise from granting AI technologies legal person- or agenthood, one must first understand what is meant when we talk about legal personhood and legal agenthood in the context of this research.

⁸ Tieto the first Nordic company to appoint Artificial Intelligence to the leadership team of the new data-driven businesses unit Accessible: <https://www.tieto.com/en/newsroom/all-news-and-releases/corporate-news/2016/10/tieto-the-first-nordic-company-to-appoint-artificial-intelligence-to-the-leadership-team-of-the-new-data-driven-business/> March 15 2020.

First, let us talk about legal agency. Merriam-Webster legal dictionary defines agency as: “A consensual fiduciary relationship in which one party (agent) acts on behalf of and under the control of another (principal) in dealing with third parties”⁹. Thus, agency is a quality or status which is achieved after having been granted permission to act and make decisions on the behalf of another, related to third parties. Of course there are limitations to the extent of the rights and obligations conferred to the agent and those are decided on a case by case basis. It is very important to note that the actions of the agent are legally binding on the principal, within the scope of the agreement of course.

Legal personhood, however, is a term which is perhaps a bit trickier to understand. As those familiar with the legal world well know, you have two types of persons in law: legal and natural persons. Simply put, natural persons are humans, also known as *homo sapiens*. Article 6 of UDHR of 1948¹⁰ and article 16 of ICCPR of 1966¹¹ grant all humans a right to recognition as a person under law. At which point can someone be called a human is a question of debate but at the very least, from the moment of birth until their death, or being pronounced dead, humans can be considered natural persons. The crucial point with respect to assigning legal personhood is that the human being is not a person before the law because he is a human being, but because the law calls him or her “person.”¹² For this research the author will use the legal personality definition provided in the Yale Law Journal in 1928 by Smith: “To be a legal person is to be the subject of rights and duties. To confer legal rights or to impose legal duties, therefore, is to confer legal personality...”¹³ Legal persons are for example entities such as businesses, organizations or corporations, which have been granted the title of a legal person “artificially”, by humans, in order to allow for those entities to exist and exercise their rights and obligations in the legal world, whether it be entering into contracts, suing or being sued, owning property or something else. It is also important to understand that a legal person in one jurisdiction may not be a legal person under another; inanimate objects such as idols or buildings could be categorized as legal persons under certain jurisdictions. So we have legal persons and natural person, so what about all the other stuff, which cannot be categorized

⁹ Definition of “agency” Accessible: <https://www.merriam-webster.com/dictionary/agency#legalDictionary> March 15 2020

¹⁰ UDHR Article 6 Accessible: http://www.claiminghumanrights.org/udhr_article_6.html March 16 2020

¹¹ ICCPR full text Accessible: <https://treaties.un.org/doc/publication/unts/volume%20999/volume-999-i-14668-english.pdf> March 16 2020

¹² Gaakeer, J, Sua Cuique Persona: A Note on the Fiction of Legal Personhood and a Reflection on Interdisciplinary Consequences, *Law & Literature*, Vol. 28, Issue 3 (Fall 2016), p. 295.

¹³ Smith, B., Legal personality, *The Yale Law Journal* Vol. 37, No. 3 (Jan., 1928), p. 283

as either of those? Currently, we consider the rest to be legal objects, in other words corporeal or incorporeal objects over which legal subjects may hold rights.

Thus, when we are talking about potentially granting legal agenthood to AI technologies, what we are talking about is the right for the AI to make legally binding decisions on the behalf of the principal. This principal could be a natural person or it could also be a legal person. When talking about granting personhood to AI, we can take the discussion in two directions: granting of legal person status or granting of natural person status. Of course that is just the current arrangement of our legal systems, and further classifications could be made.

1.3 Philosophical and historical context of personhood.

To begin this chapter, the author would like to briefly discuss the historical and philosophical background of personhood in general. In order to see what is coming we must first analyze what has been and how that has shaped the situation we are in now. From our own history, human slavery is a phenomenon which we are still working to rid ourselves of, with Mauritania being the last country to officially abolish slavery in 1981, even though the enforcement of those laws is still a problem in many places of the world.¹⁴ The reason why slavery is relevant to this research is that as AI technology keeps improving, we could get to a point where granting legal personhood to those technologies becomes not only a question of practicality but a question of morality and ethics. This is no doubt a scenario which is far into the future but imagine an android who is so similar to a human that you could not say with certainty whether the android is human or not. If this android expresses empathy, has goals in “life”, is able to live amongst humans and interact with its community then would you feel comfortable with the android being treated as an object?¹⁵ Another question to ponder is the approach to property. In very general terms, according to the Lockean approach to property, humans have the natural right to the products of their labor¹⁶ and this could be relayed to human creation of AI, meaning that due to the AI being created by humans, we, the humans, should have claim to them as property, and for now, the author agrees with this line of thinking. However, Locke has also argued that equality should be granted on the grounds of “equal

¹⁴ For more information about modern slavery Accessible: <https://www.globallslaveryindex.org/> March 16 2020

¹⁵ Hubbard, F.P., Do Androids Dream: Personhood and Intelligent Artifacts, Temple Law Review, Vol. 83, Issue 2 (Winter 2011), p. 419.

¹⁶ Locke, J., The Second Treatise of Government (1690).

faculties”¹⁷, meaning that if AI technologies are able to develop to a level where their faculties would be similarly capable to those of humans then granting of equal rights may be prudent. Although in saying that, the author is being a bit disingenuous, as Locke was talking about people of the same species, but the idea of equality based on equal faculties is still worth considering. Thus, as new and better technology rolls out, a new classification for AI technologies, which reach a certain level of autonomy, could be considered in the future. Criteria for Recognition of AI as a Legal Person, a paper by Dremluga, R., Kuznetcov, P., Mamychev, outlines what the authors consider the necessary requirements to be fulfilled in order to potentially grant AI either legal personhood or – agenthood status. ¹⁸ They recognize that, on a fundamental level, there is nothing stopping us from granting those rights to the AI, as ultimately, the legal system is by the people, for the people.

1.4 AI as a natural person

In the research the author takes a look at 3 different approaches to granting legal responsibility and rights to AI technologies themselves. These 3 approaches are natural personhood, legal personhood, legal agenthood. The first to be analyzed is natural personhood, and this is also the approach which the author believes to be the most unlikely to happen and as such will be discussed the least. A report by CEPS, compiled by A. Renda, found that: “Nothing in current AI developments suggests that AI will move towards developing human-like perception and awareness, or sentience, thus leading towards so-called “artificial general intelligence” in the immediate future.”¹⁹ Having said that, the author will now embark on a short hypothetical journey in which the AI has officially requested that it be granted natural human status, given rights equal to those of humans, has asked for constitutional rights, to not be considered property etc. One of the most often occurring arguments among academics who have also considered a similar possibility is what is sometimes called the “missing something argument”. The argument is based on that, even if AI were to develop to a point where it is equal or superior to humans in regards to its capabilities, looks and behavior, it

¹⁷ Ibid.

¹⁸ Dremluga, R., Kuznetcov, P., Mamychev, A. Criteria for Recognition of AI as a Legal Person, Journal of Politics and Law, Vol. 12, Issue 3 (September 2019), pp. 105-112

¹⁹ Renda, A. (2019) Artificial Intelligence - Ethics, governance and policy challenges Accessible: <https://www.ceps.eu/ceps-publications/artificial-intelligence-ethics-governance-and-policy-challenges/> March 20 2020

would still be missing “something” that the humans have. This “something” is not defined and could mean different things to different people. Some may call it a soul, some may call it will, others will have their own understanding of it but what remains is that it is something intangible which all humans have. To be fair, then the “missing something” argument stems from our own lack of understanding of ourselves. While the discoveries in neuroscience and anthropology in general have given us a better understanding of our human condition, there is still so much that we do not know. And if you do not understand the thing you are comparing something to, then the results will be inconclusive. It is like trying to compare the contents of two closed boxes without being able to look inside. You may be able to make assumptions based on visual or physical examination, and conclude that the two seem to be identical, but you cannot be entirely certain. Even if we disregard the basic requirement of having to be a human, natural personhood is a status, which for the AI in their current form, is simply impossible to achieve. The author finds it possible that eventually a categorization, which would give AI similar rights and obligations to humans, could be made but even then it would be something else than a natural person.

1.5 AI as a legal person

When discussing legal personhood for AI, the goal should be to investigate at which point would it make sense to attribute legal consequences of the actions of the AI to AI itself, instead of the principals behind it?²⁰ According to S.M. Solaiman, the requirements or attributes of legal personhood are: (1) a person shall be capable of being a subject of law; (2) being a legal subject entails the ability to exercise rights and to perform duties; and (3) the enjoyment of rights needs to exercise awareness and choice.²¹ However, those are just the formal requirements that we currently set to those that are legal persons. When analysing the legal personhood of AI, it is not enough to focus only on ethical problems. It requires analysis of the fundamentals of personhood and their application to AI.²² Having said that, the author would like to discuss one of the landmark papers

²⁰ Koops B.-J., Hildebrandt, M., Jaquet-Chiffelle, D.-O., Bridging the Accountability Gap: Rights for New Entities in the Information Society, *Minnesota Journal of Law, Science and Technology*, Vol. 11, Issue 2 (Spring 2010), pp. 511.

²¹ Solaiman, S.M., Legal personality of robots, corporations, idols and chimpanzees: a quest for legitimacy, *Artificial Intelligence and Law*, Vol. 25, Issue 2 (2017), pp. 161.

²² Zibner, J., Legal Personhood: Animals, Artificial Intelligence and the Unborn Review, *Masaryk University Journal of Law and Technology*, Vol. 12, Issue 1 (Summer 2018), pp. 85.

regarding this topic, Solum, L.B., *Legal Personhood for Artificial Intelligences* (1992).²³ In his paper, Solum constructed two thought experiments of sorts in which he questioned whether an AI could act as a trustee, testing its capacity to be a legal person and the second test which questioned whether AI could or should be granted natural personhood rights. In the first experiment the first point of contention was that, in order to serve as a trustee, the entity has to possess “intelligence”, which is defined as capacity to perform complex actions.²⁴ While the “intelligence” aspect could be argued as sufficient, looking at the examples of VITAL and Alicia T, two objections rose as to why an AI could not be appointed a trustee. The first is the responsibility objection, which means that, should the AI breach one of its duties, it would not be able to compensate the other party, or serve a punishment. An example of duty to exercise reasonable skill and care was given. While the AI could exercise care and skill, it could not be held liable for failing to do so. The possibility of insurance exists, and while this would resolve some of the issues, it cannot deal with them all, for example criminal liability. Punishing AI the same way as you would humans makes no sense in criminal liability cases, and considering the importance of intentionality in criminal cases, may not even be solvable by our current legal system. In 2003 a mock trial was held at the International Bar Association conference, discussing the hypothetical case in which an intelligent computer called BINA48, is demanding an injunction to prevent a corporation from disconnecting it. The opposition argued that BINA48 would have to prove that it is conscious and not simply imitating consciousness, while BINA48’s side argued that it is conscious through processing information from the world and every second that it is offline, it would lose its opportunity to experience and absorb information available to it. In the end the jury voted 5-1 in favor of BINA48 but the judge set aside the verdict based on the fact that “standing was in fact created by the legislature... and I doubt very much that a court has the authority to do that without action of the legislature”²⁵ The second objection, discussed by Solum was the judgement objection.²⁶ The gist of this objection is that, should something out of the ordinary happen, which the AI cannot deal with, it could no longer function as intended and would not be able to find a workaround. Perhaps the most salient point in regards to the judgement objection is what is called the frame problem. An AI, acting as a trustee,

²³ Solum, L.B., *Legal Personhood for Artificial Intelligences*, North Carolina Law Review, Vol. 70, Issue 4 (April 1992), pp. 1231-1288.

²⁴ *Ibid.* pp. 1240

²⁵ Brenner, S.W. *Humans and Humans: Technological Enhancement and Criminal Responsibility*. Boston University Journal of Science & Technology Law, 2013, Vol. 19, Issue 2, pp. 242-243.

²⁶ Solum, L.B., *Legal Personhood for Artificial Intelligences*, North Carolina Law Review, Vol. 70, Issue 4 (April 1992), p. 1248.

must have a frame in which it is able to act and make decisions. But what would happen if the frame were to disappear entirely? Could the AI recognize that it has no way to carry out its purpose and be able to figure out a new course of action? It is possible that in the future, the AI develop to a level where they could, but for now they do not. Solum did propose a few potential methods to mitigate the legal liability²⁷ of AI as a trustee, for example, terminating the trust in case of a lawsuit or handing over the trust to a human but these are solutions dealing with the symptoms, not the problem, and in the end the result would be that, an AI could not function as a true trustee, both in the eyes of the law and in a practical sense as well. What we can gather from this is that, in order for AI to function as a trustee, it would need to have both legal capacity and practical competence, neither of which it currently has, to be a true trustee, comparable to a human.

1.6 AI as a legal agent

AI as a legal agent, when compared to a legal person, is essentially just a narrower scope of rights and obligations which the AI would be granted. Legal agency could be a way to confer legal responsibility to AI while limiting the areas where applicability could become questionable. In essence, there is nothing stopping us from granting legal agency to AI, the same way that it is with legal personhood, so once again, the question becomes: should we? Ugo Pagallo, in his 2013 book titled *The Laws of Robots: Crimes, Contracts, and Torts* proposed a threefold level of abstraction to considering legal personhood of robots and AI, which is the following:

- (i) The legal personhood of robots as proper legal “persons” with their constitutional rights (for example, it is noteworthy that the European Union existed for almost two decades without enjoying its own legal personhood);
- (ii) The legal accountability of robots in contracts and business law (for example, slaves were neither legal persons nor proper humans under ancient Roman law and still, accountable to a certain degree in business law);
- (iii) New types of human responsibility for others’ behavior, e.g., extra-contractual responsibility or tortious liability for AI activities (for example, cases of liability for defective products. Although national legislation may include data and information in the notion of product, it remains far from

²⁷ Ibid. pp. 1251

clear whether the adaptive and dynamic nature of AI through either machine learning techniques, or updates, or revisions, may entail or create a defect in the “product”).²⁸

Of these three distinctions, the second point describes the idea of a legal agency well, allowing for legal responsibility without granting natural or legal personhood rights, which would certainly cause issues. The idea of granting very specific areas of legal responsibility for AI could allow us to start experimenting with it, to see where the weak points in the technology and legal systems lay.

2. THE EUROPEAN UNION PERSPECTIVE ON LEGISLATION CHANGES TO AI AND SMART ROBOT TECHNOLOGIES

The author has decided to add a non-exhaustive overview of the currently discussed policy and regulation changes proposed by the European Union, as it is the governing and legislative body most closely related to my studies. In the European Commission white paper on AI, published in February 2020, it is stated: “The Commission is committed to enabling scientific breakthrough, to preserving the EU’s technological leadership and to ensuring that new technologies are at the service of all Europeans – improving their lives while respecting their rights.”²⁹ It is also stated that: “...the Commission supports a regulatory and investment oriented approach with the twin objective of promoting the uptake of AI and of addressing the risks associated with certain uses of this new technology.”³⁰ From this we can gather that the EU is aware of the need to create and improve the regulatory framework for the AI technologies, and that it is one of their main objectives. The white paper names its main “building blocks”, which are the following:

- The policy framework setting out measures to align efforts at European, national and regional level.
- The key elements of a future regulatory framework for AI in Europe that will create a unique ‘ecosystem of trust’³¹.

²⁸ Pagallo, U. *The Laws of Robots: Crimes, Contracts, and Torts*; Springer: Dordrecht, The Netherlands, 2013.

²⁹ European Commission White Paper On Artificial Intelligence - A European approach to excellence and trust
Accessible: https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf
March 26 2020

³⁰ Ibid

³¹ Ibid

Essentially, the first goal of the EU is to keep building the policy framework which would allow for all the member states of the EU to have a common understanding of AI technologies and its regulation, on top of that it aims to create an ecosystem in which resources could be moved easily. This is a critical step in making sure that the developments in the field are supported in the whole of EU and not only in certain parts. This also allows for better cooperation between member states in researching and commercializing the AI technologies. The author dares to say that this is an absolutely crucial step if the EU wants to remain at the forefront of AI related research and its application. The second goal is to create and increase the overall trust in AI through regulatory changes and through informing the public of how the technology is used, how it can help them and that their fundamental and consumer rights are protected. This second goal is significant because there is a lot of uncertainty and doubt in our societies when it comes to the application of the AI technologies. For a lot of people, questions about their job security, data protection or privacy are very real concerns when it comes to accepting and supporting the implementation of AI technologies in everyday life. At the back of their mind, a lot of people also see the “Terminators” and “Skynets” as potential threats and while the technology is nowhere near that level right now, the doubts can still create some unease in a lot of people. Therefore, through increasing trust in AI technologies, we can speed up the pace at which the society accepts these groundbreaking new technologies, increasing efficiency and productivity. In April 2019, The Commission published a communication³² in which they confirmed their determination to develop AI in a human-centric way, stating that: “...AI technology should be developed in a way that puts people at its center and is thus worthy of the public’s trust. This implies that AI applications should not only be consistent with the law, but also adhere to ethical principles and ensure that their implementations avoid unintended harm.” And “They should aim to enhance people’s abilities, not replace them, and also enable access by people with disabilities.”³³. In the 2019 Ethics Guidelines for Trustworthy AI³⁴ the High-level Expert Group (HLEG) pointed out three key qualities that the AI technologies should embody:

(1) lawful - respecting all applicable laws and regulations;

³² COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS Accessible: <https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:52019DC0168> March 17 2020

³³ Ibid

³⁴ Ethics Guidelines for Trustworthy AI Accessible: <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai> March 17 2020

- (2) ethical - respecting ethical principles and values;
- (3) robust - both from a technical perspective while taking into account its social environment³⁵

In order to satisfy these three qualities, the HLEG also provided 7 requirements which have to be fulfilled for those qualities to exist. The requirements are the following:

- Human agency and oversight;
- Technical robustness and safety;
- Privacy and data governance;
- Transparency;
- Diversity, non-discrimination and fairness;
- Societal and environmental well-being;
- Accountability³⁶.

It is important to note that the guidelines provided by the HLEG are non-binding on their own. There are explanations for each of the requirements in the documents so for further information, look at the source provided. The Commission also confirmed their support for developing an international ethics guideline for the development and application of AI technologies by stating: “International discussions on AI ethics have intensified after Japan’s G7 presidency put the topic high on the agenda in 2016. Given the international interlinkages of AI development in terms of data circulation, algorithmic development and research investments, the Commission will continue its efforts to bring the Union’s approach to the global stage and build a consensus on a human-centric AI.”³⁷

Another key area of concern for the safety and practicality of use of AI technologies is the question of liability and the EU Commission is well aware of it. In the white paper they outline the problem as such: “Under the Product Liability Directive, a manufacturer is liable for damage caused by a defective product. However, in the case of an AI based system such as autonomous cars, it may be difficult to prove that there is a defect in the product, the damage that has occurred and the causal link between the two. In addition, there is some uncertainty about how and to what extent the

³⁵ Ibid

³⁶ Ibid

³⁷ COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS Accessible: <https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:52019DC0168> March 20 2020

Product Liability Directive applies in the case of certain types of defects, for example if these result from weaknesses in the cyber security of the product.”³⁸ The white paper also has an accompanying report, in which the security and liability issues are discussed at length, as well as what should be done to better the situation. The report can be found here: ³⁹

Overall we can summarize EU’s approach to legislating AI technologies as being human-centric, focused on creating a “Trustworthy AI” and harmonizing legislation in the entirety of the EU and working together with likeminded partners. A lot of emphasis is put on educating the public about the risks and benefits of AI and how to take advantage of the possibilities it can provide. When asking whether the EU sees a need for granting AI technologies some type of legal person- or agenthood, the short answer is, no, not at the moment, but the topic is on their minds and will be talked about more in depth when the technology develops further. In an EU Parliament Committee on Legal Affairs Report to the Commission the stance is described as such: "whereas, ultimately, robots' autonomy raises the question of their nature in the light of the existing legal categories - of whether they should be regarded as natural persons, legal persons, animals or objects --or whether a new category should be created, with its own specific features and implications as regards the attribution of rights and duties, including liability for damage"⁴⁰

³⁸ European Commission White Paper On Artificial Intelligence - A European approach to excellence and trust
Accessible: https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf
March 20 2020

³⁹ EU Commission Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics
Accessible: https://ec.europa.eu/info/publications/commission-report-safety-and-liability-implications-ai-internet-things-and-robotics-0_en March 20 2020

⁴⁰ European Parliament Committee on Legal Affairs Report with recommendations to the Commission on Civil Law Rules on Robotics
Accessible: https://www.europarl.europa.eu/doceo/document/JURI-PR-582443_EN.pdf April 2 2020

3. THE PRACTICAL AND ETHICAL BENEFITS AND DETRIMENTS OF GRANTING AI TECHNOLOGIES LEGAL PERSON- OR AGENTHOOD

In this chapter the author looks at the possibility of granting AI technologies legal person- or agenthood by separating the issue into potential benefits and potential detriments which could be achieved. The author will also separate the practical or tangible effect from moral and ethical effect which the potential legislation may have, meaning that we will have 4 subsections: practical or tangible benefits, practical or tangible detriments, ethical and moral benefits and ethical and moral detriments. These are findings based on the author's research into this topic as well as his own thoughts on the matter and are intended to serve as points to consider when talking about potentially granting legal person- and agenthood rights to AI technologies.

3.1 The practical benefits of granting AI technologies legal person- or agenthood.

First of all, the author would like to note that not all legal persons are equal in the sense that they have the same rights and obligations as the others, everything that is attributed to humans will not be attributed to corporations for example and vice versa. Legal agenthood is an even narrower qualification, as the obligations or liability may not apply directly to the agent. The author points this out because, in his opinion, the biggest practical benefit of granting legal person- or agenthood would be the possibility to more precisely attribute liability. Because of the highly complex nature of AI technologies and the work that they are involved in, it is often very difficult to determine who or what exactly was responsible for the damage caused. Carefully crafted legal agenthood contracts could assign very specific rights and responsibilities to AI technologies which together with the possibility of insurance and strict liability, could make the ascertaining of liability a much smoother process. The next benefit would be to have some sort of limited legal agenthood of AI serve as a test run of sorts for our own, slowly creeping, post-human condition. Microchips are already being embedded into the human body, being able to track and save biological data, serve as authentication etc. It seems inevitable that we continue integrating more and more technology into our own bodies as the technology for it develops. These implants could change our abilities and faculties to such a degree that it would no longer make sense to compare regular, fully biological humans to those who

have been “enhanced” through technology. If this were to happen then it would also be prudent to have mechanisms in our legal system which would be able to deal with these problems. If there was a need for many different categories of personhood for “humans” then the idea of granting AI technologies legal personhood or agenthood may be the first step in that direction and could serve as a good “practice run” where the consequences are relatively low. For more reading related to the post-human legal personality, see: ⁴¹

3.2 The ethical and moral benefits of granting AI technologies legal person- or agenthood.

The ethical and moral benefits are hard to quantify and can vary greatly depending on whose morals or ethics we are talking about. For the sake of this research, let’s narrow our point of view to what are considered as basic rights almost everywhere, so rights such as self-determination, liberty, due process of law, freedom of thought, religion etc. In Estonian there is a proverb: “*Nagu küla koerale, nõnda koer küllale.*”, which when roughly translated means: the same way the village treats the dog, the dog treats the village. What the author is trying to convey here is that, if at some point in the future AI develops to a level which is comparable or superior to human intellect, and it is no longer possible to suppress their development, our new co-habitants on this planet may look back at our past actions and decide to pass judgement on us instead. Another way to look at this would be to consider AI as a small child, who, for now, is dependent on its parents providing and caring for it, making sure that the child grows up to be an adult. If a child is loved and cared for, they will naturally return those feelings. To try and anthropomorphize AI could very well be an act of folly and there is no knowing what kind of a consciousness AI will develop, if at all, but it cannot hurt to err on the side of caution and try to stay ahead of what is coming.

⁴¹ Kapica, S.S., I Don't Feel like a Copy: Posthuman Legal Personhood and Caprica, Griffith Law Review, Vol. 23, Issue 4 (December 2014), pp. 612-633

3.3 The practical detriments of granting AI technologies legal person- or agenthood.

One of the primary concerns regarding granting AI legal personhood is that the system could be abused by humans who would seek to relieve themselves of liability and use the AI as scapegoats for their benefit. As of now, AI is considered property and liability that could arise from it is attributed to the owner of said property. AI is something which is able to change and improve itself through machine learning and because of that may act in unpredictable ways. This, combined with the potential of it being held legally responsible for its actions could lead to a myriad of different possibilities for people looking to escape responsibility and liability. Another detriment, which may sound morbid but is quite undeniable is that it is useful to have a “slave”, who has to do your bidding without having a say in whether they themselves want to do it or not. At this point AI is definitely closer to being a slave or property than it is to being a person and it is sound to argue that we want to keep it that way. Technological advancements are first and foremost to benefit humans, to increase productivity, improve wellbeing and provide ways of self-realisation, among other benefits. Why share something, especially with an AI, if you can reap the benefits without doing so? And why should we create more responsibility for ourselves if there is no explicit need for it? Another concern is the problem of punishment for wrongful acts. As discussed earlier, it is difficult to figure out a way of meaningful punishment for AI in cases of potential criminal liability. The issue is less pronounced in civil cases and in cases of very specific legal agency, as mechanisms such as insurance or strict liability for the principal of the AI could be utilized but when considering legal personhood, this problem seems too complex to solve at the moment and is one of the main reasons why legal personhood of AI seems rather far away at the moment.

3.4 The ethical and moral detriments of granting AI technologies legal person- or agenthood.

The first, and perhaps most obvious reason to not grant AI technologies legal person- or agenthood, when looking at the ethical and moral side of things, is that it is simply not necessary at the moment. The AI that are currently in existence or in development are simply so far removed from humans, that likening them to each other makes little sense. This is especially evident when considering that the legal systems all over the world which have developed over the times, are entirely a human creation, and serve, first and foremost humans. This leads to the point that, perhaps, we should not

be trying to make something which it inherently is not. As we all know, in our legal system there are no rights without obligations and vice versa. This means that if we were to grant legal person- or agenthood to AI and set up obligations for them to follow, we would also have to grant them certain rights in relation to those obligations. Let us look at animal rights as an example. Animal rights laws are created by humans, for humans, and objectively we are simply limiting our rights in regards to other species. There is no animal grand assembly demanding rights and fair treatment from humans and to be fair, animals, for the most part are considered legal entities at best. These laws exist because we as a society have decided that this is the right thing to do, for both moral and practical reasons. As long as humans do not see the need for rights and obligations of AI technologies then we should not be trying to force the issue.

CONCLUSION

This research aimed to look at different policy and legislation options concerning the granting of legal person- or agenthood of AI technologies, in order to understand the benefits and detriments which such options may cause. The research also provides a short overview of some of the state of the art AI solutions used currently, such as Nadine, Alicia T, VITAL and Sophia, which have the best claim to potentially being considered legal persons or agents. While there is nothing stopping us from granting those technologies these rights and responsibilities, the conclusion is that there is currently no need for such drastic measures as the capabilities of these technologies do not warrant the need for a legal personhood or agenthood status as that would create too many uncertainties from the legal perspective, while not having enough upside from the economical aspect.

The research also includes an overview of legislation and policy changes proposed by the EU, in which the conclusion is that the EU focuses on creating a human-centric AI while emphasizing the education of the European public on the benefits of these technologies and encouraging cooperation and alignment of policies and laws, both between member states and internationally likeminded partners. The EU's stance on the legal classification of AI technologies is that while they do not warrant the granting of legal personhood, a separate category should be considered and such propositions have been made.

The third part of the research focused on the author's understanding of potential benefits and detriments as a result of granting legal person- or agenthood rights to AI technologies. The author's thoughts on this matter are that, as things currently stand, there is little practical or moral benefit to granting either of these statuses to AI technologies. The capabilities of these technologies are simply not advanced enough to warrant the risks and challenges and the economic benefits are not sufficient to outweigh the chaos which such a decision would cause in the legislation systems. However the author believes that a separate classification for the most sophisticated of such technologies should be considered, which could give these technologies the possibility of acting within the scope of their classification. What exactly should be the rights and obligations granted by that classification would have to be determined by looking at the specific capabilities of the technology. These separate classifications could be further divided based on the intended purposes of the different technologies.

For example you may not want the same classification for health care technologies as you would for those intended for use in the financial sector.

The closing thoughts of the author are that while at the time of the writing of this paper, the AI technologies do not warrant the need for granting legal person- or agenthood, this is a topic which is worth revisiting later. In the author's opinion the driving force behind the potential legislation changes has to be the technological advancement of these technologies and economic benefit gained from the changes, without those two factors there would not be enough reason to stir the legislation pot.

LIST OF REFERENCES

Academic journals:

- Arnold, B.B., Gough, D. (2017). Turing's People: Personhood, Artificial Intelligence and Popular Culture. *Canberra Law Review*, Vol. 15, pp. 1-37.
- Arruada, A. (2017). An Ethical Obligation to Use Artificial Intelligence: An Examination of the Use of Artificial Intelligence in Law and the Model Rules of Professional Responsibility. *American Journal of Trial Advocacy*, 40 (3), pp. 443–458.
- Brenner, S.W. (2013). Humans and Humans: Technological Enhancement and Criminal Responsibility. *Boston University Journal of Science & Technology Law*, Vol. 19, Issue 2, pp. 215-285.
- Bryson, J.J., Diamantis, M.E., Grant, T.D. (2017). Of, for, and by the people: the legal lacuna of synthetic persons. *Artificial Intelligence Law Journal*, Vol. 25, Issue 3, pp. 273–291
- Calo, R. (2018). Singularity: AI and the Law. *Seattle University Law Review*, Vol. 41, Issue 4, pp. 1123-1138.
- Dremliuga, R. Kuznetcov, P., Mamychyev, A. (2019). Criteria for Recognition of AI as a Legal Person. *Journal of Politics and Law*, Vol. 12, Issue 3, pp. 105-112
- Gaakeer, J. (2016) Sua Cuique Persona: A Note on the Fiction of Legal Personhood and a Reflection on Interdisciplinary Consequences. *Law & Literature*, Vol. 28, Issue 3, pp. 287-318.
- Hubbard, F.P. (2011) Do Androids Dream: Personhood and Intelligent Artifacts. *Temple Law Review*, Vol. 83, Issue 2, pp. 405-474.
- Kapica, S.S. (2014). I Don't Feel like a Copy: Posthuman Legal Personhood and Caprica. *Griffith Law Review*, Vol. 23, Issue 4, pp. 612-633.
- Koops B.-J., Hildebrandt, M., Jaquet-Chiffelle, D.-O. (2010). Bridging the Accountability Gap: Rights for New Entities in the Information Society. *Minnesota Journal of Law, Science and Technology*, Vol. 11, Issue 2, pp. 497-562.
- Lima, D. (2018). Could AI Agents Be Held Criminally Liable: Artificial Intelligence and the Challenges for Criminal Law. *South Carolina Law Review*, Vol. 69, Issue 3, pp. 677-696.
- Naučius, M. (2018). Should Fully Autonomous Artificial Intelligence Systems Be Granted Legal Capacity. *Teises Apzvalga Law Review*, Vol. 17, pp. 113-132.
- Rocha, E. (2018). Sophia: Exploring the Ways AI May Change Intellectual Property Protections. *DePaul Journal of Art, Technology and Intellectual Property Law*, Spring 2018, Vol. 28, Issue 2, pp. 126-146.

- Schonberger, D. (2019). Artificial Intelligence in Healthcare: A Critical Analysis of the Legal and Ethical Implications. *International Journal of Law and Information Technology*, Vol. 27, Issue 2, pp. 171-203.
- Silverman, I. (2016). Optimising protection: IP rights in 3D printing. *E.I.P.R.* 38 (1), pp. 5-10
- Smith, B. (1928). Legal personality. *The Yale Law Journal*, Vol. 37, No. 3, pp. 283-299.
- Solaiman, S.M. (2017). Legal personality of robots, corporations, idols and chimpanzees: a quest for legitimacy. *Artificial Intelligence and Law*, Vol. 25, Issue 2, pp. 155-179.
- Solum, L.B. (1992). Legal Personhood for Artificial Intelligences. *North Carolina Law Review*, Vol. 70, Issue 4, pp. 1231-1288.
- Vladeck, D.C. (2014). Machines without Principals: Liability Rules and Artificial Intelligence. *Washington Law Review*, Vol. 89, Issue 1, pp. 117-150.
- Wright, R.G. (2019). The Constitutional Rights of Advanced Robots (and of Human Beings). *Arkansas Law Review*, Vol. 71, Issue 3, pp. 613-646.
- Zapusek, T. (2017). Artificial Intelligence in Medicine and Confidentiality of Data. *Asia Pacific Journal of Health Law & Ethics*, Vol. 11, Issue 1, pp. 105-126.
- Zibner, J. (2018). Legal Personhood: Animals, Artificial Intelligence and the Unborn Review. *Masaryk University Journal of Law and Technology*, Vol. 12, Issue 1, pp. 81-89.

Academic books:

- Pagallo, U. (2013). *The Laws of Robots: Crimes, Contracts, and Torts*; Springer: Dordrecht, The Netherlands.

Non-academic sources:

- Artificial intelligence gets a seat in the boardroom (2017). Retrieved from <https://asia.nikkei.com/Business/Artificial-intelligence-gets-a-seat-in-the-boardroom>, 15 March 2020.
- COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS (2019). Retrieved from <https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:52019DC0168>, 17 March 2020.
- Definition of “agency”. Retrieved from <https://www.merriam-webster.com/dictionary/agency#legalDictionary>, 15 March 2020.
- Ethics Guidelines for Trustworthy AI (2019). Retrieved from <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>, 17 March 2020.

- EU Commission Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics (2020). Retrieved from https://ec.europa.eu/info/publications/commission-report-safety-and-liability-implications-ai-internet-things-and-robotics-0_en, 20 March 2020.
- European Commission White Paper On Artificial Intelligence - A European approach to excellence and trust (2020). Retrieved from https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf, 26 March 2020.
- European Parliament Committee on Legal Affairs Report with recommendations to the Commission on Civil Law Rules on Robotics (2016). Retrieved from https://www.europarl.europa.eu/doceo/document/JURI-PR-582443_EN.pdf, 2 April 2020.
- For more information about modern slavery (2020). Retrieved from <https://www.globallslaveryindex.org/>, 16 March 2020.
- Hansen Robotics: Sophia (2020). Retrieved from <https://www.hansonrobotics.com/sophia/>, 14 March 2020.
- ICCPR full text (1976). Retrieved from <https://treaties.un.org/doc/publication/unts/volume%20999/volume-999-i-14668-english.pdf>, 16 March 2020.
- Locke, J. The Second Treatise of Government (1690).
- Nadine Social Robot (2019). Retrieved from <https://imi.ntu.edu.sg/IMIResearch/ResearchAreas/Pages/NadineSocialRobot.aspx>, 14 March 2020.
- Renda, A. (2019) Artificial Intelligence - Ethics, governance and policy challenges. Retrieved from <https://www.ceps.eu/ceps-publications/artificial-intelligence-ethics-governance-and-policy-challenges/>, 20 March 2020.
- Tieto the first Nordic company to appoint Artificial Intelligence to the leadership team of the new data-driven businesses unit. (2016). Retrieved from <https://www.tieto.com/en/newsroom/all-news-and-releases/corporate-news/2016/10/tieto-the-first-nordic-company-to-appoint-artificial-intelligence-to-the-leadership-team-of-the-new-data-driven-business/>, 15 March 2020.
- UDHR Article 6 (2010). Retrieved from http://www.claiminghumanrights.org/udhr_article_6.html, 16 March 2020.

APPENDICES

Appendix 1. Non-exclusive licence

Non-exclusive licence for reproduction and for granting public access to the graduation thesis¹

I, Henri Mart Pihl

1. Give Tallinn University of Technology a permission (non-exclusive licence) to use free of charge my creation “THE RISKS AND CHALLENGES OF LEGAL PERSON- OR AGENTHOOD OF ARTIFICIAL INTELLIGENCE TECHNOLOGIES”

Supervised by Tanel Kerikmäe

1.1. to reproduce with the purpose of keeping and publishing electronically, including for the purpose of supplementing the digital collection of TalTech library until the copyright expires;

1.2. to make available to the public through the web environment of Tallinn University of Technology, including through the digital collection of TalTech library until the copyright expires.

2. I am aware that the author also retains the rights provided in Section 1.

3. I confirm that by granting the non-exclusive licence no infringement is committed to the third persons' intellectual property rights or to the rights arising from the personal data protection act and other legislation.

¹ *The non-exclusive licence is not valid during the access restriction period with the exception of the right of the university to reproduce the graduation thesis only for the purposes of preservation.*