

Why Different Types of Algorithms Require Different Legal Rules

Tjaša Zapušek, Faculty of Law, University of Copenhagen

INTRODUCTION

The involvement of artificial intelligence (AI) in our everyday life will call into question the efficiency of existing legal rules. Inner reorganization of legal codes is crucial, since modern machines will no longer be considered simple tools in the hands of their user, owner etc. Therefore, who will be held liable for the harm caused by robots' actions? In this paper, there will be (briefly) presented differences between four different 'decision-making approaches (by the AI)' that come as a consequence of four different types of systems, AI can be equipped with (deterministic algorithm, supervised/unsupervised learning algorithm and artificial neural network). Through this systematization I will try to visualize why we need different legal approaches for different kinds of assignments. In the second part I will expose legal barriers for the imposition of legal personhood and try to ground them on the concept of dignity and philosophical interpretation of personhood itself. Moreover, the paper will reveal the arguments, why legal rules do not allow us to treat AI solely as simple tools and in the conclusion there will be presented a recommended framework for legal reorganization and adjustment of existing legal rules.

The question I am trying to answer in this paper is 'how to regulate advanced robots and how to legally evaluate their actions and potential harms that can derive by their actions.'

1. SYSTEMATIZATION OF AI

1.1. DETERMINISTIC ALGORITHM¹

¹ Mireille Hildebrandt, *Smart Technologies and the End(s) of Law* (Edward Elgar, 2016), 23.

This algorithm is constructed to prevent the machine from making autonomous decisions, creating new patterns and consequently creating unpredictable outcomes.² Robots equipped with deterministic algorithms follow predefined paths.³ The major significance that distinguishes the first group of algorithms from the others is that the final action/machine's response/steps have already been predicted by a human being.⁴ In the case of a deterministic algorithm, we will always find someone liable for a robot's actions - as it will be presented in the text below, existing legal rules⁵ will suffice.

1.2. SUPERVISED LEARNING ALGORITHMS⁶ (enable machine learning)

More advanced forms of algorithms can enable machine learning.⁷ In the case of supervised learning algorithms, a machine is fed with a so-called training set by human supervisors, which consists of particular data and predefined patterns that provide some sort of a background⁸ on what counts as a desired and satisfactory solution/outcome.⁹ A simple example of machine learning (in general) is photo sorting where we are not certain which sorting criteria will be chosen by the machine (hair color, gender, skin color, height etc.)¹⁰ If the machine was equipped with su-

² Zapušek, T. "Artificial intelligence in medicine". *The Asia Pacific Journal of Health Law & Ethics*, Vol. 11., p. 117.

³ Mireille Hildebrandt, *Smart Technologies and the End(s) of Law* (Edward Elgar, 2016), 23.

⁴ Ibid.

⁵ Existing legal rules regarding the liability of a producer for defective product, the liability of employer for dangerous machine, owners liability for his/her property etc.

⁶ Hildebrandt, 24, 25.

⁷ Hildebrandt, M. (2016). *Smart Technologies and the End(s) of Law*. Edward Elgar Publishing, p. 24 – 25.

⁸ Ploj, B. (2013). »Metoda mejnih parov za učenje umetnih nevronske mreže«. *Doktorska dizertacija*, Fakulteta za elektrotehniko, računalništvo in informatiko Maribor, p. 7.

⁹ Zapušek, T. "Artificial intelligence in medicine". *The Asia Pacific Journal of Health Law & Ethics*, Vol. 11., p. 117.

¹⁰ Zapušek, T. (2017). »Umetna inteligenca in njena pravna ureditev«. *Analiza*, Društvo za analitično filozofijo in filozofijo znanosti.

pervised learning algorithm, we would know which criteria were used.¹¹ Regular training may equip the machine with the ability to recognize new data, while using enhanced pattern recognition skills. As a consequence, the agent may come up with similar or even better solutions than those that would have been provided by its human supervisor, if he had concluded the task by himself.¹² As we will see in the paragraphs that present legal evaluation, this ‘*sense of unpredictability*’ prevents us to use existing legal rules¹³.

1.3. UNSUPERVISED LEARNING ALGORITHMS¹⁴ (*enable machine learning*)

The other advanced form of algorithms is an unsupervised learning algorithm, which in comparison with a supervised learning algorithm, makes correlations between obtained data without previously hypothesizing them.¹⁵ Without the provision of a training set, consisting of data and predefined patterns, the machine cannot familiarize itself with information and what counts as a desirable solution.¹⁶ A machine that is in the process of unsupervised learning finds input data and classifies them (according to its own criteria) in different categories.¹⁷ Such machines are capable of mining new data and creating novel, unexpected patterns. Many scientific articles hold the opinion that the result of unsupervised learning is completely unpredictable which stigmatized this type of machine learning as undesirable.¹⁸ On this point I want to emphasize the difference

¹¹ Ploj, B. (2013). »Metoda mejnih parov za učenje umetnih nevronske mreže«. *Doktorska dizertacija*, Fakulteta za elektrotehniko, računalništvo in informatiko Maribor, p. 8 – 10.

¹² Mireille Hildebrandt, *Smart Technologies and the End(s) of Law* (Edward Elgar, 2016), 23.

¹³ Existing legal rules regarding the operation of the machine.

¹⁴ *Ibid.*

¹⁵ Zapušek, T. “Artificial intelligence in medicine”. *The Asia Pacific Journal of Health Law & Ethics*, Vol. 11., p. 117.

¹⁶ *Ibid.*

¹⁷ Ploj, B. (2013). »Metoda mejnih parov za učenje umetnih nevronske mreže«. *Doktorska dizertacija*, Fakulteta za elektrotehniko, računalništvo in informatiko Maribor, p. 6.

¹⁸ Zapušek, T. (2017). »Umetna inteligenca in njena pravna ureditev«. *Analiza*, Društvo za analitično filozofijo in filozofijo znanosti.

that exist between defining the final goal and defining the most desirable outcome/result. We draw a wrong conclusion when we claim that the absence of strictly defined steps 'how to solve a problem' results in complete autonomy of robot and we, as individuals, do not have any influence regarding the outcome.¹⁹ The fact is that robot (while operating) tends to satisfy the goal set by a human.²⁰ Example of unsupervised machine learning is photo sorting where we are not certain which sorting criteria will be chosen by the machine (hair color, gender, skin color, height etc.)²¹ Here, we can not talk about the absence of predictability of a final goal. We wanted to have our photos sorted and we have them. Indefinableness of desirable result does not automatically mean the absence of a goal (set by a human being).²² However, even though the goal has been set by a human, it does not exclude the '*sense of unpredictability*' regarding machine's steps during its operation which prevents us to use existing legal rules²³.

1.4. ARTIFICIAL NEURAL NETWORKS

Artificial neural network is slightly different, since it presents a computing system where central concept was borrowed from the analogy of biological neural networks. Human brains are consisted of numerous nerve cells - neurons that presents a switch with information input and output.²⁴ This switch is activated if there are enough stimuli of other neuron that are hitting the information input. Incoming signal from other nerve cels is transferred to a neuron by connec-

¹⁹ Ibid.

²⁰ Ibid.

²¹ Ploj, B. (2013). »Metoda mejnih parov za učenje umetnih nevronske mreže«. *Doktorska dizertacija*, Fakulteta za elektrotehniko, računalništvo in informatiko Maribor, p. 8-10.

²² Zapušek, T. (2017). »Umetna inteligenca in njena pravna ureditev«. *Analiza*, Društvo za analitično filozofijo in filozofijo znanosti.

²³ Existing legal rules regarding the operation of the machine.

²⁴ Kriesel, D. (2017). *A brief introduction to Neural Networks*. dkriesel.com, p. 15 - 20.

tions, the synapses²⁵.²⁶ They can be found at the dendrites²⁷ of a neuron or sometimes directly at the soma²⁸.²⁹ The goal of the emulation of human brains (copying a function of biological nervous systems into software³⁰) is the creation of systems that imitate real biological organisms in terms of behavior and inner causal structures. Robots, equipped with emulated brains would likely be able to feel pain and have emotions.³¹ Brain emulation “is different from entirely artificial software in that it deliberately tries to be as similar as possible to morally considerable biological systems, and this should make us more ethically cautious than with other software.”³² According to Ploj, use of artificial neural networks may, due to their enhanced capabilities and independence in ‘decision making’ process, lead to the situation when robot’s knowledge outdoes the human’s one.³³ As we will see in the paragraphs that present legal evaluation, artificial neural networks are, according to my interpretation, the only one to which we could potentially ascribe legal personhood.

2. ARTIFICIAL INTELLIGENCE AS A LEGAL PERSON

²⁵ Synapses are connections between the axon (they serve like cables through which neurons send the information) and other dendrites. *Tutorials Point, Artificial Neural Network, 2017, 3.*

²⁶ Kriesel, D. (2017). *A brief introduction to Neural Networks*. dkriesel.com, p. 15 - 20.

²⁷ Dendrites are tree-like branches that are responsible for receiving the information from other neurons it is connected to. In other words, they are the ears of neuron. *Tutorials Point, Artificial Neural Network, 2017, 3.*

²⁸ Soma is the cell body of the nerve cell and is responsible for processing of information, they have received from dendrites. *Tutorials Point, Artificial Neural Network, 2017, 3.*

²⁹ Kriesel, D. (2017). *A brief introduction to Neural Networks*. dkriesel.com, p. 15 - 20.

³⁰ Sandberg, A. (2013). »Ethics of brain emulations (draft)«. Future of Humanity Institute, Oxford University, p. 1.

³¹ Sandberg, A. (2013). »Ethics of brain emulations (draft)«. Future of Humanity Institute, Oxford University, p. 8.

³² Sandberg, A. (2013). »Ethics of brain emulations (draft)«. Future of Humanity Institute, Oxford University, p. 12.

³³ Ploj, B. (2013). »Metoda mejnih parov za učenje umetnih nevronske mreže«. *Doktorska dizertacija*, Fakulteta za elektrotehniko, računalništvo in informatiko Maribor, p. 8 – 10.

First, I want to briefly evaluate the potential option of imposition of legal personhood on robots. It is just like with every other problem we are dealing with in our lives, if we want to understand the concept, namely we need to go back to the past and see what the reason was for its emergence.

2.1. Dignity as the basis of personhood

Theories of human dignity have their origin in the antique. Roman politician and lawyer Cicero believed that every single human being is endowed with dignity for the sole fact of his existence.³⁴ Pursuant to his interpretation, “superior mind” that is possessed exclusively by human beings gave them the capacity to think and to be self-aware.³⁵ Apart from his theories, the major public opinion that prevailed in the ancient Rome was completely different.³⁶ In Roman times the word *dignitas* was typically ascribed to the holder of a high office because of his outstanding achievements that were publicly recognized. Hence, the reason for the recognition of dignity was related to an individual’s social or political status.³⁷ Dramatic turn in people’s thinking occurred in the Renaissance, also known as the age of enlightenment. This was the era that brought forth major progress in the sense of the interpretation of dignity as such. It was a time of great philosophers, among which was also the Italian humanist Giovanni Pico Della Mirandola³⁸. His view differed from traditional Christian views, since he believed that a conception of human dignity encompasses not only the idea of being created in God’s image, but also the idea of the creation of

³⁴ Glensy, R. D. (2011). “The right to dignity”. *Columbia Human Rights Law Review*, p. 74.

³⁵ Ibid.

³⁶ Ibid.

³⁷ Stancioli, B. Ribeiro, D. (2016). “Dignity as constraint and as freedom: On the meanings of human dignity for transhuman world”. *New Approaches to the Personhood in Law*, Peter Land Edition, p. 28.

³⁸ “*Thou, constrained by no limits, in accordance with thine own free will, in whose hand We have placed thee, shalt ordain for thyself the limits of thy nature. We have set thee at the world's center that thou mayest from thence more easily observe whatever is in the world. We have made thee neither of heaven nor of earth, neither mortal nor immortal, so that with freedom of choice and with honor, as though the maker and molder of thyself, thou mayest fashion thyself in whatever shape thou shalt prefer. Thou shalt have the power to degenerate into the lower forms of life, which are brutish. Thou shalt have the power, out of thy soul's judgment, to be reborn into the higher forms, which are divine.*”

one's own unique destiny in a life of an individual.³⁹ Ascription of dignity to a person due to his unique freedom of will as a gift from God⁴⁰ and more importantly, his power to shape himself according to his freedom of will was the core idea behind Mirandola's thoughts, which due to their importance present a milestone in the history of the idea of human dignity. To sum up, a human being is special and unique, since he can become whatever he wants. Even to this day, religion is an important source for the interpretation of human dignity. As it has been stated by the Second Vatican Council, human dignity is defined as being enriched by reason and free will. This has created the inherently human need to seek truth and to abide by the truth.⁴¹ Pope Benedict XVI once said that human dignity is the locus of human rights which are the source of higher dimensions "based on the natural law inscribed on human hearts to which the law must yield".⁴² The other significant milestone was achieved by the German philosopher Immanuel Kant, who said:

*"In the kingdom of ends everything has either a price or a dignity. What has a price can be replaced by something else as its equivalent; what on the other hand is raised above all price and therefore admits of no equivalent has a dignity."*⁴³

In his work *Groundwork of the Metaphysics of Morals*, he claims that human beings have dignity, which leads him to the conclusion that they may never be treated merely as means, and vice versa that non-human beings, which have a price, may be treated exclusively as means. Because things have a price, they can be replaced by something equivalent.⁴⁴ Morality on the other hand that is exclusively attached to a human being is the condition under which a rational being can be

³⁹ Stancioli, B. Ribeiro, D. (2016). "Dignity as constraint and as freedom: On the meanings of human dignity for transhuman world". *New Approaches to the Personhood in Law*, Peter Land Edition, p. 28.

⁴⁰ Glensy, R. D. (2011). "The right to dignity". *Columbia Human Rights Law Review*, p. 74.

⁴¹ Glensy, R. D. (2011). "The right to dignity". *Columbia Human Rights Law Review*, p. 75.

⁴² Pope Benedict XVI, supra note 17.

⁴³ Kant, I. (1998). *Groundwork of the Metaphysics of Morals*. Cambridge University Press, p. 42.

⁴⁴ Atwell, J.E. (1986). *Ends and principles in Kant's moral thought* Springer Science and Business Media, p. 153.

an end in itself. The reason for it is the fact that it is possible only through morality to become a legislative member.⁴⁵ Both Mirandola and Kant come to the conclusion that what makes human beings dignified and special is their unique ability to act freely.⁴⁶ We can decide between right or wrong. Robots on the other hand are victims of their own perfectionism. Even though we are not able to predict their further steps, we can be sure that their actions will be optimal according to the given conditions.

Today, one version of the understanding of the concept of human dignity is that it is special, intrinsic and the absolute value that all human beings equally possess. It demands a respect of humanness and is grounded in properties that by nature belong exclusively to human beings.⁴⁷ According to the Nordenfelt there exist four different notions/identifications of dignity as such⁴⁸: dignity as an intrinsic value of all humans; dignity as a quality that indicates individual merit; dignity as a moral reputation of an individual; and dignity as an identity of an individual or a respect that an individual has his own past, his future and set of relationships with others.⁴⁹ These four identifications of dignity possess some common elements⁵⁰, (1) Dignity refers to a special dimension of value, (2) dignity of an individual is worthy of respect of others and of himself, (3) dignity has its own ground, constituted by a set of properties that belong to the subject.⁵¹ As we can see, there is no common consensus regarding the meaning of *dignity*. In my further analysis I will interpret dignity as a distinguishing feature that is shared by human beings.

⁴⁵ Stancioli, B. Ribeiro, D. (2016). "Dignity as constraint and as freedom: On the meanings of human dignity for transhuman world". *New Approaches to the Personhood in Law*, Peter Land Edition, p. 29.

⁴⁶ Ibid, 29.

⁴⁷ Ibid, 27.

⁴⁸ Nordenfelt, L. (2004). "The varieties of dignity". *Health Care Analysis*, Vol. 12, No. 2., p. 69-81.

⁴⁹ Stancioli, B. Ribeiro, D. (2016). "Dignity as constraint and as freedom: On the meanings of human dignity for transhuman world". *New Approaches to the Personhood in Law*, Peter Land Edition, p. 27.

⁵⁰ Nordenfelt, L. (2004). "The varieties of dignity". *Health Care Analysis*, Vol. 12, No. 2., p. 70.

⁵¹ Stancioli, B. Ribeiro, D. (2016). "Dignity as constraint and as freedom: On the meanings of human dignity for transhuman world". *New Approaches to the Personhood in Law*, Peter Land Edition, p. 27.

2.1. Transition from human dignity to concept of personhood

Even though the concept of humanity and a moral and legal notion of personhood are not synonymous⁵², it is still rooted, especially in Western societies that “*the concept of human dignity as relational empowerment exalts and protects exactly those human features that enable personhood to emerge, for example, self-consciousness, being able to act autonomously, and flourishing in social-communicative mediums.*”⁵³ Thus, according to the interpretations by numerous philosophers, it seems that the core of humanity is the human propensity to become a person, to act as a moral agent, and most importantly to possess not only a sense of self-identity yet also a sense of self-purpose that is not determined by circumstances but rather emerges from them.⁵⁴ Pursuant to the standard legal account of the term “personhood”, a person is recognized as an entity that is capable to hold right and obligations⁵⁵ and is therefore protected by law⁵⁶. We comprehend humanity as the state of being human, which is inextricably linked to the legal status of being a person.

2.2. Concept of personhood

The philosophical tradition has provided us with two concepts that clarify two main ideas behind the term *person*. This distinction will help me support my arguments and ideas regarding legal

⁵² Singer, P. (2001). *Practical Ethics*. Cambridge University Press, p. 18-19.

⁵³ Stancioli, B. Ribeiro, D. (2016). “Dignity as constraint and as freedom: On the meanings of human dignity for transhuman world”. *New Approaches to the Personhood in Law*, Peter Land Edition, p. 30.

⁵⁴ Stancioli, B. Ribeiro, D. (2016). “Dignity as constraint and as freedom: On the meanings of human dignity for transhuman world”. *New Approaches to the Personhood in Law*, Peter Land Edition, p. 31.

⁵⁵ Pietrzykowski, T. (2016). “*Law, personhood, and the discontents of juridical humanism*”. *New Approaches to the Personhood in Law*, Peter Land Edition, p. 13.

⁵⁶ Stancioli, B. Ribeiro, D. (2016). “Dignity as constraint and as freedom: On the meanings of human dignity for transhuman world”. *New Approaches to the Personhood in Law*, Peter Land Edition, p. 25.

regulation⁵⁷ of artificial intelligence. The first one is the *ontological concept* (also called substantial), which assumes that every human being is a person on the basis of his membership in the group within the human species.⁵⁸ The second concept, which is called *psychological* or non-substantial, assumes that only a being, therefore, only a human or any other being, which has free will and is equipped with full rationality and full self-consciousness, can be described as a person.⁵⁹ According to the former concept, a person is *substantia*, whereas to the latter, rationality is the distinguishing feature of a human.⁶⁰ Some scholars hold the opinion that “*each person in the psychological sense is a person in the ontological sense, but not vice versa, the adherents of the psychological conception will refuse to recognize dignity of human beings who do not actually fulfill the three criteria of personhood (e.g., human beings who are mentally ill, in a vegetative state, with advanced Alzheimer illness etc.)*.”⁶¹ My interpretation of the basis for a legal regulation of artificial intelligence does not support this opinion. I do not equate dignity with personhood. I do not recognize dignity as a logical consequence of the ascription of personhood. In my opinion, the concept of human dignity is a precondition for the ascription of personhood.

Although the meaning of human dignity remains more or less vague I support the idea of Immanuel Kant, who claims that the concept of dignity perceives all distinguishing features that are shared by human beings in this stage of evolution and are not shared with the other beings.⁶² Dignity is therefore possessed by all human beings, regardless their capacities. (If we equated dignity and legal personhood as Zaluski did, people with advanced Alzheimer disease would not possess dignity – this conclusion is wrong). For personhood, on the other hand, which is in my

⁵⁷ Option about the imposition of legal personhood on robots.

⁵⁸ Zaluski, W. (2016). “The concept of person in the light of evolutionary theory and neuroscience”. *New Approaches to the Personhood in Law*, Peter Land Edition, p. 36.

⁵⁹ Ibid.

⁶⁰ Ibid.

⁶¹ Zaluski, W. (2016). “The concept of person in the light of evolutionary theory and neuroscience”. *New Approaches to the Personhood in Law*, Peter Land Edition, p. 37.

⁶² Stancioli, B. Ribeiro, D. (2016). “Dignity as constraint and as freedom: On the meanings of human dignity for transhuman world”. *New Approaches to the Personhood in Law*, Peter Land Edition, p. 30.

opinion a descendant - derives from the concept of dignity itself, certain predefined criteria should be fulfilled. To sum up, my interpretation supports the idea of the psychological concept of personhood that requires the possession or fulfillment of three criteria: free will, full rationality, and full self-consciousness of the possessor.

An agent has *free will* if he, for any time (t-1), had at time (t-1) the ability to make a different decision from the one he (actually) made at a time (t-1).⁶³ “It means that at time t-1 all options considered by the agent were, so to speak, authentic, i.e., the past did not determine the decision that was actually made. Freedom of agent’s capacity by virtue of which his/her decisions are not elements of causal chains but are initiated by the agent in an absolute or quasi absolute manner.”⁶⁴ We can call this capability ‘*the ability to do otherwise*’. This capability led me to the conclusion that it is acceptable for a human being to choose an option which is not the most optimal one (even though he knows the optimal one). Robots on the other hand are equipped with algorithms that constantly strive to find (after taking into account all surrounding circumstances) an optimal solution. Even though the outcome is not always predictable, the optimal solution is a constant. Due to this reason it is difficult to say that robots have the capacity to do otherwise. The other criterion is *full rationality* which is described as the quality of being able to think logically or sensibly. Moreover, with this criterion we also describe the quality of being endowed with the capacity to reason. The last criterion is full self-consciousness. Despite the fact that scientific results have shown that the overwhelming part of a human mind’s activities are unconscious (automatic) in character.⁶⁵ It means that we automate our mental activities which present the most effective solution since consciously made activities are very energy consuming^{66,67} The

⁶³ Zaluski, W. (2016). “The concept of person in the light of evolutionary theory and neuroscience”. *New Approaches to the Personhood in Law*, Peter Land Edition, p. 38.

⁶⁴ Ibid.

⁶⁵ Ibid, 43.

⁶⁶ Wilson, T. D. (2002). *Strangers to ourselves. Discovering the adaptive unconscious*. The MIT Press: Cambridge, England.

⁶⁷ Zaluski, W. (2016). “The concept of person in the light of evolutionary theory and neuroscience”. *New Approaches to the Personhood in Law*, Peter Land Edition, p. 43.

answer to the question why our brains “choose” to automate the activities of our minds lies in the process of evolution. If we look at mental processes of our animal ancestors we can notice that the overwhelming majority of their decisions and activities were unconscious in character. Since we share with them a large part of the brain it is not surprising that the overwhelming part of the human mind’s activities are unconscious in character as well.⁶⁸ Does a robot possess these three criteria? Is it possible for a machine to think, to act consciously?⁶⁹

Rene Descartes was the first who occupied himself with the question whether it would be possible for a machine to think⁷⁰. The latest scientific attempts are focusing on the creation of a computer model of the brain. As presented in the beginning, the artificial neural network is a computing system central concept of which was borrowed from the analogy of biological neural networks. The goal of attempts to emulate human brain, i.e., copying the function of biological nervous systems to software is to achieve software intelligence. According to Sandberg, successfully emulated brain would probably behave in the same way as real brain.⁷¹ Is the structure of our brain the key that enables the ability to have a will, to have desires, the ability of being self-conscious etc. An excerpt from the Robert A. Heinlein’s book *The Moon is a Harsh Mistress* says

“(...) Am not going to argue whether a machine can “really” be self-aware. Is a virus self-aware? Nyet. How about oyster? I doubt it. A cat? Almost certainly. A human? Don’t know about you, tovarishch, but I am. Somewhere along the evolutionary chain from macromolecule to human brain awareness crept in. Psychologists assert it hap-

⁶⁸ Ibid, 43.

⁶⁹ Solum, L. (2008). “Legal personhood for artificial intelligences”. *Illinois Public Law and Legal Theory Research Papers Series No. 09-13.*, p. 1234.

⁷⁰ “For we can easily understand a machine’s being constituted so that it can utter words, and even emit some responses to action on it of a corporeal kind, which brings about a change in its organs; for instance, if touched in a particular part it may ask what we wish to say to it; if in another part it may exclaim that it is being hurt, and so on. But it never happens that it arranges its speech in various ways, in order to reply appropriately to everything that may be said in its presence, as even the lowest type of man can do.”

⁷¹ Sandberg, A. (2013). »Ethics of brain emulations (draft)«. Future of Humanity Institute, Oxford University, p. 8.

*pens automatically whenever a brain acquires certain very high number of associational paths. Can't see it matters whether paths are protein or platinum.(...)*⁷²

If this is the case, that all these abilities derive from the structure of a human brain, does it really matter whether paths are protein or platinum? Ethical concerns regarding the recognition of robots (who will due to the emulation presumably feel pain, have emotions etc.⁷³) that are equipped with such 'software' as a human being, exceed the scope of this term-paper. However, if science had no restraint and no legal impediment with regard to creating a "human being" that is not made of flesh and blood, it would be immoral to discriminate someone who is just like us on the basis of their structure. Pursuant to my opinion and argumentation presented above, I do not support the idea of the ascription of legal personhood to artificial intelligence, however, in the case this kind of robots are created, I want to keep this possibility open.

2.3. Is artificial intelligence really that similar to legal entities that are recognized as legal persons?

It does not matter how they are created and structured—legal systems and laws as such are incapable of self-realization. Their *spiritus agens* (motive force) is a legal person, or the holder of not only legal actions, but also legal changes and legal consequences.⁷⁴ In the world of law, an individual can be described as a "legal actor" to whom a "legal system" acknowledges a "legal role." The "legal role" we play is composed of the rules and obligations that enhance us with the ability to enter legal relationships with the main purpose of achieving acceptable legal aims. At the end of the day, a legal actor is always a human being who plays the acknowledged legal role either as a physical person or as a member of legal entity, a non-human entity that is treated as a person for certain legal purposes. Human beings as the final decision maker must be aware of the consequences of his or her actions. Decisions with all of the attached legal consequences return to

⁷² Heinlein, R. (1966). *The moon is a harsh mistress*. G. P. Putnam's Sons.

⁷³ Sandberg, A. (2013). »Ethics of brain emulations (draft)«. Future of Humanity Institute, Oxford University, p. 8.

⁷⁴ Marijan Pavčnik, *Teorija prava* (GV Založba, 2011), 136.

him or her afterward, with all of the attached rights and obligations —*Hominum causa omne ius constitutum*.⁷⁵ The main reason for the assignment of legal personhood is a value-based decision to assign certain rights and obligations to the legal actor.⁷⁶ The assignment of such quality acknowledges legal actors as holders of rights and obligations, capable of making decisions and at the same time being aware of legal consequences. However, in order to become a legal actor, there needs to be not only a human, but also his or her personal decision behind “the mask.” Meaning, the legislatures would never be able to assign legal personhood to legal entity which in this case serves as “the mask,” if there were not real people with ambitions and visions behind it. The same can be claimed for robots. The attachment of legal personhood on robots⁷⁷ (for robots equipped with Artificial neural network I leave this possibility open) would undermine the idea of legal personhood. Even though we will get “the mask”/robot, there will not be a person standing behind robot’s actions (like it is the case for legal entities). Laws of the world traditionally distinguish between two kinds of persons. Natural or physical persons are placed in the first group while so-called organizational entities such as associations, corporations, municipalities and states fall in the second group.⁷⁸ Organizational entities that constitute the second group were organized and created to fulfill human goals.⁷⁹ The recognition of a legal status to a legal entity does not automatically present its legal personification. It presents an acknowledgement of legal capability that allows participation in legal processes. A legal entity cannot be treated like a “being” that has besides its own intellect also an ability to create and express its will (as it can be done by a human being). The classic discussion concerning the idea of legal personhood can be

⁷⁵ *Ibid.*, 137.

⁷⁶ *Ibid.*

⁷⁷ Robots equipped with algorithms that give us a sense of unpredictability during their operation (supervised learning algorithms, unsupervised learning algorithms).

⁷⁸ Pietrzykowski, T. (2016). “*Law, personhood, and the discontents of juridical humanism*”. *New Approaches to the Personhood in Law*, Peter Land Edition, p. 13.

⁷⁹ Stancioli, B. Ribeiro, D. (2016). “Dignity as constraint and as freedom: On the meanings of human dignity for transhuman world”. *New Approaches to the Personhood in Law*, Peter Land Edition, p. 25.

found in *The Nature and Sources of the Law*⁸⁰, written by John Chipman Gray, where he criticized the idea that inanimate things might be recognized as a legal person. He wondered what the purpose is behind making a thing the subject of legal rights and duties, if it can neither understand the law and legal rules nor can act on it. Moreover, he claimed that at the end of the day even legal entities (corporations) are reducible to the legal relations between individuals, who own stocks.⁸¹ Legal systems of the world ascribe legal entities particular human qualities in order to enable their participation in legal processes as subjects of legal rights and obligations. In legal relationships we are using some kind of analogy that equates a legal entity with a human being.⁸² Every legal entity will be held accountable for its duties and obligations with the assets owned by the enterprise and not with the personal assets of the owner. This fundamental principle can be legally defined as the principle of *total liability*. Law constitutes besides the system of total liability also the liability of members of the enterprise. The latter is based either on a legal code (like separate legal personality) or a social contract.⁸³

The principle of “separate legal personality” was established in the case of *Salomon v A. Salomon & Co. Ltd.* In the mentioned case a shoe manufacturer, Mr. Salomon sold his business to a limited liability company, in which he and his family were the shareholders and directors.⁸⁴ Mr. Salomon owned 20,001 of the total of 20,007 company shares. The other remaining shares were distributed equally among his wife and children. The company ran into financial difficulties and sought a loan of 5,000£ from one Mr. Edmund Broderip who granted the loan. Subsequently the company was confronted with even bigger financial difficulties and was unable to pay the debt. As a result, the action for liquidation was carried out. Mr. Edmund Broderip instituted an action

⁸⁰ Gray, J. C. (1909). *The Nature and Sources of the Law*. New York, The Columbia University Press.

⁸¹ Solum, L. (2008). “Legal personhood for artificial intelligences”. *Illinois Public Law and Legal Theory Research Papers Series No. 09-13.*, p. 1239.

⁸² Ivanjko, Š. Kocbek, M. (2003). *Korporacijsko pravo: Pravni položaj gospodarskih subjektov*. GV Založba, p. 97.

⁸³ Ivanjko, Š. Kocbek, M. (2003). *Korporacijsko pravo: Pravni položaj gospodarskih subjektov*. GV Založba, p. 139.

⁸⁴ Teacher, Law. (November 2013). *Separate Legal Personality*. Retrieved from <https://www.lawteacher.net/free-law-essays/company-law/separate-legal-personality.php?cref=1?cref=1>

with which he was claiming Mr. Salomon's personal liability. The High Court and the Court of Appeal held Mr. Salomon liable.⁸⁵ After the appeal the House of Lords reversed the ruling, and unanimously held that since the company was duly incorporated it now presents an independent person with its rights and obligations and that *"the motives of those who took part in the promotion of the company are absolutely irrelevant in discussing what those rights and liabilities are"*⁸⁶.

We may conclude that the legal fiction of "corporate veil" between the company and its owner was established in the Salomon case. This legal fiction states that a company has a legal personality which is independent and separate from the identity of its members. Hence, the incorporation of a limited liability company restricts the liability of its members, where the latter will not be held personally liable if an act is done in the name of a company.⁸⁷ The primary purpose to create this fiction was to enable individuals to pursue an economic purpose as a single unit, without exposing themselves (and indirectly their families) to risks or liabilities in one's personal capacity. Consequently, a company can be an owner of a property, it can execute contracts, raise debt, can assume other rights and obligations which are independent of its members and can sue and be sued in its own name. The problem will arise where members of the legal enterprise have taken advantage of the separate legal status, when they were acting fraudulently or in a manner which seems unjust. In this kind of situation *"the court acts with caution and depending on the fact and surrounding circumstances the law may go after the individual who has acted dishonestly to hold him liable."*⁸⁸

We may conclude that legal entity is only a manifestation of its member's will. If their actions are legally acceptable, a legal entity will serve its purpose to enable safer participation in busi-

⁸⁵ Ibid.

⁸⁶ Teacher, Law. *Case Summery: Salomon V A Salomon And Co Ltd [1897] AC 22*. Retrieved from <https://www.lawteacher.net/cases/salomon-v-salomon.php>.

⁸⁷ Teacher, Law. (November 2013). *Separate Legal Personality*. Retrieved from <https://www.lawteacher.net/free-law-essays/company-law/separate-legal-personality.php?cref=1?cref=1>.

⁸⁸ Ibid.

ness. However, if actions of its members will be unjust, the corporate veil will be lifted and individuals who manifested their will through the corporation will be held liable. Robots equipped with algorithms that act unpredictably⁸⁹ cannot be described as a manifestation of actions ‘taken’ (ordered) by a human in order to pursue a goal.

3. ARTIFICIAL INTELLIGENCE AS A ‘THING’

3.1. LIABILITY OF THE PRODUCER

Someone might claim that in this case the burden of liability should be attached to the producer instead of experimenting with the attachment of legal personhood. For example, Council Directive 85/374/EEC (Directive)⁹⁰ sets the rules concerning producer’s liability for a defect product. According to the Directive, term “product” presents a material thing that is run by electricity. The product can be described as defective when it does not provide the safety that a person is entitled to expect when he/she takes into account all of the relevant circumstances, such as presentation of the product, its use and the time when the product was put into circulation.⁹¹ Article 4 of the mentioned Directive requires from an injured person the proof about the damage, the product’s defect and the causal relationship between the two.⁹² Pursuant to Article 3 of the Directive, the word producer means the manufacturer of a finished product, of a raw material, of a component part and also a person who attaches his name, trademark or some other distinguishing feature on the product, as a “label,” presenting him as its producer.⁹³ In a situation when a producer is unidentifiable, any supplier of the product shall be treated as product’s producer. However, neither the producer nor the supplier (in a case of un-identifiability of a product) can be held liable

⁸⁹ And independently.

⁹⁰ European Parliament, *Council Directive 85/374/EEC of 25 July 1985* (on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products).

⁹¹ *Ibid.*, Article 6.

⁹² *Ibid.*, Article 4.

⁹³ *Ibid.*, Article 3.

for the damage/injury caused by the defect product if he proves (among the exculpatory reasons, listed in the Article 7 of the Directive) that the product's defect that led to the damage did not exist at the time when it was put into circulation. When applied to the case of robots, since we are talking about machines equipped with such algorithms that can, through observation of the surrounding areas, learn from their own experiences and create their own patterns about how they will conclude a particular task. Here, the producer can reasonably claim that the defect which caused the harm did not exist at the time when the machine was put into circulation. However, this legal framework would be sufficient for machines equipped with deterministic algorithm⁹⁴.

3.2. LIABILITY OF THE EMPLOYER

Employers are the ones who organize and coordinate working processes that fall within the scope of their main activity. The organization of working processes requires, besides general coordination of his employees, also the provision of sufficient working space, machines, necessary facilities and other conditions crucial to maintain a good and healthy working environment. The nature of his work gives an employer a great sense of autonomy, since his organization and coordination affects how his employee carries out every particular action he is assigned. When creating a way in which a particular action will be carried out, he bears in mind the risks and possible consequences of different options. The inclusion of advanced types of robots into the working process presents to the employer a sense of uncertainty, since he cannot predict the risks and possible consequences.⁹⁵ On the other hand, at the moment the inclusion of advanced types of robots does not impose legal risks, since there is no legal framework that would predict who the responsible person for the harm caused by robots would be.⁹⁶ When we ask ourselves, who is responsible for a harm occurred during or in relation to an employer's working processes, we should keep in mind that in relation to a third party, an employer will be held accountable for all harmful

⁹⁴ Deterministic algorithm follows predefined paths. It is constructed to prevent the machine from making autonomous decisions, creating new patterns and consequently creating unpredictable outcomes.

⁹⁵ Zapašek, T. (2017). "Artificial intelligence in medicine". *The Asia Pacific Journal of Health Law & Ethics*, Vol. 11.

⁹⁶ Ibid.

commitments by his *personal substrate*. In a situation when harm comes as a consequence of the operation of a machine, the law recognizes an employer as a strictly liable person, since it presumes that the machine was used by an employee, hence by someone who is a part of an employer's *personal substrate*.⁹⁷ To conclude, the law does not regulate emergent (not automated) reactions, performed by the machine, yet it treats the machine as something that needs to be used (as a simple tool).⁹⁸ Consequently, if we are using a machine, we stand behind its action. What is important here is the aspect of predictability—hence if we give the machine a specific command, we know exactly what to expect.

Strict liability, also known as absolute legal responsibility, is the legal responsibility for damages and injuries that does not require guilt as a prerequisite for its imposition. An employer can be relieved of his strict liability for a harm that occurs as a consequence of the operation of the machine (in existing legislation recognized as simple tool), if he proves the actual reason for a harm is an action or an occurrence that fall outside his sphere.⁹⁹ In order to define an action as the action that falls outside the employer's sphere¹⁰⁰, it must cumulatively satisfy two requirements.

First, the action must be *uncontrollable*, meaning that the employer was not able to prevent its occurrence nor deter its negative consequences. Second, the action must be *unpredictable*, in other words an employer was not able to anticipate its occurrence¹⁰¹ Therefore, according to current legislation, the employer will always be relieved of his strict liability for harms that occur as a consequence of the operation of the machine, since he will not be able to anticipate the occur-

⁹⁷ N. Plavšak, M. Juhart, and R. Vrenčur, *Obligacijsko pravo* (GV Založba 2009), 568.

⁹⁸ Zapušek, T. (2017). "Artificial intelligence in medicine". *The Asia Pacific Journal of Health Law & Ethics*, Vol. 11.

⁹⁹ N. Plavšak, M. Juhart, and R. Vrenčur, *Obligacijsko pravo* (GV Založba 2009), 580.

¹⁰⁰ According to current legislation there exist three reasons that fall outside of the employer's sphere (the reason for the harm was) : (1.) natural events, (2.) action committed by harmed party, (3.) actions committed by third party.

Recommendation: Since we are confronting with more and more autonomous machines that are capable of emergent reactions, I would add another reason - (4.) action committed by advanced robots (advanced algorithms - sense of unpredictability)

¹⁰¹ N. Plavšak, M. Juhart, and R. Vrenčur, *Obligacijsko pravo* (GV Založba 2009), 581.

rence of harm (even the best computer scientists do not have an insight into emerging correlations in advanced algorithms) and consequently he will not be able to prevent it.

3.3. LIABILITY OF THE OWNER OF THE PROPERTY

The majority of legal systems still treats animals the same as it treats material things. However, legal rules regarding animals have been adjusted (due to their nature - domesticated animals / dangerous animals) on some points. This paragraph only briefly presents recommendation for regulation of robots, designed for personal use (not as a part of a work process)¹⁰². Even though robots are considered machines it makes no sense to treat them all as dangerous objects. Criterion “*presenting higher risks for the outside world*” (level of unpredictability, possible consequences of malfunction, potential risks to other people (number), etc.) classifies robots, designed for personal use in two different groups, i.e., higher-risk robots and lower-risk robots. This recommendation was inspired by legal rules regulating liability for animals (dangerous animal and domesticated animal)¹⁰³ where,

1. the owner shall be held *strictly liable* for the damage, caused by a *dangerous animal*¹⁰⁴ / *higher risk robots*¹⁰⁵

¹⁰² Such as self-driving cars, cleaning robots, entertaining robots etc.

¹⁰³ N. Plavšak, M. Juhart, and R. Vrenčur, *Obligacijsko pravo* (GV Založba 2009), 589.

¹⁰⁴ Ibid.

¹⁰⁵ Preconditions for the applications of **strict liability rules** are:

1. *Unlawfulness* and the fact that *the cause* for the damage derives from a **higher-risk robot**
2. *Harm* (physical, mental)

2. the owner shall be held *culpably liable* (with reversed burden of proof) for the damage caused by a *domesticated animal*¹⁰⁶ / *lower-risk robots*¹⁰⁷, unless (he/she proves) he provided sufficient protection and supervision (presumption of *guilt/culpa*).

This recommendation definitely requires further explanation, however, for now it seems that existing legal rules would suffice for regulation of robots, designed for personal use. Due to the fact that both robots (with advanced algorithms) and animals have unpredictable behavior we will not need additional adjustments of existing legal rules, since the latter have already taken into account the presence of *'the sense of unpredictability'*.

4. LEGAL RECOMMENDATIONS

4.1. RE-REGULATION OF EMPLOYMENT RELATIONS FOR ROBOTS THAT ARE INCLUDED IN WORKING PROCESSES

In the two previous sections I have evaluated two potential solutions: the first was recommended by European Parliament, where I concluded that an imposition of legal personhood on robots would shake the cornerstones on which this legal fiction was built. The analysis of existing rules regarding product liability, I have observed that they are clearly not sufficient (except for deterministic algorithms¹⁰⁸). In this section I will present the other potential solution, my creation of re-organized employment relationships. With the latter, the existing law will be able to find a

¹⁰⁶ N. Plavšak, M. Juhart, and R. Vrenčur, *Obligacijsko pravo* (GV Založba 2009), 590.

¹⁰⁷ Preconditions for the application of **culpability rules** are:

1. *Unlawfulness* and the fact that *the cause* for the damage derives from a **lower-risk robot**

2. *Harm* (physical, mental)

3. Presumption of *guilt/culpa* (negligent behavior)

Exculpation: if the owner proves, he/she provided sufficient protection and supervision (application of the higher standard)

¹⁰⁸ Concept of predictability, cannot create its own paths.

person responsible for robot's unpredictable actions and therefore effectively protect the harmed party.

I suggest that we split the workforce in two different groups. The first would be the digital workforce and the second will be human workforce.¹⁰⁹

The human workforce is capable of autonomous, yet still unpredictable decision-making, just as it seems to be the case (in the sense of unpredictability) for the machines of new generation.¹¹⁰ This characteristic prevents us from using existing legal rules. The attachment of the word "mechanical" next to the word "workforce" enables *the adjustment of legal rules* (or the permission of discriminatory rules, for instance that robots will not need sick leave, etc.) regarding the human workforce and to use them for the machines of new generation that are equipped with the ability to perform activities and tasks that used to be exclusively attached to the human workforce.¹¹¹ To sum up, the aim is to include the mechanical workforce into employer's *substrate* in order to protect a third party against the situation when the law does not predict an individual responsible for harmful consequences.

Now, I will present how this legal regulation would work in practice.

When an employee/human workforce causes harm to the third party, we should first look at the strict liability rules. Here, certain prerequisites need to be met in order to treat a harmful behavior committed by an employee as a behavior of an employer. The first requirement is that an employee needs to be a part of employer's *personal substrate* and the second one requires that an employee, as a part of the *personal substrate*, causes a harm while working for his employer, or when harm is in some kind of relation to his work.¹¹²

¹⁰⁹ Zapušek, T. (2017). "Artificial intelligence in medicine". *The Asia Pacific Journal of Health Law & Ethics*, Vol. 11.

¹¹⁰ Ibid.

¹¹¹ Zapušek, T. (2017). "Artificial intelligence in medicine". *The Asia Pacific Journal of Health Law & Ethics*, Vol. 11.

¹¹²N. Plavšak, M. Juhart, and R. Vrenčur, *Obligacijsko pravo* (GV Založba 2009), 568.

Secondly we have to check if we can apply culpability rules. When an employee causes harm to a third party (injured party) on purpose, the latter is allowed to require a compensation also from an employee.¹¹³ In such situation, we define tortuous liability by an employer and his employee as a *liability in solido*. The law always protects third party (injured party) and for this reason an injured party may require compensation either from an employee or its employer. If he/she requires it from an employer, the latter also has a legal claim (same amount as he/she paid to an injured party) against his employee.

In a case when the digital workforce causes harm to the third party, we should also first look at the strict liability rules. Here, just like in the previous case, certain prerequisites need to be met in order to treat harmful behavior committed by an employee as the behavior of an employer. A mechanical worker needs to be a part of employer's *personal substrate* and digital workforce, because as a part of the *personal substrate*, the machine causes a harm while working for his employer or when a harm is in some kind of relation to his work.¹¹⁴

When we try to apply culpability rules we observe that a mechanical worker, due to its nature, is incapable of causing harm on purpose. Therefore, in case of digital workforces, employers will always be held strictly liable for digital workforce' harmful (unforeseeable) actions or omissions, since they will not be able to transfer responsibility to digital workforce.¹¹⁵

CONCLUSION

¹¹³ *Ibid.*, 570.

¹¹⁴ Zapušek, T. (2017). "Artificial intelligence in medicine". *The Asia Pacific Journal of Health Law & Ethics*, Vol. 11.

¹¹⁵ Zapušek, T. (2017). "Artificial intelligence in medicine". *The Asia Pacific Journal of Health Law & Ethics*, Vol. 11.

As it has been presented and analyzed in the previous sections it is obvious that law falls behind the technological innovations. I believe, we should not legally regulate ‘whole artificial intelligence’ with the same legal rules, yet we should focus on characteristics possessed by individual algorithms. The latter are namely (according to my interpretation) the main reason for differences in legal regulation. Moreover, rapid decisions regarding the imposition of legal personhood on robots, in order to fill the gap that arose in the past few years, seem imprudent. Personhood is in my opinion a descendant of the concept of dignity. As we could see, dignity as such suffers from an inherent vagueness at its core and common agreement is definitely a desirable goal - to define this concept with substance and precision¹¹⁶. Its clarification would definitely help legislators to make wiser decisions when adjusting existing legal rules / creating new ones¹¹⁷.

¹¹⁶ Glensy, R. D. (2011). “The right to dignity”. *Columbia Human Rights Law Review*, p. 73.

¹¹⁷ According to my opinion this imposition would be reasonable only in the case of robots that are equipped with emulated brains.